

A strategy of Mapping Polish WordNet onto Princeton WordNet⁺

Ewa Rudnicka¹, Marek Maziarz¹, Maciej Piasecki¹, Stanisław Szpakowicz^{2,3}

1) Institute of Informatics, Wrocław University of Technology

2) School of Electrical Engineering and Computer Science, University of Ottawa

3) Institute of Computer Science, Polish Academy of Sciences

{rudnicka,maziarz,piasecki}@pwr.wroc.pl, szpak@eecs.uottawa.ca

ABSTRACT

We present a strategy and the early results of the mapping of *plWordNet* – one of the largest such language resources in existence – onto Princeton WordNet. The fundamental structural premise of *plWordNet* differs from those of most other wordnets: lexical units rather than synsets are the basic building blocks. The addition of new material to *plWordNet* is consistently informed by semantic relations and by various analyses of large corpora. The mapping is difficult because of the subtly distinct structures and because of WordNet's focus on synsets. We have designed a set of inter-lingual semantic relations and an effective mapping procedure. In the course of mapping, we have discovered a range of systematic differences between *plWordNet* and WordNet, and proposed ways of accounting for such differences.

Strategia rzutowania polskiego WordNetu na WordNet princetoński

STRESZCZENIE

Przedstawiamy strategię i wstępne wyniki rzutowania *plWordNetu* (Słownosieci) – jednego z największych takich zasobów językowych na świecie – na WordNet princetoński. Struktura *plWordNetu* różni się zasadniczo od struktury większości innych wordnetów: najmniejszym elementem sieci jest w nim nie synset, tylko jednostka leksykalna. Nowy materiał wprowadza się do *plWordNetu* po konsekwentnym i systematycznym rozpoznaniu relacji semantycznych, wynikającym z wielostronnej analizy dużych korpusów tekstu. Subtelne różnice w strukturze i specjalne miejsce synsetu w WordNecie sprawiają, że rzutowanie jest zadaniem trudnym. Zaprojektowaliśmy zbiór międzyjęzykowych relacji semantycznych i skuteczną procedurę rzutowania. W toku prac nad rzutowaniem wykryliśmy szereg systematycznych różnic między *plWordNetem* i WordNetem, po czym zaproponowaliśmy sposoby opisywania i wyjaśniania takich różnic.

KEYWORDS: wordnet, bilingual wordnet, wordnet-to-wordnet mapping, synset, lexical unit

SŁOWA KLUCZOWE: wordnet, wordnet dwujęzyczny, rzutowanie wordnetów, synset, jednostka leksykalna

⁺ Work financed by the EU, the European Innovative Economy Programme Project POIG.01.01.02-14-013/09

1 Introduction

We present a strategy and the preliminary results of the mapping of Polish WordNet [*plWordNet*] onto Princeton WordNet [PWN] (Fellbaum 1998). There have been many attempts to build such mappings for wordnets, including EuroWordNet [EWN] (Vossen 1998, Vossen 2002), MultiWordNet (Bentivogli, et al. 2000; Bentivogli & Pianta 2000), AsianWordNet (Robkop et al. 2010) and IndoWordNet (Sinha, et al. 2006, Bhattacharyya 2010). Those projects usually took advantage of EWN's transfer-and-merge method, which largely consisted in the translation of most of PWN's structure and content into the target language. In contrast with this, *plWordNet*'s design and construction are independent of EWN or PWN, though inevitably substantially influenced by both. A unique corpus-based method was employed (Maziarz et al. 2012, Piasecki et al. 2009). Synsets in *plWordNet* are merely groups of similarly interconnected lexical units [LUs], and it is the LU that is the basic element of the network. We aim at linking two largely independent lexical systems. An inter-lingual mapping procedure connects *plWordNet* synsets with PWN synsets via an ordered set of inter-lingual semantic relations. Mapping is manual, but it is very strongly supported by automatic prompting and bookkeeping. Nouns are by far the most numerous class in PWN and in *plWordNet*, so we decided to test our procedure by the mapping of *plWordNet* nouns in specific domains: people*, artefacts*, places*, family relationships, food, drinks, time units, illnesses, economic vocabulary*, scientific disciplines and names connected with thinking and communication*. (The domains marked with * have been covered selectively.)

