

The Reversed Dutch-Slovene database: Shortcomings and Some Contrastive Linguistic Issues

Anita Srebnik

University of Ljubljana
Aškerčeva 2, 1000 Ljubljana, Slovenia
E-mail: anita.srebnik@guest.arnes.si

Abstract

In 2007 a Dutch-Slovenian dictionary was compiled for the purpose of decoding. The paper describes the process of reversing it and its result; the converted Slovenian-Dutch database and the subdivision of entries into distinct groups. A general review of the converted database addresses some contrastive relations between Dutch and Slovene; valuable encoding information for the new dictionary, two-way translation dynamics, and the discovery of some inconsistencies of the source Dutch-Slovenian dictionary. The shortcomings of the reversing process, like the absence of multi-word headwords and the parts of headwords inside brackets will have to be resolved in the next conversion round. The main problem to be solved is insufficient lemmatizing and POS-tagging of Dutch separable verbs used separately in a context – which proved to be a universal issue of tagging Dutch texts. The analysis of the converted database is only one step towards a reversed dictionary. Prior to the inclusion of the converted and post-edited database into a reversed dictionary, a monolingual analysis of Slovenian has to be carried out by various monolingual sources and tools.

Keywords: bilingual dictionaries; reversed database; contrastive analysis; two-way translation dynamics

1. Introduction

There are a very few lexicographic resources for the language pair Dutch/Slovene. Next to two pocket-size dictionaries and a Slovene-Dutch dictionary on the internet, the only more comprehensive dictionary is a Dutch-Slovene dictionary (Srebnik, 2007) in paper form. Accordingly, a new Slovene-Dutch dictionary is still waiting to be compiled.

The main purpose of reversing a dictionary is to “maximize the abundance of information” (Krek *et al.*, 2008) in a source dictionary, and to reuse the already established cross-linguistic equivalents. It therefore seemed reasonable to reverse the existing Dutch-Slovenian dictionary before starting to compile a companion Slovenian-Dutch dictionary to gain maximum benefit from the information already contained in a monodirectional dictionary. The reversal was done automatically by Amebis, a software company for language technologies involved in NLP.

The purpose of reversing is rather twofold; not only can a reversed database serve as a database for a new Slovenian-Dutch dictionary, but at this stage of our research it is a valuable source for the analysis of Slovenian as seen from the Dutch perspective, taking into account the lexicographic context. The reversed database represents only “the mirror image” (ibid.) of the Dutch-Slovenian contrastive relation. The analysis of contrastively relevant aspects can contribute to well-founded decisions when compiling a new dictionary.

2. The source

The source of the dictionary data that has been reversed is an XML database of the printed Dutch-Slovene Dictionary (Srebnik 2007). It is monodirectional and primarily intended for Slovenian users, learners of Dutch,

to help them with understanding and translating from Dutch into Slovenian. The main function is thus passive or receptive, the secondary function, however is rather active or productive; by means of grammatical information and illustrative examples a more skilled user could produce written or spoken texts in Dutch. The dictionary is not a corpus-based work, but has been compiled by means of numerous available paper and electronic resources involving Dutch, mainly bilingual and monolingual dictionaries and Dutch learning methods. The macrostructure comprises 11,200 of the most frequent and common lexemes, which belong to the basic vocabulary of the Dutch. Therefore, no specialized, archaic, or dialectal variants of language are included. In the microstructure stress is laid on numerous examples of usage, and on significant context in which a headword appears. As the author of the dictionary, I based the decision about the inclusion of a fair amount of pregnant context on the experience that for a learner it is not enough to know only the meaning of the word, it is equally important to be able to see it in a broader context – and in an ideal case to use it actively.

The mono-directional concept dictated the selection of Dutch headwords and illustrative examples that were often based upon expected difficulties in the Slovene translations. It is more interesting to observe what kind of “mirror image” of Slovenian would show up in the reversed database.

The dictionary was compiled with the help of an editor called ADICTED (Anita's DICTIONary EDITor), which was developed for the Dutch-Slovenian dictionary at the Institute for Dutch Lexicology (INL) in Leiden with an underlying XML format.

