

# Vocabulary Development in Finland

*Sari Maamies*

*Research Institute for the Languages of Finland*

Language planning, i.e., the deliberate development and maintenance of the Finnish standard language, is one of the tasks of the Research Institute for the Languages of Finland, founded in 1976. It is a government institute under the auspices of the Finnish Ministry of Education. The Research Institute studies and serves to maintain Finnish, Swedish, Saami, Finnish sign language and Romani. It is specialised in lexicography, language planning, onomastic research and the research into certain cognate languages of Finnish. In addition to that, the Research Institute has an extensive archive and a library specialised in the research into Finno-Ugric languages.

One of the main tasks of the Research Institute, besides lexicography, is language planning, including name planning. The Language Planning Department of the Research Institute follows and guides the Finnish standard language and issues statements on lingual norms, as well as recommendations and information to the general public. An important part of language planning is, of course, the follow-up and development of vocabulary. The activities of the Language Planning Department are supported by the Finnish Language Board, which makes decisions on lingual norms and takes part in discussions on language policy. – Finnish language guidance no longer has to answer comprehensive normative questions and issue new norms. Rather, at present, the questions have to do with the controlling of existing norms, often loosening the norms that are too tight or specifying those that are difficult to apply. The operations have lately focused on active participating in discussions on language policy.

## A BRIEF HISTORY: VOCABULARY DEVELOPMENT AND PURISM

Finnish has had a written language since the 16th century, but the conscious development of the Finnish literary language was not started until the 19th century, at the time of the national awakening and building of the national identity. In order to better understand the lingual situation and the tradition of language guidance in Finland, it is necessary to explain the history of the country with a few words. Geographically, Finland is located in northern Europe in Scandinavia, between Russia and Sweden. For centuries, Finland was part of the kingdom of Sweden, until Russia defeated Sweden in a war in 1809, and Finland became an autonomous grand duchy of Russia. The linguistic situation was that Finnish was spoken by the majority

of people, i.e. by common people in their everyday lives. The language of the authorities, administration, education and culture was Swedish.

The new situation, separating Finland from Sweden, affected the language policy. The officials were Swedish-speaking, but under the Russian reign, it was no longer self-evident that Swedish would maintain its status. It was suggested that Finnish, the language of the people should be adopted as the language of the entire nation. Thus, there was a social need for the development of literary language. Finnish was consciously developed towards a language of civilisation usable in all fields of society. This meant the deliberate building of vocabulary, as well as developing the linguistic structures. From the present-day perspective, the effort was huge: it could be said that Finnish was really modified into a language of science, arts, technology and administration – a language of civilisation – in approximately fifty years, between 1820 and 1870. During that period, the vocabulary was supplemented with thousands of new words, most of them core vocabulary used every day. For example the following concepts were named in Finnish through conscious developing of the vocabulary: *arts, science, literature, independent, democracy, state, religion, citizen, industry, cell, artery, varicose vein, hospital, future*. The examples show how central the words were.

The vocabulary development in the 19th century and the early 20th century was characterised by a certain purism, which reflected the political and spiritual situation of the time. Swedish has influenced Finnish profoundly, which can be seen both in the vocabulary and in the structures of the language. As the political connection to Sweden was broken, the new national identity and the position of the Finnish language were strengthened by e.g. eliminating the foreign influences in the language, especially those of Swedish. The purism at that time was national in nature, typical of a phase where the language is consciously constructed and revived. At such a stage, the expressive potential of the language is adapted to function as a communicative tool of society, and the language that has held the dominating position until then is rejected. The attitude towards loan words is negative, and the new vocabulary is created using the reserves of the nation's own language, either by inventing words or adapting dialectal words and old lingual forms. (Thomas 1991: 79, Rintala 1998: 55.)

