Walenty – Extending Valence Dictionary with Semantics

Tomasz Bartosiak and Bartłomiej Nitoń

Institute of Computer Science, Polish Academy of Sciences,
ul. Jana Kazimierza 5, 01-248 Warsaw, Poland

Abstract. The resource described here is a syntactico-semantic valence dictionary of Polish. While syntactic part of *Walenty* is nearly finished, semantic layer has been lately deeply analyzed and its construction is now in progress. First, syntactic layer and phenomena represented in it will be described. Secondly, basic ideas of approach used in *Walenty* will be presented. Thirdly, representation of semantic layer in *Walenty* will be portrayed. Fourthly, possible and implemented improvements of process of dictionary creation will be shown. Last, but not least, possible future use of *Walenty* will be discussed.

1 *Walenty* – syntactic layer

Walenty (`http://zil.ipipan.waw.pl/Walenty`) is a comprehensive valence dictionary of Polish developed at the Institute of Computer Science, Polish Academy of Sciences (ICS PAS)([1, ?]). Initially it contained only syntactic information for verbs, but it has been extended with phraseological component ([3]) and nonverbal entries. At current stage of work it contains 12028 verbs, 2000 nouns, 950 adjectives and 217 adverbs.

Syntactic layer of *Walenty* contains schemata – sets of syntactic positions that are opened by a given word. Those positions are occupied by a set of phrase types. Usual phrase types are considered, such as nominal phrases (`np`), prepositional phrases (`prepnp`), adjectival phrases (`adjp`), clausal phrases (`cp`), etc. Phrase types are further parameterized by corresponding grammatical categories, e.g., `np` and `adjp` are parameterized by information concerning case. For some grammatical categories underscore symbol '_' denotes any of possible values.

There are some special grammatical categories and labels concerning linguistic phenomena occurring in Polish. Most common one is structural case. It is strongly connected with special labels that are additionally assigned to some of arguments – `subj` and `obj` – denoting subject and passivizable direct object correspondingly. For `subj`, structural case corresponds to nominative in active

voice, prepositional phrase `prepnp(przez,acc)` in passive voice and either genitive or `prepnp(przez,acc)` if the verb is nominalized. Otherwise, structural case corresponds to accusative in non-negated active voice, genitive in negated active voice (so-called Genitive of Negation in Polish ([4])), nominative in passive voice (if applicable) and genitive when the verb is nominalized (see Examples 1-5).

(1) *Jan* *aranżuje* *mieszkanie*
John.NOM arrange flat.ACC

'John arranges a flat'

(2) *Jan* *nie* *aranżuje* *mieszkania*
John.NOM not arrange flat.GEN

'John doesn't arrange a flat'

(3) *Mieszkanie* *jest* *aranżowane* *przez* *Jana*
flat.NOM is arrange by John.ACC

'A flat is being arranged by John'

(4) *Jana* *aranżacja* *mieszkania*
John.GEN arrangement flat.GEN

'Johns arrangement of a flat'

(5) *aranżacja* *mieszkania* *przez* *Jana*
arrangement flat.GEN by John.ACC

'arrangement of a flat by John'

Other (less complex) special categories include: partitive case (structural case that additionally can correspond to genitive if noun is uncountable (see Examples 6 and 7)) for nouns, predicative case (either nominative or instrumental (see Examples 8 and 9)) for adjectives and agreeing case (same as noun it is connected to) for both adjectives and nouns.

(6) *Jan* *kupił* *chleb*
John.NOM buy bread.ACC

'John bought (a loaf of) bread'

(7) *Jan* *kupił* *chleba*
John.NOM buy bread.GEN

'John bought (some) bread'

(8) *Jan* *wydawał się* *piękny*
John.NOM appear beautiful.NOM

'John appeared to be beautiful'

(9) *Jan* *wydawał się* *pięknym*
John.NOM appear beautiful.INSTR

'John appeared to be beautiful'

Walenty uses coordination test to distinguish syntactic positions – if two morphosyntactically different phrases can be coordinated, they are considered to be different realizations of single syntactic position.

(10) `lubić: subj{np(str)} + {np(str); cp(kiedy)}`

Lubię *dobry* *alkohol* *i* *kiedy* *patrzysz* *na* *mnie*
like good alcohol and when look at me

'I like a good drink and when you look at me'

In Example 10 we see coordination within a single syntactic position. It consists of a nominal phrase in structural case (`np(str)`; *dobry alkohol* 'a good drink') and a clausal phrase with conjunctive *kiedy* 'when' (`cp(kiedy)`; *kiedy patrzysz na mnie* 'when you look at me'). This schema also contains a position for a subject, which is not explicitly present in the sentence. This is a typical situation for Polish - some (sometimes all) arguments (phrase types) can be omitted in a sentence.