2 The mapping procedure

Our mapping procedure has three steps: recognize the sense of a source language synset *S*, search for candidate target-language synset(s) to link *S* with, and select the target-language synset and the appropriate inter-lingual relation [*I*-relation]. The mapping goes from *plWordNet*, so our source synsets are Polish synsets. The relations are applied in the following order: synonymy, hyponymy, hypernymy, meronymy, holonymy, near-synonymy, inter-register synonymy (Rudnicka et al. 2012). Once the highest possible relation has been established, others are no longer searched for and applied.

The procedure's first step is the proper identification of the source synset's sense. While very few *plWordNet* synsets have glosses, the considerably more frequent comments partly make up for the absence of glosses. Still, *plWordNet* is largely relation-based, so the key (sense) denominator will be the position of the given set of synonymous LUs in the overall wordnet structure. Nevertheless, the *plWordNet* editor begins with reading all LUs in the synset, plus the glosses or comments if there are any. For example, consider the Polish synset {zagranica 1, obczyna 1, obce terytorium 1} (countries abroad, foreign lands, foreign territory):

(**Example 1**) {zagranica 1, obczyna 1, obce terytorium 1} —*I*-holonymy→ {foreign country 1}
 {zagranica 1} —hypo→ {strefa 2} {foreign country 1} —hypo→ {state 4}
 {zagranica 1} —meronymy→ {świat 2}

The editor now considers the wordnet structure: the immediate hypernyms/hyponyms and meronyms and holonyms, if there are any. These are *strefa* and *świat* (zone, world). In case of doubts or difficulties with determining the synset sense, the editor considers the direct and indirect hypernyms (or other relations). Once the sense of the analysed synset has been established ('area located beyond the borders of a given country'), the editor can move to the

next stage: seek the equivalent target synset in PWN. First, automatic prompts are checked if they are present. We re-implemented an automated mapping algorithm described in (Daudé et al. 2003, Daudé et al. 2000). If there is no prompt, the editor’s language intuitions help select among target-language LUs one or two candidates which share the sense of the source-language synset (‘foreign country’). These candidate LUs are located in PWN and their synsets are analysed with respect to their sense and position in the wordnet structure (hypernym *state*). Special attention must be paid to their immediate hypernym(s) and hyponyms (or other relations if there are any), since these are going to be juxtaposed with the equivalent relations of the target synset. The editor must check if there already exist, or are likely to be posited, inter-lingual synonymy links between any of the immediate relations of the source and the target synset. When such links exist or are likely to be established for most of the immediate relations, and the gloss of the target synset also matches the sense of the source synset, the inter-lingual synonymy is granted between the two synsets in question; otherwise, the next candidate is considered.

When the editor has exhausted the list of candidates to test, the previously chosen candidates are checked for their potential of linking via other relations. In Example 1, we could try linking our source synset with {world 4, earth 9, Earth 1, globe 1} and then {terrestrial planet 1}; or with {solar system} via *I*-meronymy, because this synset can be a synonym of {świat 2}, a meronym of our source synset. That is not correct: the source synset {zagranica 1, obczyzna 1, obce terytorium 1} is in the domain of political organization, while the target synset is in the domain of geography, so the link must be dismissed. Next, we check the potential for linking of the candidate target synset {foreign country 1} —hyper→ {state 4} and decide that the source synset can be linked to this target synset via *I*-holonymy.

Since the start of our project in March 2012 we have introduced 28061 *I*-relation instances, see Table 1. The frequency of specific relations almost ideally agrees with the proposed ranking, based on our intuitions, concerning meaning closeness and the identity and inclusion of *denotata* sets. Surprisingly, *I*-hyponymy and *I*-hypernymy account for half of all inter-lingual relations. This suggests that the structures of *plWordNet* and PWN differ non-trivially.

