

This document is downloaded from DR-NTU, Nanyang Technological University Library, Singapore.

Title	Mapping wordnets from the perspective of inter-lingual equivalence
Author(s)	Rudnicka, Ewa; Piasecki, Maciej; Piotrowski, Tadeusz; Grabowski, ukasz; Bond, Francis
Citation	Rudnicka, E., Piasecki, M., Piotrowski, T., Grabowski, ., & Bond, F. (2017). Mapping wordnets from the perspective of inter-lingual equivalence. <i>Cognitive Studies</i> , 2017(17), 1373-.
Date	2017
URL	http://hdl.handle.net/10220/44639
Rights	© 2017 The Authors. This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 PL License (http://creativecommons.org/licenses/by/3.0/pl/), which permits redistribution, commercial and noncommercial, provided that the article is properly cited.

Citation: Rudnicka, E., Piasecki, M., Piotrowski, T., Grabowski, L., & Bond, F. (2017). Mapping wordnets from the perspective of inter-lingual equivalence. *Cognitive Studies | Études cognitives*, 2017(17). <https://doi.org/10.11649/cs.1373>

EWA RUDNICKA^{1,A}, MACIEJ PIASECKI^{1,B}, TADEUSZ PIOTROWSKI^{2,C}
ŁUKASZ GRABOWSKI^{3,D}, & FRANCIS BOND^{4,E}

¹ Wrocław University of Technology, Wrocław, Poland

² University of Wrocław, Wrocław, Poland

³ University of Opole, Opole, Poland

⁴ Nanyang Technological University, Singapore

^Aewa.rudnicka@pwr.edu.pl; ^Bmaciej.piasecki@pwr.edu.pl; ^Ctadeusz.piotrowski@uwr.edu.pl

^Dlukasz@uni.opole.pl; ^Ebond@ieee.org

MAPPING WORDNETS FROM THE PERSPECTIVE OF INTER-LINGUAL EQUIVALENCE

Abstract

“Equivalence is a very simple idea. Unfortunately it becomes quite complex in its applications.”
(Pym, 2007, p. 273)

This paper explores inter-lingual equivalence from the perspective of linking two large lexico-semantic databases, namely the Princeton WordNet of English and the plWordnet (*pl. Słowosieć*) of Polish. Wordnets are built as networks of lexico-semantic relations between words and their meanings, and constitute a type of monolingual dictionary cum thesaurus. The development of wordnets for different languages has given rise to many wordnet linking projects (e.g. EuroWordNet, Vossen, 2002). Regardless of a linking method used, these projects require defining rules for establishing equivalence links between wordnet building blocks, known as *synsets* (sets of synonymous lexical units, i.e., lemma-sense pairs). In this paper an analysis is carried out of a set of inter-wordnet relations used in the mapping of the plWordNet onto the Princeton WordNet, and an attempt is made to relate them to equivalence taxonomies described in specialist literature on bilingual lexicography and translation.

Keywords: Princeton Wordnet; Polish wordnet (*Słowosieć*); wordnet mapping; equivalence; translation; bilingual lexicography.

1 Introduction

Equivalence is a concept originating from translation and, it lies at the heart of bilingual lexicography. It defines the relation between a source language item and a target language item when

they convey the same meaning. This study extends this concept into the area of computational linguistics, where the term equivalence is used to describe relations between component parts of different languages' wordnets. Clearly, the three areas share the common goal of creating a link between a source language item and a target language item (possibly of a remarkable strength). However, while lexicography and translation studies have always been intertwined, computational lexicography has so far developed independently. This paper aims to bridge this gap. The methodology and results of mapping the plWordNet (henceforth 'plWN', cf. Maziarz, Piasecki, Rudnicka, & Szpakowicz, 2013) onto the Princeton WordNet ((henceforth 'WN', cf. Fellbaum, 1998a, 1998b) will be analyzed from the perspective of theories of equivalence developed in bilingual lexicography and translation studies.

Wordnets are large-scale, electronic lexico-semantic databases combining the input of traditional monolingual dictionaries, thesauri and networks of lexico-semantic relations. Their main building blocks are *synsets*, sets of synonymous *lexical units*, lemma — sense pairs. Both lexical units and synsets are linked by a dense network of lexico-semantic relations. The original WN for English was built mainly on the basis of dictionary data, yet more and more of its successors exploit corpus data (e.g. plWN). The construction of wordnets for languages other than English has often entailed their immediate linking to WN, since in many cases they consist of translations of WN synsets and their relation structures into one of the national languages (cf. EuroWordNet, Vossen, 1998, 2002, and others). Such translated wordnets were sometimes extended with additional target language data. Wordnets built independently of WN require a special *mapping* procedure.

The task of linking wordnets is closer in spirit to bilingual dictionary construction than to the translation of texts. It occurs at the level of words rather than sentences or text fragments. Nevertheless, one key difference is that bilingual dictionaries link the specific meanings of headwords, while linking wordnets takes place at the level of synonym sets. On the other hand, it is now hard to imagine the process of translation without bilingual lexicography, and linked wordnets also function as resources for translators. Therefore, it is crucial to verify the import of the (recently established) inter-wordnet relations in view of the theories of equivalence developed for the purposes of bilingual lexicography and translation. Hopefully, this study will be the first attempt to bridge the gap between bilingual lexicography and multilingual lexico-semantic databases.

The paper is organised as follows. Section 2 offers a review of equivalence theories in bilingual lexicography and translation studies. Sections 3 and 4 provide background on wordnets and wordnet mapping. In Section 4 plWordnet-WordNet inter-wordnet relations are also analysed from the perspective of the selected typology of equivalence adopted from bilingual lexicography and translation studies.

2 Equivalence in translation and bilingual lexicography

This section will present the main lines of research into the phenomenon of equivalence in translation studies and bilingual lexicography.

2.1 Equivalence in Translation Studies

Interlingual equivalence is the essence of translation. Two texts, the original in one language and a translation in another language, are “similar” for a competent bilingual, i.e., they are perceived as “saying roughly the same”. In other words, they are perceived to be linked by a certain relation, known as equivalence. This is a relation between texts. In principle, any text in any language can be translated into a text in any other language, even if their systems are extremely different. However, proceeding from these assumptions quite a number of scholars have concluded that equivalence between texts must be based on equivalence between language systems; hence, providing solutions

to structural problems in translation has been one of the main areas of research (Baker, 2005, pp. 298–299).

In structural linguistics a text (an object of parole) can only be produced if the text producer knows the underlying language system (langue). Accordingly, equivalence between language systems was thought to be of primary importance and to be essential when producing texts. Most theoretical writing on the principles of translation and most theories of contrastive linguistics are based on this assumption: it is thought that a translator starts from some (implicit) knowledge about system equivalence, which is described explicitly by contrastive linguistics or translation theorists, and adjusts this equivalence to the given context. Hence the notions of shifts, translation procedures, techniques etc.

This attitude has given rise to a multitude of typologies of equivalence, based on the dichotomy between system and text (e.g. Nida, 1964; Catford, 1965; Koller, 1979; Pym, 2010), developed in the field of Translation Studies. To provide only two examples, Catford (1965) distinguishes equivalence between systems, which he calls *formal correspondence* and *textual equivalence*; in a similar vein, Koller (1979) distinguishes between *correspondence*, operating at the level of a language system, langue, and *equivalence*, operating on the textual level (cf. also Pym, 1997; Kenny, 2001, p. 78; Munday, 2001, p. 47). Instead of the term *text*, *discourse* is now often used by cognitivists (e.g. Lewandowska-Tomaszczyk, 2015).