There is also a special type of a phrase – `xp` ([1]). It tends to be adverbial in its nature, but allows multiple different realizations (e.g., adverbial phrase, prepositional phrase) and is parameterized by a type that is rather semantic then grammatical in its nature (e.g., location, duration, path).

Another important linguistic phenomenon considered in *Walenty* is control ([1, **?**]). Some predicates have infinitival phrase as an argument. Two of such verbs can be seen in Examples 11 and 12.

(11)
```
kazać: subj{np(str)} + controller{np(dat)} +
         + controllee{infp(_)}
```

Jan *kazał* *Piotrowi* *czytać*
John order Peter read

'John ordered Peter to read'

(12) `obiecać: subj,controller{np(str)} + {np(dat)} + + controllee{infp(_)}`

Jan obiecał Piotrowi czytać
John promise Peter read

'John promised Peter to read'

Both verbs (KAZAĆ *'order'* and OBIECAĆ *'promise'*) have schemata with 3 syntactic positions – subject in nominative (*Jan*), nominal phrase in dative (*Piotrowi*) and infinitive phrase (*czytać*). In both cases the infinitival phrase has its covert subject controlled (this information is marked with `controllee` label), but the two schemata differ in what controls it (in this example – who is going to reed; marked with `controller`).

In such control constructions the controller is defined independently of the controllee. This is main difference between control verbs (such as those above) and raising verbs, such as ZACZYNAĆ *'begin'*. The subject of a raising verb is simply whatever would count as the subject of its infinitival phrase (or would not take any subject, if infinitive disallows one). Such raised subjects are marked as `E`, as in the following example:

(13) `zaczynać: subj,controller{E} + controllee{infp(_)}`

Zaczyna padać
begin rain

'It begins to rain'

Similar connection can also occur between other schema elements. In Example 14, while there is no surface connection between *Jan* and *analfabeta*, the second one obviously describe the first one. Such connections are also marked with `controller` and `controllee` labels.

(14) `być: subj,controller{np(str)} + controllee{np(inst)}`

Jan jest analfabetą
John be illiterate

'John is illiterate'

2 Semantic valence dictionary – basic concepts

Semantic valence dictionary can be created in various ways. A good example of a concurrent approaches can be seen in English syntactico-semantic valence dictionaries – *FrameNet* (`https://framenet.icsi.berkeley.edu/fndrupal/home`;

[6, ?,?,?]) and *VerbNet* (`http://verbs.colorado.edu/~mpalmer/projects/verbnet.html`; [10, ?,?,13]).

Walenty is being created on the basis of the approach proposed by E. Hajnicz in [14]. Basic aspects of this approach will be discussed in this section.

2.1 Predicate-arguments structure

Predicate-arguments structure constitutes a basis of a formal representation of an utterance. It consists of a predicate (analyzed word, phrase, etc.), obligatory positions (called arguments; sometimes including time and location) and facultative positions. Those positions appear on syntactic level as dependant phrases (such as those appearing in *Walenty*; see Section 1).

In texts such arguments are typically represented by some lexical resource (selectional preference; see Section 2.3) and play a certain role with respect to predicate (semantic role; see Section 2.2).

In this article semantic argument will be defined as a pair: <semantic role, selectional preference>. A set of arguments opened by a predicate will be called a semantic frame (or simply a frame).

2.2 Semantic role

The conception of semantic roles (or thematic roles) has been developed on basis of works of Fillmore ([15–17]). He considered that it is important not only to analyze the number of arguments a predicate can take, but also what semantic content do they provide.

It is usually considered that semantic roles should come from a limited set, but some approaches create role sets for whole classes of verbs (e.g., *FrameNet*). Due to general disagreement about set of possible semantic roles, *Walenty* uses its own set of roles which will be shortly discussed in Section 4.2.

2.3 Selectional preferences

Selectional preference ([18, 19]) is a semantic dependence between a predicate and its arguments, indicating what concepts may appear as a realization of a semantic role. In Examples 15 and 16) polysemic word PRZEJECHAĆ has its meanings distinguished due to selectional preferences.

(15) *Przejechał psa*
run over dog

'He ran over a dog'

(16) *Przejechał dwie przecznice*
pass two block

'He passed two blocks'

Utterances that do not fulfill such limitations (Examples 18 and 17) are considered senseless. This phenomenon is called classification error or selectional violation.

(17) *Przejechał grom*
run over thunder

'He ran over a thunder'

(18) *Przejechał grom*
pass thunder

'He passed a thunder'

There are two main approaches to selectional preferences. The first one splits a possible arguments into two sets – one with all arguments that fulfill selectional preference and one with arguments that do not fulfill them. In this approach selectional preferences are often called selectional restrictions. The other one estimates to what extend an arguments fulfills selectional preferences.

3 *PlWordNet* – Polish wordnet

A resource strictly connected with semantic layer of *Walenty* is the biggest of Polish wordnets. *PlWordNet* (*Słowosieć*; [20, **?**,**?**]) is a lexico-semantic network which reflects the lexical system of the Polish language.