<i>I</i> -synonymy	<i>I</i> -hyponymy	<i>I</i> -hypernymy	<i>I</i> -meronymy	<i>I</i> -holonymy	<i>I</i> -near-synonymy	<i>I</i> -inter-register synonymy
11173	12092	2622	927	332	649	266

Table 1. The number of inter-lingual relation instances

3 Mapping dilemmas and their solutions

In the course of mapping, we have faced dilemmas resulting both from the differences in the conceptual and lexico-grammatical structure of English and Polish, and from different methodological assumptions which underlie the construction of *plWordNet* and PWN.

3.1 Lexico-grammatical differences

The existence of lexical gaps is an obvious problem: concepts either are not lexicalised in one of the languages or do not exist in its extra-linguistic reality and conceptual structure (cultural gaps). An example of the former is the English word *chantry* meaning “a chapel endowed for singing Masses for the soul of the donor” (adopted from PWN’s definition of {chantry 2}). The concept is not lexicalised in Polish, though it exists in its extra-linguistic reality, so {chantry 2} is linked

to its closest Polish equivalent {kaplica wotywna 1} via *I*-near synonymy which signals partial correspondence in meaning and/or structure:

(Example 2) {chantry 2} —hypo→ {chapel 1} —hypo→ {place of worship 1}
{kaplica wotywna 1} —hypo→ {miejsce kultu 1} {place of worship 1} —*I*-hypo→ {miejsce kultu 1}
{chantry 2} ←*I*-near-synonymy→ {kaplica wotywna 1}

Cultural gaps can be the names of occupations or administrative functions never present in the other language's reality, thus not lexicalised. An apt example is {kaowiec 1}, a Polish term denoting an institution's employee responsible for the organization of cultural and recreational events in the Communist times. It is linked to the PWN synset {organizer 1 ...} meaning "a person who brings order and organization to an enterprise" via the *I*-hyponymy relation, which is the most often used relation in such cases:

(Example 3) {kaowiec 1} —hypo→ {pracownik oświaty 1}; {organizator 1}
{organizator 1} —*I*-hyper→ {organizer 1 ...} {kaowiec 1} —*I*-hypo→ {organizer 1 ...}

The last type of lexical gaps is a mismatch resulting from different structuring of information, as in the case of English and Polish family relation hierarchies. Polish lexicalizes the distinction between the brother of one's father (*stryj* or *stryjek*) and one's mother (*wuj* or *wujek*), although the former term is marked and slowly becomes obsolete. Both terms are present in *plWordNet*. The unmarked term {wujek 2} is linked to its English equivalent {uncle 1} via the *I*-synonymy relation, while the marked term {stryj 1} is linked to {uncle 1} via *I*-hyponymy:

(Example 4) {stryj 1} —hypo→ {wujek 2}
{wujek 2} —*I*-near-synonymy→ {uncle 1} {stryj 1} —*I*-hypo→ {uncle 1}

The contrast can be expressed in English using the premodifying adjectives *paternal* and *maternal*, but the phrases *paternal uncle* and *maternal uncle* are not LUs in PWN. It is important to distinguish all these gaps from dictionary-content gaps due to differences in sources or methodology of building the two wordnets. (We repair most dictionary-content gaps in *plWordNet* and catalogue such gaps in PWN for possible future use.) Clearly, our most preferred *I*-synonymy relation cannot be used in either instance. Still, most of these cases can be handled by the *I*-hyponymy/hypernymy relation which we treat as the second option. Occasionally, we resort to *I*-meronymy/holonymy and *I*-near-synonymy.

Another type of dilemma is to do with the divergent degree of gender lexicalisation in English and Polish. Polish feminine nominal forms are frequent, while most of English nouns are not marked for gender, e.g., the English word *cousin* and Polish *kuzyn* 'cousin_{masc}' and *kuzynka* 'cousin_{fem}'. The most natural strategy to adopt here is again to resort to *I*-hyponymy, making the English {cousin 1} the hypernym of both Polish {kuzyn 1} and {kuzynka 1}, which can easily be construed as two sub-types of a more general concept. Interestingly, there are also mixed English synsets consisting of feminine and masculine forms (and sometimes also unmarked forms), as in {bondswoman 1, bondsman 2} or {chairman 1, chairwoman 1, chairperson 1}. *I*-hypernymy links such synsets to the corresponding Polish synsets lexically differentiated for gender.

