A golden mean? Compromises between quantity of information and user-friendliness in the bidirectional Norwegian-Lithuanian Dictionary

Aurelija Griškevičienė & Sturla Berg-Olsen

Keywords: *bilingual lexicography, user-friendliness, bidirectional dictionaries, Norwegian, Lithuanian*.

Abstract

This paper explores the concept of user-friendliness in the context of bidirectional bilingual dictionaries, presenting and discussing some of the choices taken by the editors of the Norwegian-Lithuanian Dictionary (NLD). The NLD is a medium-sized paper dictionary compiled by a joint group of lexicographers from the Universities of Vilnius and Oslo. The dictionary is intended both for native speakers of Norwegian and of Lithuanian. Designing a user-friendly bidirectional dictionary necessarily involves making compromises between the needs of different target groups. User-friendliness in lexicography is a problematic concept, because a feature that enhances the user-friendliness of a dictionary for one group of users often reduces it correspondingly for other groups. This is especially acute in the case of bidirectional dictionaries. The amount of information given and the degree of linguistic precision must be balanced against the danger of information overload. Thus, designing the structure of a dictionary is largely a matter of seeking compromises between quantity of information, precision and user-friendliness. The paper shows concrete examples of how the editors of the NLD have tried to maintain this balance. Many elements in the NLD are based on another bilingual dictionary (Berkov et al. 2003), but the system for information on the target language, Lithuanian, is designed by the editors of the NLD. The paper shows the steps taken to make the dictionary user-friendly from two angles: 1) adapting and improving the lemma list and information on the source language and 2) designing the system for providing information on the target language. In this context it also discusses problems arising from the wish to re-use data from one bilingual dictionary when compiling another dictionary with a different target language.

1. Introduction: What is a user-friendly dictionary?

In principle, a maximally user-friendly dictionary would be a dictionary which is available anywhere at any time and which provides the user with the exact information he or she is looking for within a minimum of look-up time. In practice, this ideal is unattainable, not least because any dictionary, no matter how specialised, will be used by a wide range of different people, with different backgrounds and different motivations for looking up information in the dictionary in a specific setting. Given that a dictionary cannot please everyone in every possible way, lexicographers are faced with the task of defining a core target group and satisfying the needs of this group to the greatest possible extent. At the same time, in the light of real-world constraints of practical and financial kind, dictionaries must nearly always aim at meeting the needs of a large target group (possibly with the exception of some projects dealing with very large language communities). Thus, designing the structure of a dictionary usually involves seeking compromises between quantity of information, precision and user-friendliness. In recent years there have been several studies dealing with different aspects of dictionary user-friendliness, discussing for example the use of grammatical coding systems, as well as criteria for selecting and presenting information (e.g. Dziemianko 2006, 2011, Kernerman 2008). This paper presents some challenges faced by the compilers of the *Norwegian-Lithuanian Dictionary* (hereafter: NLD) in their attempt to maximise the user-friendliness of this dictionary.

2. The Norwegian-Lithuanian Dictionary project

The NLD is a joint project of lexicographers from the Universities of Vilnius and Oslo, initiated in 2005. The main part of the work – the translation of the lemmas and examples from Norwegian to Lithuanian – has been completed, but some editing work still needs to be done. The dictionary corpus consists of approximately 48.000 entries. The NLD is primarily intended as a paper dictionary; however, an electronic version might be issued in the future, as the dictionary is composed in XML.

Since both Norwegian and Lithuanian are middle-sized languages in the European context, the available resources do not make it feasible to produce two separate dictionaries for Lithuanians and Norwegians which would contain different information needed for the understanding and production of Norwegian and Lithuanian respectively. So the NLD is intended both for native speakers of Norwegian and of Lithuanian. However, in certain respects the NLD favours the Lithuanian user; this choice seems justified because currently the number of Lithuanians who are in a situation where they need to understand and produce Norwegian is substantially higher than the opposite.

The basis for the NLD was adapted from *Stor norsk-russisk ordbok* (*The Big Norwegian-Russian Dictionary*, Berkov et al. 2003), but the system for providing information on the target language, Lithuanian, was designed by the editors of the NLD.