The basic units in *PlWordNet* are meanings (represented by lexical units). This means that there is no single entry for a lexeme (including grammatical category). For example there are 7 entries for word *zamek*, e.g. ZAMEK(NOUN)-1 *'castle'*, ZAMEK(NOUN)-2 *'lock'*, ZAMEK(NOUN)-6 *'zipper'* and 7 entries for word *piec*, e.g. PIEC(VERB)-2 *'to bake'*, PIEC(NOUN)-2 *'furnace'*. Lexical units exist not only for words, but also for multiword expressions (e.g., PIEC KUCHENNY-1 *'stove'*) and idioms (e.g., DRZEĆ KOTY-1 *'squabble'* (literally '*tear cats*')).

Lexical units are connected into sets of synonyms, representing basic concepts, called synsets. Those synsets along with relations of hypernymy and hyponymy create most basic structure of *PlWordNet*.

Hypernymy is a relation connecting a more general concept with a less general one. For example DRZEWO *'tree'* is hypernym of SOSNA *'pine'*. Hyponymy is a relation opposite to hypernymy.

While hypernymy and hyponymy are somehow most important relations in *PlWordNet*, there are many more syntactic and semantic relations connecting

synsets and lexical units ([23, **?**]). Structure of those relations is useful for supervised creating of semantic layer of *Walenty*, but also provides some problems due to its complexity (see Sections 4 and 5).

Current version of *PlWordNet* (`2.3`) contains about 176 000 nouns, verbs and adjectives, 255 000 word senses, and 600 000 relations.

4 *Walenty* – semantic layer

In Section 2 basic elements of a semantic valence dictionary have been described. In this section their representation in *Walenty* will be discussed. More details about semantic layer of *Walenty* can be found in [25].

4.1 Predicate

As the first part of creating semantic valence dictionary predicates for *Walenty* entries have to be defined. To some extent this information is provided by *PlWordNet*. Sadly, lexical units provided by *PlWordNet* are sometimes too granulated. Similarly, *PlWordNet* sometimes misses a word meaning that should be included (for example, because there is a syntactic schema suggesting its existence).

Due to granularity problem, semantic frames are connected to a list of lexical units (with common base form) constituting a concept. This also include verbs with different reflexivity that are describing same situation (diathesis alternations; see Section 4.4).

Missing meanings (as additional lexical units) are added to *Walenty* with additional information about possible location in *PlWordNet* structure.

4.2 Semantic role

It has been decided that in *Walenty* strictly limited number of semantic roles is preferable to creating many predicate-meaning specific roles (like it is done in the *FrameNet*). The semantic roles set in *Walenty* has been modelled mainly on the *VerbNet*.

While in the *VerbNet* roles are connected in the hierarchy of roles ([26]), in *Walenty* semantic roles are split into two levels – *basic roles* and *role attributes*. This solution is flexible, as every basic role may be equipped with any attribute (but only one). On the other hand, the relations between roles are apparent without any external knowledge.

Those roles are connected only to the meaning of a predicate and are independent from actual realization in a sentence. For sentences in Examples 19 and 20, *wino* plays the same semantic role for the predicate *napełniać*, even though its deep syntactic position has changed (in one of those sentences it is subject).

(19) *Wino napełnia kieliszki*
wine fill glass

'Wine is filling glasses'

(20) *Kelner napełnia kieliszki winem*
waiter fill glass wine

'Waiter is filling glasses with wine'

A set of 17 semantic roles and 4 role attributes has been selected.

Basic semantic roles are divided into two groups: main roles (*Initiator*, *Theme*, *Stimulus*, *Experiencer*, *Instrument*, *Factor*, *Recipient*, *Result*) representing situation participants and auxiliary roles (*Condition*, *Attribute*, *Manner*, *Location*, *Path*, *Time*, *Duration*, *Measure*, *Purpose*) representing its circumstances. This set is supposed to cover both required verb dependents (arguments) considered in the dictionary and free modifiers (actual adjuncts). As definitions of roles are quite complex, they will be omitted here. More details can be found in [27].

Roles can be supplemented with attributes organized into pairs. *Foreground* and *Background* are used in symmetric situations, when only sentence focus distinguish objects (e.g., $Theme^{Foreground}$, $Theme^{Background}$ in *'exchanging something for something'*). *Source* and *Goal* are used when some sort of direction can be seen (e.g., $Initiator^{Source}$, $Initiator^{Goal}$ in *'someone buys (something) from someone'*, or $Location^{Source}$, $Location^{Goal}$ in *'to go from somewhere to somewhere'*).

4.3 Selectional preferences

Unlike many other dictionaries (e.g., [28]), *Walenty* does not use a fixed set of qualifiers (like *abstract/concrete, solid/liquid/gaseous*, etc.). *PlWordNet* synsets (represented by lexical units) and relations are used to represent selectional preferences.

