Tiina Leemets, Institute of Language and Literature,
Estonian Academy of Sciences

# A link between language planners and language speakers: a dictionary

ABSTRACT: This paper deals with the Estonian language planning diction-
ary, especially with the problem how to forward language planners' cur-
rent positions to language speakers as clearly and concisely as possible.
A seeming inconsistency may serve the purpose. It also describes prob-
lems in dictionary-making that arise from political and economic changes
in the society.

I would like to share some thoughts and problems of the Estonian language planning[1] dictionary compiling team. That dictionary is being made in the Estonian Institute of Language and Literature since 1988, with Tiiu Erelt as the editor-in-chief. Over 1/2 of the manuscript is ready by now. In a way our dictionary is the successor of ÕIGEKEELSUS-SÕNARAAMAT, the orthological dictionary published in 1976, edited by Rein Kull and Erich Raiet. However, the construction of these two is different. Our aim is to give less headwords (the orthological dictionary has about 120000 of them) and more commentaries. The basic features of this type of dictionary were worked out by Henn Saari in 1983/84 (he presents them at this Euralex congress). A considerable amount of his ideas have been used.

## 1. Choice of material. Headwords

The compilers have been left a rather free hand in choosing the material, including the headwords, and in building up the entries. All the more we must follow the principle: present the information which the user wants to know most of all, using as little space as possible. In his Instruction to the compilers H. Saari has said:

> The one-volume dictionary of the present-day Estonian literary language is meant for a wide range of users who write, edit, correct and read texts in Estonian: fiction, newspapers, hand-books, handwritten works in schools and universities etc.

To guess the needs and wishes of this kind of readers, we use two card files. Since 1947 it has been possible to get linguistic consultations in our institute, mostly by telephone. (At

this moment everyone who is making our dictionary is giving consultation as well.) Since 1966 all questions and answers have been registered and though this file contains much accidental material, it still shows what the main problems have been in different periods and what the current problems are. Another file contains excerpts, mostly from newspapers and magazines. They have been written out for more than 10 years to show developments in language, appearance of new words and meanings, and the most wide-spread deviations from linguistic standards and rules.

No list of entry words has previously been made, each compiler must make his or her choice mainly on the basis of the orthological dictionary. For instance, very rare foreign words and narrow special terms are left out. Of course, these are just the words that the reader wants to look up in a dictionary, but a lexicon of foreign words and numerous terminological dictionaries have been published, and they can give more adequate infor-mation. We also leave out most of the verbal nouns. The user can find information about how to form them in the morphological appendix of the dictionary under each pattern verb. Generally every language speaker can form them automatically by the model: laul/ma 'to sing' – laul/mine 'singing' – laul/ja 'singer'. In this case though *laulja* (singer) is a separate entry because it is a part of several compounds. One or two com-pounds formed with verbal nouns can be shown under the verb itself: ihu/ma 'to whet', ihu/mis‿kivi(< ihu/mine + kivi) 'whetting stone'. Even an agent noun can be in the entry of the action noun. E.g. *kulla‿kaevandamine* 'mining gold' is a headword. The verb here would be *kulda kaevandama*, which is not treated separately. Under *kulla‿kaevan-damine* we find *kulla‿kaevur* 'gold miner', that is more common than the automatically formed *kulla‿kaevandaja*. Verbal nouns are separate headwords when their meaning has to be specified or when they are not formed by the automatic model (näge/ma 'to see', but nägi/ja 'one who sees or can see').

The 1976 orthological dictionary contains a large number of compounds. In some cases this has left an impression that if a certain compound cannot be found in the dictionary, it does not exist, and therefore it has to be written as separate words. How-ever, it is impossible to write down all Estonian compounds in any dictionary. So we choose for headwords only those "about which there is something to say". Many com-monly known, automatically formed compounds with one obvious meaning are left out altogether and we give a short selection of others under the attributive or modified noun.

In addition to the vocabulary from the orthological dictionary, we give neologisms of recent years, most of which have been published in Tiiu Erelt's VÄIKE UUDISSÕNASTIK (Little dictionary of neologisms, Tallinn, 1989) and/or are registered in the two files I mentioned. We also add more colloquial words and slang, with a corresponding usage label of course. They can be found from the manuscript of the defining dictionary of literary Estonian, which the authors kindly let us use, and even from our own language practice. We want to stress the fact that literary language does not have to be dry and official, and to encourage the reader to compare different styles and use them flexibly. Among others there is a special usage label *paberlik* 'bookish, clerky' to mark a group of words and expressions which should be avoided in most texts; we give "more humane" synonyms instead. Unfortunately the present-day Estonian mass media give the im-pression that many authors either have a very dry and clerky style or they try to be very popular and forget all language rules. Of course we do not know if such a writer ever opens a dictionary or cares about what is written in it.