3. Adapting and improving a good basis: The lemma list and information on the source language

As mentioned, the basis for the NLD (the lemma list, examples and various information on the source language) was taken from the BNRD. The BNRD was chosen among several alternatives for the following reasons: 1) It is bilingual; 2) it is explicitly bidirectional and consequently provides a large amount of different information about the Norwegian lemmas and their Russian equivalents (grammatical, semantic and stylistic information, stress marking, pronunciation, plenty of examples, idioms and quotations, etc.); 4) the structure of Lithuanian is in many ways quite similar to that of Russian and 5) BNRD is generally considered to be a very professional and exhaustive work, cf. Nesset and Trosterud (2005).

The NLD editors initially intended to reuse the structure of the entries from the BNRD and all of the Norwegian part of the entries – the Norwegian lemmas as well as the information about them – in the NLD. However, in the process of writing the NLD, it turned out that although the basis chosen was quite comprehensive and well designed, much of the information from the BNRD could not be used directly without adaptation and change. In fact, the NLD in the end turned out to differ substantially from the BNRD both in its structure and in how information is presented. As evidenced by the experience from the Scandinavian LEXIN project, even the compilation of a ‘neutral’ lemma list to be used for several target languages is far from unproblematic (Gellerstam 1999: 8–10, Dahl 1999). In the case of the NLD, the original lemma list was designed with the target language Russian in mind, and this fact alone meant that certain changes could not be avoided (ways of differentiating between meanings, the selection of examples, etc.). The number of lemmas and examples was reduced because the NLD was planned to be a much smaller dictionary than the BNRD. On the other hand, several types of additional information about the Norwegian lemmas and their Lithuanian equivalents were included in the NLD. Many of the changes made were motivated by the editors’ goal of making the NLD more user-friendly. Some selected examples of these changes are presented in the following.

3.1. *Compounds*

In the BNRD compounds and some inflectional and derivational forms of the headwords are placed within the entries of the main component of the compound. In the NLD all lemmas of this kind are expanded to fully fledged headwords and treated in separate entries.

3.2. *The cross-reference system*

In the BNRD cross-references are made on various levels of the article, also from one part of an entry to another entry or to a part of another entry. This system appears quite user-unfriendly, especially in cases where the user is referred to a particular meaning in large entries treating polysemous headwords. In the NLD only cross-references between absolute synonyms headwords are used. The more frequent word is provided with target language equivalents, while the less frequent headword is given in a reference entry.

A type of cross-reference with the new function ‘see also’ was added in the NLD. This reference tells the user that additional information about the headword can be found in another entry. Thus, for example, the article **hammer** ‘hammer’ contains a reference to the compound *smihammer* ‘blacksmith’s hammer’.

3.3. *Reflexive verbs*

The BNRD uses seven different ways of presenting reflexive forms of Norwegian verbs. The NLD reduces this number to three: A reflexive form of a verb can be presented 1) as an example, 2) as a separate part of the entry when the verb has both reflexive and non-reflexive forms and the forms differ in their meanings, and 3) as a separate lemma if the verb does not have a non-reflexive form.

3.4. *Homonyms*

In the BNRD homonyms which are related to the same meaning(s) are rendered in a single entry. In the NLD, all homonyms belonging to different parts of speech are presented in separate entries.

3.5. *Meaning differentiation*

Many of the semantic paraphrases used to differentiate between meanings in the BNRD have been replaced by new ones in the NLD. The purpose of these changes is to use paraphrases which are more easily understood by most users – that is such ones which are shorter, less abstract, less scientific, not consisting of a reference to another headword, etc.

3.6. *Dead and live examples*

The BNRD provides a large number of so-called dead examples, that is examples constructed with codes for grammatical information. Thus, the Norwegian verb *inspirere* ‘inspire’ might be given in an example only with the information that it is used with the preposition *til* ‘to’, corresponding to the Russian preposition *na*, taking the accusative case. Such examples are avoided in the NLD. They are mostly replaced by 'live' ones, that is examples providing not only grammatical, but also semantic information, for example in the form of common collocates. With the verb *inspirere*, the mentioned example from the BNRD is replaced by *inspirere til nye handlinger* 'inspire to new actions'.