Using *PlWordNet* concepts instead of semantic markers has two main reasons. First, it allows to specify constraint with any arbitrary accuracy (e.g., we can write only with certain inanimate objects). Secondly, it allows us to verify whether and to what extent given interpretation (morphosyntactic, syntactic, semantic) of utterance fulfills them.

Therefore, only *dogs* can BARK, we DRINK only *beverages* (not all *liquids*), and only *bandages* BANDAGE (not every *cloth*). The complete frame including selectional preferences for POCHODZIĆ-1 *'come from'* is presented in Example 21. Since *country* is connected neither to *administrative district*, nor to *geographical region*, it has to be considered separately

(21) *Theme* $Location^{Source}$
OSOBA-1 *'person'* {JEDNOSTKA ADMINISTRACYJNA-1 '*administrative district*',
OBSZAR-1 *'region'*,
PAŃSTWO-1 *'country'*}

Due to the complex structure of *PlWordNet* some predefined sets of synsets have been formulated. Synsets from those sets are frequently used as selectional preferences, e.g. PODMIOTY *'subjects'* being set of {OSOBA-1 *'person'*, GRUPA LUDZI-1 *'group of people'*, PODMIOT-3 *'institutions, organisations and firms'*}.

For similar reasons a *PlWordNet* relations to another synset have been introduced as a way of representing selectional preferences. For instance, an *Instrument* for PISAĆ *'write'* could be *a pen, a ballpen, a pencil*, etc. However, in *PlWordNet* their direct hypernym is ARTYKUŁ PAPIERNICZY-1 *'paper article'* which is evidently too wide. They are correctly joined by the *holonymy (collection)* relation to PRZYBORY DO PISANIA-1 *'writing implements'* (the reason is that this term is used in Polish only in plural).

For some predicates, arguments considered separately represent a wide class of entities, but actually they are closely related to each other. For instance, SKŁADAĆ SIĘ-1 *'consist of'* may concern concrete objects, e.g., various devices, groups of people composed of concrete people, events having their phases, artistic pieces etc. Classic selectional preferences give us no information about such a relationship. Therefore, selectional preferences determined by means of relations to another argument have been introduced. The frame for *SKŁADAĆ SIĘ-1* is presented in Example 22. *RELAT* is a predefined symbol meaning any close *PlWordNet* relation between lexical units realizing arguments in text.

(22) `składać się-1`

$Theme^{Foreground}$	$Theme^{Background}$
ALL	$\{CZĘŚĆ\text{-}3,\ RELAT \longrightarrow Theme^{Foreground}\}$

Therefore, the selectional preferences are represented as list of any number of following:

1. a *PlWordNet* synset (including predefined sets),
2. a *PlWordNet* relations to another synset,
3. a *PlWordNet* relations to another argument.

4.4 Connection with syntactic layer

In *Walenty*, syntactic and semantic valence information are represented separately. Nevertheless, they are closely connected, but this relation is many-to-many. On one hand, one semantic frame can be syntactically implemented by several schemata (diathesis alternation). On the other hand, one schema can be used in several frames.

We directly link semantic arguments with corresponding syntactic positions. We assume that the fact that all phrase types composing a single position can coordinate means that they cannot represent different semantic arguments. This does not mean that all phrase types composing a position must be connected with a corresponding argument. Similarly, not all positions must be connected with all frames adequate for a particular lemma.

