Virpi KALLIOKUUSI, Tekniikan Sanastokeskus (The Finnish Centre for Technical Terminology)
Krista VARANTOLA, University of Tampere

# From general dictionaries to terminological glossaries. User expectations vs editorial aims

## Motto

'I have to look in the dictionary to find out what a virgin is. I know the Mother of God is the Virgin Mary and they call her that because she didn't have a proper husband, only poor old St. Joseph. In the *Lives of Saints* the virgins are always getting into trouble and I don't know why. The dictionary says, Virgin, woman (usually a young woman) who is and remains in a state of inviolate chastity.
Now I have to look up inviolate and chastity and all I can find here is that inviolate means not violated and chastity means chaste and that means pure from unlawful sexual intercourse. Now I have to look up intercourse and that leads to intromission, which leads to intromittent, the copulatory organ of any male animal. Copulatory leads to copulation, the union of sexes in the art of generation and I don't know what that means and I'm too weary going from one word to another in this heavy dictionary which leads me on a wild goose chase from this word to that word and all because the people who wrote the dictionary don't want the likes of me to know anything.'

**Angela´s Ashes** by Frank McCourt 1996 (Flamingo Edition from 1997, p. 333)

## Abstract

This paper explores the types of questions for which the users of general dictionaries and terminological glossaries try to find answers when consulting lexical reference sources. We claim that users often attempt to access data that does not fall within the scope of the information that is normally considered to belong to dictionaries or glossaries. We shall first illustrate this problem by means of a real-life example from a general dictionary context, and then we shall project the problem further into a special field context. We shall argue that, if we wish to better address user needs, then it might make sense to adopt a more permissive and less orthodox approach at all levels of dictionary compilation so as to better meet user expectations.

Keywords: user expectations, presentation of knowledge structures, enclyclopedic information and dictionary entries

## Introduction

This study investigates why the users of lexical reference sources such as dictionaries or terminological glossaries are often disappointed by the information they find in those sources. We shall tackle this issue by contrasting user skills and expectations with editorial aims and presuppositions. First, we shall briefly outline of our premises and then elaborate on the main points by means of a case study which is based on a real event in the general dictionary domain. This case will then be projected into a hypothetical situation in the specialized domain of terminology, where it will be used to illustrate the problems that present-day terminologists face.

We shall discuss user expectations and editorial aims from the angle of general and specialized sources of lexical information. These expectations and aims will then be analyzed in the network of intersecting continua or overlapping areas such as user skills, needs, expectations, knowledge structures, cultural background, and the scope of the information provided.

## Background

It is well-known that dictionary/glossary[1] users often are frustrated when consulting reference sources. Atkins and Varantola (1997:1) maintain that "there are two direct routes to more efficient dictionary use: the first is to radically improve the dictionary; the second is to radically improve the users". Yet what happens if both sides feel that the fault lies with the other side and each refuse to mend their ways?
In other words, have we reached a dead-end if users, in their attempt to find answers to simple lexical queries, are uninterested in the editors' application of complicated compilation and presentation principles? On the other hand, will an impasse also be reached, if editors refuse to stray away from their proven practices and theoretically-sound guidelines in order to satisfy what they consider to be whimsical and random requests made by the users?
We hope that neither scenario proves to be true, and we attempt to closely examine the sources of frustration so as to suggest reasons why it might be necessary to re-evaluate editorial standpoints in order to reach higher customer satisfaction.
We are thus of the opinion that the initiative lies with the dictionary maker and not with the users, as the latter do not normally have the necessary expertise. The users should nevertheless be taken seriously when they express complaints about the product and suggest improvements to it.

## Research issues

To learn to understand dictionary users' needs and behaviour the following questions could be relevant:

What is the target group that we expect to use the particular lexical source?
What kind of information does the user comprehend?

What type of information does the editor provide?
What type of information does the user want to find?

What does the editor think is the proper coverage for a particular dictionary?
What does the user think is the proper coverage for a particular dictionary?

How can a "context-free" editor cater to the needs of a "context-bound" user?

Where does the editor place the different units of information?
Where does the user look for the different units of information?