The Finnish purism in the building of vocabulary was not so strong that it would have led to the eliminating of already adopted loan words. Instead, the lexicographical work aimed at developing Finnish-based expressions for concepts that had not been named before. Translated loan words were not rejected, either. However, other types of loan words were objected to, partly because it was felt that they made it difficult for the ordinary people to understand and use the language. This was justified by saying that literary language should maintain at the possession of the entire nation and that civilisation should be distributed to all social layers. This view has also guided the development of specialised vocabularies in Finland. (Rintala 1998: 55–56; Haarala 1989: 261.)

The puristic tendency in the vocabulary building and language planning continued to some extent until the 1950s, but less restricted views were introduced from the early 20th century onwards. Ideas about language as a tool for human actions – a tool that could be improved and planned – were introduced, and the situational variation of language was taken into account. Today, language planning acknowledges that

different discourses are used in different contexts for varying communication needs, and this applies to vocabulary as well. (Rintala 1998: 58–.)

## WAYS TO FORM NEW VOCABULARY

The vocabulary of the Finnish language accumulates largely in the same way as that of any other language: borrowing words, translating, and making use of existing vocabularies. However, the structure and orthography of the Finnish language limit the ways of forming new words. Finnish has a solid status, and the normative system of its standard language is established. The normal vocabulary is created and developed spontaneously in lingual communities. Conscious planning of the standard language is necessary when a word is not created naturally, or when its meaning remains somewhat unclear, or its compliance to the norm is problematic. As for the developing of jargons, it requires deliberate, organized terminological efforts.

Finnish is not an Indo-European language. It is part of the Finno-Ugric language group. These languages have a tendency towards agglutination, i.e. the grammatical relationships, such as the number, person, tempus and modus of a verb, are expressed through suffixes attached to the stem of the word. Both verbs and nouns can be inflected; a Finnish noun has fifteen cases, and the conjugation of verbs is even more varied. Inflecting can also affect the word stem. The new words therefore have to be adapted into the inflection system of the language. In addition to inflecting, derivation is very characteristic of the Finnish language. This means that Finnish is morphologically a very rich language, offering many alternatives for word formation.

Another factor affecting the structure of the vocabulary is the phonological structure and orthography of a language. The equivalence of pronunciation and writing is high in Finnish: one phoneme is equivalent to one letter in writing. It could be said that Finnish is written as it is pronounced, with the exception of a few specific instances. For example in English, the pronunciation and writing of a word often differ from each other to a great degree, and this makes it often problematic to introduce English loan words into Finnish.

## LOAN WORD OR ORIGINAL FINNISH WORD?

The question of whether we should borrow a new word along with a new concept or whether we should create a word form based on our own language, is not simple. The Finnish language planning tradition has had a rational attitude towards loan words: if it is clear that it is advantageous to borrow a word, it can be regarded as a good way to enrich the language. A loan word can be more suited for its purpose e.g. in certain special fields, linking the word to its equivalents in other languages and thus more comprehensible in its own field. What is decisive in such a case is how the word is suited to the system of the language. Whilst evaluating the appropriateness of a word in the system, we have to take into account its length, how it can be distinguished from other words in the same thematic field, grammatical correctness, and ease of pronunciation, spelling and inflecting.

The Finnish language community seems to consider Finnish-based vocabulary as some kind of an ideal. There is a clear tendency towards creating Finnish-based

expressions for new concepts, as far as that is possible. Thus, it is not solely the view of deliberate language planning. For example, public language debates often have to do with the Finnish equivalents of English words, and language planning officials are frequently asked how loan words that seem foreign could be replaced by Finnish ones. It is perhaps the old tradition of purism in relation to vocabulary that is in the background of the pursuit of Finnish-based forms. Yet, it is also obvious that the structure of Indo-European vocabularies does not easily fit into the Finnish language.

### *Loan words*

Finnish, as any other language, has thousands of loan words from different periods; most of them belong to the core vocabulary. They have been incorporated into the language in different ways. Part of the loan words no longer have any phonological or structural features that would associate them to a foreign origin; some have partly adapted to the language, in such a way, however, that they have become part of everyday vocabulary, and part of them have maintained their foreign form. In Finnish the share of loan words of the basic vocabulary – i.e. of words that are neither compounds nor derived – has been estimated at 20 per cent. Although this is a relatively rough estimate, it proves that Finnish vocabulary is basically very Finnish in origin. However, it should be noted that creating new vocabulary consists largely of giving new names to concepts that have been adopted from elsewhere, so the vocabulary may be phonetically original, and yet the meanings may be borrowed. (Hakulinen 1979: 479–481.)