3. The reversal process

3.1 Other languages

Given a vast amount of dictionary compilation software available nowadays, the reversal of a bilingual dictionary is technically much easier than ever before. Several reversal projects have been completed so far, with various software and diverse outcome (see e.g. Geisler, 2002; Honselaar and Elstrodt, 1992; Tamm, 2002; Krek *et al.*, 2008; Maks, 2007; Prinsloo and de Schryver, 2002; Newmark, 1999; Veisbergs, 2004; Martin and Tamm, 1996). Seen from the results that the reversion of the monodirectional Oxford DZS English-Slovenian Dictionary has yielded, the reversal can be rewarding and provide the lexicographer with useful information for the new dictionary and relevant contrastive data for further analysis.

In some cases, especially in the Dutch experience of reversing dictionaries with the OMBI, a tool for creating and editing bilingual dictionaries, the reversal is planned from the outset and should be anticipated in the design of a bilingual dictionary, e.g. Dutch-Finnish v.v., Dutch-Estonian v.v., Dutch-Turkish v.v., Dutch-Arabic v.v., Dutch-Hungarian v.v., Dutch-Polish v.v., Dutch-Italian v.v., and Dutch-Swedish v.v. (Martin and Tamm, 1996; Laureys, 2007).

The reversibility doesn't need to be integrated into the compilation process from the outset, and can nonetheless yield successful results, as has been proved by Krek's team who reversed the English-Slovene dictionary.

3.2 Dutch-Slovenian database

Also reflections upon the converted Dutch-Slovenian database reveal that the underlying XML structure forms a flexible base for a conversion process even if the reversing has not been planned from the beginning when compiling the L1 (source language) → L2 (target language) dictionary. The XML structure allows arbitrary sequencing of different information categories. Before reversing the lexicographer is supposed to define the sequence of XML elements with information categories which are coded according to a specific system, and consequently the previous elements are labelled anew.

The most obvious reversal is between a Dutch headword in a source dictionary which now becomes a translation equivalent of the previous Slovenian translation equivalent and vice versa and between illustrative example in Dutch which now becomes a translation of the Slovenian example.

However, the conversion process enables the tracing back of every new element to its position in a source dictionary, especially previous Slovenian translation equivalents of new Dutch translation equivalents.

Here is an example of an illustration of a reversed dictionary entry for *alarm* (alarm) in an XML format:

```
<article>
<hw>alarm</hw><hwx>alarm</hwx>
<pron><IPA/></pron>
<prio>01</prio>
<gen>het</gen>
<tText>alarm</tText>
<tText2>al<L>a</L>rm</tText2>
<pl>ed.</pl>
<Ann>
<trx>alarm, preplah</trx>
<part>1</part>
</Ann>
<e>
<eMarker/>
<eFromL>lažni alarm</eFromL>
<eToL>loos alarm</eToL>
<Ann>
<hwe>loos</hwe>
<parte></parte>
</Ann>
</e>
</article>
```

Figure 1: Slovenian-Dutch database, XML-format

In the next phase both the Slovenian and Dutch texts had to be POS-tagged and lemmatized. For the Slovenian, this has been done fully automatically by the proprietary tagger owned by the Amebis company. It was also necessary to POS-tag and lemmatize the Dutch part of dictionary examples in order to be able to detect the existence of translation equivalents. This has been done at the INL in Leiden by means of FROG (Dutch morpho-syntactic analyzer and dependency parser).

After that the dictionary was ready to be converted following the routine of XML elements defined by the lexicographer. This operation was done automatically by the Amebis company.

Additionally, an XSL file was created to enable the lexicographer to visualize the data chosen to be seen in a user-friendly form. The final result is the following:

bližina
 › **group 01**
de buurt [*buurten*]
bližina, sosesčina

- Stanujemo v bližini rdeče četrtri. Wij wonen in de buurt van de rosse buurt.

de omgeving [*ed.*]
okolica; bližina

- Jo stanuje v bližini Leuvna. Jo woont in de omgeving van Leuven.