**The present framework**

If we approach the aforementioned questions from the point of view of knowledge structures, we can assume that a «general information level» is the typical starting point for learner's dictionaries, school level and compact bilingual dictionaries.

A semi-general + (a more or less random) special information level is the level aimed at by larger general bilingual dictionaries and by some monolingual dictionaries, at least this is true of Finnish dictionaries.

The coverage of semi-general and systematically selected special fields is the level reached by such large monolingual dictionaries as the *Collins English Dictionary* and the *Random House Unabridged Dictionary*.

Dictionaries of the above types are normally consulted for general domain issues. Mono- and bilingual special dictionaries have a more comprehensive field-specific coverage which naturally makes them useful sources of information for field-specific queries, but at the same time this limits their general usefulness. Terminological glossaries represent a further step in specialization. They tend to have a relatively restricted scope, particularly if they are based on a stringent conceptual analysis which is not feasible in large scale glossaries.

The clashes that occur between expectations and the information provided are the following:

- when expectations and aims do not meet
- when there is an extralinguistic, often cultural, factor that hampers judgement
- when users do not know how to use the dictionary or what to expect from it
- when the dictionary maker provides insufficient or misleading information (e.g. an equivalent in a bilingual dictionary that is semantically too restrictive or too wide in scope)
- when the user does not have enough background knowledge to understand the definition or to apply it
- when the user uses the source without due caution or has an uncritical and excessively respectful attitude towards the dictionary

**Case studies**

We shall discuss clashes of the above type in light of a case study which begins at the general level. This is an example of a general dictionary use that went wrong and generated a long discussion in a Finnish newspaper. We shall then transfer this example to the level of specialized terminology and subject it to a hypothetical terminological treatment and analysis. We shall discuss this example in terms of an orthodox terminological approach and the problems embedded in it. Let us first, however, take a look at the issues a terminogist has to face when compiling a special-field glossary.

In the specialized domain, the terminologist has to decide how detailed and specialized the information should be. If the glossary primarily addresses experts or skilled craftsmen of a certain subject field, the information given about the concepts of the field needs to be extremely detailed and accurate. But how could a non-expert of the field make use of this glossary?

Or conversely, if the glossary is meant to be used by non-experts, the information given about the concepts would need to be adapted to the knowledge level of the "average" user. How

could the glossary then satisfy the needs of the expert searching for scientific or technical information?

In addition to the expected user group(s), the editor of a specialized glossary thus has to analyze and decide upon the information level of the glossary and the expected knowledge level of the users, any special characteristics of the field that need to be taken into account, the state of the field and the status of its vocabulary in the language community, language planning and standardization aims involved, etc. A clearly defined target group and purpose do not, however, necessarily guarantee user-friendliness or user satisfaction.

Users may find that certain areas of vocabulary which they expected to be covered are either not covered or only superficially covered, usually the information given is not relevant to them or that the description of the concepts is inadequate.


## Part I. The lexicographical frame

This concrete example belongs to the field of "forensic lexicography"; it is a case of a mistaken identity, an inaccurate equivalent, and a cultural clash. The object is a deadly but delicious mushroom that is considered to be an essential part of the Finnish food culture but not necessarily a part of the general EU food culture.

The case history of what I would describe as "layman behaviour in dictionary use" is as follows:

A Letter to the Editor in the largest Finnish daily (Helsingin Sanomat) on 23 July 1995:

A Frenchman and a permanent resident in Finland writes how Finland used to be an idyllic place for a foreigner to live in until he bought fresh *korvasieni* at a supermarket. He recognised the mushroom but checked the French equivalent in a Finnish-French dictionary, found *morel* and prepared lunch for his party.

Extralinguistically, it was a case of erroneous identification. Linguistically, the dictionary equivalent was inaccurate. Fortunately, the mistake was discovered before it was too late and everybody survived.

This incident resulted in a cultural clash during which following types of opinions were expressed in the ensuing discussion on the Letters to the Editor section.

The victim claimed that the shopkeeper and the authorities approached were indifferent to the sufferings of the party because he was a foreigner.
He would have understood a warning if he had seen one. Furthermore, he learned later that another foreigner had narrowly escaped the same fate when he was told at the last minute how poisonous *korvasieni* is.

