
The development of scholarly lexicography of the Estonian Language as a Second Language in an historical and a theoretical perspective

Jelena Kallas

Institute of the Estonian Language, Tallinn University, Estonia

This paper aims to provide an overview of the development of scholarly lexicography of the Estonian language as a second language in an historical and a theoretical perspective. The paper describes what kind of information is presented traditionally in dictionary entries on the level of morphology, derivation, syntagmatic relationships and paradigmatic relationships. In addition, taking into consideration theoretical and practical viewpoints of modern lexicography on what kind of information should be presented in a dictionary entry so that the dictionary could be classified as a production dictionary (Apresjan (ed.) 2006; Atkins & Rundell 2008; Bo Svensén 2009; Novikov 2001; Siepmann 2006), the author is going to illustrate what kind of information should be added into the entries of a learners' dictionary of the Estonian language as a second language so that they could be used as production dictionaries.

In an historical perspective the analysis of the learners' dictionaries, which were published during the last 160 years, indicated that dictionary compilers provide dictionary users mostly with information about inflectional formation; meanwhile, the information about word formation (derivatives, compounds), syntagmatic and paradigmatic relationships is almost neglected. On the other hand, learners' dictionaries meant for speakers of Estonian as a first language provide much more information: the information about inflectional formation, word formation, synonyms, antonyms, paronyms is presented explicitly. The information about syntagmatic relationships is presented mostly implicitly by means of examples at the level of phrases, clauses and sentences.

The author puts forward detailed proposals for what kind of formal (inflectional formation, derivatives, compounds), semantic (mostly content-paradigmatic information) and syntagmatic (syntactic valency, collocations, idioms) characteristics should be given in a dictionary of the Estonian language as a second language and demonstrates practical implementations of explicit systematic description of syntactic valency and collocations of different parts of speech (nouns, adjectives, adverbs, verbs, quantifiers).

1. Introduction

The main objects of the study are semasiological Estonian-Russian learners' dictionaries meant for speakers of Estonian as a second language. Additionally, monolingual learners' dictionaries meant for speakers of Estonian as a first language were investigated.

The first aim of the study was to analyze and get a systematic overview of what kind of formal, syntagmatic and semantic information is offered in the dictionary entries and how the amount and type of information has been changing diachronically. The second aim was to characterize dictionaries from the typological point of view according to their functional characteristics: whether they can be classified as production dictionaries or as reception dictionaries. As a result of this analysis, it became possible to make practical suggestions on what kind of information should be provided for learners' dictionaries meant for speakers of Estonian as a second language so that they could be used as production dictionaries.

The main preconditions for the development and formation of Estonian-Russian scholarly lexicography were the translation activities during the Protestant Reformation period beginning in the mid-1520s, the publication of the first Estonian language grammar in the year 1637 (Stahl 1637) and the compilation of German-Estonian dictionaries in the XVIII and XIX centuries (Hupel 1780, Wiedemann 1869). The first known Estonian-Russian dictionary came into existence in the year 1853 (Lunin 1853). As a basis for the data of the dictionary the author has assumed the data and compilation principles presented in Hupel's (1780) dictionary. This fact allows the assertion that the formation of Estonian-Russian scholarly lexicography is rooted in the Baltic German lexicographic traditions.

To achieve the aim of the study nine scholarly Estonian language dictionaries were investigated: six Estonian L2 scholarly dictionaries (Lunin 1853; Kuusik 1916; Käbin, Pekarsky 1923; Neumann 1925; Stalnuhhin, Välja 1998, EVÕ 1984) and three Estonian L1 scholarly dictionaries (ÕpÕS 2004; TEA 2004; TEA 2008).

Below, the analysis of the kind and manner of formal, syntagmatic and semantic information given in analyzed dictionaries is presented.

2. Inflectional information

The presentation of inflectional information in general Estonian language dictionaries has been analyzed by different Estonian linguists (Langemets et al. 2005: 79–84, Viks 2000: 486–495, Viks 2008: 258). For general dictionaries of the Estonian language it is typical to present inflectional information as outside matter by means of systematic surveys in the form of model paradigms with numeral codes.

Since the Estonian language has a complex system of nominal inflection (e.g. nouns and adjectives decline in fourteen cases), this type of information is presented in all analyzed scholarly dictionaries as well. As a rule, inflected forms of lexemes presented under lemma form explicitly, but the amount of information varies from some forms to full form database.