(Example 5) {bondswoman 1, bondsman 2} —*I*-hyper→ {gwarant 1, poręczyciel 1}; {poręczycielka 2}

Apart from lexically marked gender, Polish has a variety of other marked forms such as diminutives and augmentatives, which either do not appear or are very rare in English. *plWordNet* has a special relation of *markedness* (*nacechowanie* in Polish) to show the links

between base forms and their derivatives. Crucially, it is a relation between LUs, not synsets. It has three variants: *istota młoda* ‘young creature’, *diminutywność* ‘diminutiveness’ and *augmentatywność i ekspresywność* ‘augmentativeness and expressiveness’. Polish LUs which denote young creatures but are not derivative forms, such as *cielę* or *cielak* ‘calf’, *prosię* or *prosiak* ‘piglet’, are linked to {młodzik, młodziak 2} ‘young animal’ via hyponymy. Analogically in PWN, synsets denoting young animals are attached by hyponymy to synsets denoting young sub-kinds of animals, such as {young mammal 1}. Now, PWN often places LUs denoting young animals with diminutive forms, when such forms exist, e.g., {kitten 1, kitty 3}, or {piglet 1, piggy 1, shoat 1, shote 1}. Since LUs denoting young creatures and diminutive LUs are not always in the same synsets, they are linked to PWN synsets via *I*-hyponymy relation, e.g.,

(Example 6) {prosiaczek 1} —dimin→ {prosiak 1}
 {kitten 1, kitty 3}, {piglet 1, piggy 1, shoat 1, shote 1} —hypo→ {young mammal 1}
 {piglet 1, ..} —*I*-hyponymy→ {prosiak 1, prosię 2}; {prosiaczek 1, prosiatko 2}

In the rare cases without direct equivalents, *I*-synonymy will be applied. If an item has no English equivalents, we opt for *I*-hyponymy to link it to its English hypernyms.

3.2 Structural differences

3.2.1 Synonymy and synsets

The different strategy of synset construction and the resulting different idea of intra-lingual synonymy have led to systematic structural discrepancies. To begin with, *plWordNet* systematically distinguishes between count and mass nouns and never places them in the same synset. Conversely, PWN often neutralises this distinction at the synset level, putting both mass and count LUs into one synset (e.g. {furniture 1, piece of furniture 1, article of furniture 1}) (Miller 1998: 36). Such cases may cause problems for mapping, because it is hard to determine which *plWordNet* synset should be linked via *I*-synonymy, if any. *I*-hyponymy could be also applied to link the count nouns and mass noun *plWordNet* synsets to such “mixed” PWN synset:

(Example 7) {mebel 1} —*I*-hypo→ {furniture 1, piece of furniture 1, article of furniture 1}
 {mebel 1} —hypo→ {element wyposażenia 1}, {sprzęt 2} {mebel 1} —meronymy→ {umeblowanie 1}
 {umeblowanie 1} —*I*-hypo→ {furniture 1, piece of furniture 1, article of furniture 1}

There also are PWN synsets with singular and plural forms of the same lemma, e.g., {dumpling, dumplings 1} with singular and plural hyponyms such as {matzo ball 1}, {wonton 1}, {gnocchi 1}. These are also linked via *I*-hyponymy to their corresponding *plWordNet* synsets:

(Example 8) {pierog 1, pieróg 2} —*I*-hypo→ {dumpling 1, dumplings 1}
 {pierogi 1} —*I*-hypo→ {dumpling 1, dumplings 1}
 {matzo ball 1}, {wonton 1}, {gnocchi 1} —hypo→ {dumpling 1, dumplings 1}