Quite a few headwords are inadvisable, in the sense that they should be avoided altogether. These are mostly direct (and sometimes distorted) borrowings from Russian, lately also from English and Finnish. In that case we refer to the advisable form, either categorically (the headword is in brackets) or preferably, with the word *parem* 'better'.

One could say that negative examples should be left out of a dictionary altogether, but we consider it important to present them, because in many cases this inadvisable word may be the only one the user has heard and he does not know where to look for the correct or better equivalent(s). But we can leave out some words which the 1976 dictionary has declared inadvisable, because they have lost their actuality.

A word combination can also be a separate entry. Thus we present some widespread phrasal verbs which can be written in one or two words depending on the form, set phrases taken from other languages and occasional combinations when it is a common error to write them as one word or with an unnecessary hyphen.

A headword may exist only as part of a compound. For example ‿*istmeline* is used in practice only in compound words like *nelja‿istmeline* 'with four seats', *pehme‿istmeline* 'with soft seats' etc. Some widespread word parts of Latin or Greek origin are presented this way. There are no words like ‿*loogia* or ‿*droom* in Estonian but there are such headwords in our dictionary, followed by compound examples like *farmako‿loogia* 'pharmacology', *aero‿droom* 'aerodrome'. Our aim in giving such rear components is to show their inflection in one place only. First components can also be headwords. There is a group of international word parts ending mostly with -o that make compounds with word parts of Latin or Greek origin as well as purely Estonian words. E.g. the word *narko* exists separately only in slang meaning 'drug addict'. As the first part of a compound, however, it can mean almost anything connected with drugs: *narko‿maania* 'drug addiction', *narko‿maan* 'drug adict', *narko‿katk* figuratively 'the plague of drugs', *narko‿politseinik* 'policeman dealing with drug dealers' etc. *euro* (without a capital letter) means 'European, concerning Europe': *euro‿raha* 'Euromoney, the ECU', *euro‿liiga* 'the European League'. A certain case of first component as a headword are the multiples of the units of measure (*milli‿*, *kilo‿*). Then there is a considerable group of adjectives where the suffix *-ne* is preceded by an extra long syllable (syllables in Estonian can have three degrees of length: short, long and extra long). E.g. maksi*maal*/ne 'maximum', *kald*/ne 'sloping, inclined'. These words make compounds without the ne-suffix: *maksimaal‿kaal* means the same as the word combination *maksimaalne kaal* 'maximum weight', *kald‿pind* is *kaldne pind* 'inclined plane'. Compounds are preferred especially as technical terms. So, if it is alphabetically possible, we present two equal headwords side by side: *maksimaal‿*, *maksimaalne*. If some other words come between them, they both open separate entries as *kald‿* and *kaldne* do. One more possibility is to construct the entry this way:

> *memuaarne* 'memorial' (morphological information and a purely Estonian synonym follow). Liitsõnus 'in compounds' memuaar‿: ‿kirjandus, ‿teos.

Our way of showing pronunciation (the degree of syllable length) in the headwords is simplified in comparison with the orthological dictionary. The latter shows exactly which sound or diphthong carries the length: r'aud 'iron', ke'lder 'cellar', koll'eeg 'colleague' etc. In some cases this way of marking only seems to be scientific and it is impossible to determine the particular sound or diphthong clearly. Anyhow, this is more a scientific problem than something that would interest a common language speaker. It

is more simple and practical to point out the extra long syllable. As the overwhelming majority of monosyllabic words are extra long, we do not have to show it. In other cases we mark the syllable with a dot before it, so the discussed words look like this in our dictionary: raud, .kelder, kol.leeg.


## 2. Construction of entries

In addition to the headword, an entry may contain the following parts: 1) grammatical, i.e. mostly morphological information; 2) a usage or subject label; 3) semantic information through an explanation, synonym(s) or antonym(s); 4) usage examples; 5) compounds containing the headword; 6) derivatives. Again the compiler has to decide which parts are needed to give information necessary to the user.

The morphological principles of the dictionary have been worked out by H. Saari. He introduces them in his paper, so I shall not discuss them here.

