What is a Dictionary?

Henning Bergenholtz, *Department of Afrikaans and Dutch, University of Stellenbosch, Stellenbosch, Republic of South Africa and Centre for Lexicography, Aarhus University, Aarhus, Denmark (hb@asb.dk)*

Abstract: By comparing different definitions in dictionaries and in theoretical papers we know that lexicographers have very different opinions on the topic: "What is a dictionary?" We see too that lexicographic tools with the same type of content sometimes are called dictionary, sometimes encyclopedia, sometimes lexicon. In reality it is not a large problem if the information in a certain tool is of a high quality as to the intended function. Nevertheless I will give not one but two additional definitions of two different types of information tools. There will also be a discussion on the relevance for the concept of dictionary in connection with the research registration in universities.

Keywords: DICTIONARY, LEXICON, ENCYCLOPAEDIA, LEXICOGRAPHY, INFORMATION TOOL, LEXICOGRAPHIC TOOL, RESEARCH REGISTRATION, LEXICOGRAPHIC DATABASE, MONOFUNTIONAL DICTIONARY, POLYFUNCTIONAL DICTIONARY

Opsomming: Wat is 'n woordeboek? Deur verskillende definisies in woordeboeke en teoretiese artikels te vergelyk, weet ons dat leksikograwe baie verskillende opinies het oor die onderwerp: "Wat is 'n woordeboek?" Ons sien ook dat leksikografiese hulpmiddels met dieselfde soort inhoud soms woordeboeke genoem word, soms ensiklopedieë en soms leksikons. In die praktyk is dit nie 'n groot probleem nie indien die inligting in 'n sekere hulpmiddel van hoë kwaliteit is met betrekking tot die beoogde funksie. Tog sal ek nie een nie maar twee bykomende definisies gee van twee verskillende soorte inligtingshulpmiddels. Daar sal ook 'n bespreking wees oor die relevansie van die konsep woordeboek met betrekking tot die registrasie van navorsing by universiteite.

Sleutelwoorde: WOORDEBOEK, LEKSIKON, ENSIKLOPEDIE, LEKSIKOGRAFIE, INLIGTINGSHULPMIDDEL, LEKSIKOGRAFIEHULPMIDDEL, NAVORSINGSREGISTRASIE, LEKSIKOGRAFIESE DATABASIS, MONOFUNKSIONELE WOORDEBOEK, MEERFUNKSIONELE WOORDEBOEK

1.   Do dictionary users know what a dictionary is?

Any native speaker of Danish will normally immediately know what is meant if someone says: *Luk døren!* (close the door!) or *Du har hul i din bluse* (there's a hole in your blouse). The addressee knows what *dør, hul* and *bluse* mean and will therefore not consult a dictionary in order to understand these sentences. It's a different story if someone says: *Det vil jeg kalde en eufemistisk omskrivning* (I would call that a euphemistic description). In this case many would ask: "What

does *euphemistic* mean?" Or they would look up *euphemistic* in a dictionary. Such consideration of user requirements was the main reason why, in the first editions, the Nudansk Ordbog (a well-known Danish dictionary) did not include definitions for ordinary words if these could be expected to be immediately understood by all potential users whose mother tongue is Danish. Consequently this dictionary contained no definitions for words such as *ordforråd* (vocabulary), *ordne* (arrange), *ordre* (order) and *ordsprog* (proverb), but the word *ordbog* (dictionary) was not defined either.