Words that have been borrowed directly from another language, but have not been adapted according to the Finnish writing system, are, in fact, rare. The reason for this is that it is difficult to pronounce, spell and inflect them; they are too different from the rest of the Finnish text flow, thus causing stylistic irregularity in the text. A quoted loan word, i.e. a foreign word is adopted as such when there is no satisfactory Finnish equivalent and when the word cannot be adapted even phonetically. It also happens that a foreign word actually has a Finnish equivalent, but it has not been adopted for some reason. Quoted loans are frequent especially in vocabularies related to technical science (*hands free* [mobile phone], *online-*), finance (*franchising*, *benchmark*, *tax free*), sports (*motocross*, *squash*), and food (*cappuccino*, *espresso*, *rata-touille*, *texmex*, *smoothie*). – The English word *toast* is an example of how fashionable English words are. We have a Finnish equivalent compound *paahtoleipä*, established for a long time. As the cafeterias have started to sell toasted bread with different fillings, the word *toast* has been borrowed again. The old Finnish word is not attractive since it does not give the desired image of the new product and therefore does not sell. Foreign words may be experienced as modern and impressive in the quickly changing superficial culture, susceptible to foreign influence; words from our own language seem too commonplace. Here, it is the images created by the words that matter.

### *Accommodation*

Nevertheless, loan words are usually adapted phonetically. The word is then also accommodated to the Finnish orthography. Finnish words do not have the letters *x*, *z*, *c* and *w*; they are replaced by *ks*, *ts*, *k* and *v*. Long vowels are marked with two vowels in writing. Finnish has a strong tendency towards nouns with a two-syllable

stem. In order to add a syllable to a one-syllable loan word ending with a consonant, we add the vowel *i* at the end of the word, which makes the word more Finnish.

If the pronunciation and orthography of a word differ from each other in the source language, we have to decide which form we should adopt. Whilst accommodating the loan words into the Finnish system, we often have to accept two or even more competing forms. The attitude of language planning officials towards this kind of variation is normally tentative: both alternatives are accepted in the beginning, until one of them clearly becomes more common. After that, the one that is used more frequently is usually recommended.

The following examples are loan words whose orthography has been accommodated.

fax	faksi ~ telekopio
business	bisnes
dealer	diileri
brand	brandi ~ brändi
blazer	bleiseri
extranet	ekstranet
to surf	surffata
pizza	pizza ~ pitsa
curry	curry ~ karri
wok	vokki
fondue	fondue ~ fondyy
nougat	nougat ~ nugaa
croissant	croissant ~ kroissantti

*Fondue*, *nougat* and *croissant* are words that can also be written in their original French form, provided that the pronunciation is correct (this, however, is often problematic). If the orthography of a word is accommodated, it is unproblematic to pronounce and easy to inflect. However, orthographically modified words are often felt to be somewhat strange, since they look odd, at least in the beginning. Somehow the foreign forms seem more convincing. Yet, the more commonplace and ordinary a word becomes, the more easily its accommodated orthography is accepted.

The word *brandi* illustrates the problem the differences in English pronunciation and spelling cause in Finnish. A Finnish equivalent to the word has not been found despite several attempts, which is why the word has been borrowed. So, should we comply with the pronunciation or with the orthography? The form *brandi* should be pronounced with an [a]: [brandi]. However, if we wanted to comply with the original English pronunciation, we should write it with an [ä]: [brändi]. – At the moment, both forms are accepted.

### *Translated loan words*

A special category of borrowing words is translating. Translated loan words are foreign in their contents but Finnish in their structure, which is why they fit into the vocabulary extremely well. Translating words has been a common way to enrich the vocabulary during the entire history of literary language in Finland. As cultural concepts are given an international name translated into Finnish, the language preserves

its own features, but yet the connection to the equivalents in other languages is maintained.