To clarify further, equivalence between language systems usually means similarity of items in the systems of two different languages. Similarity is understood as similar value (a term of de Saussure, cf. Pym, 2007, 2010), which is described in structural linguistics as similar relations in the language system. These relations are intralinguistic, e.g., they exist between items in the system of a language, they do not depend on relations between a linguistic item and the external world. The latter type of relations are labelled extralinguistic. There are two types of relations: syntagmatic, which actually occur in texts (e.g., *rancid butter*, *rancid* can be used only with words denoting fats, cf. Polish *zjełczwały*, which has similar syntagmatic requirements), and paradigmatic, which are potential, and which can be activated by a native speaker to substitute one word for another in a sentence (*rancid* can be substituted by the more general word *bad*, similarly Polish *zjełczwały* by *zepsuty*). *Rancid* and *zjełczwały* are systematic equivalents, because they have similar values in their respective linguistic systems. In contrast, extralinguistic relations fall into two categories: denotation and reference. In structural semantics, intralinguistic relations are primary; Cruse (1986), in his well-known book on lexical semantics, does not mention any extra-linguistic relations; they are even said to determine its denotation (Griffiths, 2006, p. 13; Lyons, 1977 in Ch. 7 discusses in some detail whether it is intra- or extralinguistic relations that are primary). With a knowledge of intralinguistic relations and denotations of words, one can easily identify relevant entities in an extralinguistic reality. This process is called reference.

That the meaning of individual words emerges from the co-text and context of their use is also one of the central claims of pragmatics, which studies how conventional semantic knowledge (i.e. meaning encoded in the lexicon of a language and in lexical patterns up to the sentence level) interacts, in the particular context of language use, with our assumptions about the world and our knowledge of it (Griffiths, 2006, pp. 1–2). According to Griffiths (2006, p. 15), semantic knowledge of meanings guides us on how to interpret those meanings in context, and, from the perspective of the producer, how to impose meaning on strings of texts. However, it is interesting to note that while the notion of similar value (position) is used in system equivalence, there has really been no attempt to identify in a systematic manner the relevant criteria that are necessary to adequately describe equivalence relations. On the other hand, there have been textbooks that have identified numerous techniques to reach equivalence, such as the influential book by Vinay and Darbelnet (1958), in which they suggest seven translation procedures to use when system equivalence fails. Additionally, most translation theorists have gradually lost interest in system relations. Since there is no general equivalence between two languages, it is created anew in each translation task. This is a view expressed as early as 1962 (Haas, 1962/1968). Therefore, in the recent history of translation studies it has come to be believed that the most important thing for the translator to

take into account when working on a translation task is adherence to the norms and conventions of text production. These norms and conventions also include the expectations of the receiver of the translated text (including the expectations of the commissioner of the translation). One can safely say that system equivalence is a concept whose significance is limited to theoretical approaches far removed from practical considerations. This would suggest that translation equivalence seems ill-suited to approaching equivalence relations in wordnets.

2.2 Equivalence in bilingual lexicography

In contrast to translation studies, in bilingual lexicography the notion of equivalence is still used. After all, bilingual dictionaries do not deal with texts but with words (lexical items in general) in two languages in isolation. Recently, lexicographers have attempted to supplement this with as much of the syntagmatic environment of a word as possible. Therefore, it seems that the theoretical approaches to equivalence developed in the field of bilingual lexicography may be more applicable to exploring the mapping between wordnets. For example, Adamska-Sałaciak (2014, p. 7) argues that the choice of equivalents in bilingual dictionaries is contingent on two distinct yet complementary criteria, that is *semantic* (two items with similar meanings — denotation — as attested in identical definitions) and *distributional* (two items used as translation equivalents in a given context). Along similar lines, Piotrowski (2011a, p. 49) claims that equivalents in bilingual dictionaries should be selected based on analysis of a large number of texts produced in both languages, so that in certain contexts, associated with particular situations of language use, a given target-language headword may be used as a substitute for a source-language item. Adamska-Sałaciak (2014, p. 7) claims that the semantic criterion seems to be more important in the L2-L1 section of a bilingual dictionary. The reason for this is that passive dictionary users primarily want to understand the meanings of target-language headwords, so they should be presented with *cognitive* (i.e. conveying general meaning of source language items/words, representing an obvious first choice) or *explanatory* equivalents (glosses, paraphrases or mini-definitions) in the first place. On the other hand, the distributional criterion seems to be more important in the L1-L2 section; this is because dictionary users would like to know which linguistic items to use in a particular situation (Adamska-Sałaciak, 2014); hence, they should be presented with *translational* equivalents (words that may differ in meaning from source language words, yet may still be acceptable in some contexts, e.g. hypernyms, hyponyms, or other related words used in translation) and/or *functional* equivalents. In practice, in certain contexts cognitive equivalents found in dictionaries may be successfully used as translational equivalents, which are, by definition, context-sensitive. The reason for this is that context-free cognitive equivalents may be general enough to suit a whole variety of contexts. It is interesting to note that traditional bilingual lexicographers have realized the significance of offering a translator (or any other user) a whole range of synonymous expressions, and they have included related words as equivalents, thus, offering a lexical field for the user, a range of items that realize paradigmatic relations. Of course, they have been severely limited by the rigid technology of the printed book. It is only quite recently that bilingual dictionaries have started to use specific translation equivalents, strictly related to the syntagmatic context, thus offering a large number of context-sensitive translational equivalents. The difference can be seen in the following example: traditional (Stanisławski, 1983): “**noble** ... 1. szlachetny; zacny; wzniosły; wielkoduszny; szczytny”, pays little attention to the syntagmatic environment, instead giving a string of paradigmatic replacements, while modern (Linde-Usiekniewicz, 2002), “**noble** ... 2. (virtuous) [spirit, sentiment, character, act] szlachetny”, pays great attention to the syntagmatic environment, and offers no paradigmatic replacements.

To summarise the discussion presented so far, one may arrive at some conclusions. Firstly, in translation the words used in context, the text elements, establish the relation of equivalence and this usually differs from the relation between the words in the lexicon (language system). Secondly, words or expressions do not have meaning by themselves; on the contrary, their meaning emerges out of the context of their use. This means that translation equivalence relations between

words or expressions in L1 and L2 (source- and target-language texts) change or vary across textual contexts, a point which was recently made by Trklja (2016). Thirdly, both in translation and in bilingual lexicography, syntagmatic relations are the most important: both translators and lexicographers want to be able to produce actual texts. They are not concerned with the comparison of potential language relations, and therefore it is syntagmatic relations that are taken into account first and foremost, rather than paradigmatic relations.

Now that an outline of equivalence in translation and bilingual lexicography has been presented, the paper will attempt to explore its relevance and implications for the pWNet to PWN mapping. These problems will be discussed in greater detail in Section 4. In Section 3, however, WN and pWNet will be presented in greater detail: their theoretical basis as well as the mapping procedure, with particular emphasis on the inter-wordnet relations that will be studied further in the context of equivalence.

3 Wordnets

This section will describe the concept of wordnet as a representation of a lexico-semantic system, based on the example of the Princeton WordNet model (Fellbaum, 1998b). Next, the different wordnet construction and alignment methods will be discussed, mainly on the basis of EuroWordNet (Vossen, 2002). Finally, the pWNet model (Piasecki, Szpakowicz, & Broda, 2009) will briefly be presented.