However self-evident the argument about the assumed common language knowledge and language ability may be, the question whether a word is known to everybody always remains a problematic issue. For example, *ordbog* (dictionary) was not defined — but *leksikon* (lexicon) was. True, in this case one word is a "Danish" word and the other a foreign word. But is *ordbog* really a word understood by all, and *leksikon* not? The answer can be both yes and no. If we answer "yes", then we are saying that all native speakers will understand if we say: *Please pass me the big blue dictionary over there on the table*. If the answer is "no", we imply that hardly any ordinary language user is able to define the word *ordbog* exactly or accurately. Ordinary language users also do not know exactly whether a lexicon and a word list are the same as a dictionary, or what exactly constitutes the difference (if any) between these three terms. And when such ordinary mother-tongue speakers say they know the answer, it will not be difficult to discover that different people offer quite different definitions. Each one will nevertheless insist that their definition is correct. In this respect they are no different from the experts, nor from the descriptions of the word *ordbog* — which are as numerous as the dictionaries themselves and often problematic or even downright incorrect to boot.

I will return to such problematic definitions later; for the moment, here is the dictionary article from the Nudansk Ordbog from the edition that defines all words:

> **ordbog** en bog med ord der er systematisk ordnet, og med oplysning om fx ordenes stavemåde, ordklasse, bøjning og betydning el. deres oversættelse til et andet sprog (= **dictionary** a book with words in a systematic order and with information about the spelling, word class, inflexion and meaning of the words or their translation to another language) (Nudansk Ordbog 2005).

This definition of *dictionary* is one of the better ones. But it is not optimal. To start with, the phrase "a book with words in a systematic order" is already unclear. Do not all books contain words in a systematic order (so that they correspond to the system that is used to represent the sounds of a language in letters)? It is probably rather a book **about** words or combinations of words. But what is a word? Is a name a word? After all, you can also find names of cities, countries and important people in dictionaries. But if defined in this way, a telephone guide would also be a dictionary, and that would not fit the normal

meaning of *dictionary*. The most serious error in the definition quoted, however, is that it excludes most of the electronic dictionaries. The words and combinations of words described are most certainly arranged systematically in a database, but this system has nothing to do with what the user sees. In many cases, the user sees only a single dictionary article when he has entered a search term and given the "search" command. And if only a single dictionary article is displayed, no system can be discerned.

2.   Is it important to know what a dictionary is?

In practice, the title of a lexicographic reference work does not give an unequivocal indication of the type of reference work it is. For example, Politikens Musikordbog (Politiken's Music Dictionary) of 1995 was called a music lexicon by the author who submitted the manuscript for publication. But the publisher changed the title, arguing that the book was being published by the publisher's dictionary department. Had it been published by the textbook department instead, they explained, the book would have been entitled "Lexicon". We cannot say whether the change of title from "lexicon" to "dictionary" had a positive or negative effect on sales. We don't think so. Similarly, there are many reference works with "dictionary" in the title which nobody would call dictionaries, and which do not fit a known definition of "dictionary" either. Moreover, there are many works without the words "dictionary", "lexicon", "word list", "encyclopaedia" or the like in the title, but which must undoubtedly be classified as such. In this regard I would like to refer to an article by Pálfi (2011), which has the meaningful title: "On dictionaries which aren't and non-dictionaries which are". It is a fine article, but although — like many thousands or more on the definition of "dictionary" — it is scientifically interesting, it is neither of immediate practical nor social relevance. In this article I would like to propose a scientifically founded definition and, in addition, discuss practical and social consequences of its terminological application. First of all, it is important to know how a dictionary can be cited. For example, can the common name under which several dictionaries are being sold be used as a title in the references? This applies in the case of the "Music dictionary", Den Danske Musikordbog, which is sold as an electronic dictionary of the Verlag Ordbogen A/S. Actually, this title comprises four dictionaries. They also differ in terms of size; the four hard copies, due to appear in 2012, range from 58, 160, 442 to 444 pages. In the case of the electronic versions, the user receives clear instructions if he clicks on the link "About the dictionary". Here he learns that there are four dictionaries, even though they are not called that in the title, and clear instructions are provided how they can be cited:

Bergenholtz, Inger i samarbejde Henning Bergenholtz. Database: Richard Almind og Martin Gyde Poulsen: *Betydning af musikudtryk.* Odense: Ordbogen.com 2011. (www.ordbogen.com). (ISBN 978-87-788-2555-1)