Translated loan words are numerous for instance in the language of IT; its core vocabulary is therefore very Finnish, even though the link to other languages is clear.

home page	kotisivu (koti 'home' + sivu 'page')
e-mail	sähköposti, s-posti (sähkö- 'electronic' + posti 'mail')
data base	tietokanta (tieto 'data' + kanta 'stock, base')

If the meaning of a translated loan remains unclear, the only way to understand it is to know the respective word or phrase in the source language. An this, of course, can't be accepted.

Semantic expansion, based on the model of a foreign language, is one example of an failed translated loan. We borrow not only the word but also its semantic features. An example of this is the English word *park*, which has a wider semantic field in English than in Finnish. In Finnish the word *puisto* 'park' has originally referred to a place where trees grow. However, according to the English pattern, we have started to use the word *puisto* in a wider sense to refer to places built for certain activities, for example *business park*. More Finnish translations have been suggested, but they have not been adopted. The new meaning is attractive.

park	puisto ('a collection of trees')
business park	"yrityspuisto" (yritys 'business, company' + puisto 'park') > yrityskeskus (yritys 'business, company' + keskus 'centre')
science park	"tiedepuisto" (tiede 'science' + puisto 'park') > teknologiakeskus ('technology centre')

### Abbreviated words

Words that have been consciously invented, not based on anything previous do not emerge in the modern language. In the 19th century, as the process of developing the language was underway, there were some instances, some of which survived: a concept as central as 'electricity' (*sähkö*) is one of them.

A category close to the consciously invented words is that of abbreviated words. At present, they are part of the vocabulary that is growing fast and colloquial in style. The abbreviated words are often created first for unofficial use within a working community, but they are often transferred to standard language, as well.

luomu < <b>lu</b> onnon <b>mu</b> kainen 'ecological'
luki(häiriö) < <b>lu</b> ku- and <b>kir</b> joitus- '(lit.) reading and writing (dysfunction), dyslexia'
sotu < <b>so</b> siaaliturvat <b>un</b> nus 'social security number'
mamu < <b>ma</b> ahan <b>mu</b> uttaja 'immigrant'

### Compounds and derivatives

As the above examples have already shown, the majority of new words of the past few years are compounds. Making compounds is an extremely productive word-forming method in Finnish; it is very easy to make them, the only limitations are semantic. The popularity of compounds can be seen in that some 65 per cent of the

words in the extensive dictionary describing the Finnish standard language (Nykysuomen sanakirja) are compounds (Häkkinen 1994: 419). New compounds are thus formed ad hoc, and most of them are only used for a short time. Only a fraction of them are established as terms. Thus, compounds are a common method of producing new words based on original Finnish vocabulary. They have been successfully used for example in the field of information technology and medicine.

computer	tietokone (tieto 'data' + kone 'machine')
chat	verkkojuttelu (verkko 'net, network' + juttelu 'chat')
placebo	lumelääke (lume 'imaginary' + lääke 'medicine')
burn out	työuupumus (työ 'work' + uupumus 'exhaustion')

Another very productive way to form words is derivation. By adding a suffix to the word stem, we can create a new meaning or nuance. Suffixes have been used to give new concepts and shades of meaning names that suit the Finnish system. The derivation of words follows already existing patterns. Words linked to the existing vocabulary for their structure are easily adopted and understood because of their lingual associations. This enhances their chances to be established in practice.

-sto: 'a collection of something'	
file	tiedosto (tiedo- 'data' + suffix -sto)
keyboard	näppäimistö (näppäin 'a key' + suffix -stö)
web site	www-sivusto (sivu 'page' + suffix -sto)
compare:	puisto 'park' (< puu 'a tree' + suffix -sto)
-kko: 'a place where something can be kept'	
menu	valikko (vali- 'to choose' + suffix -kko)
	compare: taulukko 'a scale, table' (taulu 'table' + -kko)

## TERMINOLOGICAL WORK ON JARGON

Standard language vocabulary often emerges spontaneously, whereas jargons do not develop without conscious terminological work. The development of technology, finance and several other fields has led to the birth of new jargons; these new forms of language now threaten to be separated from the standard language understood by everybody.