3.1 WordNet

The Princeton WordNet (Fellbaum, 1998b) was initially created as a result of a psychological experiment on modelling human lexical memory. One of the aims of this experiment was to investigate the potential acquisition and representation of lexical meanings by children, especially during first language acquisition. Since its development, WN has been used in a number of research projects and in numerous applications, becoming an important semantic language resource for English. It is often treated as a reference set of descriptions of lexical meanings. The fact that it is available free of charge has helped it achieve such widespread use.

WN is organised around lexical meanings rather than words, as is the case with traditional alphabetic dictionaries, or ontologies, as is the case with thesauri such as Roget's. Lexical meanings are called *word senses* in PWN, or, in short, senses. Senses are described with the help of *synsets*, which are sets of words with a specific meaning. More specifically, a synset is intended to be “a set of synonyms that serve as identifying definitions of lexicalised concepts” (Miller, Beckwith, Fellbaum, Gross, & Miller, 1990, 1993) or a set of words which refer to the same *lexicalised concept*. Importantly, lexicalised concepts are not described overtly. Instead, they implicitly underlie synsets. In addition to synsets, the second most significant means of description in WN are the relations linking synsets, called “conceptual-semantic relations” (Fellbaum, 1998a, p. 210). In other words, the most important type of meaning is that of intralinguistic lexical relations.

The criteria for building synsets were always rather underspecified in WN and in most of its descendants. The idea of a synset, by default, refers to the notion of synonymy, or to a lexical field; one has to note the close connection of the WN theory to structural semantics. However, no WN-specific lexical or semantic theory has been developed, especially in an operational form that could guide the decisions of wordnet editors. Miller et al. (1993) rely on Leibniz's perspective on synonymy, often used in structural semantics: when a word in a sentence is substituted by its synonym the truth value of that sentence does not change. According to this strict interpretation, the construction of synsets could result in just several instances. A weaker criterion of meaning preservation in some contexts was therefore used (Miller et al., 1993), but the properties of contexts were not defined. In the EuroWordNet (henceforth EWN, Vossen, 2002, p. 5), synonymy of synset members is defined to be that of word denotation above all, which is difficult to employ in the

case of many word types, e.g. abstract ones. It should be noted, however, that this is a significant change of perspective from that of the original PWN: denotation is abstract lexical meaning, a sort of generalized relation of a linguistic unit to the external world via a concept. Intra-linguistic relations are evidently thought to be less important.

In the practice of WN building, words that can intuitively replace each other in some linguistic contexts (regardless of some explicitly listed differences, e.g. for verbs in a valency frame; Fellbaum, 1998b), or which refer to the same concept in one of their meanings in terms of their denotation, (Vossen, 2002) were grouped into a synset. EWN creators proposed the use of substitution tests to check the appropriateness of exchanging synset members in the same linguistic context (Vossen, 2002, p. 18). However, it is not clear to what extent the substitution tests were actually used in the development of EWN wordnets and WN.

Words that have several meanings are included in more than one synset, one for each meaning. Thus, the number of synsets equals the number of different lexical meanings that can be distinguished in one wordnet. All of the meaning differences that cannot be expressed by a combination of synsets and relations are neglected. In this way, a wordnet represents only an approximation of the totality of lexical meanings found in the lexicons of natural languages. It is entirely possible that a given lexical meaning occurs in the lexicon of a given language, yet is it not represented in wordnets. Also, in the case of mapped wordnets (e.g. plWN and WN), the language-pair specific gaps and mismatches may result from differences in the structures and methodologies used to compile the wordnets, the specificity of the mapping procedure, systemic differences between languages, or cross-cultural differences (Rudnicka, Witkowski, & Grabowski, 2016).

Synsets are linked by relations intended to connect their underlying lexicalised concepts. The relations inherit names from selected lexico-semantic relations and clearly originate from them, for instance, hypernymy/hyponymy or holonymy/meronymy. While relations are the primary means of semantic description in PWN, synsets are also described by glosses (short definition-like textual explanations) and by *use examples*. Obviously, use examples include individual words from a synset, not the whole synset, and should illustrate a particular lexical meaning represented by the given synset.

As there are many lexico-semantic relations that cannot be attributed to synsets as groups of words, like antonymy or derivational relations, several relations linking words in particular senses — henceforth called *lexical units* — have already been introduced in WN. Such relations are called *lexical relations* and they have been significantly expanded in wordnets for other languages.

3.2 EuroWordNet and other wordnets

3.2.1 Wordnet construction methods

Following the success of the Princeton WordNet, similar lexical resources for other languages started to be built, e.g. GermaNet for German (Hamp & Feldweg, 1997), EuroWordNet, (Vossen, 2002), or IndoWordNet for Hindi and related languages (Bhattacharyya, 2010). Their constructors adopted the PWN model with its main assumptions, yet few could afford a construction process parallel to that of WN (that is start from scratch, employ a team of linguists and programmers, and consult a variety of traditional language resources for a given language). Among exceptions to this are GermaNet and plWordNet. The majority of wordnets, however, usually due to limited funds, were built by means of either the *expand* or *extend* method (EuroWordNet, Vossen, 2002). Both methods take as their input WN's synsets and their relation structures and try to find their equivalences in the target languages. This is where the issue of equivalence appears. In the expand method, English synsets are basically translated to the target language (in some cases even fully or semi-automatically). In the extend approach, the translated WN network of synsets and relations is further extended with synsets and relations specific for a given language. Still, one of the consequences of the translation phase is that very often in the wordnet of the target language one can find a large number of empty synsets, i.e. WN synsets for which translators could

not find equivalents. As a result, the empty synsets are present in the target wordnet structure and the relation structure is a copy of the WN one to a very large extent. Thus, the goal of reflecting a particular language’s lexical system can no longer be pursued. Only a few wordnets have been constructed by the *merge* method, e.g. GermaNet (Hamp & Feldweg, 1997) and DanNet (Pedersen et al., 2006). In the merge method, the lexical resources for a given language (mostly electronic monolingual dictionaries) are explored semi-automatically in order to extract language-specific wordnet structure. In the case of plWN, a large wordnet for Polish, the creators went a step further. It was built with the help of a proposed *corpus-based wordnet development process* (Piasecki et al., 2009) in which a large corpus is the primary source of linguistic information. The information is extracted with the help of several language tools and supplemented by other sources, such as dictionaries, encyclopaedias or shared linguistic intuition (Maziarz, Piasecki, Rudnicka, & Szpakowicz, 2013). In this way, the corpus-based method can be considered a sub-type of the merge method, placing a very strong emphasis on linguistic knowledge extracted directly from large language corpora.

Irrespective of the differences, wordnets share several common properties. A wordnet should be a resource focused on lexical meanings, represented by synsets grouping words sharing the same meaning. Synsets, as well as word senses, should be connected by lexico-semantic relations. One synset represents one lexical meaning which is the same (from the given wordnet perspective) for all the words in the given synset. Polysemous words belong to more than one synset. Individual differences between wordnets, especially those built by the merge method, are inevitable due to the significant underspecification of the original WN model.