Bergenholtz, Inger i samarbejde Henning Bergenholtz. Database: Richard Almind og Martin Gyde Poulsen: *Viden om musikudtryk.* Odense: Ordbogen. com 2011. (www.ordbogen.com). (ISBN 978-87-788-2299-4)

Bergenholtz, Inger i samarbejde Henning Bergenholtz. Database: Richard Almind og Martin Gyde Poulsen: *Find et musikudtryk.* Odense: Ordbogen. com 2011. (www.ordbogen.com). (ISBN 978-87-788-2556-8)

Bergenholtz, Inger i samarbejde Henning Bergenholtz. Database: Richard Almind og Martin Gyde Poulsen: *Fremmedordbog med musikudtryk.* Odense: Ordbogen.com 2011. (www.ordbogen.com). (ISBN 978-87-788-2557-5)

Thus a "dictionary" is not always a "dictionary". Accordingly, we will subsequently propose two different definitions of the term "dictionary":

1. When we refer to the Music Dictionary, "dictionary" means a collection of dictionaries, each of which has completely different dictionary articles.

2. For the individual dictionaries as referred to above, "dictionary" is used in the conventional sense, as shown in the next chapter of this article. That these definitions are often problematic and imprecise is less important in this context; the user senses that this is a dictionary and that it should be cited as shown above.

Far more socially relevant, and certainly relevant to the university, is the use of "dictionary" in the research registration, which all researchers must submit. On the one hand the researchers must document that they are fulfilling their research obligations. On the other, the individual universities have an interest in these research results being submitted because the ministry pays the respective university for each newly published scientific contribution or book, provided that it is a scientific dictionary. Until recently, it was customary to extract only a single dictionary from a database, but in principle 5, 10, 20, 100 or even thousands of dictionaries can be extracted from one and the same database. We have actually done this in several cases at the Centre for Lexicography in Denmark. As mentioned earlier, four dictionaries were extracted from a music database; from an accounting database containing Danish, English and Spanish accounting terms, we have to date published 16 different accounting dictionaries, and this number will grow to 27 in a few years. Whether the publisher sells these as a package or separately is of no importance. But for the registration of the research it does matter. When we register, we have so far submitted 16 different accounting dictionaries which the ministry must recognise. Certainly, the ministry would prefer a definition of "dictionary" which would result in fewer editions.

Something similar would apply if one received a library fee for electronic dictionaries. Such a fee is paid, for example, in Denmark and Germany as a

function of the number of books bought by the respective libraries. Oddly enough it is paid only for printed dictionaries, even though libraries buy more electronic than printed dictionaries. However, if one were to be paid the library fee for electronic dictionaries, the same dilemma would arise as with the allocation of ministerial funds to the universities. I propose emphatically that electronic dictionaries should be dealt with in the same way as printed dictionaries, in other words when the works are completely different, a separate royalty or library fee should be paid.

3.   Existing definitions of "dictionary"

If all the scientific contributions related to this problem were gathered together, there would be enough material for several voluminous tomes. I will not quote from the multitude of scientific contributions here, as their definitions do not differ from those I have found on the internet and in some printed dictionaries. Below I will quote a small selection from the many thousands of definitions and mention some points of criticism against each of them, more or less in the same way as those raised against the definition in the Nudansk Dictionary mentioned above. We will see that this definition — despite its problematic imprecision — is better and clearer than most of the other definitions. I did not take particular trouble to look for really good or really poor definitions. I did a Google search with the search string "What is a dictionary?" and then took some of the definitions that appeared on the first three pages. The following discussion therefore does not offer a representative picture of all definitions on the internet. But then that was not the idea either. I want to show what existing definitions reflect as important and in what way they contain problematic or even incorrect statements. Against this background, I want to suggest a definition that is applicable not only to printed or certain types of dictionaries, but takes into account that a dictionary is an information tool and that this tool can be an electronic or printed dictionary. This the first definition does not do; it is brief and easy to understand, but says almost nothing:

(1)   Dictionaries are alphabetical lists of words or entries. They differ in the kind and in the volume of information they hold. (http://www.lib.uct. ac.za/infolit/dict1.htm, found November 2011)

The statement that a dictionary is an alphabetical list is, of course, not correct. It is incorrect for at least two reasons: (A) There are dictionaries with systematically arranged dictionary articles; we call this a systematic macrostructure. This means that the dictionary articles are actually not alphabetic; as in a thematic dictionary, they are arranged in such a way that specific topics are presented in a thematic order. Such dictionaries usually have an alphabetical index which enables the user to look up the dictionary articles more easily. (B) The database of an electronic dictionary is specifically not arranged in alphabetic order at all,

but is provided with its own address internally in the database. This may be a number, with each new dictionary article being given a new serial number. But this numbering has nothing to do with the alphabet. It is an internal allocation in a database which has nothing to do with the presentation of the dictionary articles. More important is that the user is usually looking for a particular word and that he receives as the result exactly one and only one dictionary article. If he receives two or more dictionary articles as a result, these dictionary articles may be arranged alphabetically, but they do not have to be. But even if some of the dictionary articles shown are in alphabetical order, the definition is still not correct. After all, the user sees only a small section of the dictionary. The definition is also inadequate because it would turn even a telephone directory into a dictionary. You could call it that, but this would conflict severely with the present terminological language usage and run counter to any existing understanding of "dictionary". Nevertheless, there are many variants of this type of definition:
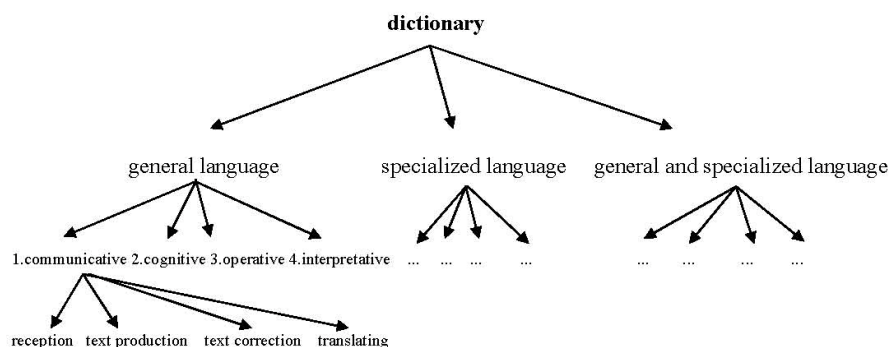
(2)   a reference book containing an alphabetical list of words with information about them (http://wiki.answers.com/Q/What_is_the_biggest_ dictionary_definition, found November 2011)

All the arguments against definition (1) could be repeated here. It can be described as the shortest and probably also as the most primitive. Nevertheless, definitions such as (1) and (2) and the definition below are those found most commonly. Here an attempt is made to define a dictionary by listing the necessary item types:

(3)   A definition for dictionary is a reference that tells you the meanings, parts of speech, sometimes a sentence using the word, and how to pronounce the word. (http://wiki.answers.com/Q/What_is_the_biggest_ dictionary_definition, found November 2011)

Such a definition excludes many information tools we normally consider to be dictionaries: a spelling dictionary normally does not contain definitions, no specimen sentences and pronunciation items; according to this definition, this is not a dictionary. A frequency dictionary, which lists only the word form and its frequency, would also not be worthy of the name; neither would a bilingual dictionary, which in most cases has no meaning items. The reason why this widespread and very inaccurate definition has been repeated in so many variants may be that what so many have said and written before is simply repeated blindly. To be a little more positive, one could regard this as the definition of a monolingual polyfunctional dictionary. Or, in other words: It is the type of dictionary that is often regarded as THE dictionary, but it is really a very special dictionary — the kind that linguists and linguists claiming to be lexicographers view as the sole object. It is a very narrow view of a dictionary, which is aimed only at solving communication problems (text production, text reception

and translation), and in this case not even at translation. Such dictionaries are also called "language dictionaries". Many lexicographers regard only diction-aries indicated as (1) for communicative dictionaries in the illustration below as the object, not the other eleven types of dictionaries:

dictionary

general language   specialized language   general and specialized language

1.communicative 2.cognitive 3.operative 4.interpretative ... .