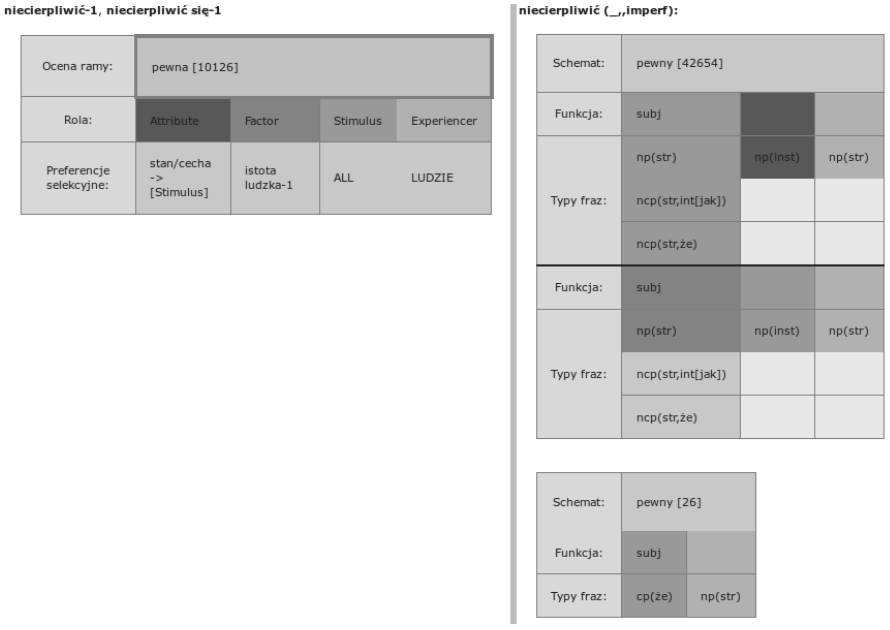

Fig. 1. A semantic frame and schemata being its syntactic realisation

For example, let us consider a verb NIECIERPLIWIĆ *'make impatient'*. The subject of the verb can be *a person* (*Factor*) who makes someone (*Experiencer*) *impatient* with their behaviour (*Stimulus*) etc. (*'Paul makes Jane impatient with his vagueness'*), or that behaviour itself (*'Paul's vagueness makes Jane impatient'*). In the first case only nominal subject is allowed, in the latter case all kinds of clausal subjects are possible as well.

The snapshot of the program visualisation of this entry is presented in Figure 1. On the left we can see a semantic frame, with *PlWordNet* lexical units connected to it on the top. The roles are distinguished by various colours that are assign to them in a fixed way. Syntactic schemata are positioned on the right. Phrase types and syntactic positions they belong to are coloured accordingly to the role they are connected to. Finally, all exemplary sentences connected to the corresponding lexical units appear at the bottom.

This entry exemplifies yet another feature of our representation. Namely, the same phrase type may be used in text in two different ways, as nominal subject of the upper schema [42654], used as a *Stimulus* and a *Factor*. We call such phenomenon *autoalternation*. Currently, only one autoalternation per schema is allowed.

5 Supervised creation of semantic frames

As semantic layer of *Walenty* is growing, it allows to use some supervised and semi-supervised methods of automatic frame creation. Currently there are few semantic entries in *Walenty* and they are located sparsely. This makes complex algorithms unpractical.

There are three aspects of a supervised frame creation that have to be taken into account. First, roles for a predicate have to be selected. Secondly, selectional preferences have to be ascribed. Thirdly, semantic arguments have to be connected with corresponding syntactic positions. Steps two and three can be taken independently, but both highly depend on precision and accuracy of the first step.

At current stage of work mostly heuristics supporting lexicographer in creation of the semantic layer of *Walenty* are being used. While they may appear simple, they provide enough support to simplify process of creating semantic frame.

To select semantic roles for a predicate *PlWordNet* relations are used. Synonymy, hypernymy, hyponymy and cohyponymy (having common hypernym) are used to suggest what semantic roles should be considered for a given predicate, as those relations suggest close semantic description. Some other *PlWordNet* relations are even more useful for supervised semantic frames creation. Those relations include aspect (perfect – imperfect; e.g., DOCENIĆ-1 *'appreciate'* and DOCENIAĆ-1 *'appreciate'*; see Figures 2 and 3) and converse (looking at same situation from different sides; e.g., KUPOWAĆ-1 *'buy'* and SPRZEDAWAĆ-1 *'sell'*; see Figures 4 and 5). In future those relations will be weighted to minimize amount of lexicographer's work. Currently corresponding frames are copied due to small size of training data (there are almost no entries connected to others by more than one of those relations).

Statistics of connections between semantic roles and phrases in syntactic positions are also analyzed. That information may not be helpful for automatic creation of semantic frame, but are something that should be considered by lexicographers while adding new role and creating connections between semantic and syntactic layers.

Similarly, statistics of connections between concepts appearing together as a selectional preference. As result some predefined sets of selectional preferences have been created for further use.

6 Future work

First, semantic layer of *Walenty* has to be finished. This includes extending it with phraseology and improving methods for supervised and semi-supervised frames creation. Once finished, it will find uses in numerous natural language processing algorithms and resources.

Składnica treebank ([29]) is planned to be extended with predicate-arguments layer (currently there are two layers – purely syntactical, created with *Świgra*

