

TOWARDS A SYNTAGMATIC BILINGUAL DICTIONARY FOR SPECIFIC PURPOSES

Alice Tihelková

Abstract

The article represents a contribution to specialized lexicography, a discipline dealing with the preparation of specialized dictionaries of various fields (law, business, engineering, medicine, etc.). The main objective is to introduce an original model of a bilingual (English – Czech) specialized dictionary based on the user-oriented approach. The model enables Czech users to consult a specialized dictionary not only for bare terms but also for their linguistic and encyclopaedic properties. The inclusion of linguistic and non-linguistic information renders the dictionary suitable for a variety of uses, including production in and translation into the foreign language.

1 Introduction

Recent years have witnessed an increase in the need for Czech-English/English-Czech specialized (terminological) dictionaries as a result of the spreading use of English among field experts, semi-experts as well as the general public. However, few of the LSP dictionaries available on the Czech market meet the requirements for a multi-purpose user-friendly reference work able to assist users in a variety of lookup tasks.

The article proposes a model of a specialized bilingual dictionary for Czech users which incorporates the principles of the monolingual learner's dictionaries (such as OALD, LDOCE, COBUILD, etc.), thus providing a wealth of information on the grammar and use of the specialized terms listed. The aim of the model (called "the syntagmatic dictionary") is to increase the user-friendliness of a bilingual LSP dictionary so that it enables users not only to receive target language texts, but also produce them with confidence. Although the model could be applied to any foreign language (with the necessary modifications), the article concentrates on English as the second language.

2 Shortcomings of current dictionaries

The discipline of specialized lexicography, defined as "the branch of lexicography concerned with design, production and evaluation of specialized dictionaries" (van Sterkenburg 2003: 414), appears to suffer from a considerable neglect in the Czech Republic; there exist almost no relevant studies on the subject (Tihelková 2006: 19). Moreover, few of the specialized dictionaries of

Czech origin are prepared by trained linguists able to assess true user needs – most are authored by field experts with some degree of knowledge of English. As a result of these as well as other facts (financial limitations due to the small Czech market for LSP dictionaries, etc.), the existing Czech-English and English-Czech dictionaries display a variety of shortcomings which could be summarized under the following three categories:

2.1 Disregarding the dictionary use

Two basic uses (functions) of a dictionary are recognized – passive use, consisting in the reception of the second language (L2), and active use, equalling the production in L2 (Bergenholtz & Tarp 2003: 24). It is obvious that the active function (corresponding to the Czech-English direction in our case) requires more information on grammar and usage than the passive function (English-Czech), as the user proceeds from the known into the unknown.

However, many of the Czech bilingual dictionaries fail to regard the two different uses, providing very little linguistic information even in the Czech-English word lists. Let us, for instance, consider the simple phrase *přípevnit přístroj na stojan*, which a student of cartography may wish to translate in his/her seminar paper in English (equally, the phrase could be encountered by a translator of a technical manual). The available Czech-English dictionary of cartography gives the following information:

přípevnit *fasten, fix* (Šíma 1993: 124)

Having obtained the two equivalents, we are left with a number of unanswered questions. First, are the equivalents absolutely interchangeable, i.e. is there total synonymy between them? Second, what kind of preposition do we use the verbs with – “to”, “on”, “onto”, “into”, etc.? And, moreover, is the preposition the same for each of the equivalents? Adding at least the preposition, if not a collocation or a usage example, would not burden the dictionary to any considerable extent, and the user could avoid having to search the valency in a general-purpose dictionary.

2.2 Disregarding the dictionary user

When compiling a dictionary, the lexicographer should keep in mind what user type (experts, semi-experts, non-experts) the dictionary is designed for, and plan the dictionary structure accordingly (Bergenholtz & Tarp 1994: 173). It is only logical that non-experts and semi-experts will require more information (collocations, usage notes, labels, etc.). However, although many of the Czech

bilingual LSP dictionaries claim in their prefaces to be intended not only for experts but also for students, translators and the interested general public, collocations and other user-friendly devices (including, among others, phonetic transcription to ensure correct pronunciation of the specialized terms) are missing in their microstructures. This disregard for the intended dictionary user can lead to reception and translation problems.

2.3 Lack of general lexicographic expertise

The third type of shortcoming consists in the lack of the necessary dictionary-making skills as well as the mastery of the English language, mostly caused by the non-involvement of a trained linguist on the lexicographic team. Examples of this shortcoming include incorrect or redundant grammar information, inconsistent use of typographical devices or insufficient treatment of equivalence. This last type is less relevant for the present argumentation than the previous two.