The shopkeeper replied that he sells korvasieni every year and every time there is a notice in Finnish (approved by the authorities) with a picture of *korvasieni* + information that fresh *korvasieni* is poisonous and must therefore either

- be boiled twice for five minutes in plenty of water (1 part *korvasieni* 3 parts water) and the water must be thrown away afterwards. After boiling the mushrooms must be rinsed well
- be dried first, soaked before using and then boiled. When drying or boiling *korvasieni* one must also see to proper ventilation

Other contributions included following comments:

- Dictionary makers should be charged, the shopkeeper is innocent
- Dictionary makers always get equivalents wrong out of sheer ignorance and laziness. Don't trust dictionaries - study mushroom books instead
- Furthermore, in a truly civilized country, *korvasieni* is not on sale

How could one then explain culture-specific phenomena/life styles in a comprehensible way in a foreign language and for an audience, which is not familiar with the phenomenon in real life, that is, outside language?

In this case, in the aftermath a year later, just before the new *korvasieni* season, was that *korvasieni* is still on sale in Finland but the authorities had published new brochures in several languages, with detailed instructions and a glossary.
The equivalents given in different languages were:
For the poisonous variety:

| | |
|---|---|
| fi | korvasieni |
| sv | stenmurkla |
| en | lorchel, false morel, lorel |
| de | Frühjahrslorchel, Giftlorchel |
| fr | gyromitre, fausse morel |
| ru | stroshók, strochók |

For the harmless variety:

| | |
|---|---|
| fi | huhtasieni |
| sv | toppmurkla |
| en | morel |
| de | Morchel |
| fr | morille |
| ru | smortshók, smortchók |

Let us first examine mushroom books.
Finnish mushroom books basically give the same information as above: *delicious but deadly.*

An American book on mushrooms from the 1970´s states:
*Gyromitra esculenta.* This odd-looking fungus, closely resembling a brown 'brain' perched on a white stalk, is a relation of the morels and, like them, it appears in the spring. *It is known to be poisonous, and yet in Eastern Europe and in Sweden it is eagerly sought and marketed in large quantities because of its superlative taste.*

A British mushroom book (Collins Nature Guides 1994, originally German from 1985) says:
*Gyromitra esculenta...* From April to May in large groups in sandy pine forests: very frequent in eastern Europe, otherwise rather rare.
The turban fungus contains gyromitrin, *a deadly poison which is retained in the cooking process,* its action is similar to that of α-amanitin. *The poison evaporates after some time so that dried fungi are harmless. This fungus was till recently sold in markets in Europe and used a lot in the canning industry, but its sale is now prohibited. Gyromitra esculenta is wrongly described as edible if cooked in some fungus books, but it must not be eaten at all.*

The first question is whether or not we are talking about the same mushroom and the same concept? The answer is yes, as the Latin name confirms that but we are referring to widely different perceptions of the same concept. What, of course, is extremely interesting is that no dictionary mentions anything about *false morels* being poisonous. This type of information is not traditionally considered to belong to dictionaries. The dictionary user thus made a wrong identification and the dictionary gave a wrong or inaccurate equivalent. In addition, the user blindly trusted the dictionary and probably also expected it to include some encyclopedic information.
On the one hand, the entire debate is an excellent example of the normative attitudes which non-professional dictionary users often have towards dictionaries. In this respect, a dictionary is like the Bible: what is printed in the dictionary is by definition correct, irrespective of the

age or the size of the dictionary. Furthermore, "a dictionary equivalent" is actually a contradiction in terms. It is rarely that a bilingual dictionary can provide a 100% equivalent in the other language. If a dictionary could do this, dictionaries would be simple code conversion tables. But languages are not one-to-one codes. This is also why one of the first lessons one has to teach about dictionary use is that «I found it in a dictionary« is not an acceptable explanation.

The conclusion is that we are dealing with a highly culture-specific phenomenon and if lexicographers had been taken to court, it is unlikely that they would have been found guilty. (The above discussion has to a large extent been adapted from Varantola 1997).