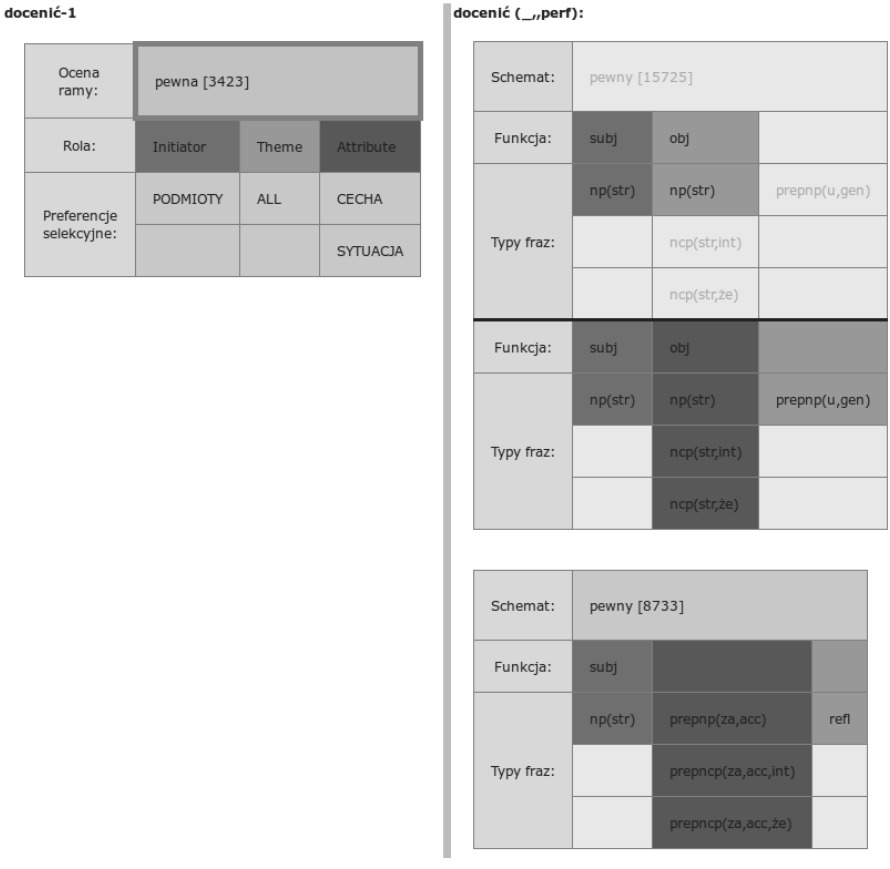

Fig. 2. A frame and schemata for verb DOCENIĆ

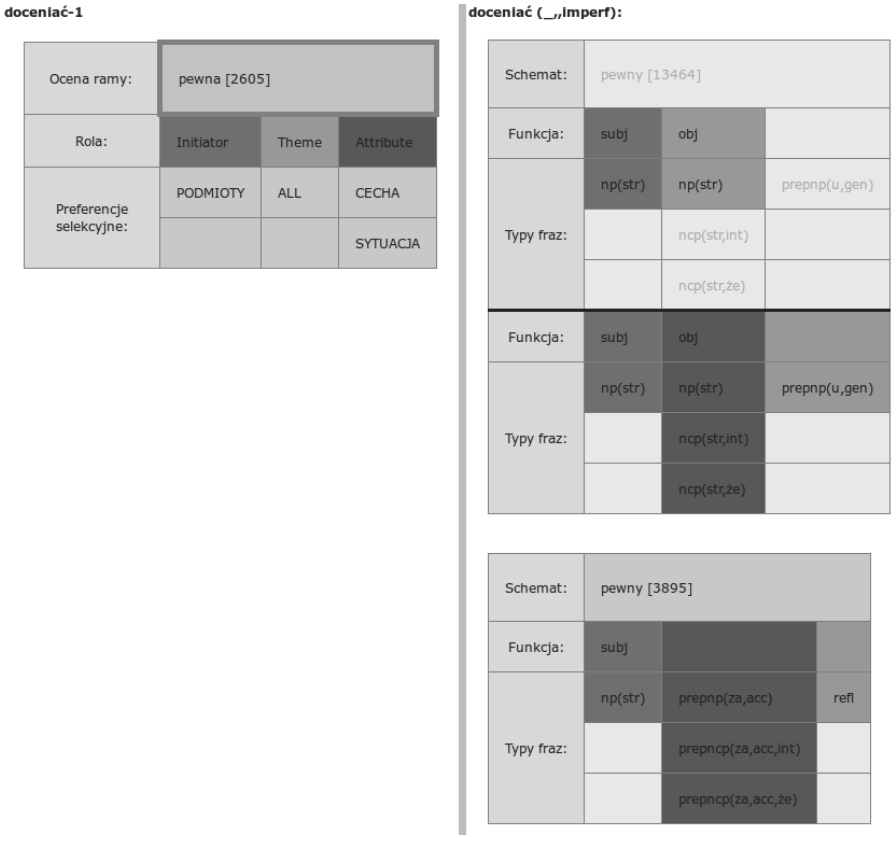

Fig. 3. A frame and schemata for verb DOCENIAĆ

parser ([30]) and syntactical, additionally marked with *PlWordNet* lexical units; [31]). *Świgra* parser is also going to be extended, so it would not only return all possible parse trees and corresponding predicate-arguments structure, but also mark relevant words with their possible meanings (with help of selectional preferences) and find one (or few) most adequate.