- **V neposredni bližini ni nobene trgovine.** *Er is hier in de onmiddellijke omgeving geen winkel.*
- › **group 02**
- de nabijheid** [ed.]
- neposredna bližina, bližnja okolica**
- **V neposredni bližini vasi je jezero.** *In de nabijheid van het dorp is een meer.*

Figure 2: Slovenian-Dutch database, XSL-file

Taking into account that the recent reversion has only been the first attempt towards a new Slovene-Dutch reversed database and that in the short time available for evaluation a few draw-backs showed up, we will briefly compare the source dictionary and its reversed database in terms of numbers:

Dutch-Slovene dictionary

headwords: 11.117
examples of usage: 9.153
translation equivalents: 13.117

Slovene-Dutch reversed database

headwords: 11.465
examples of usage: 39.716
translation equivalents: 26.192

The high number of headwords is expected to be higher after the next reversion round because of the absence of multi-word headwords and those headword parts, which are in between brackets. The latter got partly lost in the process, so that for example the Slovenian translation of the Dutch noun *berouw* (remorse) *kes(anje)* which is a compressed form of two variants of the same lemma, *kes* and *kesanje*, was only registered as one word (*kes*).

The high number of examples of usage in the new database stands out most since the Slovene example is automatically listed under a new Slovene headword if it contains that headword. Still, this was done consistently during the automatic process, which will be illustrated later on.

3.3 Organizing principles of the converted database

Krek *et al.*, (2008) invented an innovative approach to the organizing principles of the reversed dictionary database which resulted in an enhanced reversing process with built-in categorization of the material. Their Slovenian-English database was a much more extensive database with a more complex and detailed structure. The research resulted in the article 'The Funny Mirror of Language,' (ibid.) where they introduced four distinct groups into which the new entries are subdivided. The same model for reversion has been applied for the Slovene-Dutch converted database. The groups are clear-cut and most helpful as a starting point for any kind of contrastive analysis. Otherwise the reversed material

would be too unorganized and less accessible, and valuable information would be more difficult to find. The categories are the following drawing on the above mentioned article by Krek *et al.*:

1. group “one to one”: the new Slovenian headword appears as a one-to-one translation of the new Dutch candidate for the translation equivalent. The corresponding Slovenian examples from the entire database where a one-to-one translation appears in the Dutch part of the example are grouped under each equivalent. E.g.:

banalen
› **group 01**
banaal
plehek; vsakdanji; banalen

laag-bij-de-gronds
puhel, plehek, banalen, prozaičen

Figure 3: Translation group 1

2. group “one to multi-word + base form”: in this case the new Slovenian headword appears as a part of the multi-word Slovenian translation equivalent in the Dutch-Slovenian dictionary and is used in its base form. E.g.:

biro
› **group 02**
de projectontwikkelaar [*projectontwikkelaars*]
gradbeno podjetje, projektni biro, projektant

Figure 4: Translation group 2

3. group “one to multi-word + inflected form”: in this case the new Slovenian headword appears as a part of the multi-word Slovenian translation equivalent in the Dutch-Slovenian dictionary and is used in one of its inflected forms. E.g.:

pločevinast
› **group 03**
de trommel [*trommels*]
pločevinasta škatla

- **Piškoti so v pločevinasti škatli.** *De koekjes zitten in de trommel.*
-

Figure 5: Translation group 3

4. group “no translation”: the new Slovenian headword, e.g. adjective *dobro* (good), is used in the Slovenian part of the example: *To pravim v tvoje dobro.* (I am saying

this for your own good.), but none of the Dutch translation equivalents from previous groups 1 to 3 (goed, lekker, knap, wel, zoet) are used in the Dutch part of the example: *Dat zeg ik om je eigen bestwil*.

According to Krek *et al.* (2008), “the last group is seen as particularly useful since it exposes contrastively interesting cases where in the English-Slovenian dictionary (in our case the Dutch-Slovenian), lexicographers had to find a solution that did not include the most common translation equivalents for a particular headword.”

4. General assessment of the first reversion round

4.1 Contrastive relations between Dutch and Slovene

As the reversion will have to be activated once again with more narrowly defined input because it did not yield expected results for the analysis, the results of the first general review which follow here can not be elaborated and complete. Below are some first reflections upon the contrastive relations in the reversed database between the Dutch and Slovenian:

4.1.1 Valuable encoding information for the new dictionary

A new dictionary is supposed to be compiled in the first place for Slovenians to actively use Dutch, so extra information is needed about grammatical and collocational behavior of a Dutch lemma. An instance of good usage examples, found under the headword *akademija* (academy) would be: *študirati na akademiji za likovno umetnost* (to study at the art academy) → *studeren aan de kunstacademie, aan/op de kunstacademie zitten*. A Slovenian speaker learns how to properly use the lemma, and that next to a formal translation (*studeren aan de kunstacademie*) he can also use a frequent non-formal combination with two different prepositions (*aan* or *op*) and the verb *zitten*.