On the other hand, one could take an opposite view and say that dictionary users have been led to believe that they can trust their dictionaries and that a user-friendly dictionary should pay attention to the users needs. And if dictionary makers want to know why users might turn to their product they will also try to determine why a user would look up a particular expression. A natural consequence is then that lexicographers will try to give users information that would help them solve context-bound problems.

In the case of mushrooms, a general dictionary entry on mushrooms (both monolingual and bilingual entries) would therefore contain information about their edibility because a major use for mushrooms in real life is their use as foodstuff. So far, dictionary makers have usually refused to think this way, and for good reasons. The possibility of incorporating cultural and world knowledge in dictionary entries is limited and the selection of relevant information poses endless questions. It is not possible to guess every individual users' reason for consulting a particular dictionary entry. And, yet in this day and age, with access to large corpora and information on collocational probabilities, perhaps it would be possible, at least to some extent, to respond to the users' needs and narrow the gap between the dictionary's context-free aims and the users' context-bound needs.

## Part II. The terminological frame

### 1. Editorial aims

#### 1.1. Representation of context-free information

The editor of a specialized glossary, working according to terminological principles and methods, aims at presenting knowledge structures of a certain subject field in such a way that they are not bound to any particular context. The aim of producing as context-free information as possible is present in at least two different principles that the editor usually tries to follow. First, when dealing with general concepts[2], the information given in the definition is a generalisation of a group of cases (referents) in the real world and the properties that are common to these cases. Therefore no information is given on the properties of various individual items that fall under the scope of the concept. Second, the information given in the definition is not intended to solve problems related to a particular linguistic or textual context. Consequently, no information is given, for example, on collocational probabilities, derivational characteristics or inflection.

The following examples are all construed as a kind of follow-ups to the preceeding real-life mushroom example and they aim at showing what essential information the definition/ explanation part of the *false morel/lorchel* entry might contain in terminological reference sources compiled according to differing principles and emphases.

*False morel, lorchel* defined as a context-free general concept :
- fungus/mushroom of the genus *Gyromitra*

This type of definition would be adequate as a generalization but it would not single out the concept adequately and place it correctly in the relevant concept system.

## 1.2. Systematic analysis and representation of "encyclopedic" information

In principle, a terminological analysis and presentation does not differ from human thinking in general. When thinking, people usually apply some kind of conceptual analysis, either consciously or unconsciously, whenever they are structuring knowlegde, for example when they are in a very pragmatic situation such as when they are trying do decide where to look for bread crumbs in a supermarket.

In a proper terminological analysis, concepts are not treated as isolated units in a vacuum, but as interdependent parts of a larger system. The encyclopedic core information is presented by means of a minimum number of characteristics that are sufficient to identify the concept and to delimit it from related concepts.

The concept system determines the criteria used to select the information (= characteristics of concepts) for the definitions, but the general approach, the target group and the aims of the glossary, determine the potential concept systems available. A number of alternatives are usually available, but when the system has been established, the information included in the definitions is restricted and limited by this system. Traditionally, the definitions do not contain redundant information. Their main purpose is to link the concepts of a certain system to each other by reference to terms used for the related concepts instead of repeating the characteristics of the related concepts.

*False morel/lorchel* defined as part of a concept system based on a biological taxonomy of fungi (all the terms used in the definition are themselves defined elsewhere in the glossary):
- an operculate unitucate (macro)fungus of the order Pezizales, of the class Ascomycetes

## 2. User expectations

## 2.1. Easy answers

A clash can arise because of the systematic nature of terminological glossary, if the user is looking for a simple answer ("quick and easy information retrieval"), because it takes time and patience to find and understand the information included in a terminological definition.

The terminologist could help by

a) relaxing some restrictions concerning the systematic structure of the definitions

b) by increasing the amount of encyclopedic information (e.g. by using notes to complement the definitions).

An «improved» systematic entry for the *false morel/lorchel* :
Definition:
- an operculate unitucate (macro)fungus of the order Pezizales, of the class Ascomycetes (fungi, in which the spores, ascospores, are formed inside a club-shaped cell,ascus)

Note:
- *false morels* closely resemble a brown 'brain' perched on a white stalk. The cap varies in form and is 3-12 cm in diameter. The fungus has a pleasant smell and a mild taste

607