To illustrate the shortcomings outlined, an entry from an English-Czech dictionary of motoring can be considered, combining all the three types:

change up – řadit nahoru
change-over – přepínání
change-over – přeplnutí
change-over valve – přepínací ventil
change-over valve – řadící ventil
changing – řadit rychlostní stupeň
channel – šachta
channel – kanál (Vlk 2001: 211)

Despite representing a rather extreme case of user-unfriendliness, the dictionary's faults (no information on grammar and use, over-simplistic macrostructure, incorrect handling of equivalence, illogical typography, etc.) are typical of a large number of Czech specialized dictionaries, especially those dealing with technical fields.

The above-discussed shortcomings were confirmed by dictionary user research we conducted in 2005. The research task involved 26 university students of ESP who were translating a specialized text using a standard Czech bilingual dictionary and then answering a number of questions about the translation process and the usefulness of the dictionary provided. The majority of the students stated some difficulty in translating the text due to the lack of linguistic information in the dictionary, most notably valency (38%), collocations (27%) and usage examples (26%). Other suggestions for the improvement of the dictionary included pronunciation, synonyms and the use of pictures. For a more detailed treatment of the research experiment see Tihelková (2006: 87-89).

3 The model of the syntagmatic dictionary

Although some recently published specialized dictionaries (mostly of humanities, authored or co-authored by a linguist) show some progress towards user-friendliness and an increased sensitivity to user needs, we intend to take the user-oriented approach further in presenting a new model of a bilingual specialized dictionary based on the explicit as well as implicit provision of linguistic and non-linguistic information. Inspiration for this model has been sought in monolingual learner's dictionaries, as they represent the absolute peak of lexicography today, combining the use of state-of-the-art computer technologies with the latest knowledge and methods of linguistics (including the use of large corpora). To illustrate the complexity and comprehensiveness of these reference works, example entries dealing with the same headword are shown below, each from a different learner's dictionary:

feed /fi:d/ *verb, noun*

■ *verb* (fed, fed /fed/) GIVE/EAT FOOD 1 ~sb/sth (on) sth— ~sth to sb/sth to give food to a person or an animal: [VN] *Have you fed the cat yet?* [...] ◇ [VNN, VN] *The cattle are fed barley.* ◇ *The barley is fed to the cattle.* 2 [V] (of a baby or an animal) to eat food: *Slugs and snails feed at night.* (Hornby: 2000, 464)

feed¹ /fi:d/ v **fed** /fed/ 1 [T (on, with)] to give food to: *We have to feed 120 guests after the wedding.* [...] 2 [I (on)] (esp. of an animal or a baby) to eat: *The horses were feeding quietly in the stable.* – Cows feed on grass.

3 [T+obj+adv/prep] to put, to supply, to provide; especially continually: *Keep feeding the wire into/through the hole.* (LDOCE: 1994, 373)

When applying the learner's dictionary approach to Czech bilingual *specialized* dictionaries, some modifications will need to be made with respect to their specific purpose. First, instead of definitions, there will be translation equivalents. Definitions remain an optional element (space permitting), the user being far more interested in the translations of the given terminology into the target language.

Second, the amount of linguistic and non-linguistic information will also differ from the learner's dictionary due to the fact that specialized language is being described. It is only logical that the LSP dictionary cannot be as exhaustive in its treatment of the linguistic properties of terms as the general dictionary. For instance, it is very difficult to imagine a specialized dictionary formally indicating all the individual valency patterns as seen in the examples above; indeed, some authors writing on specialized lexicography argue against the inclusion of any valency information in terminological dictionaries (Opitz 1999: 163). Nevertheless, for a specialized dictionary to be truly user-friendly we propose the incorporation of the linguistic information items listed below.

3.1 SPELLING INFORMATION, comprising three basic information types: spelling variation (e.g. *arrowhead/arrow-head*) regional spelling differences (e.g. BrE *honor*/AmE *honour*) and specific spelling features (e.g. *acquit – acquitted*).

3.2 PRONUNCIATION. Although very rare in specialized dictionaries, there are a number of reasons why its inclusion is recommendable. A great deal of LSP vocabulary is highly technical and the pronunciation cannot be found in general dictionaries. In addition, LSP dictionaries are increasingly used for active purposes (including the production of spoken texts), the knowledge of pronunciation thus aiding LSP communication. In English-Czech/Czech-English dictionaries, two systems of indicating pronunciation come into question:

a) the “czechisized” version – Czech characters are used where possible while the sounds not present in Czech are borrowed from the International Phonetic Alphabet (IPA).

b) the exclusive use of the IPA, similar to learner’s dictionaries. For the appropriateness of either type see Tihelková (2006: 116-117).