A different, but again a very useful type of usage examples which can be directly applied to a new dictionary was found under the headword *bencin* (petrol). The Slovenian sentence *Avto porabi liter bencina na osem kilometrov. / Avto porabi 12,5 litrov bencina na sto kilometrov* (The car does 8 km to the litre.), the first sentence being a literal translation of Dutch and the second a commoner version in Slovene (still the use of a decimal number in this context is unusual, because it is an exact translation), is rendered into Dutch with: *De auto loopt één op acht*. A sentence which would never occur to a Slovene speaker if he translates more or less literally from his own language. The Dutch sentence could be a pragmatic formula, because of its frozen structure and transparent meaning, which makes it a good candidate for a reversed dictionary.

In the database there is an abundance of examples with an encoding function. It is interesting and quite confronting to see, as the only author of the source dictionary, why so many examples of usage are so readily insertable in a new, reversed dictionary while we have converted a primarily passive dictionary and are going to work on an active one. The reason lies in the already mentioned secondary, active or encoding function, and, admittedly, in a personal urge of the lexicographer not merely to explain but also to teach how to use Dutch.

4.1.2 A two-way translation dynamics

As Šorli (2009) and Krek *et al.*, (2008) emphasize in their findings about the reversed database, there is the “simple truth about the dynamics of the translation process: *L1 (Source Language) → L2 (Target Language) does not equal L2 (Source Language) → L1 (Target Language)*.” The lexicographers should take into account some specific issues “if they are to avoid falling into traps set continually by the reversed perspective. Typically, the L1 content will be a self-contained semantic unit and, ideally, rendered into L2 with an equally natural and/or lexically frozen semantic unit. However, in many cases the levels of this naturalness and fixedness differ, sometimes considerably. The key problem is not so much that of semantic equivalence, but rather that of equivalence in terms of typicality/frequency.” Below are some examples:

- multi-word lexical units as entry headwords:

umazana posoda (a washing up) = *de vaat, imeti slavnostno večerjo* (to have a dinner, a formal meal or to celebrate something) = *dineren* in one of its meanings, *ostati doma* (to stay at home) = *thuisblijven*. The Slovenian multi-word lexical units are quite acceptable as a translation equivalent in a source dictionary, but the lexical ties between their individual items are not strong enough to be used as headwords.

- insufficiently contextualised lexical strings as illustrative examples:

jata škurhov (a flight of whimbrels) = *een vlucht regenwulpen*

The context in the Dutch-Slovenian dictionary is determined by the Dutch perspective. In the reversed dictionary this example is not suited to the needs of a Slovene speaker. Firstly, this sort of migratory birds is extremely seldom in Slovenia, secondly, the corpus analysis reveals that it is mainly used with adjectives *veliki* (big) or *mali* (small), and never with *jata* (a flight), and thirdly, the example of usage is quite acceptable in the explanatory source dictionary since it does not only refer to a common bird in the Netherlands but also to a title of a very famous Dutch novel:

- idiomatic lexical strings:

An idiom can often be translated by an idiom in another language: *het topje van de ijsberg* (the tip of the iceberg) = *vrh ledene gore*, *To nima ne repa ne glave*. (There's neither rhyme nor reason to it.) = *Daar is geen touw aan vast te knopen.*, *Malo je čez les*. (He's got a screw loose.) = *Hij ziet ze vliegen*. But quite often, the translations in a source dictionary can be too generic or explanatory; *prenehati s čim* (make an end of something) = *een punt achter iets zetten*, *zmesti se* = *de kluts kwijt zijn*, *obvladati kaj* (to master something) = *iets onder de knie krijgen* to be considered as the source idioms for the new dictionary.

4.1.3 Shortcomings of the source dictionary

"The mirror image" of Slovenian can confront the lexicographer with inconsistencies in the source dictionary or simple mistakes which he sometimes could not have been able to notice if the base had not been reversed. In this way he can improve the source dictionary. A quick look at the headword *antologija* (anthology), which is rendered into Dutch by the near synonyms *anthologie* and *bloemlezing*, reveals that in the source dictionary *anthologie* had been translated by *antologija*, *izbor*, and *bloemlezing* by *zbirka*, *antologija* which means that both strings of translation equivalents lack one more, *zbirka* must be added to the equivalents of *anthologie* and *izbor* to those of *bloemlezing*.