3.2.2 EuroWordNet equivalence relations

For the purposes of aligning different language wordnets to WN within EWN, a set of *Equivalence Relations* between synsets was defined (cf. Vossen, 2002). This term implies a connection to the notion of equivalence as it appears in the lexicographic and translation literature (see Section 2), yet it can only be deplored that no such explicit reference is made. The main *Equivalence Relations* are defined in the following way (cf. Vossen, 2002, pp. 38–39):

Equivalence Relations in *EuroWordnet*

- EQ_SYNONYM — “only holds if there is a 1-to-1 mapping between synsets”
- Complex Equivalence Relations:
 - “eq_near_synonym when a meaning matches multiple ILI-records simultaneously, when multiple synsets match with the same ILI-record, or when there is some doubt about the precise mapping
 - eq_has_hyperonym when a meaning is more specific than any available ILI-record
 - eq_has_hyponym when a meaning can only be linked to more specific ILI-records

The definition of EQ_SYNONYM entails that for a source language wordnet synset there can only be a maximum of one target language wordnet synset linked by such a relation.

Importantly, a rich network of *Equivalence Relations* postulated in the EuroWordNet project has never been fully implemented in any of its component wordnets. Most of them link their synsets to ILI only via EQ_SYNONYM relation, and when this one cannot be established no other type of link is introduced. Such a state of affairs became evident during the construction of the Open Multilingual Wordnet (Bond & Foster, 2013).

In view of the theories of equivalence presented in Section 2, the question arises whether, and to what extent, EWN *Equivalence Relations* can be related to different perspectives on equivalence described in the lexicographic and translation literature. It is clear that they cannot be analysed from the perspective of translational equivalence, as this can only be attributed to texts, i.e. language expressions used in a particular context, cf Sec. 2. Yet another problem is that wordnet and inter-wordnet relations, such as EWN *Equivalence Relations*, are established on the level of

synsets, which are, in fact, labels for lexicalised concepts in WN. In fact, concepts are ontological rather than linguistic constructs, while the relation of equivalence is defined in linguistics primarily for units of language, be it as small as morphemes or as large as texts (see Section 2). Nonetheless, in the case of two synsets linked by a EQ_SYNONYM relation, it cannot be denied that their component lexical units should be translation equivalents of some kind to each other.

The next question, therefore, is at which level we should look for equivalents between wordnets. So far, all the mappings between wordnets have been carried out at the level of synsets (cf. mapping projects such as GermaNet, EWN, Open Multilingual WordNet, to name just a few), but questions arise as to whether this approach should be reconsidered. However, an attempt may still be made to relate the known perspectives on equivalence from bilingual lexicography to inter-wordnet relation types. For example, one may want to check to what extent inter-wordnet relation types fulfill semantic and distributional equivalence criteria (Adamska-Sałaciak, 2014), cf. Section 2. Also, one may want to see what kind of correspondence can be found between inter-wordnet relations and equivalence criteria. All of the above questions are relevant, and gain additional significance in the case of the plWN to WN mapping which will be described in Section 4.

3.3 plWordNet

plWordNet is strongly based on language data, that is word uses rather than concepts. Therefore, *lexical units*, not synsets, were selected as the basic building blocks of plWN (cf. Derwojedowa, Szpakowicz, Zawisławska, & Piasecki, 2008; Maziarz, Piasecki, & Szpakowicz, 2013). A lexical unit was triply defined as a: ⟨part of speech, lemma, sense number⟩, e.g. ⟨N, *piec* ‘stove (WN: “any heating apparatus”)’, 1), ⟨N, *piec* ‘guitar amplifier’, 4) or ⟨V, *piec* ‘to bake’, 1). Importantly, all lexico-semantic relations are defined for pairs of lexical units, even those that form the basis for WN conceptual relations, like hyponymy or meronymy.

To avoid problems with formulating a theory of synonymy focused on wordnet development, an operational definition of synsets was introduced in plWN. More specifically, a synset includes such lexical units that share a set of selected lexico-semantic relations, called *constitutive relations* and described in greater detail by Maziarz, Piasecki, and Szpakowicz (2013), in which any two lexical units from a synset are linked to the same third lexical unit by a certain constitutive relation. Constitutive relations encompass, for example, hyper/hyponymy, holo/meronymy, type/instance, inter-register synonymy, inhabitant, inchoativity, cause, or presupposition, cf. Maziarz, Piasecki, and Szpakowicz (2012). Most of them follow lexico-semantic relations well-known from linguistic literature (cf. Cruse, 1986, 2006; Matthews, 2007), or are inspired by such relations. In this way, relations linking synsets are notational abbreviations for lexico-semantic relations linking the lexical units included in them. Thus, relations linking synsets are in fact proper lexico-semantic relations. plWN 3.0 includes more than 20 constitutive relations (~90 subtypes) and more than 20 relations linking lexical units (non-constitutive, as they are not shared), e.g. antonymy and semantic relations based on derivational relations. Lexical units clustered in one synset are considered to be synonymous, but synonymy is rather a derived notion than the primary concept in plWN.

In many cases, constitutive relations appear insufficient for separating groups of lexical units. This is why *constitutive features* were introduced as a supplementary device for discovering synsets. In plWN 3.0, these features include stylistic register, aspect, semantic classes of verbs (represented as a hypernymy tree of synsets representing artificial lexical units, i.e. non-lexicalised) and semantic classes of adjectives. The features are referred to in the definitions of constitutive relations and influence the relation structure that results in limitations on the structure of synsets. Thus, in plWN a synset is a kind of a class of equivalence over the topological structure of lexico-semantic relations constrained by constitutive features, cf formal analysis in Maziarz, Piasecki, and Szpakowicz (2013).

Every plWN relation is described by a proper textual definition and a set of carefully defined substitution tests consisting of sentential patterns with slots for two lexical units and requiring a

yes/no answer. As a result, in the construction of plWN efforts were made to use a minimum set of theoretical assumptions to make the construction process maximally transparent and unbiased in relation to any theory of meaning. As mentioned earlier, thanks to the adopted theoretical model, and to the fact that the resulting structure is based on corpus data, plWN is actually derived from language data (that is, from contexts of use attested in large language corpora) rather than from secondary sources, e.g. dictionaries.

Glosses and examples of usage are not necessary in the lexico-semantic representation based on the relational paradigm, but after the initial phase of the project adding them to the wordnet was found to be useful. Resembling editors' comments, glosses were added to a large portion of lexical units (and to certain synsets too) in order to make the identified lexical meaning clearer to users, as well as to explain the rationale behind particular sense delimitation decisions. The examples of usage illustrate the ways lexical units are used in sentential contexts, and in this way they can be potentially useful in various plWN applications in language processing.

During the plWN development, every effort was made to avoid any influence from the PWN structure, e.g. no form of translation was used and wordnet editors were even asked not to consult PWN during their work. The aim was to build plWN as an independent resource so that it can accurately reflect the system of word senses and lexico-semantic relations of the Polish language.

To summarise, the combination of development based on knowledge extracted from a large language corpus, consistent use of substitution tests and the organisation of collaboration among the team of wordnet editors (coordinated and supported by the group working system) enhanced the level of consistency the editors' decisions and the resulting plWN structure.

4 Equivalence in wordnet mapping

This section will describe the process of mapping the plWordNet synsets onto the Princeton WordNet synsets, its main challenges, the strategy employed, and the final results. A set of inter-wordnet plWN-WN relations will be related to EWN *Equivalence Relations* and to equivalence relations as defined in translation and lexicography literature.