4. Designing a system: Information on the target language

In a truly bidirectional dictionary, the target language part should be designed to meet the requirements of users for whom the target language is L2, although much of this information will be superfluous to users for whom the target language is L1. In fact, information on for example the pronunciation or declension of the equivalents is useful only when the dictionary is used for productive purposes by people who are not proficient in the target language. Probably for this reason, information on the target language is often given only limited attention.

In the Lithuanian lexicographical tradition, bilingual dictionaries *to* Lithuanian have as a rule been designed with Lithuanian native speakers in mind. This would seem to be a natural consequence of the fact that the number of Lithuanians learning foreign languages is larger than the number of foreigners learning Lithuanian. To some extent the NLD represents a break with this tradition, in that it aims to provide its Norwegian users with at least *some* information on the target language. Still, the compilers of the NLD do not take quite as radical a stance as Berkov (1996: 109, our translation):

> 'As a rule, a bilingual dictionary should provide both the active and the passive user […] with the information necessary to form any element of the paradigm (i.e. any form) of the word – the lemma and the equivalent – as well as to use that word in a grammatically correct way in a context.'

This may be an intellectually attractive position as well as a desirable one from some users' point of view. However, in a paper dictionary a solution along these lines – where both lemmas and their equivalents are provided with maximal information – runs the risk of overwhelming many users, making it more difficult for them to find the information they are seeking and thus all in all lowering the user-friendliness of the dictionary. This has been pointed out by Svensén (2004: 311); cf. also Dziemianko's (2006: 5) remarks on the need to find compromises between ease of accessibility and accuracy of description. The editors of the NLD have tried to follow a middle way, including only what is deemed to be the most important information about the equivalents.

4.1. *Information on pronunciation*

While Lithuanian standard orthography for the most part is quite consistent and through a limited set of rules guides the learner to the correct pronunciation, there is one important aspect which is not reflected by the orthography, namely accentuation. The language has free and mobile accent, as well as an opposition between two tonemes in long accented syllables. Nominals and verbs are divided into several accentuation classes according to how the accent is distributed across their paradigms, and it is impossible to determine the accentuation class only from the basic form. In other words, learners wishing to read or speak Lithuanian correctly have to learn the accentuation class of each lemma separately.

Dictionaries to Lithuanian which do not provide information on the accentuation of the equivalents effectively force foreign users seeking this information to look elsewhere, for example in monolingual dictionaries. The NLD aims at amending this to some extent, by marking accentuation on equivalents. However, translations of Norwegian examples are not accentuated, and the accentuation classes are not provided on nominals and verbs. Thus, the Norwegian learner of Lithuanian will see how the equivalent is accentuated in the basic form (infinitive or nominative singular (masculine)). This information is not sufficient for him or her to be able to correctly accentuate all forms in the paradigm. However, if the user has some knowledge of Lithuanian accent patterns, the basic form will often provide enough information to form at least some of the other forms.

4.2. *Information on declension and conjugation*

For Lithuanian nouns, the declensional paradigm is for the most part predictable from the basic form in which the equivalents are given. One notable exception is formed by nouns ending in -*is*, which could be declined according to two different patterns – one masculine and one feminine (the latter being the less frequent of the two). Information on which pattern is the correct one is given in a fairly unobtrusive way in the form of a superscript F on feminine nouns in -*is*, like this: **slott** n1 pilìs^{F} (‘castle’).

When it comes to Lithuanian verbs, their conjugation patterns are much less predictable from the basic form (the infinitive). To give the user enough information to be able to form the whole paradigm, two other forms would in most cases have to be given in addition to the infinitive (or, possibly, a complicated system of grammatical codes could be developed to indicate conjugations). Although a similar system is followed for Norwegian verbs in the Lithuanian-Norwegian dictionary (Jakaitienė and Berg-Olsen 2001), the editors of the NLD decided against this, mainly on the grounds that it would reduce the user-friendliness of the dictionary for the largest group of users (that is for Lithuanian native speakers). However, in many cases the Lithuanian translations of Norwegian examples contain verb forms which will help Norwegian users form (part of) the paradigm for the verb in question.