3.3 USAGE LABELS, conveying information on deviations from the unmarked majority of the lemma stock found in the dictionary. The most common labels comment on regional usage (*AmE*, *BrE*, etc.), formality (*fml*, *infml*) and specialized field (*med.*, *arch.*, *zool.*, etc.), but there can also be other label types (Bergenholtz & Tarp 1994: 132).

3.4 GRAMMAR INFORMATION. In dictionaries of English, this information should be provided for the categories of the noun and the verb where needed (although the lexicographer may also wish to extend it to the categories of adjectives and adverbs). In nouns, the crucial information category will be number, most likely to be indicated by stating the irregular plural of terms or by showing the label *pl.* with pluralia tantum as well as any other terms listed in the plural form:

filamentous algae (pl.) vláknité řasy (Hájková et al. 1998: 143)

A category that is almost never indicated in Czech bilingual dictionaries is countability. Yet, for an active dictionary intended for the production of texts, its inclusion may prove beneficial due to its relatedness to the use of articles. Countability may either be stated by some Czech indicators (e.g. *poč./nepoč.*) or, preferably, by the well-established English abbreviations *C/U* as seen in the following example:

bacteria /bæk'tɪəriə/ n[C] pl (sg bacterium /-iəm/) bakterie
(Tihelková 2006: 163)

On the other hand, the categories of gender and determination will probably not bring the user greater benefit when indicated in a LSP dictionary, except for some very specific instances (e.g. *the police*, etc.).

In verbs, it is recommended to explicitly state the categories of tense (in irregular verbs) and valency. While the former is more or less straightforward, the latter deserves some elaboration. The information on valency can be presented either explicitly, using a separate information item, or implicitly by means of usage examples. Probably the most space-saving way of stating valency information in an explicit manner is using simple valency formulae which can or need not be accompanied by information on transitivity (to be marked *I/T* for intransitive and transitive verbs, respectively), e.g.:

attach /ə'tætʃ/ **v[TT]** ~st to/on st připojit, připevnit **co k čemu/na co**
(Tihelková 2006: 163)

Space permitting, the most user-friendly solution is to combine the explicit information category with implicit exemplification as featured in a majority of learner's dictionaries. However, in dictionary projects that do not allow the inclusion of usage examples, stating the most basic valency patterns should be sufficient in enabling the user to produce target language syntagms containing the verb in question. Naturally, valency information need not be restricted to verbs only – prepositional patterns can also be stated for nouns and adjectives.

3.5 SYNTAGMATIC INFORMATION, in the form of collocations related to the individual terms listed. Collocations represent a crucial information item in both passive and active dictionaries, especially where semi-expert and non-expert users are anticipated. Collocations can be indicated implicitly in usage examples or form an explicit information category of its own, as seen below:

- bit** *m* bit (binary digit)
- ~, **adresové značky** address bit mark
- ~, **bezdatový** non-data bit
- ~, **doplňkový** *tlgr.* overhead bit [...]
- **bitově orientovaný** (*terminál, formát*) bit-oriented
- **počet bitů za sekundu** bits per second (bps)
- **prokládání po bitech** bit interleaving (Malinová 1993: 493)

3.6 PARADIGMATIC INFORMATION, primarily in the form of synonyms (optionally also antonyms). This information can be addressed to the lemma as well as the equivalent (Nielsen 1994: 277). Synonyms addressed to the equivalent are an integral part of most Czech bilingual dictionaries. On the other hand, synonyms addressed to the lemma are provided very rarely. However, their inclusion can only be recommended, as terminologies are not directories of isolated words, but display a network of relations, of which synonymy is one. The ways of marking synonymy addressed to the lemma are numerous, as seen on the example of the entry term *hand-feeding*, which has a *synonym hand-rearing*:

hand-feeding ruční dokrmování **hand-rearing**
hand-feeding SYN ruční dokrmování **hand-rearing**
hand-feeding/hand-rearing ruční dokrmování
hand-feeding *též* **hand-rearing** ruční dokrmování
hand-feeding ruční dokrmování = **hand-rearing** (Tihelková 2006: 127)

3.7 USAGE EXAMPLES, representing a linguistic information item that can help convey implicit information on the morphological, syntactic, pragmatic and, possibly, encyclopaedic properties of a lemma (Schaefer & Bergenholtz 1994: 422). They have a much broader scope than collocations, which primarily convey syntagmatic information and are simpler in structure (although the two information items can overlap to some degree and combine in one dictionary entry). Most often, usage examples will have the form of complete sentences, although they can be presented as phrases as well.