4.1 Mapping between plWordNet and Princeton WordNet

The result of the corpus-based plWN construction process is a faithful description of the Polish lexical system, not a connection of plWN to Princeton WordNet. Thus, for the purposes of linking plWN to WN, a special mapping strategy had to be formulated (cf. Rudnicka, Maziarz, Piasecki, & Szpakowicz, 2012). To stay consistent with earlier mapping and linking projects (see Section 3.2), it was decided to perform the mapping on the synset level and to look for pairs of plWN and WN synsets potentially close in the conveyed meaning. The latter was not an easy task for a variety of reasons.

Firstly, the two resources are based on slightly different methodological assumptions (see Sections 3.1 and 3.3). To reiterate, a synset is viewed as a representation of a lexicalised concept in WN, while it is defined as a set of lexical units sharing constitutive lexico-semantic relations and features in plWN. This affects the average synset granularity. In general, WN synsets contain more lexical units than plWN synsets, whose component lexical units have more finely discriminated senses. Secondly, the key denominator of meaning in a wordnet is a network of lexico-semantic relations, both synset and lexical unit ones. Again, plWN and PWN relation sets partly differ. The degree of their overlap varies across parts of speech. There is much more correspondence between noun relation structures than between adjective and adverb relation structures.

For this reason, noun synsets were chosen to be mapped first. A detailed mapping procedure and a set of intra-lingual relations were defined. The procedure is focused on finding synset graph fragments in the two wordnets that have similar relation structures and describe similar meanings. It consists of three basic stages: an analysis of the sense and relation structure of a source synset,

the selection of candidates for a target synset, and the choice of a target synset and an inter-lingual relation the source and target synsets will be linked with.

The proposed set of inter-wordnet relations was inspired by EWN *Equivalence Relations* and by some intra-wordnet relations in plWN. Currently, the set includes synonymy, inter-register synonymy, partial synonymy, hyponymy, hypernymy, meronymy, holonymy, and cross-categorical synonymy. In each case, the relations are considered in a predetermined order of a decision list. The first relation decided as appropriate for the given pair of synsets is selected and the decision is made. The top-most and top-priority relation is inter-wordnet synonymy. Similarly to EWN EQ_SYNONYM, there can be only one such relation established for a given synset. By definition, *inter-wordnet synonymy* should represent a (very) close correspondence in synset relation structures and senses rendered by them. Only if one cannot find a target synset that fulfills such criteria, does one turn to other inter-wordnet relations such as inter-register synonymy, partial synonymy and hyponymy. *Inter-register synonymy* is used in cases of considerable register differences, usually when a stylistically marked plWN synset is linked via an intra-plWN near-synonymy to a stylistically unmarked plWN synset, already holding an inter-wordnet synonymy relation to some WN synset. *Inter-wordnet partial synonymy* is applied in cases of a substantial overlap of meanings and structures of relations between a source synset and usually two target synsets. Thus, it matches the EQ_NEAR_SYNONYM to a certain extent.

In the remaining cases the first relation to consider is *inter-wordnet hyponymy*, established for source synsets whose meanings and relations graphs form a subset of those of a target synset. The *inter-wordnet hypernymy* goes in the opposite direction. Again, these relations are (partly) related to EQ_HAS_HYPONYM and EQ_HAS_HYPERONYM. If none of the above inter-wordnet relations can be applied, *inter-wordnet meronymy* and *holonymy* are considered for noun synsets. In the case of very general inter-wordnet hyponymy links for adjectives and adverbs, supplementary inter-wordnet cross-categorical synonymy links to PWN noun synsets are added.

At this point, one may ask again how EWN and plWN-WN inter-wordnet relations can be related to the equivalence criteria described in the lexicographic and translation literature (see Section 2), and whether all the differences between them can be accounted for in terms of those criteria, and how the order of their application is related to the decreasing extent to which they fulfil the criteria. One may expect that the highest degree of fulfillment of equivalence criteria would appear in the inter-wordnet relations located at the beginning of the decision list. It is also worth analyzing whether the inter-wordnet relations represent all the kinds of translation equivalents or near equivalents that can be identified. Moreover, it is also worth considering whether the relations altogether provide a description that may facilitate the application of wordnets in multilingual language processing, lexicography or translation.

4.2 Inter-wordnet mapping and the level of equivalence

It might be tempting to describe relations between wordnets for different languages, which, one might suppose, show differences in lexical relations between various languages to some degree (the creators of EWN express this hope, cf. Vossen, 2002, pp. 5–6) — in terms of equivalence relations as used in bilingual lexicography and translation studies. After all, the two fields of study have been examining relations, including lexical relations, between two or more languages for some time now. There are, however, significant differences between the three fields. It was mentioned earlier that wordnets were originally based primarily on intralinguistic relations, rather than on denotations, and that plWN is almost exclusively founded on them. This at once restricts the scope of comparison, because in translation there is one aspect of meaning that is usually not taken into account, and that is the one determined by intralinguistic relations. Items in the original and in the translation quite often do not agree in their part-of-speech assignment (this is partly a semantic property) or in their intension or extension. In contrast, in lexicography and in wordnets these relations are relevant. However, if one goes back to earlier studies of translation (see Section 2), there are some approaches which can be used (even though they are not considered

to be feasible at present): the view that translation starts from correspondence between lexical systems in two languages, and that the task of a translator is to find the equivalent item in the other language which corresponds as closely as possible in its intralinguistic relations to the intralinguistic relations of the item in the original (cf. Catford, 1965; Koller, 1979; also in Pym, 1997 and Munday, 2001). This is precisely what is done when one maps one wordnet to another (at least this was done in the WN — pWN mapping): the intralinguistic relations of the two synsets are supposed to match. This shows that wordnet mapping is rank-bound, established at a given level of a language system. In translation studies literature, such an approach is referred to as formal correspondence (Catford, 1965), in which a formal correspondent is

“any TL category which may be said to occupy, as nearly as possible, the «same» place in the economy of the TL as the given SL category occupies in the SL” (Catford, 1965, p. 32, cited in Ivir, 1981, p. 53).

Thus, it may be argued that in the wordnet context, formal correspondence between source and target language synsets exists if the relations between their ranks have approximately the same configuration.

The following section of the paper will analyse the set of pWN-WN inter-wordnet relations, with an eye to their connection to equivalence types described in the lexicographic and translation literature (see Section 2). I-relations will be discussed in a hierarchical order according to which they are applied in the mapping process.

As described in Section 3, wordnet internal relations are modelled on and ‘borrow’ their names from a set of well-known lexico-semantic relations. In linguistics these relations are determined for word meanings. In WN most of them are treated as conceptual relations, linking lexicalised concepts. In contrast, in pWN, all relations are established on the lexical unit level and relations between synsets are only abbreviations for relations shared by synset members (i.e. lexical units). In linguistics all lexico-semantic relations are used to describe meaning relations between words of a single language. Still, hyponymy and hypernymy do appear in a bilingual context in lexicographic literature (cf. Svendsen, 2009), because it is common practice to resort to a target language hypernym in a dictionary gloss, especially if no direct translation equivalent exists in the target language. Again, we are discussing relations between individual words and their meanings, rather than between abstract concepts.