3. Word formation (the presentation of derivatives and compounds)

There is no general system for presenting information about derivatives and compounds of lemma in the Estonian language lexicography.

Among Estonian L2 scholarly dictionaries there were no dictionaries which would explicitly present this kind of information.

On the other hand, this type of information is presented in Estonian L1 dictionaries: compounds and derivatives are presented as sublemmas of lemma and they are highlighted by means of a specific structure indicator or as individual entries with a component-internal cross-reference to a connected word.

4. Syntagmatic information (syntactic valency and collocations)

At the moment there are no syntagmatic specialized dictionaries of Estonian language which would provide syntactic and semantic valency of lexemes of different parts of speech. There is only one specialized dictionary which explicitly provides information about the verbal government (Pool 1999).

As Bo Svensén (2009: 143) states, information about constructions is an important element in most types dictionary, but above all in dictionaries intended for production (in the native language as well as in a foreign language) or for translation into a foreign language, i.e. in L1 dictionaries and L1→L2 dictionaries.

A brief overview of the kind and manner of syntactic information in Estonian language dictionaries is presented in the article ‘Syntactic information in dictionaries: problems and solutions’ (Langemets et al. 2005). The authors emphasize that syntactic constructions have been described in several Estonian dictionaries, but mostly implicitly, by means of case

examples, as no general syntactic encoding system has yet been elaborated (Langemets et al. 2005: 72).

The same situation occurs within Estonian L1 and L2 scholarly dictionaries: all relevant information is presented mostly implicitly on the level of examples. Probably the reason is behind the general attitude that grammatical information should be presented in grammar textbooks, not in dictionaries.

The analysis of Estonian language grammar textbooks and studies (EKG II 1993: 8, 114-119, Rätsep 1978: 64, Langemets et al. 2005: 91–93, Kerge 2000: 18–19, 22-44, Muischnek 2006) indicated that from a grammatical point of view, in order to provide the users of the Estonian L2 dictionary with all necessary constructional and collocational information, there is a need to include into the dictionary entries the following information:

- for noun entries – case, adposition and infinitive government and *noun+noun*, *adjective+noun*, *noun (subject)+verb* and *noun (object)+verb* collocations;
- for adjective entries – case, adposition, infinitive government and *adverb+adjective* collocation;
- for verb entries – object, case, adposition, infinitive, utterance government and *verb+adverb* collocations;
- for adverb entries – case government and *adverb+adverb* collocations;
- for quantifier sentries – case government.

5. Semantic relationships

In Estonian scholarly lexicography there is no tradition to present semantical relationships explicitly in the individual entries. The information about hyperonyms, hyponyms and synonyms typically occur in the definition format and the information about antonyms is mostly neglected. There are also some dictionaries which present antonyms as individual entries with cross-references to connected words, e.g. TEA 2008.

The study of Estonian L2 dictionaries showed that dictionaries provide mostly inflectional information and Russian equivalent. Syntagmatic and semantic information, which play a significant role in the second language acquisition process and production of grammatically correct Estonian texts, is almost neglected. For that reason, analyzed Estonian L2 dictionaries can be classified mostly as reception dictionaries. They do help to understand Estonian text, but they do not offer sufficient amount of information needed for production of Estonian text. Diachronically, the amount of information on the microstructural level did not change significantly: mostly inflectional information was presented explicitly.

On the other hand, learners' dictionaries meant for speakers of Estonian as a first language provide much more information: the information about inflectional formation and word formation is presented explicitly. The information about syntagmatic relationships is presented mostly implicitly by means of examples at the level of phrases, clauses and sentences.

The analysis of the information, which is typically presented in Estonian L2 dictionary entries, indicated that lexicographers of the Estonian language as a second language need to elaborate new format for explicit presentation of word formation, syntagmatic and semantic information in dictionary entries.