The differently defined synonymy affects the definition of hyponymy in *plWordNet* and PWN. In PWN, singular and collective nouns (*pluralia tantum*) may be hyponyms/hypernyms of each other. This is impossible in *plWordNet*: {dumpling 1, dumplings 1} ‘small balls or strips of boiled or steamed dough’ is a hypernym of synsets {gnocchi 1} ‘(Italian) a small dumpling made of potato or flour or semolina that is boiled or baked and is usually served with a sauce or with grated cheese’, {matzo ball, matzoh ball, matzah ball} ‘a Jewish dumpling made of matzo meal; usually served in soup’ and {won ton, wonton} ‘a Chinese dumpling filled with spiced minced

pork; usually served in soup'. A somewhat drastic, though maybe not unmotivated, case of using a broad notion of synonymy in PWN is the synset {monte 1, four-card monte 1, three-card monte 1} 'a gambling card game of Spanish origin; 3 or 4 cards are dealt face up and players bet that one of them will be matched before the others as the cards are dealt from the pack one at a time'. It is obvious that a *four-card monte* is not a synonym of a *three-card monte*, they are just both hyponyms of *monte*. In Poland *monte* is not so popular. There only is a three-card monte – *trzy karty* (literally 'three cards'). The synsets were joined by inter-language hyponymy, since the English equivalent *three-card monte* of the Polish LU is in the PWN synset:

(Example 9) {trzy karty 1} —*I*-hypo→ {monte 1, four-card monte 1, three-card monte 1}

To sum up, we consistently use *I*-hyponymy in all cases of mixed PWN synsets.

3.2.2 Differently defined relations

There is a lot of correspondence between the set of linguistic relations employed by PWN and *plWordNet* and their respective construction, but there are differences. They are reflected in the structure of both wordnets and may have consequences for the mapping. To give an example, PWN uses the conjunction *or* in its definitions, thus allowing for the hypernymy *and/or*, while *plWordNet* restricts its hypernymy to *and*. For example, the PWN synset {musical 1, ...} was given the gloss 'a play or film whose action and dialogue is interspersed with singing and dancing'; it received the following relational description in PWN (two instances of hyponymy):

(Example 10) {musical, musical comedy, musical theater} —hypo→ {movie, film, picture, ...},
 {musical, musical comedy, musical theater} —hypo→ {play 2}.

The word *musical* gained a similar definition in (Dubisz 2004): "a theatre or film spectacle with comedic or melodramatic content, consisting of oral, sung or danced parts". We had to split the concept into theatrical musical and musical film in order to avoid *or*-hyponymy:

(Example 11) {musical 1, komedia muzyczna} —hypo→ {film 1, obraz 6} 'movie, picture',
 {musical 2} —hypo→ {przedstawienie 7} 'play'

Or-hyponymy was banned from *plWordNet* in order to preserve the transitivity of hyponymy. For example, the English synset {musical 1, ...} also contains synonyms *musical comedy* and *musical theatre*. The first is a synonym of *musical* (Merriam-Webster Dictionary Online). The second clearly refers to theatrical musical (Oxford English Dictionary). In fact, the latter LU should be a hyponym, not a synonym, of *musical* in the broader sense. This leads to a paradox: two synonyms of the synset have both hyponymy relations (to a play and to a film), while *music(al) theatre* has only one (to a play). The opposite could be noted in *plWordNet* where the LU *komedia muzyczna* could be found in the meaning *film musical*. It is linked to {film 1, obraz 6} 'movie, picture' with hyponymy and is, of course, a synonym of Polish *musical 1*. It seems that in PWN hyponymy is only partly transitive and in some cases synonymy captures cases of hyponymy. *Musicals* from PWN and *plWordNet* had to be, naturally, linked with *I*-hyponymy:

(Example 12) Pol. {musical 1, komedia muzyczna} —*I*-hypo→ Eng. {musical 1, ...}
 Pol. {musical 2} —*I*-hypo→ Eng. {musical 1, ...}

Relations are not the only source of difficulty. Glosses also pose dilemmas during mapping. A case of *thriller* 'a suspenseful adventure story or play or movie' is somehow similar to *musical*. Here the connective *or* appears twice, surprisingly followed by only one hyponymy:

(Example 17) {kaplica 1} ‘chapel, autonomous building’ —hypo→ {świątynia} ‘temple’
 {kaplica 1} —hyper→ {kaplica przycementarna, ...} ‘cemetery chapel’
 {kaplica 2} ‘chapel, part of another building’ —hypo→ {pomieszczenie 3} ‘room’
 {kaplica 2} —mero:place→ {klasztor 1} ‘monastery’ {kaplica 2} —mero:place→ {kościół 2} ‘church’

The two senses do have different lexical neighbourhoods, so we assume that they should stay separate. PWN shows an alternative way of describing the concept ‘chapel’. Instead of splitting the sense, it was kept intact and linked to a higher hypernym {place of worship, ...}. At a first glance the two approaches appear justified. Unfortunately, the hypernym {place of worship, house of prayer, house of God, house of worship} was itself linked to {building, edifice} and was given too narrow a definition ‘any building where congregations gather for prayer’, although {chapel} has two hyponyms which clearly are not buildings: {lady church} ‘a small chapel in a church; dedicated to the Virgin Mary’ and {side chapel} ‘a small chapel off the side aisle of a church’. Despite this inconsistency we decided to link our {kaplica 1} and {kaplica 2} with *I-hyponymy* with {chapel 1}, assuming that it has both meanings:

(Example 18) {kaplica 1} ‘building’ —I-hypo→ {chapel 1} ‘a place of worship’,
 {kaplica 2} ‘room’ —I-hypo→ {chapel 1} ‘a place of worship’.

3.2.4 Dictionary-content mismatches

Mapping is also made more difficult by *dictionary content gaps*. We have decided that, though we could improve *plWordNet*, we were not supposed to make any changes inside PWN. What is a dictionary gap? Lexical gaps are caused by specificities of the two languages, dictionary gaps are produced by limitations of any dictionary/thesaurus/wordnet size. For example, in PWN names of artists are restricted to only one domain of art even in cases when they apply quite systematically to more than one domain. For example, {impressionist 1} ‘a painter who follows the theories of Impressionism’ has one hypernym relation instance to {painter 1}, although there is a clear evidence that the word could be used also to indicate impressionist musicians (see the entry in (Procter 1978)) or poets (see *impressionism* in (Myers, Wukasz 2003)). Polish *impresjonista* ‘impressionist painter, musician or poet’ is defined using two *and*-hypernyms *artysta* ‘artist’ and *przedstawiciel* ‘exponent (of an artistic trend)’. We cope with the lexical database mismatch between PWN and *plWordNet* simply using *I-hyponymy* between more specific English {impressionist 1} and broader Polish {impresjonista 1}.

Conclusion and perspectives

The system of inter-lingual relations and the mapping procedure proposed in this paper have been shown to work successfully. We have managed to map about 28000 *plWordNet* synsets onto PWN synsets. All edited *plWordNet* synsets have been linked to PWN’s synsets by one of the proposed inter-lingual relations. The manual mapping was enhanced by an automatic prompt system, which turned out to be useful. The created mapping is especially valuable in that we have been linking two completely independently created large-scale wordnets. It enabled a systematic comparison of *plWordNet*’s and PWN’s structure and content, but also *plWordNet*’s verification and correction. We have encountered mapping dilemmas which boil down to lexico-grammatical differences between English and Polish and to structural incompatibilities resulting from different methodologies which underlie the construction of the two wordnets; we have proposed systematic solutions.