Inter-wordnet relations, called *Equivalence Relations* in EWN, are established between synsets of two wordnets. Such a move was probably motivated by the fact that synsets are considered basic component parts of WN, the lexical units which form them are treated as strongly synonymous, and each synset is additionally attributed with a gloss specifying its meaning. However, as already noted in Section 2 and Section 3, in linguistics equivalence is defined as a relation between units of language, not between concepts, presumably represented by synsets. Introduced as relations between synsets, inter-wordnet relations are partly modelled on wordnet intra-lingual synset relations such as near-synonymy, inter-register synonymy, hyponymy and meronymy, as well as synonymy — a lexical unit relation (cf. Vossen, 2002; Rudnicka, Witkowski, & Kaliński, 2015a). Such a strategy raises a number of questions, already signalled in Section 3. In linguistic tradition and literature, the term *synonymy* only functions in monolingual contexts. In bi- and multilingual contexts, the term equivalence is used to signal substantial meaning correspondence in translation (and potential replaceability in similar contexts). To be even more precise, in lexicographic and translation literature the equivalence relation is determined with respect to different aspects of language, such as form, meaning, and use.

If one understands ‘form’ as encompassing such categories as syntactic level, grammatical class, number, gender, and countability, then one may hypothesise that in many respects inter-wordnet relations express agreement in form between member lexical units of the linked synsets. The great majority of EWN EQ_Relations and of enWN I-relations are relations between the same parts of speech, and in the case of nouns usually of the same number and countability. Yet EWN allows for the use of EQ_NEAR_SYNONYM across different parts of speech when they express similar meanings (cf. Vossen, 2002, pp. 38–40). Similarly, inter-wordnet cross-categorial synonymy was

employed to code meaning relations between pLWN adjective synsets and WN noun synsets (cf. Rudnicka et al., 2015a). However, the terms ‘formal equivalence’ or ‘formal correspondence’ as used by early theoreticians of translation (cf. Catford, 1965; Koller, 1979) do not simply entail the agreement in features of the above mentioned categories (see Section 2). On the contrary, they signal a similar position in a language system, and especially in a kind of abstract lexical system, which brings us back to meaning.

Wordnet is a relational network and there are lexico-semantic relations that define the meanings of its nodes - synsets. Thus, it can be viewed as an approximation of the lexical system of a given language (provided it was constructed by means of the merge method, see Section 3). Consequently, linking two ‘merge’ wordnets involves an attempt at linking the nodes of the lexical systems of two languages. This was done in the case of the pLWN-to-WN mapping (see Section 4). Clearly, the process appears to be putting Catford (1965) and Koller’s (1979) ideas into practice. Inasmuch as those postulates became inapplicable in translation practice, the pLWN-to-WN mapping resulted in the creation of a huge bilingual lexico-semantic database and revealed many interesting facts about the lexico-semantic systems of English and Polish, and about the construction methodologies of the two wordnets (cf. Rudnicka et al., 2016). For the purposes of mapping, a decision list of inter-wordnet relations was defined (see Section 4). In view of the above considerations, it seems justified to analyse them in terms of the ‘formal correspondence’ of Catford (1965) and ‘correspondence’ of Koller’s (1979). An additional argument in favour of this proposal comes from the mapping procedure itself, which focuses on aligning fragments of relations graphs of possibly similar internal structures, which constitute their key sense denominator. Still, in the process of determining the sense of a source Polish synset, the glosses of its member lexical units are also taken into consideration as an additional sense denominator. This directs the attention to a denotational aspect of synset meaning as well. As such, glosses present examples of how lexical units are used in a whole variety of linguistic contexts.

Moreover, to be able to find the corresponding graph fragments between the two wordnets one also needs to go down to the level of lexical units for which inter-lingual English equivalents are sought, e.g. in the available bilingual dictionaries. At this point, both sense and denotation are taken into account. Next, their lemmas are looked up in WN and candidate target synsets are selected. However, the choice of the target synset is again mainly based on its relational structure. Denotations are also taken into account in cases of sense mismatches which require the use of inter-wordnet partial synonymy, inter-register synonymy, hypernymy and hyponymy relations. To determine them one has to refer to denotations.

4.3 Inter-wordnet relations types vs. equivalence types

The following section of the paper will verify the strength and specificity of the pLWN-WN relation links (Rudnicka’s et al., 2012) with respect to the equivalence types of Piotrowski (1994, 2011a) and Adamska-Sałaciak (2014).

The top-priority relation on the pLWN-WN mapping decision list is that of inter-wordnet synonymy used in cases of: “large correspondence in sense and position in the source wordnet structure, combined with many indirect inter-lingual links between the source and target synsets” (Rudnicka et al., 2012).

The term *sense* is understood here as the meaning conveyed by the source and target synset. Therefore, the essence of mapping wordnets lies in linking the most similar, in terms of relation structure, fragments of the two networks, since these are relations, not glosses, which largely determine meaning in a wordnet. Consequently, the most suitable candidates for inter-lingual synonyms are synsets whose direct (and/or indirect) hypo- and hypernyms, and possibly meronyms and holonyms, are already linked by inter-wordnet relations (or the lexicographer sees the potential for such linking). This is why lexical units’ glosses and synsets’ definitions serve merely as additional meaning determinants.

Bearing in mind the necessary preconditions for the introduction of inter-wordnet synonymy, one can hypothesise that among Polish-English synset pairs linked by this relation there may be selected pairs of Polish-English lexical units that will function as cognitive equivalents, that is to say the most natural, ‘first choice’ equivalents (cf. Adamska-Sałaciak, 2014; Piotrowski, 2011a; also Pym, 2007). Such a situation is most likely to happen in the case of single lexical unit synset pairs (their senses are the most finely distinguished). An example is the link between the plWN synset *kaptur* 1 and the WN synset *hood* 8. However, such a possibility cannot be excluded in the case of multiple lexical unit synset pairs either, although cognitive equivalence will probably only exist between some member lexical units of the Polish-English synset pair. The majority of the remaining lexical unit pairs will probably fulfill the requirements for translational equivalence. This means that they will be replaceable in most contexts, although their semantic import will not be identical.

Occasionally, the main contrast in the semantic import between the source and target synsets rests in the stylistic register of their lexical units. In plWN, lexical units strongly marked for a specific register are put into different synsets than their unmarked synonyms. When no target language synonym of the equivalent register can be found, the next resort is an inter-register synonymy relation. It is established for plWN synsets connected via intra-lingual near-synonymy relation to their unmarked counterparts which already have inter-wordnet synonymy links to WN synsets. It links the stylistically marked plWN synset to its unmarked equivalent WN synset. An example is the link between the plWN synset *kapuca* 1 ‘*old fashioned, hood*’ — and the WN synset *hood* 8. In such cases, the unmarked English lexical units can usually function as translational equivalents for the marked Polish lexical units. Often, when no equivalent of the same register exists in the target language, lexical units of the general register are presented as equivalents in a dictionary, and are also used in translation.

A more complicated situation arises when there is a significant overlap in senses and relation structures between two source synsets and one target synset, or the other way round. We use inter-wordnet partial synonymy to link such trios of synsets. An example is the relation between the plWN synsets *warsztat* 1 and *pracownia* 1 and the WN synset *workshop* 1 ‘small workplace where handcrafts or manufacturing are done’. There are potentially many translational equivalents between the lexical units of such trios, because they are very similar in meaning. For the same reason, we cannot exclude the possibility of cognitive equivalent pairs as well.