References

Dictionaries

- [EVÕ 1984] *Eesti-vene õppesõnastik* comp. by Kaarma, G.; Külmoja, I.; Lagle, T.; Leemets, H. et al. Tallinn: Valgus, 1984.
- [Hupel 1780] Hupel, A. W. *Ehstnische Sprachlehre fuer beide Hauptdialekte, den revalschen und doerptschen; nebst einem vollstaendigen Woerterbuch. Herausgegeben von August Wilhelm Hupel.* Riga und Leipzig.
- [Kuusik 1916] Kuusik, T. *Keeletark. Eestlastele vene keele õppimiseks ja pruukimiseks igapäevases elus. Venelastele eesti keele õppimiseks ja pruukimiseks igapäevases elus.* Revel, 1916.
- [Käbin, Pekarsky, 1923] Käbin, T.; Pekarsky, N. *Eesti-vene sõnastik grammatiliste vormide näitamiseks.* Narva, 1923.
- [Lunin 1853] *Эстонско-русский словарь, составленный по Гупелю Иваномъ Лунинымъ.* Дерптъ, 1853.
- [Neumann 1925] Neumann, M. *Eesti-Wene sõnaraamat. M. Neumanni süstemaatilised sõnaraamatud.* Tallinn, 1925.
- [Pool 1999] Pool, R. *Eesti keele verbirektsioone.* Tartu, 1999.
- [Stalnuhhin, Välja 1998] Stalnuhhin, M.; Välja, E. *Eesti-vene õppesõnastik: 10 800 sõna.* Narva: ElMi, 1998.
- [TEA 2001] *Eesti keele raskete sõnade sõnaraamat: kuidas sõnu õigesti kirjutada ja kasutada?* comp. by Mägi, R.; Raudvere, K.; Vaba, M. Tallinn, 2001.
- [TEA 2008] *TEA koolisõnastik: eesti keel* comp. by Koit, K.; Kütt, A.; Maandi, E.; Mets, M.; Mägi, R.; Sagen, A. et al. Tallinn.
- [ÕpÕS 2004] *Õpilase ÕS.* 2004. Comp. by Erelt, T.; Leemets, T.. Tallinn: Eesti Keele Sihtasutus.
- [Wiedemann 1869] Wiedemann, F. J. *Ehstnisch-deutsches Wörterbuch.* St. Petersburg: Kaiserliche Akademie der Wissenschaften.

Other references

- [Aprésjan (ed.) 2006] Апресян, В. Ю.; Апресян Ю. Д. et al. (2006). *Языковая картина мира и системная лексикография.* Москва.
- Atkins, B. T. S.; Rundell, M. (2008). *The Oxford Guide to Practical Lexicography.* Oxford: Oxford University Press.
- Bo Svens én (2009). *A Handbook of Lexicography. The Theory and Practice of Dictionary-Making.* Cambridge: Cambridge University Press.
- [EKG II 1993] Erelt, M.; Kasik, R.; Metslang, H.; Rajandi, H. et al. (1993). *Eesti keele grammatika II.* Tallinn: Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut.
- Kerge, K. (2000). *Eesti keele süntaks võõrkeeleõppe praktikule.* Tallinn.
- Langemets, M.; Mägedi, M.; Viks, Ü. (2005). 'Süntakiline info sõnastikus'. In *Eesti Rakenduslingvistika Ühingu aastaraamat 1 (2004).* Tallinn.
- Muischnek, K. (2006). *Verbi ja noomeni püsiühendid eesti keeles.* Tartu.
- [Novikov 2002] Новиков, Л. А. (2002). 'Типология учебных словарей'. In *Современный русский язык: Лексикология. Фразеология. Лексикография.* Санкт-Петербург: Санкт-Петербургский государственный университет.
- Rätsep, Huno (1978). *Eesti keele lihtlausete tüübid.* Tallinn: Valgus.
- Siepmann, D. (2006). 'Collocation, Colligation and Encoding Dictionaries. Part II: Lexicographical Aspects'. In *International Journal of Lexicography* 19. 1-39.
- Stahl, H. (1637). *Anführung zu der Estnischen Sprach auff wolgemeinten Rath und bittliches Ersuchen / publiciret von M. Henrico Stahlen.* Reval.
- Viks, Ü. (2000). 'Kuidas tekib sõnastikukirjesse grammatika'. In *Keel ja Kirjandus* 7. 486-494
- Viks, Ü. (2008). 'Eesti-X-keele sõnastik ja grammatika'. In *Eesti Rakenduslingvistika Ühingu aastaraamat IV.* Tallinn: Eesti Keele Sihtasutus.