Usage examples are an integral part of foreign learner's dictionaries. Analogically, they can be recommended for inclusion in Czech LSP dictionaries where viable, the reason being their helpfulness regarding production in and translation into the second language, as seen in the following extract from an English-Czech business dictionary, concerning the verb *agree*:

agree [...]
the accounts a. účty souhlasí
the figures were a-d údaje byly schváleny
the statement does not a. with the facts prohlášení neodpovídá skutečnosti
(Kalina et al. 2001: 45)

In simpler dictionaries where space-saving may be a concern or where the author is a field specialist rather than a linguist, the inclusion of usage examples may prove beyond the means of the project. In such instances, care should be taken to present the relevant information explicitly within the other information items included, i.e. morphological information, valency, collocations, etc.

3.8 ENCYCLOPAEDIC INFORMATION. In a bilingual LSP dictionary, encyclopaedic information serves two distinct purposes: meaning discrimination and the provision of additional information. The former usually comes in the form of brief explanatory notes (from one or two words to one or two sentences) to specify the meaning of highly technical, polysemous or culture-dependent terms. The latter goes beyond the purpose of a bilingual dictionary, offering subject-related information in the form of definitions, citations from ISO norms or other LSP literature, etc.

The proposed syntagmatic dictionary regards the inclusion of encyclopaedic information for meaning discrimination as user-friendly and thus recommendable, while the "additional" function is considered highly optional and, in fact, unlikely

to appear in Czech specialized dictionaries in any case. The significance of the meaning-specifying explanatory notes can be illustrated by the example shown below:

lord [...] **L~s Temporal** brit. svěťští lordi, lordi Sněmovny lordů, kteří svůj titul dědí nebo je jim udělen po dobu jejich života
(Chromá 1995: 189)

Without the explanatory note, the term “svěťští lordi” would remain unclear to anyone but a user well-acquainted with the British political system.

Having discussed the contents of the proposed syntagmatic dictionary in general terms, let us now demonstrate the points raised on a concrete dictionary project. As there is currently no Czech-English/English-Czech specialized dictionary comparable in format to the English learner’s dictionaries and/or containing all of the information categories stated above, an original mini-dictionary was designed to illustrate the individual aspects of the model. The specialized subject of the dictionary is the discipline of parrot-keeping (psittaculture) due to the author’s lifetime involvement in keeping and breeding exotic birds. Two extracts from the dictionary will be shown, one for each direction. Let us first consider a number of entries from the English–Czech direction:

clutch /klatʃ/ **n[C]** 1 snůška vajec 2 mláďata z jedné snůšky = **brood**
lay a clutch of fertile eggs snést snůšku oplozených vajec; *the ave rage clutch size is 4-6 eggs* průměrná velikost snůšky je 4-6 vajec;
the number of eggs per clutch počet vajec na snůšku
cock /kɔk/ **n[C]** samec, sameček u ptáků ≠ **hen**
use a young Rosella cock for breeding použít mladého samečka rozely k chovu
Congo African Grey /'kɒŋəʊ 'æfrɪkən greɪ/ žako šedý kongo poddruh žaka šedého *Psittacus erithacus erithacus*
command¹ /kə'mɑ:nd/ **n[C]** povel, rozkaz
fly to one’s arm on command přiletět na povel na ruku
command² /kə'mɑ:nd/ **v[T]** ~sb to do st přikazovat, dávat povel
command the bird to step up přikázat ptákovi, aby nasedl na ruku
crop /krɒp/ **n[C]** *anat.* vole
empty the crop vyprázdnit vole; *suffer a crop burn* utrpět popálení volete

From the short extract, the above-discussed user-friendly devices are apparent: pronunciation, grammar information (countability, valency), syntagmatic information (collocations/usage examples), paradigmatic information (synonyms, antonyms) and encyclopaedic information (field labels, explanatory notes, Latin names of species). A similar model can be applied to the opposite direction, as seen below on a different extract:

přilnout ke komu bond /bɔnd/ **v[I/T]** ~ with/to sb
mládě přilne k člověku a baby bird will bond to a human; *pár k sobě ihned přilnul* the pair have bonded immediately
psitakóza med. psittacosis /,psitə'kəʊsɪs/ psitakóza **u[U]** = **chlamydióza**
psitakóza je přenosná na člověka psittacosis is transferable to humans
rozela Pennantova Crimson Rosella /krɪmzn rɔ'zelə/ *Platyercus elegans*
ručně dokrmovat mláďata handfeed /,hænd'fi:d/ (handfed –handfed /fed/) **v[T]**, hand-rear /,hændr'ɪə/ **v[T]**
ručně dokrmovat mláďata injekční stříkačkou/lžičkou hand-feed babies with a syringe/spoon; *ručně dokrmený pták je vázán na člověka* a handreared bird is bonded to a human

Naturally, the entry type shown above only represents an “ideal construct”; the possibilities of its modification or, indeed, improvement, are almost endless, depending of the nature of the lexicographic project in question. The expertise of the author(s), the material resources, the expected users, the space-saving concern, the number and scope of the LSP fields treated – all these factors will have an impact on the viability of the syntagmatic model.