Acknowledgements

The study is cofounded by the European Union from resources of the European Social Fund. Project PO KL "Information technologies: Research and their interdisciplinary applications", Agreement UDA-POKL.04.01.01-00-051/10-00.

kupować-1

Ocena ramy:	pewna [11789]					
Rola:	Measure	Location	Initiator, Source	**Recipient**	**Theme**	Initiator, Goal
Preferencje selekcyjne:	waluta-1	kraj-1	PODMIOTY	ISTOTY	DOBRA	PODMIOTY
	pieniądze-1	miejscowość-1		PODMIOTY		
		obiekt handlowy-1				

kupować (_,,imperf):

Schemat:	pewny [20]		
Funkcja:	subj	obj	
Typy fraz:	np(str)	np(str)	np(inst)

Schemat:	pewny [43573]					
Funkcja:	subj	**obj**				
Typy fraz:	np(str)	**np(str)**	**np(dat)**	prepnp(u,gen)	xp(locat)	distrp

Schemat:	pewny [43571]					
Funkcja:	subj	**obj**				
Typy fraz:	np(str)	**np(str)**	**np(dat)**	prepnp(od,gen)	prepnp(za,acc)	xp(locat)

Schemat:	potoczny [1234]	
Funkcja:	subj	obj
Typy fraz:	np(str)	np(str)
		ncp(str,że)

Fig. 4. A frame and schemata for verb KUPOWAĆ

sprzedawać-1, sprzedawać się-1

Ocena ramy:	pewna [11804]					
Rola:	Manner	Location	Measure	Initiator, Goal	**Theme**	Initiator, Source
Preferencje selekcyjne:		kraj-1	waluta-1	PODMIOTY	DOBRA	PODMIOTY
		miejscowość-1	pieniądze-1			
		obiekt handlowy-1				

sprzedawać (_,imperf):

Schemat:	pewny [43550]		
Funkcja:	subj		
Typy fraz:	np(str)	xp(mod)	lex(np(str),sg,XOR('krew','skóra','życie'),atr1({possp}))

Schemat:	pewny [25769]			
Funkcja:	subj	**obj**		
Typy fraz:	np(str)	**np(str)**	xp(adl)	distrp

Schemat:	pewny [3164]			
Funkcja:	subj	**obj**		
Typy fraz:	np(str)	**np(str)**	prepnp(za,acc)	xp(adl)

Schemat:	pewny [43578]				
Funkcja:	subj	**obj**			
Typy fraz:	np(str)	**np(str)**	np(dat)	prepnp(za,acc)	xp(locat)

Fig. 5. A frame and schemata for verb SPRZEDAWAĆ

References

1. Przepiórkowski, A., Skwarski, F., Hajnicz, E., Patejuk, A., Świdziński, M., Woliński, M.: Modelowanie własności składniowych czasowników w nowym słowniku walencyjnym języka polskiego. Polonica **XXXIII** (2014) 159–178
2. Przepiórkowski, A., Hajnicz, E., Patejuk, A., Woliński, M., Skwarski, F., Świdziński, M.: Walenty: Towards a comprehensive valence dictionary of Polish. In Calzolari, N., Choukri, K., Declerck, T., Loftsson, H., Maegaard, B., Mariani, J., Moreno, A., Odijk, J., Piperidis, S., eds.: Proceedings of the Ninth International Conference on Language Resources and Evaluation, LREC 2014, Reykjavík, Iceland, ELRA (2014) 2785–2792
3. Przepiórkowski, A., Hajnicz, E., Patejuk, A., Woliński, M.: Extended phraseological information in a valence dictionary for NLP applications. In: Proceedings of the Workshop on Lexical and Grammatical Resources for Language Processing (LG-LP 2014), Dublin, Ireland, Association for Computational Linguistics and Dublin City University (2014) 83–91
4. Przepiórkowski, A.: Long distance genitive of negation in Polish. Journal of Slavic Linguistics **8** (2000) 151–189
5. Witkoś, J., Cegłowski, P., Snarska, A., Żychliński, S.: Minimalist Facets of Control. An English-Polish Comparative Overview of Gerunds and Infinitives. Minimalist Facets of Control. An English-Polish Comparative Overview of Gerunds and Infinitives. Wydawnictwo Naukowe UAM, Poznań (2011)
6. Baker, C.F., Fillmore, C.J., Lowe, J.B.: The Berkeley FrameNet project. In: COLING-ACL '98: Proceedings of the Conference, Montreal, Canada (1998) 86–90
7. Fillmore, C.J., Wooters, C., Baker, C.F.: Building a large lexical databank which provides deep semantics. In Tsou, B., Kwong, O., eds.: Proceedings of the 15th Pacific Asia Conference on Language, Information and Computation, Hong Kong (2001)
8. Baker, C.F., Fillmore, C.J., Cronin, B.: The structure of the FrameNet database. International Journal of Lexicography **16**(3) (2003) 281–296
9. Fillmore, C.J., Johnson, C.R., Petruck, M.R.: Background to FrameNet. International Journal of Lexicography **16**(3) (2003) 235–250
10. Dang, H.T., Kipper, K., Palmer, M., Rosenzweig, J.: Investigating regular sense extensions based on intersective Levin classes. In: Proc. Coling-ACL98, Montreal CA (August 11-17, 1998)
11. Dang, H.T., Kipper, K., Palmer, M.: Integrating compositional semantics into a verb lexicon. In: Proc. COLING-2000 Eighteenth International Conference on Computational Linguistics, Saarbrucken, Germany (July 31 - August 4, 2000)
12. Kipper, K., Dang, H.T., Palmer, M.: Class-based construction of a verb lexicon. In: Proc. AAAI-2000 Seventeenth National Conference on Artificial Intelligence, Austin, TX (July 30 - August 3, 2000)
13. Kipper, K., Korhonen, A., Ryant, N., Palmer, M.: Extending verbnet with novel verb classes. In: Proc. Fifth International Conference on Language Resources and Evaluation (LREC 2006), Genoa, Italy (June, 2006)
14. Hajnicz, E.: Automatyczne tworzenie semantycznego słownika walencyjnego. Problemy Współczesnej Nauki. Teoria i Zastosowania: Inżynieria Lingwistyczna. Akademicka Oficyna Wydawnicza EXIT, Warsaw (2011)
15. Fillmore, C.J.: The case for the case. In Bach, E., Harms, R.T., eds.: Universals in linguistic theory. Holt, Rinehart and Winston, New York, NY (1968) 1–88