reception   text production   text correction   translating

The definition thus aims to fulfil non-cognitive functions in particular, i.e. information tools which aim to transmit knowledge. Dictionaries of this kind are often called lexicons or encyclopaedias. This type accounts for almost two-thirds of all lexicographic works. We will attend to possible differences between the different names of lexicographic information tools later. First we want to quote one more example with a variant of the third definition:

(4)   a book consisting of an alphabetical list of words of a language showing their forms and meanings and etymologies (http://wiki.answers.com/ Q/What_is_the_definition_of_a_dictionary, found February 2011)

Almost all the critical arguments against definition (3) can be advanced against this definition as well. Yet it is interesting that only the meaning and etymology item types are deemed decisive for the definition of a dictionary here. Besides this, it is also striking that none of the definitions quoted so far even hint that a dictionary is produced as an aid to people having specific information needs. This also applies to the next definition:

(5)   A dictionary is a book of words of a particular language and their accepted definitions, origins, parts of speech, pronunciation, spelling and in some cases a sample of their use. Depending on the age and target audience, it may also contain cultural slang and/or other non-traditional words as well. A "language translation dictionary" lists the words of one language and their equivalent words in another language. (http://wiki. answers.com/Q/What_is_a_dictionary, found November 2011).

When it comes to the pinch, one could say that definition (5) takes into account

in one point that a dictionary is an information tool, since "language translation dictionary" is mentioned. But even so, it would in this case be a relatively basic translation dictionary, since only equivalence items are mentioned; no translated collocations, no translated examples and no contrast items. But still. Also excluded in this case is the type that is usually called a biscopal dictionary, a dictionary with words from language A translated into language B and simultaneously from language B into language A. It could also be argued that a biscopal dictionary is not a single dictionary, but two dictionaries. But that is not common practice. It should also be noted that in this definition the expression "a book of words" is used. That's not what it is, of course; it is "a book of dictionary articles". There are lemmas for which different data are provided. The misunderstanding could also be due to the fact that the English term "entry" is used to refer to the lemma as well as to the "articles" in a dictionary. But this does not fully explain the misunderstanding. A dictionary entry ("article") is never called a "word" in English. We will quote one last example, but with the observation that this example, as well as the preceding five examples, is quite typical of all existing definitions:

(6)   A dictionary is a reference book that focuses on defining words and phrases, including multiple meanings. The most frequently used dictionary is a language dictionary that includes the majority of frequently used words in a language. (http://www.wisegeek.com/what-is-a-dictionary.htm, found November 2011)

Everything that has been said about (4) and (5) can be repeated with regard to this definition. In addition, there is the statement that a dictionary should always only contain words which are used frequently. This is not correct, of course. There are some dictionaries that do this, e.g. learners' dictionaries, school dictionaries and other general-language dictionaries with fewer than 20,000 lemmas which contain almost only frequently used words. But specialist dictionaries do not do this. For example, a music dictionary does not contain the words most frequently used in music texts; most of the words in texts about music are everyday words. Music dictionaries do not only list the most frequent music terms, but also systematically include those terms which are important in the language of music. Whether these are the most frequent in music texts is not of primary importance. In general-language dictionaries too, the user frequently gets information about less commonly used words, and such words are also found in larger common-language dictionaries. And with good reason, as rarely used words can more often cause problems of reception than frequently used words. Another problem is that this definition focuses exclusively on the idea that a dictionary provides meanings. This excludes bilingual dictionaries, in fact all other dictionaries without items of meaning.