To test the usefulness of the proposed model, a modest user survey experiment was conducted, drawing on the above-described research from 2005. The format remained the same: ten students of ESP were asked to translate a specialized text (dealing with aviculture this time, with respect to the subject of the model dictionary) and report on the difficulties arising from the task as well as on the un/helpfulness of the dictionary provided.

As opposed to the earlier experiment, where a great deal of user dissatisfaction emerged, an overwhelming majority of the participants found the syntagmatic dictionary helpful. Nine students graded it as “very good”, one as “good”. Especially the inclusion of collocations was welcome by the students, while no major suggestions for improvement were made, the information offered by the dictionary being described as “sufficient” (in the 2005 experiment, the suggestions were plentiful).

The only difficulty having emerged was the confusion of one student over the abbreviation “**v[I]**” for intransitive verbs; the participant concerned regarded it as the indication of an irregular verb and demanded the corresponding irregular forms. However, this can be put down to the common reluctance of users to read the user’s guide containing the explanations of the symbols employed. A number of students stated that the source of their translation difficulty had been their imperfect English (not the dictionary’s fault). In any case, the survey can only be regarded as preliminary; more extensive research is needed to test the strengths and weaknesses of the syntagmatic model, its individual information items, the indicators of grammatical information, etc. This represents our future work.

4 Conclusion

In conclusion, it can be stated that despite the possible constraints regarding the full implementation of the syntagmatic dictionary in the Czech environment, the model remains an attempt at increasing the helpfulness of Czech LSP dictionaries by presenting specialized terms in use. As a result of the proposed increase in the amount of linguistic and non-linguistic information (pronunciation, irregularities, countability, valency, the provision of collocations or usage examples, the inclusion of explanatory notes in highly technical or culture-dependent terms, etc.), LSP dictionaries could play a far more active role than they have done so far, aiding not only reception but also production of both written and spoken specialized texts in English or any other foreign language. The benefit of such multi-purpose dictionaries to field experts, students, translators and interpreters as well as the interested general public would be obvious.

It is to be hoped that any subsequent debate on the model of the syntagmatic dictionary as presented above will facilitate its move towards yet greater user-friendliness.

Acknowledgments

The funds for the research have been provided by the development project MŠMT 281 (POSTDOC-07).

References

- Bergenholtz, H., Tarp, S. (1962) *Manual of Specialized Lexicography. The Preparation of Specialized Dictionaries. Words*. Amsterdam: John Benjamins.
- Chromá, M. (1995) *Anglicko-český právníký slovník*. Praha: LEDA.
- Hornby, A. S. (2000) *Oxford Advanced Learner's Dictionary of Current English*, Cowie, A. P. (ed.) Oxford: Oxford University Press.
- Longman Dictionary of Contemporary English* (1994) New Edition. Harlow: Longman.
- Malinová, L. (1993) *Anglicko-český a česko-anglický elektrotechnický a elektronický slovník*. Brno: SNTL.
- Nielsen, S. (1994) *The Bilingual Dictionary. Principles and Practice for Legal Language*. Tübingen: Narr.
- Opitz, K. (1999) 'The Terminological/Standardized Dictionary.' In: Hartmann, R. R. K. (ed.) *Lexicography: Principles and Practice*. London: Academic Press. 163-180.
- Schaeder, B., Bergenholtz, H. (1994) *Fachlexikographie. Fachwissen und seine Repräsentation in Wörterbüchern*. Tübingen: Narr.
- Sterkenburg, P., van, (ed.) (2003) *A Practical Guide to Lexicography*. Amsterdam: John Benjamins.
- Šíma, J. (1993) *Česko-anglický slovník pro geodety a kartografy*. Zdíby: VÚGTK.
- Vlk, F. (2001) *Anglicko-český odborný slovník motorových vozidel*. Brno: Nakladatelství Vlk.
- Tihelková, A. (2006) *Theory and Practice of Czech Bilingual Specialized Dictionaries*. Unpublished Ph.D. dissertation. Praha: Univerzita Karlova.