4.3 *Information on valence and other syntactical properties*

Information on the valence of Lithuanian verbs and adjectives is very useful for Norwegians in the production of Lithuanian. This information is provided in parentheses next to equivalents. As noted in 3.6 above, information on valence is often also available from examples. Also certain other syntactic properties of lemmas are indicated, for example in cases where a Norwegian collective singular noun has as its equivalent a Lithuanian pluralis tantum: **hvètemel** -et kviečių̃ {kvietìniai} mìltai^{PL} (‘wheat flour’). Which syntactic properties of Lithuanian lemmas should be provided in dictionaries is discussed by Jakaitienė (2006).

5. Can a bidirectional dictionary be user-friendly?

User-friendliness in lexicography is a problematic concept, because a feature which enhances the user-friendliness of a dictionary for one group of users often reduces it correspondingly for other groups. This is especially acute in the case of bidirectional dictionaries. We have shown some relatively small concrete steps taken to make the NLD – a traditional paper dictionary – more user-friendly both for Lithuanian and Norwegian native speakers.

Generally, we are of the opinion that lexicographers should try to resist their often-felt urge to provide as much information as possible in as little space as possible. However, when an information type is deemed to be especially useful for a particular group of users, one should strive to include it, at least if it can be done in a relatively simple and unobtrusive way. Finally, it should be noted that electronic dictionaries provide considerably wider possibilities for increasing user-friendliness, in that one database can be presented through different interfaces for different groups of users. However, we must not forget that paper dictionaries still account for a large share of the bilingual dictionaries used and produced.

References

A. Dictionaries

Berkov, V. P. et al. 2003. *Stor norsk-russisk ordbok*. Oslo: Kunnskapsforlaget.

Griškevičienė, A. et al. (to appear). *Norvegų-lietuvių kalbų žodynas*. (NLD)

Jakaitienė, E. and S. Berg-Olsen 2001. *Lietuvių-norvegų kalbų žodynas / Litauisk-norsk ordbok.* Vilnius: Baltos lankos.

B. Other literature

Berkov, V. P. 1996. *Dvujazyčnaja leksikografija*. St. Petersburg: Izdatel'stvo S. Peterburgskogo universiteta.

Dahl, A. 1999. 'Lexinlexikonen og översättaren.' *Lexiconordica* 6: 47–53.

Dziemianko, A. 2006. *User-friendliness of Verb Syntax in Pedagogical Dictionaries of English*. (Lexicographica Series Maior, Volume 130.) Tübingen: Max Niemeyer.

Dziemianko, A. 2011. 'User-friendliness of Noun and Verb Coding Systems in Pedagogical Dictionaries of English: A Case of Polish Learners.' *International Journal of Lexicography* 24.1: 50–78.

Gellerstam, M. 1999. 'LEXIN – lexikon för invandrare.' *Lexiconordica* 6: 3–17.

Jakaitienė, E. 2006. 'Sintaksinių duomenų laukas aiškinamajame žodyne.' In R. Petrokienė et al. (eds.), *Aiškinamųjų bendrinės kalbos žodynų aktualijos*. Vilnius: Lietuvių kalbos instituto leidykla, 92–104.

Kernerman, A. 2008. 'Improving Dictionaries.' In E. Bernal and J. DeCesaris (eds.), *Proceedings of the XIII Euralex International Congress: Barcelona, 15-19 July 2008*. Barcelona: L'Institut Universitari de Lingüistica Aplicada, Universitat Pompeu Fabra, 1259–1263.

Nesset, T. and T. Trosterud 2005. 'Ny norsk-russisk ordbok: Ei leksikografisk storhending.' *Lexiconordica* 12: 273–284.

Svensén, B. 2004. *Handbok i lexikografi. Andra, omarbetade og utökade upplagan*. Stockholm: Norstedts.