References*

- Bentivogli, L., Pianta, E., Pianesi, F. (2000). *Coping with lexical gaps when building aligned multilingual wordnets. Proceedings of LREC 2000*. Athens, Greece. 993-997.
- Bentivogli, L., Pianta, E. (2000). *Looking for lexical gaps. Proceedings of Euralex 2000*, Stuttgart, Germany. multiwordnet.fbk.eu/paper/wordnet-euralex2000.pdf
- Bhattacharyya, P. (2010). *IndoWordNet. Lexical Resources and Evaluation Conference LREC 2010*, Malta. www.cse.iitb.ac.in/~pb/papers/lrec2010-indowordnet.pdf
- Daudé, J., Padró, L., Rigau, G. (2003). *Making Wordnet mappings robust. Proceedings of the 19th Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural*. Madrid, Spain. 47-54.
- Daudé, J., Padró, L., Rigau, G. (2000). *Mapping Wordnets Using Structural Information. Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics ACL'00*. Hong Kong. www.aclweb.org/anthology-new/P/P00/P00-1064.pdf
- UDP = Dubisz, S. (Ed.). (2004). *Uniwersalny słownik języka polskiego [Universal Dictionary of Polish Language]*, electronic version 1.0. Wydawnictwo Naukowe PWN.
- Hamp, B., Feldweg H. (1997). *GermaNet – a Lexical-Semantic Net for German. Proceedings of the ACL Workshop on Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*. Madrid. 9-15.
- Fellbaum, Ch. (Ed.). (1998). *WordNet: An Electronic Lexical Database*. MIT Press: Cambridge, Massachusetts.
- Maziarz, M., Piasecki, M., Szpakowicz, S., Rabeiga-Wiśniewska, J. (2011). *Semantic Relations Among Nouns in Polish WordNet Grounded in Lexicographic and Semantic Tradition. Cognitive Studies, 11*, 161-182.
- Maziarz, M., Piasecki, M., Szpakowicz, S. (2012). *Approaching plWordNet 2.0. Proceedings of the 6th Global Wordnet Conference*, Matsue. 189-196.
- Miller, G.A. (1998). *Nouns in WordNet*. In Fellbaum, Ch. (Ed.). *WordNet. An Electronic Lexical Database*. The MIT Press. 23-46.
- Myers, J, Wukasz D. C. (2003). *Dictionary of Poetic Terms*. University of North Texas Press.
- Piasecki, M., Marcińczuk, M., Musiał, A., Ramocki, R. and Maziarz, M. (2010), *WordnetLoom: a Graph-based Visual Wordnet Development Framework*, Proceedings of International Multiconference on Computer Science and Information Technology - IMCSIT 2010, Wisla, Poland, 18-20 October 2010, 469-476.
- Piasecki, M., Szpakowicz, S., Broda, B. (2009). *A Wordnet from the Ground Up*. Oficyna Wydawnicza Politechniki Wrocławskiej: Wrocław.
- Piotrowski, T., Saloni, Z. (2002). *Słownik angielsko-polski, polsko-angielski*.
- Polish Wikipedia: pl.wikipedia.org

* (Rudnicka et al. 2012) is a much more complete version of this paper.

Polish Wiktionary: pl.wiktionary.org

Procter P., ed. (1978) Longman Dictionary of Contemporary English. www.ldoceonline.com/

Robkop, K., Thoongsup, S., Charoenporn, T., Sornlertlamvanich, V., Isahara, H. (2010). *WNMS: Connecting the Distributed WordNet in the Case of Asian WordNet. The 5th International Conference of the Global WordNet Association (GWC-2010)*. Mumbai, India, 31st Jan. - 4th February.

Rudnicka, E., Maziarz, M., Piasecki, M., Szpakowicz, S. (2012) "Mapping *plWordNet* onto Princeton WordNet". www.nlp.pwr.wroc.pl/pl/slowosiec-20/130/show/publication (Technical Report).

Sinha, M., Reddy, M., Bhattacharyya, P. (2006). *An Approach towards Construction and Application of Multilingual Indo-WordNet. 3rd Global WordNet Conference (GWC 06)*. Jeju Island, Korea. www.cse.iitb.ac.in/~pb/papers/gwc06_IITB_IndoWN.pdf

Vossen, P. (2002). *EuroWordNet General Document*. EuroWordNet Project LE2-4003 & LE4-8328 report. University of Amsterdam.

Vossen Piek (Ed.). (1998). *EuroWordNet: a multilingual database with lexical semantic networks for European Languages*. Kluwer, Dordrecht.