In order to show that somewhat better definitions can also be found on the internet, a final example is quoted. However, such "better" definitions are rare. Although the phrase "a collection of words" is also used in the definition below,

it is also mentioned (besides the usual list of certain item types) that a dictionary is a tool intended to help the users when they have certain information needs:

(7)   A **dictionary** (also called a **wordbook**, **lexicon** or **vocabulary**) is a collection of <u>words</u> in one or more specific languages, often listed <u>alphabetically</u>, with usage information, <u>definitions</u>, etymologies, phonetics, pronunciations, and other information; or a book of words in one language with their equivalents in another, also known as a <u>lexicon</u>. According to <u>Nielsen</u> (2008) a dictionary may be regarded as a <u>lexicographical</u> product that is characterised by three significant features: (1) it has been prepared for one or more functions; (2) it contains data that have been selected for the purpose of fulfilling those functions; and (3) its lexicographic structures link and establish relationships between the data so that they can meet the needs of users and fulfil the functions of the dictionary (Wikipedia, found November 2011).

At the end of my article I will contribute my own definition — two, in fact — showing that "dictionary" has two different meanings. First, however, I would like to summarise two points of uncertainty which prompted the criticism of the definitions quoted (and those not quoted) above. Furthermore, there are some points the quoted definitions do not mention, but which do matter. This also applies to the first point, which does not turn up in any of the definitions quoted:

(1)   How many dictionary articles must a dictionary have before it can be called a dictionary? There are many printed dictionaries with fewer than 1,000 dictionary articles. On the internet I found a dictionary with only 157 dictionary articles. But how low can the number be? I would say: two. Admittedly, I don't know any dictionaries that small, but in principle this should be the limit if one argues that a dictionary must contain a collection of individual dictionary articles.

(2)   A second issue was referred to indirectly in definition (7), which says: "dictionary, also called a wordbook, lexicon or vocabulary". This is also the solution I would propose. I see "dictionary" as a general term for all kinds of lexicographic reference works — as communication, cognitive as well as information tools which aim to fulfil both main functions. It is right that there is a trend towards terminological use, but not more than a trend. In Den Danske Netordbog (2012) under the dictionary article *ordbog* (dictionary), we describe this trend as follows in a note (translated from Danish):

> A distinction can be drawn between lexicographic reference works which offer assistance in case of problems with a concrete text and other reference works where general or specific knowledge can be

obtained. The first type is often called a dictionary and the second type a lexicon or encyclopaedia, but there is no generally agreed distinction between and naming of these two types.

(3)   Are there dictionaries in dictionaries? This question has several aspects. What is interesting is that the question is not asked in any of the definitions of "dictionary" found or otherwise known to me. But dictionaries have many more parts than the lists of lemmas, namely outer texts. This is known and has been described in detail. Some of these outer texts contain brief dictionary articles which correspond directly to the definition of "dictionary", e.g. an alphabetical list of grammatical terms with short explanations of the meanings of these terms. Similarly, lists of currency units or masses and weights also fit any current definition of "dictionary". In other words, you can have several dictionaries within one dictionary. Granted, such outer texts are not usually called "dictionaries". But even biscopal dictionaries are normally referred to as not two dictionaries, but one dictionary if they are bound in the same printed volume, e.g. in a dictionary with two lists of lemmas: Danish–English and English–Danish. Such a dictionary therefore often has a double-barrelled name: Danish–English and English–Danish Dictionary.