We do not use subject labels very consistently. It would be a separate task to determine which word is a special term and which is not. After all, *hand* or *leg* are also anatomical terms. Besides, one term is often used in more than one special field. A label may serve the purpose when there is no sense in defining the term, in this case it sometimes gives a necessary hint. On the other hand, when we say that calcium is a chemical element, this should show clearly enough that it belongs to chemistry. But a label is certainly needed if specialists and common language speakers use different words for the same thing as it happens quite often in biology, medicine or technical science. Here we have to stress that the headword is a special term and add the colloquial word for comparison (or the other way round). We try to give different usage labels as our principle is to offer the reader a choice of equivalents in different styles under one headword.

The semantic information must be as concise as possible. We keep telling ourselves that it is not our task to compete with an encyclopedia. Often it is enough to give a subject label, synonyms (including inadvisable ones with a proper commentary) or antonyms. But when words are used in the wrong meaning, we always make an exception. Then we give a thorough explanation: what the word really means and what words should be used instead. Such is the case with paronyms or when foreign influence has caused unacceptable shifts of the meaning. Of course, it will take years before the dictionary gets published. Such common errors should be (and are) explained currently in other forms.

The rest of the entry contains illustrative material. In the beginning of our work we followed strictly the pattern: word combinations or short sentences containing the headword, then compounds, then derivatives. But now we are not so strict any more and one may find, for instance, a sentence with a derivative or compound in the entry of its parent word. In this part of the entry we also use prescriptive commentaries like "better", "in most cases better", "in special language better", "more precisely", "usually superfluous" or something else like that, depending on the concrete needs. In a few cases, when changes in the language are going on at the present moment, parallel forms can be given with a commentary "still also" or "already also" about one of them.

## 3. Dictionary-making in a changing society

At first glance it would seem that linguistics and politics are rather distant fields. But the turbulent political and economic changes which have been going on in Estonia for the past few years are constantly causing extra work and new problems in language planning and dictionary-making. Above all, the vocabulary of economy and politics needs updating. As the society is developing towards democracy and market economy new (or well forgotten) concepts have to be denoted, terms must correspond to international standards. A lot of necessary words that exist in the language already, can be found in previous Soviet-time dictionaries, marked as historisms or followed by a scornful commentary "in capitalist countries". Now they are used again as ordinary contemporary words. Thus, for instance, the administrative-territorial units of the state are now called as they were before 1940, the militia has been named police together with all the respective ranks etc. At the same time, a lot of previous everyday words like militia, the Komsomol (the Soviet communist mass organization of the youth) etc. have become history. We have not excluded them from the dictionary, but we treat them as rare words that need a brief explanation, nothing more. Many political terms have to be redefined. They have had their compulsory dogmatic definitions which could not be avoided before (e.g. internationalism 'solidarity of the working people of different countries in their struggle with capitalism' etc.) and which can be replaced now with the common explanation of the concept.

In the present stage, our manuscript bears clear marks of different political periods. For example, some military ranks (*kapral* 'corporal' etc.) have a scornful commentary "in bourgeois Estonia" in the 1976 dictionary. In 1988, when we started our work, we thought it would be more polite to say "in the Republic of Estonia" instead. Since May 8, 1990, Republic of Estonia is the official name of the state again and we should probably say "in 1918-1940" about the period of the first independent statehood. Now, Estonia's own army with all it ranks will come into being again and who knows, maybe we shall have corporals, instead of yefraytors, again by the time the dictionary is finished. It means we shall have to bring the whole manuscript up to date from this point of view in the final stage of the work. It is, however, a rather pleasant task.

In conclusion I would like to mention some of our problems and hesitations. Our rather loose form requires that the user reads the entry through to the end, only then can he or she get all the necessary information. On the other hand, our entries sometimes turn out very long, a few may even take up a whole column. Are we asking too much from our user? Entries become so large mostly if we enter long strings of compounds or if we want to write down and explain all idiomatic expressions which contain the headword. It seems proper, but if they cause no linguistic problems, should we not trust our user's intelligence and exclude most of them?

The dictionary was originally planned in one volume. In the present stage of the work it seems there will be at least two volumes. This already means it will be less comfortable for the user to handle.

And finally, however thoroughly we explain critical cases in a dictionary, it is not enough to catch the language speaker's attention. A dictionary can be just one link between lanquage specialists and speakers.

## Endnote

1    I am using the term "language planning" in the narrower sense, denoting the aspect which
     has also been called corpus planning, language treatment or language maintenance. The cor-
     responding word in German is *Sprachpflege.*