When differences in senses and relation structures between source and target synsets are too substantial to establish any of the above inter-wordnet synonymy relations, it is necessary to turn to inter-wordnet hyponymy and hypernymy relations. The first step is always to look for a target language hypernym first. Hypernyms are commonly used in dictionary definitions, so they can be said to fulfill the role of explanatory equivalents. They also very often function as translational equivalents, especially in cases of lexical gaps between languages. In cases when a hypernym synset is located much higher in the wordnet hierarchical relation structure than the target language hyponym synset(s), the source synset is linked to the latter. An example is the I-hypernymy relation between the plWN synset *palec* 1 ‘a digit of a hand or foot’ and the PWN synsets *finger* 1 and *toe* 1.

In the rare cases when no close target hypernym or hyponym synset can be found and there existed a closer holonym or meronym synset on the target language side, a link was created with the latter. An example is the I-meronymy link between the plWN *anima* 1 ‘anima’ — and the PWN synset *subconscious mind* 1. Inter-wordnet holonymy and meronymy links can be argued to pertain to the kind of information which is often part of an explanatory equivalent. Obviously, in this case the equivalence relation exists on the extra-lexical, rather than on the lexical level, contrary to the previously discussed cases.

The last type of inter-wordnet relation is cross-categorial synonymy. It is introduced between plWN adjective synsets and PWN noun synsets as a kind of supplementary relation in cases when only very distant inter-wordnet hyponymy links could be established for the former. Similarly to inter-wordnet meronymy and holonymy relations, the semantic import of this relation can be

treated as an element of an explanatory equivalent. An example is the pIWN synset *warszawski* 3 ‘Varsovian, Warsaw-based’ — linked by the I-hyponymy to *Polish* 1 and by I-cross-categorical synonymy to *Warsaw* 1.

5 Conclusions

In this paper, we have analysed the import of inter-wordnet relations linking pIWordNet and Princeton WordNet synsets from the perspective of equivalence relation, as defined in bilingual lexicography and translation studies. The conclusions are as follows.

Equivalence relation can be determined for different language levels and categories. The key aspects taken into consideration are form, meaning, and the use of language items for which it is established. In wordnet mapping, the key denominator for establishing equivalence relations between synsets are their relation structures, both intra-lingual and inter-wordnet. Such an approach resembles the concept of formal correspondence developed by translation theoreticians in the 1960s and 70s (Catford, 1965; Koller, 1979). Although its applicability in actual translation is largely superseded by textual equivalence, it allows the similarities and differences between two languages’ lexical systems to be captured. This was also demonstrated by generalizations drawn on the basis of the results of the pIWN-WN mapping (cf. Rudnicka et al., 2012, 2016).

However, by its very nature, wordnet linking is much closer in spirit to bilingual dictionary construction than it is to text translation. Therefore, we have decided to compare the import of inter-wordnet relations with the equivalence types specified for equivalents in bilingual dictionaries by Piotrowski (2011a) and Adamska-Sałaciak (2014). The latter include cognitive, translational, explanatory and functional equivalents. Of these four types, the last one, functional equivalent, has been found non-applicable in wordnet mapping, mainly because it goes beyond the lexical level at which the mapping takes place. Interestingly, to determine the connection between equivalence types and inter-wordnet relations, it was necessary to move from the synset to the lexical unit level. Equivalence relations in a bilingual dictionary hold between specific word senses, not lexicalised concepts, as in a wordnet. This raises the question of the necessity of rethinking the strategy of inter-wordnet mapping. If it is to become of greater use and help to translators, the mapping should be performed at the lexical unit level, in order to achieve greater equivalence accuracy.

Returning to equivalence types, it would appear that cognitive, translational and explanatory equivalents can be tracked within inter-wordnet relations. Cognitive equivalents can be found within inter-wordnet synonymy and partial synonymy, because both assume close sense correspondence. The role of explanatory equivalents is often fulfilled by inter-wordnet hypernymy and hyponymy, supplemented by cross-categorical synonymy, meronymy and holonymy. Hypernyms are usually the main part of dictionary definitions, hyponyms appear as illustrations, while meronyms and holonyms appear as differentiating information. A similar role is fulfilled by cross-categorical synonyms. Translational equivalents appear in cases of synonymy, partial synonymy, inter-register synonymy, hypernymy and hyponymy. The first three assume close sense correspondence, while the remaining two are often used when the former are missing, e.g due to differences in lexicalisation between languages.

Nevertheless, to transform inter-wordnet relations into specific equivalence links, the mapping would have to be done at the lexical unit level. This is a challenge that is certainly worth taking on in future research.

References

- Adamska-Sałaciak, A. (2014). Bilingual lexicography: Translation dictionaries. In P. Hanks & G.-M. de Schryver (Eds.), *International Handbook of Modern Lexis and Lexicography* (pp. 1–11). Berlin: Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-45369-4_6-1

- Atkins, B., & Rundell, M. (2008). *The Oxford guide to practical lexicography*. Oxford: Oxford University Press.
- Baker, M. (Ed.). (2000). *Routledge encyclopedia of translation studies*. London: Routledge.
- Baker, M. (2005). Linguistic models and methods in the study of translation. In H. Kittel, A. Frank, N. Greiner, T. Hermans, W. Koller, J. Lambert, & F. Paul (Eds.), *Übersetzung – Translation – Traduction* (pp. 285–294). Berlin: Walter De Gruyter.
- Bhattacharyya, P. (2010). IndoWordNet. In *Lexical Resources Engineering Conference 2010 (LREC 2010), La Valetta, Malta*. Retrieved 10 June 2017, from www.cse.iitb.ac.in/~pb/papers/lrec2010-indowordnet.pdf
- Bond, F., & Foster, R. (2013). Linking and extending an open multilingual wordnet. In *51st Annual Meeting of the Association for Computational Linguistics: ACL-2013, Sofia* (pp. 1352–1362). Association for Computational Linguistics.
- Catford, J. (1965). *A linguistic theory of translation: An essay in applied linguistics*. Oxford: Oxford University Press.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.
- Cruse, D. A. (2006). *A glossary of semantics and pragmatics*. Edinburgh: Edinburgh University Press.
- Derwojedowa, M., Szpakowicz, S., Zawisławska, M., & Piasecki, M. (2008). Lexical units as the centrepiece of a wordnet. In M. Kłopotek, A. Przepiórkowski, S. Wierchoń, & K. Trojanowski (Eds.), *Proceedings of the 16th International Conference Intelligent Information Systems* (pp. 351–358). Retrieved 10 June 2017, from <http://iis.ipipan.waw.pl/2008/proceedings/iis08-34.pdf>
- Fellbaum, C. (1998a). A semantic network of English: The mother of all WordNets. *Computers and the Humanities*, 32(2), 209–220. <https://doi.org/10.1023/A:1001181927857>
- Fellbaum, C. (Ed.). (1998b). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.
- Griffiths, P. (2006). *An introduction to English semantics and pragmatics*. Edinburgh: Edinburgh University Press.
- Haas, W. (1968). The theory of translation. In G. Parkinson (Ed.), *The theory of meaning* (pp. 86–108). Oxford: Oxford University Press. (Reprinted from *Philosophy*, 1962, 37(141), 208–228). <https://doi.org/10.1017/s0031819100061957>
- Hamp, B., & Feldweg, H. (1997). GermaNet — a lexical semantic net for German. In *Proceedings of ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*. Madrid.
- Ivir, V. (1981). Formal correspondence vs. translation equivalence revisited. In I. Even-Zohar & G. Toury (Eds.), *Theory of translation and intercultural relations* (pp. 51–59). Tel-Aviv: The Porter Institute for Poetics and Semiotics.
- Kenny, D. (2001). Equivalence. In M. Baker (Ed.), *Routledge encyclopedia of translation studies* (pp. 77–80). London: Routledge.
- Koller, W. (1979). *Einführung in die Übersetzungswissenschaft*. Heidelberg: Quelle & Meyer.
- Lewandowska-Tomaszczyk, B. (2015). Equivalence. In Ł. Bogucki, S. Goźdz-Roszkowski, & P. Stalmaszczyk (Eds.), *Ways to translation* (pp. 11–54). Łódź: Wydawnictwo UŁ.
- Linde-Usiekiewicz, J. (Ed.). (2002). *Wielki słownik angielsko-polski PWN-Oxford*. Warszawa: Wydawnictwo Naukowe PWN.
- Lyons, J. (1977). *Semantics* (Vol. 1). Cambridge: Cambridge University Press.
- Matthews, P. H. (2007). *The concise Oxford dictionary of linguistics*. Oxford: Oxford University Press.
- Maziarz, M., Piasecki, M., & Szpakowicz, S. (2012). Approaching plWordNet 2.0. In *Proceedings of the 6th Global Wordnet Conference, Matsue, Japan* (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. <https://doi.org/10.1007/s10579-012-9209-9>
- Maziarz, M., Piasecki, M., Rudnicka, E., & Szpakowicz, S. (2013). Beyond the transfer-and-merge WordNet construction: plWordNet and a comparison with WordNet. In *Proceedings of RANLP, Hissar, Bulgaria*. Retrieved 10 June 2017, from <http://aclweb.org/anthology/R/R13/R13-1058.pdf>
- Miller, G., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. (1990). Introduction to WordNet: An online lexical database. *International Journal of Lexicography*, 3(4), 235–244. <https://doi.org/10.1093/ijl/3.4.235>