The officials setting norms for jargon must have the sufficient authority in their own field so that the recommendations could be applied in practice. This is why practical terminology work is, to a great degree, the responsibility of each special field. The opportunities of general language planning to influence a greatly specialised language are limited, but it is valuable to have a language planning official in the specialist team working on the vocabulary.

The need for terminological work in the field of jargons is widely accepted; however, we often hear people questioning the use of Finnish in a special field. Since the world of science, research, technology and trade is English-speaking, people tend to think that there is no need to discuss these topics in our own language. However, if we accept this view, we go back to the situation of a hundred years ago when Finnish was used merely at home and in free-time activities and was not considered suitable for describing abstract themes. In other words, there is a danger of our language losing ground again as regards its scope of usage.

There is an association in Finland specialised in terminological work, The Finnish Centre for Technical Terminology, TSK. Its task is to coordinate the terminological work in the different specialised fields and to publish multilingual word lists. The operations of this association are financed by its member organisations, representing the commercial and industrial life. The cooperation between the Research Institute for the Languages of Finland and TSK is very close, and it enables the general language planning officials to be involved in the terminological work of the specialised fields. There are always representatives of the various special fields participating in the terminology projects of TSK.

The language of information technology is a good example of how we develop the vocabulary of one central jargon. IT language is not merely a jargon of one special field – ordinary people are faced with it every day. There have been conscious efforts to develop IT language in Finland since 1966, as The Finnish Information Processing Association published its *Dictionary of Information Technology*. Eleven editions of the dictionary have been published so far.

The Centre for Technical Terminology has been active in the field of information technology as well. Since 1999 it has coordinated an project which aims to create terminology and recommendations. The recommendations are given by a group of IT, language and media specialists after thorough consideration and a wide comments round. The work is done mainly over the Internet, which makes it possible for a great number of people to participate and give comments. Each recommendation is followed by a description of the underlying concept and advice on how to use the term in practice. The finished terminology proposals are published on the web site of the project ([www.tsk.fi/termitalkoot/](http://www.tsk.fi/termitalkoot/)). This helps the recommendations to be adopted more quickly and by a wider user group.

The work is carried out with the help of the Internet and e-mail, which is a new way to work. It has proved to be successful. So far, that is, in two years, the project has yielded some 130 terms and their definitions.

Medicine is another example of a branch of science where the Finnish terminology is being consciously developed. The Finnish Medical Society Duodecim was founded in 1881, and one of its principal tasks was, at least during the first few decades of its existence, to create a Finnish medical language. Medical terminology is handled by the vocabulary board of the Society, with a membership consisting of doctors from various branches of medicine and language planning officials. The Finnish Medical Society and its publications have had a great influence on the Finnish medical terminology, its development and establishment.

## DICTIONARIES OF THE FINNISH STANDARD LANGUAGE

New vocabularies are interesting not only from the point of view language planning officials, but also to lexicographers. Language planning and the edition of the modern standard language dictionary are closely linked at the Research Institute. During the past hundred years, two large standard language dictionaries have been published in Finland: the six-volume *Dictionary of Modern Finnish* (Nykysuomen sanakirja) was published between 1951–61 and the *Basic Dictionary of the Finnish language* (Suomen kielen perussanakirja) with three volumes between 1990–94; the

CD-ROM version of the Basic Dictionary, the *Basic CD Dictionary of the Finnish language* (CD-Perussanakirja) was published in 1997.

The Dictionary of Modern Finnish is based on data collected from various sources before 1938. The dictionary has over 200,000 entries. It describes the standard language of the early 20th century; it also includes dialectal words and words that were rare already at the time. The Dictionary of Modern Finnish is normative, that is, it recommends expressions that should be used instead of other incorrect ones. It was never published as a revised edition, which means that it no longer describes modern Finnish vocabulary; rather, it is an important description of the language of its time. It has also been considered an authority as regards orthographical norms.