That a dictionary can contain a dictionary becomes even clearer when the tradition followed in Japan and other Asian countries is examined. There one can buy a small microcomputer with a number of individual dictionaries installed on it; these are often previously printed dictionaries which have been digitised. Such a computer with up to 400 individual dictionaries is also called a dictionary or, more accurately, an electronic dictionary in Japanese:

電子辞書
ELECTRONIC   DICTIONARY
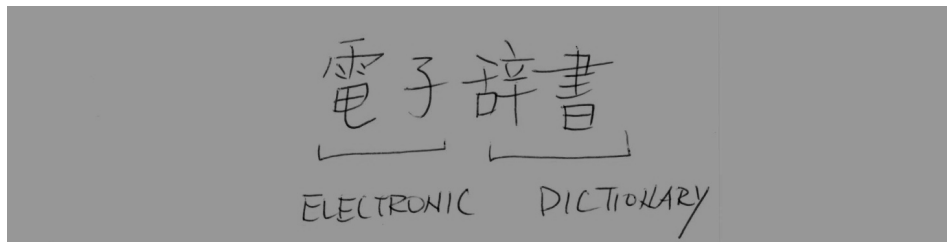
4.   Proposed definition of "dictionary"

After all the criticisms of existing definitions of "dictionary", it might appear somewhat presumptuous to propose one. I will propose one nevertheless — and not just one, but two, as it has become clear that "dictionary" has not just one, but two distinctly different meanings. It is crucial to explain that a dictionary is an information tool that must satisfy specific requirements:

Dictionary

1. Lexicographic reference work containing dictionary articles related to individual topics or elements of language, and possibly several outer texts as well, which can be consulted if someone needs assistance with text reception, text production or translation or would simply like to know more about a word, part of a word or a combination of words.

2. Lexicographic reference work consisting of several dictionaries, each of which corresponds to the definition of an individual dictionary, i.e. a reference work containing dictionary articles related to individual topics or elements of language, and possibly several outer texts as well, which can be consulted if someone needs assistance with text reception, text production or translation or would simply like to know more about a word, part of a word or a combination of words.

5.   References

Bergenholtz, Inger i samarbejde Henning Bergenholtz. Database: Richard Almind og Martin Gyde Poulsen: *Betydning af musikudtryk.* Odense: Ordbogen.com 2011. (www.ordbogen.com).

Bergenholtz, Inger i samarbejde Henning Bergenholtz. Database: Richard Almind og Martin Gyde Poulsen: *Find et musikudtryk.* Odense: Ordbogen.com 2011. (www.ordbogen.com).

Bergenholtz, Inger i samarbejde Henning Bergenholtz. Database: Richard Almind og Martin Gyde Poulsen: *Fremmedordbog med musikudtryk.* Odense: Ordbogen.com 2011. (www.ordbogen. com).

Bergenholtz, Inger i samarbejde Henning Bergenholtz. Database: Richard Almind og Martin Gyde Poulsen: *Viden om musikudtryk.* Odense: Ordbogen.com 2011. (www.ordbogen.com).

Den Danske Netordbog. 2012. = Bergenholtz, Henning under medvirken af Filip Odgaard Bodilsen, Kathrine Brosbøl Eriksen, Helene R. Gudmann, Stine Busk Hedegaard, Helene Halkjær Jensen, Jesper Halkjær Jensen, Christian Kjølhede, Anders Kroman Liin, Jane Nguyen, Heidi Agerbo Pedersen. 2012. Database: Richard Almind, Rasmus Theodor Styrk, Peter Christensen: *Den Danske Netordbog.* Odense: Ordbogen.com. (www.ordbogen.com) (7th edition)

Nudansk Ordbog. 2005. = Politikens Nudansk Ordbog. 2005. 19th edition. Copenhagen: Politikens Forlag.

Pálfi, Loránd-Levente. 2011. Om ordbøger, som ikke er ordbøger, og ikke-ordbøger, som er or ordbøger. Arboe, Torben and Inger Schoonderbeek Hansen (Eds.). 2011. *Jysk, ømål, rigsdansk mv. Studier i dansk sprog med sideblik til nordisk og tysk — Festskrift til Viggo Sørensen og Ove Rasmussen udgivet på Viggo Sørensens 70-års dag den 25. februar 2011:* 325-335. Aarhus: Aarhus University.

Politikens musikordbog = Inger Bergenholtz. 1995. *Politikens musikordbog.* Copenhagen: Politikens Forlag.