- Miller, G., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. (1993). Introduction to WordNet: An online lexical database (Revised August 1993). Retrieved 10 June 2017, from <http://wordnetcode.princeton.edu/5papers.pdf>
- Munday, J. (2001). *Introducing translation studies* (2nd ed.). London: Routledge.
- Nida, E. (1964). *Toward a science of translating*. Leiden: E. J. Brill.
- Pala, K., & Smrż, P. (2004). Building Czech WordNet. *Romanian Journal of Information Science and Technology*, 7(1–2), 79–88.
- Pedersen, B., Nimb, S., Rasmussen, N., Sørensen, N., Trap-Jensen, L., & Lorentzen, H. (2006). DanNet — A WordNet for Danish. In *Proceedings from Third International Conference on Global Wordnets* (pp. 329–331). Jeju, South Korea.
- Piasecki, M., Szpakowicz, S., & Broda, B. (2009). *A Wordnet from the ground up*. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej.
- Piotrowski, T. (1994). *Problems in bilingual lexicography*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Piotrowski, T. (2011a). Ekwiwalencja w słownikach dwujęzycznych. In W. Chlebda (Ed.), *Na tropach translatów: W poszukiwaniu odpowiedników przekładowych* (pp. 45–70). Opole: Wydawnictwo Uniwersytetu Opolskiego.
- Piotrowski, T. (2011b). Tertium comparationis w przekładoznawstwie. In P. Stalmaszczyk (Ed.), *Metodologie językoznawstwa: Od ontologii do pragmatyki* (pp. 175–192). Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Pym, A. (1997). Koller's Äquivalenz revisited. *The Translator*, 3(1), 71–79. <https://doi.org/10.1080/13556509.1997.10798989>
- Pym, A. (2007). Natural and directional equivalence in theories of translation. *Target*, 19(2), 271–294.
- Pym, A. (2010). *Exploring translation theories*. London: Routledge.
- Rudnicka, E., Maziarz, M., Piasecki, M., & Szpakowicz, S. (2012). A strategy of mapping Polish WordNet onto Princeton WordNet. In *Proceedings of COLING 2012*. Retrieved 10 June 2017, from www.aclweb.org/anthology/C12-2101
- Rudnicka, E., Witkowski, W., & Kaliński, M. (2015a). A semi-automatic adjective mapping between plWordNet and Princeton WordNet. In P. Král & V. Matoušek (Eds.), *Text, speech, and dialogue* (pp. 360–368). Cham: Springer. (*Lecture Notes in Computer Science*, 9302). https://doi.org/10.1007/978-3-319-24033-6_41
- Rudnicka, E., Witkowski, W., & Kaliński, M. (2015b). Towards the methodology for extending Princeton WordNet. *Cognitive Studies / Études cognitives*, 2015(15), 335–351. <https://doi.org/10.11649/cs.2015.023>
- Rudnicka, E., Witkowski, W., & Grabowski, Ł. (2016). Towards a methodology for filtering out gaps and mismatches across wordnets: The case of noun synsets in plWordNet and PrincetonWordNet. In B. Barbu Mititelu, C. Forascu, C. Fellbaum, & P. Vossen (Eds.), *Proceedings of the 8th International Global WordNet Conference 2016, Bucharest, Romania* (pp. 344–351). Retrieved 10 June 2017, from <http://gwc2016.racai.ro/proceedings.pdf>
- Stanisławski, J. (Ed.). (1983). *Wielki słownik angielsko-polski*. Warszawa: Państwowe Wydawnictwo Wiedza Powszechna.
- Svensen, B. (2009). *A handbook of lexicography: The theory and practice of dictionary-making*. Cambridge: Cambridge University Press.
- Trklja, A. (2016). Corresponding lexical domains: A new resource for onomasiological bilingual dictionaries. *International Journal of Lexicography*, ecw019. <https://doi.org/10.1093/ijl/ecw019>
- Vinay, J.-P., & Darbelnet, J. (1958). *Stylistique comparée du français et de l'anglais: Methode de traduction*. Paris: Didier.
- Vinay, J.-P., & Darbelnet, J. (1995). *Comparative stylistics of French and English: A methodology for translation*. (J. C. Sager & M.-J. Hamel, Trans.). Amsterdam: John Benjamins. (Original work published 1958)
- Vossen, P. (Ed.). (1998). *Euro WordNet: A multilingual database with lexical semantic networks*. Dordrecht: Kluwer.
- Vossen, P. (Ed.). (2002). *EuroWordNet general documentation* (Version 3). Retrieved 10 June 2017, from <http://www.vossen.info/docs/2002/EWNGeneral.pdf>

Acknowledgment

This paper is the result of work carried out within the project funded by the National Science Centre, Poland, under the grant agreement No UMO-2015/18/M/HS2/00100.

The authors declare that they have no competing interests.

The authors' contribution was as follows: concept of the study: Ewa Rudnicka, Maciej Piasecki, Tadeusz Piotrowski, Łukasz Grabowski, Francis Bond; data analyses: Ewa Rudnicka, Maciej Piasecki, Tadeusz Piotrowski, Łukasz Grabowski, Francis Bond; the writing: Ewa Rudnicka, Maciej Piasecki, Tadeusz Piotrowski, Łukasz Grabowski, Francis Bond.

This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 PL License (<http://creativecommons.org/licenses/by/3.0/pl/>), which permits redistribution, commercial and non-commercial, provided that the article is properly cited.

© The Authors 2017

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