The Basic Dictionary of the Finnish Language has ca. 100,000 entries, half as many as in the Dictionary of Modern Finnish. In spite of this, it contains ca. 20,000 words that cannot be found in the old dictionary. This illustrates well the pace at which the vocabulary is reformed.

The Basic Dictionary of the Finnish Language aims at describing the core vocabulary of present-day Finnish. In addition to standard language vocabulary, the dictionary contains central terms from specialised fields, colloquial expressions, slang words and dialectal variants, especially those that have been frequently used in the mass media. The Basic Dictionary includes recommendations on orthography, derivation, pronunciation, style, meaning, as well as use in clauses and phrases. The recommendations are based on the recommendations issued by the Finnish Language Board of the Research Institute.

A machine cannot replace human work in the follow-up of vocabulary development. Lexicographers at the Research Institute follow the development daily by reading newspapers, magazines and books and by following other mass media; nowadays the Internet also offers lots of source material. New words and new ways of using expressions are stored in a database. Most of the new words are short-lived, e.g. words related to a phenomenon of the day, disappearing together with the phenomenon.

We still need traditional word collecting – gathering data from papers and books – but we also need wider textual material. The electronic corpora offer us a reliable way of obtaining data on the modern meaning of words and their generality, syntactic features, typical collocations, style, phrases, etc. The Research Institute for the Languages of Finland has been actively involved in the compilation of electronic text corpora, and presently it has a corpora containing over 40 million words. The corpus consists mainly of newspaper and magazine articles from the 1990s, as well as of certain other genres.

### *The Basic CD Dictionary of the Finnish Language*

The CD-ROM version of the Basic Dictionary of the Finnish Language was published in 1997. The electronic data for the dictionary was therefore structurally modified in accordance with SGML coding, which enables many-sided search options. In addition to searching words by using the alphabetical index, searches can be done by giving just the end of a word, or the search can be directed to the various structural parts of an article, e.g. the explanation or the example. A central feature of

Finnish words is that they can be inflected, and it is a great challenge for an electronic dictionary to produce and identify inflections. Language technology was used in the compilation of the CD-ROM Basic Dictionary so that the dictionary can produce all the inflections of inflectable words, even compounds, if necessary. Moreover, the entire text in the dictionary has gone through a morphological analysis program, and consequently, in addition to the uninflected form of a word, the dictionary program finds even its inflected forms in the text. This feature has so far been rare in the Finnish electronic dictionaries.

Compared to the printed Basic Dictionary, the CD-ROM version has 2,000 more new entries, and as many updated old entries. A new revised version of the CD-ROM dictionary is almost ready (in October 2001), and there will be 3,000 new words along with other changes in the dictionary. The dictionary will also include new features, such as the comparatives of adjectives. There will also be 20,000 Finnish place names – names of cities, town, communities – together with information on their inflection; even appellatives are inflectable in Finnish, and that may cause problems in some cases. So, the traditional dictionary is gradually changing into a multi-purpose program combining the needs of language planning and lexicography, serving the purpose of solving problems of language usage in a many-sided way.

## SOURCES

- HAARALA, Risto (1989). *Sanat tiedon ja taidon palveluksessa*. In *Nyky-suomen sanavarat*. Edited by Jouko Vesikansa. Porvoo–Helsinki–Juva: WSOY.
- HAKULINEN, Lauri (1979). *Suomen kielen rakenne ja kehitys*. Fourth, revised and added edition. Helsinki: Otava.
- HÄKKINEN, Kaisa (1994). *Agricolasta nykykieleen. Suomen kirjakielen historia*. Porvoo–Helsinki–Juva: WSOY.
- RINTALA, Päivi (1998). *Kielikäsitys ja kielenohjailu*. – Sananjalka. Suomen kielen seuran vuosikirja 40. (The 40th annual of the Society for the Finnish Language). Turku, pp. 48–65.
- THOMAS, George (1991). *Linguistic Purism*. London & New York: Longman.