5. Shortcomings of the reversion process

As already mentioned, there have been some shortcomings discovered so far, which need to be improved during the next reversion round. Unfortunately, not all of them are due to the reversion process.

One of the most conspicuous drawbacks is the complete absence of Dutch illustrative examples containing separable verbs, when the prefix is used separately from the verb. Those sentences together with their Slovene translations do not appear under the new Slovenian headword, which is a translation of that compound verb. Sometimes they appear under another lemma in the sentence, but never under a lemma of that verb. For example, the Dutch sentence with the verb *nadoen* (imitate, copy) *Marja svojo starejšo sestro v vsem posnema*. (Marja copies her older sister in everything.) = *Marja doet haar oudere zusje in alles na*. does not, but should appear under the headword *posnemati* (imitate). The reason for that is, that lemmatizing of the Dutch text was carried out word for word. "The parts of a separable compound verb /.../ are therefore each allocated their own lemma." (Van Eynde, 2004). The prefix in a sentence like above is placed at the end of the sentence and tagged as a preposition in a final position. In this way the verb is not recognized anymore as *nadoen*, but only as *doen*. In Van Eynde's *Protocol for POS tagging and lemmatizing* we can read the following: "'Separability' is not included because allocating the values needs a full syntactic analysis." This is one of the

major reversion drawbacks and at the moment it is not clear yet, how to fix it so as to avoid time-consuming manual work.

It also has to be found out why the sentences with separable verbs are arbitrarily distributed under some lemma's and why the others are simply left out. For instance, the usage example in the previous paragraph appears as an illustrative sentence under the headword *star* = *oud* (old), *vse* = *alles* (everything), and *v* = *in* (in), but not under the headword *sestra* = *zus* (sister) and, as already mentioned, under the separable verb *posnemati* (imitate).

Another issue to be resolved during the next conversion round is the use of brackets containing affixes especially in Slovene verbs, but also other word classes. During the conversion only parts of the words outside brackets were considered new headword candidates, consequently a lot of headwords were left out. The use of brackets is an economical way of spelling Slovene imperfective/perfective verb pairs. Translation equivalents of the Dutch *knetteren* are listed as: *(za)prasketati*, *(za)pok(lj)ati*, and *(za)hřeščati*. Written out fully, they would yield eight translation equivalents contributing to eight new headwords in a reversed database: *zaprasketati*, *prasketati*, *zapokljati*, *pokljati*, *zapokati*, *pokati*, *zahřeščati*, and *hřeščati*. Verb prefixes together with the verb root convert the imperfective verb into a perfective one.

Because the parts (usually verb prefixes) in brackets were not merged with the rest of the words, illustrative examples got lost as well. The sentence *V kuhinji sta vse prepleskala na rumeno*. (They painted the whole place yellow in their kitchen) = *Ze hebben in hun keuken de hele boel geel geschilderd*. appears four times: under the headwords *kuhinja* (kitchen), *v* (in), *na* (on) and *ves* (all), but not under the headword *prepleskati* (to paint). The reason lies in the spelling of the Slovenian translation equivalent *(pre)pleskati*, where *(pre)* has been ignored by the conversion programme. This issue is going to be resolved by means of regular expressions which will enable merging parts of the word into one.

Another issue touches upon entry division into the above four categories, but not accurately implemented in the conversion process. Only a few examples of usage that fell into the fourth category do actually belong there. What exactly has triggered the incorrect categorization remains to be discovered so that the next conversion round will yield better results.

A more detailed and comprehensive analysis of contrastive relations can only be carried out after the next reversal, when the data will be more complete – with more headwords due to the inclusion of multi-word translation equivalents and due to merging parts of words inside brackets into new words. The issue of the fourth category "no translation" must be resolved, and

the Dutch lemmatizing and POS tagging of separable verbs which are used separately in usage examples, must be corrected.

6. Conclusion

A general review of the converted Slovene-Dutch database addresses some contrastive relations between the languages and the shortcomings of the reversing which will be resolved in the next conversion round. An analysis of a reversed database represents only one step towards a new reversed dictionary, and can only be done thoroughly after the next conversion round. It remains to be seen to what extent are the results of an automatic reversion applicable for the production of the L2-L1 dictionary. All this taken into account, the sources and tools for the analysis of Slovenian, like the Gigafida Reference Corpus of Slovenian (<http://demo.gigafida.net/>), the Slovenian Lexical Database (LBS), the Word Sketch Engine, a corpus tool that analyses a word's grammatical and collocational behaviour, may adjust the "distorted image of Slovenian" (Krek *et al.*, 2008) determined by the Dutch. The monolingual analysis must be done prior to the use of the converted Slovenian-Dutch database in a post-editing phase.