16. Fillmore, C.J.: Some problems for Case Grammar. In: Report of the 22nd Annual Roundtable Meeting on Linguistics and Language Studies. Georgetown University Press, Washington, DC (1971) 35–56
17. Fillmore, C.J.: Frame semantics and the nature of language. In: Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech. Volume 280. (1976) 20–32
18. Katz, J.J., Fodor, J.A.: The structure of a semantic theory. In Fodor, J.A., Katz, J.J., eds.: The Structure of Language. Prentice Hall (1964) 479–518
19. Resnik, P.: Selection and Information: A Class-Based Approach to Lexical Relationships. PhD thesis, University of Pennsylvania, Philadelphia, PA (December 1993)
20. Piasecki, M., Szpakowicz, S., Broda, B.: A Wordnet from the Ground Up. Oficyna Wydawnicza Politechniki Wroclawskiej, Wroclaw (2009)
21. Piasecki, M., Maziarz, M., Szpakowicz, S., Rudnicka, E.: PlWordNet as the Cornerstone of a Toolkit of Lexico-semantic Resources. In: Proc. 7th International Global Wordnet Conference. (2014) 304–312
22. Piasecki, M., Szpakowicz, S., Broda, B.: Toward plWordNet 2.0. In Bhattacharyya, P., Fellbaum, C., Vossen, P., eds.: Principles, Construction and Application of Multilingual Wordnets. Proceedings of the 5th Global Wordnet Conference, Mumbai, Narosa Publishing House (2010) 263–270
23. Piasecki, M., Maziarz, M., Szpakowicz, S., Rudnicka, E.: Registers in the System of Semantic Relations in plWordNet. In: Proc. 7th International Global Wordnet Conference. (2014) 330–337
24. Maziarz, M., Piasecki, M., Szpakowicz, S.: An Implementation of a System of Verb Relations in plWordNet 2.0. In: Proceedings of the 6th Global Wordnet Conference. Volume Accepted for publishing. (January 2012)
25. Hajnicz, E., Andrzejczuk, A., Bartosiak, T.: Semantic layer of the valence dictionary of Polish *Walenty*. submitted to LREC 2016 conference (2015)
26. Center for Computational Language and Education Research, University of Colorado Boulder, CO: VerbNet Annotation Guidelines. (2012)
27. Andrzejczuk, A., Hajnicz, E.: Zasady opracowania ram semantycznych w słowniku walencyjnym *Walenty*. Manuscript (2015)
28. Polański, K., ed.: Słownik syntaktyczno-generatywny czasowników polskich, t. I–V. Zakład Narodowy imienia Ossolińskich, Wrocław · Warszawa · Kraków · Gdańsk (1980–1992)
29. Woliński, M., Głowińska, K., Świdziński, M.: A preliminary version of Składnica—a treebank of Polish. In Vetulani, Z., ed.: Proceedings of the 5th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics, Poznań, Poland (2011) 299–303
30. Woliński, M.: Komputerowa weryfikacja gramatyki Świdzińskiego. PhD thesis, Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland (2004)
31. Hajnicz, E.: Lexico-semantic annotation of *Składnica* treebank by means of PlWN lexical units. In Orav, H., Fellbaum, C., Vossen, P., eds.: Proceedings of the 7th International WordNet Conference (GWC 2014), Tartu, Estonia, University of Tartu (2014) 23–31