7. Acknowledgements

The reversion would never have been done without the encouragement and continued support of Simon Krek and his professional expertise. I am also grateful to the Amebis company and its director Miro Romih for reversing my dictionary and Iztok Kosem from the Trojina Institute for all his support and advice.

8. References

- Geisler, C. (2002). Reversing a Swedish-English dictionary for the Internet. In L. Borin (ed.) *Language and Computers, Parallel Corpora, Parallel Worlds*. Amsterdam: Rodopi, pp. 123-133.
- Honselaar, W., Elstrod, M. (1992). The electronic conversion of a dictionary: from Dutch-Russian to Russian-Dutch. In H. Tommola, K. Varantola, T. Salmi Tolonen & J. Schopp (eds.) *EURALEX '92 Proceedings I-II; Papers submitted to the V EURALEX International Congress on Lexicography in Tampere, Finland. Part I*. University of Tampere: Department of Translation Studies. *Studia Translatologica*, pp. 229-237.
- Krek, S., Šorli, M. & Kocjančič, P. (2008). The Funny Mirror of Language: The Process of Reversing the English-Slovenian Dictionary to Build the Framework for Compiling the New Slovenian-English Dictionary. In E. Bernal, J. DeCesaris (eds.) *Proceedings of the XII EURALEX International Congress*. Barcelona: Universitat Pompeu Fabra, pp. 535-542.
- Laureys, G. (2007). Optimizing Procedures for the Making of Bilingual Dictionaries and the Concept of Linking Contrastive Lexical Databases. *International Journal of Lexicography*, 20(3), pp. 295-311.
- Maks, I. (2007). OMBI: The Practice of Reversing Dictionaries. *International Journal of Lexicography*, 20(3), pp. 259-274.
- Martin, W. (2007). Government Policy and the Planning and Production of Bilingual Dictionaries: the 'Dutch' Approach as a Case in Point. *International Journal of Lexicography*, 20(3), pp. 221-237.
- Martin, W., Tamm, A. (1996). OMBI: An editor for constructing reversible lexical databases. In M. Gellerstam, J. Järborg, S.-G. Malmgren, K. Norén, L. Rogström, & C. Røjder Papmehl (eds.) *EURALEX '96 proceedings I-II. Papers submitted to the VII EURALEX international congress on lexicography in Göteborg, Sweden*. Göteborg: Göteborg University, pp. 675-687.
- Newmark, L. (1999). Reversing a One-Way Bilingual Dictionary. *Dictionaries*, 20, pp. 37-48.
- Prinsloo, D.J., de Schryver, G.-M. (2002). Reversing an African-Language Lexicon: *the Northern Sotho Terminology and Orthography No. 4* as a case in point. *South African Journal of African Languages*, 2., pp. 161-185.
- Srebnik, A. (2007). *Nizozemsko-slovenski slovar / Nederlands-Sloveens woordenboek*. Ljubljana: DZS.
- Šorli, M. (2009). Pridobivanje podatkov o slovenščini za izdelavo slovensko-tujejezičnih slovarjev. In M. Stabej (ed.) *Infrastruktura slovenščine in slovenistike. Simpozij Obdobja 28*. Ljubljana: Znanstvena založba Filozofske fakultete, pp. 359-369.
- Tamm, A. (2002). Reversing the Dutch-Estonian Dictionary to Estonian-Dutch. In A. Braasch, C. Povlsen (eds.) *Proceedings of the Tenth EURALEX International Congress. Vol I*. Copenhagen: CST, pp. 389-399.
- Van Eynde, F. (2004). Protocol for POS tagging and lemmatizing. Accessed at: http://www.tst.inl.nl/cgndocs/doc_English/topics/annot/poos_tagging/tg_prot_en.pdf (accessed: 29 September 2011)
- Veisbergs, A. (2004). Reversal as Means of Building a New Dictionary. In G. Williams, S. Vessier, (eds.) *Proceedings of the XI EURALEX International Congress*. Lorient: UBS, pp. 327-332.