Henning Bergenholtz / Heidi Agerbo (Aarhus)

Lexicographical structuring: the number and types of fields, data distribution, searching and data presentation[1]

Abstract: This contribution will not describe the structure in existing dictionaries. Instead, it will focus on the decisions that lexicographers make when they draw up the concept for and carry out the production of one or more new dictionaries, or when they consider making changes in the data presentation in an existing dictionary. This part of the lexicographical work is what we call structuring, which encompasses a number of various lexicographical decisions. One of these is choosing the fields that a database should contain. Typically, for some of these field types, it will be easy to distribute data, but for other fields it will require much consideration as there are several distribution options with different outcomes of varying usefulness. A second type of lexicographical decision to be made by the lexicographer is the predefined searching, which involves in what order searches are to be made in the different database fields and how these searches should result in a certain presentation of fields in the dictionary. An essential part of the argumentation in this contribution is based on the distinction between polyfunctional and monofunctional dictionaries. Most printed dictionaries are polyfunctional dictionaries, which are close to useless on electronic devices such as tablets and smartphones as they contain a vast amount of data. Only by producing monofunctional dictionaries is it possible to avoid information overload. In the case of monofunctional dictionaries, lexicographical structuring becomes particularly important as these dictionaries are derived from the same database but contain completely different structures.

---

1 Lexicographical structure and lexicographical structuring

In this article, we will refer very little to the existing literature about lexicographical structures. We will mainly refer to Tarp (2008, 101–113), who makes a general critique of the continuous discussion about structure, and to his conclusions (2008, 102):

1. The lexicographical relevance of all the structures that have been identified has not been explained sufficiently.
2. The structures are described formally – not in relation to dictionary functions.
3. No clear structure theory that is common to both printed and electronic dictionaries has been formulated.

To this critique, we will add that most scholars make contemplative lexicography (also called retrospective lexicography), analysing finished dictionaries. Like Tarp (2002), we are more interested in transformative lexicography. This type of lexicography focuses on theoretical analyses of potential types of user situations for certain types of user needs for a certain user type. Incorporating and working with such analyses, the lexicographer makes proposals for new dictionary concepts. This main interest in transformative lexicography does not mean that structures in finished printed and electronic dictionaries are not scientifically relevant or interesting, only that we are primarily interested in the process of dictionary making, especially with the development of concepts for e-dictionaries. With the exception of the selection of fields in the database, we can talk about a running process, e.g. if we look at the distribution of data to the fields in the database, the searching possibilities and the presentation(s) of data in one or more dictionaries. Thus, we will not focus on **structure**, but on **structuring**. In our use of the term **structure**, we follow a revised version of the definition in Tarp (2008, 109). Tarp only talks about data in a dictionary; to this, we add the distribution of data to data fields in the database:

A **lexicographical structure** is a set of relations that exist between the data in a database or in a concrete dictionary in terms of its form, content or mutual position in the dictionary article.

Lexicographical structuring is a set of decisions made by the lexicographer. The decisions lead to the selection of certain item types from the database in a concrete dictionary and to the planned set of relations between this data in a database or in a concrete dictionary in terms of its form, content or mutual position.

Lexicographical structuring leads to certain structures in one or more dictionaries. This presupposes that we can extract one or many dictionaries from one database, and the dictionaries extracted from one database can be published in printed or

in electronic versions. Lexicographical structuring is not one single process; instead, it involves several distinct phases, the following six phases as a minimum:

1. Selection of fields for the database, i.e. fields containing different items types. You cannot present an item type in a dictionary if this type does not have its own field in the database - this is banal, but must be considered in the planning phase. The lexicographer should often not present the data from all the database fields in a specific dictionary (see chapter 2).
2. Searching in the database fields. This has nothing to do with a concrete user's search in a concrete dictionary or in a dictionary article. Instead, it is the electronic search in the database based on the assumed user needs in a certain situation type (see chapter 3). This search can be predefined by the lexicographer or selected by the user. It is important to notice that a search can be made in a certain database field even though this field is not presented in the concrete dictionary. The search intends to find the cards (lemmas) that could be relevant for a certain user need.
3. Deciding which fields should be presented in a certain dictionary for a certain type of user in a certain situation type (see chapter 3.3).
4. Deciding in what order the selected fields with a certain item type should be presented (see chapter 3.3).
5. Deciding what kinds of markers should indicate how each item type should be used so the user can easily identify the type. We do not discuss this part of the structuring in this paper (see Almind/Bergenholtz 2000).
6. Decisions about the layout for printed dictionaries or the design for e-dictionaries (see Almind 2005a, 2005b).

2 Data distribution between the meaning fields

2.1 The selection and use of data fields and the distribution of data

When structuring the database, the lexicographer has to decide on the number of fields with which to work (phase 1 in the list above). However, these fields do not necessarily make up the number of different data types that will also occur in the extracted dictionaries, but these are the potential fields that may be presented in them. This distinction between database fields and dictionary fields is based on the fact that there is a difference between a database and a dictionary though many lexicographers do not distinguish between the two (Bergenholtz/Skovgård Nielsen, 2013). A possible difference in the number of database data fields and dictionary data fields is a consequence of the distinction between monofunctional and polyfunctional dictionaries.

In the database, a lexicographer works with a predefined number of fields, which have been conceptualized by the lexicographer himself and produced by an IT programmer. Some fields are relevant for some lemmas, while other fields are relevant for other lemmas, and some types of data are relevant for some functions, while other

types of data are relevant for other functions, though with a certain amount of overlap. Thus, the data that the lexicographer works with in the database is not necessarily the same data that will be extracted and shown to the dictionary user in the dictionaries that are extracted from this database.

Most dictionaries are polyfunctional and therefore they have not been designed to help dictionary users in specific user situations; they simply provide all the data that has been added to the current database, which may cause a situation of information overload if a dictionary user is looking for something specific in a dictionary, but cannot find it due to the enormous amount of data it contains. Contrary to this type of dictionary, a monofunctional dictionary is aimed at helping a certain type of user with a certain type of problem in a certain type of situation. In order to help the dictionary user as quickly and precisely as possible, a monofunctional dictionary only contains the type of data that is needed for one specific type of user situation. This means that even though a database may contain for instance 10, 30, 100 or 500 different data fields, only a selected number of these – a selection made by the lexicographer – will be extracted and presented in the specific dictionary, that is, the data fields that are relevant for the specific type of user situation. The types of fields to include in a database depend on the individual lexicographer, and this is also the case concerning the number of fields – both of these will depend on what dictionaries will be produced on the basis of these fields.

The Danish Internet Dictionaries, which are six Danish general language e-dictionaries extracted from one and the same database, are based on this distinction between mono- and polyfunctional dictionaries. The number and types of fields forming the structure of this specific database have been decided on and defined by the lexicographers working in the project. These fields were selected based on their use in the dictionaries that were initially expected to be extracted from this database. Many more fields could have been included, and, indeed, some have been added after the initial construction of the database – this, however, should be avoided as it is not always an easy task to include them in the dictionaries after the conception phase. For example, the database did not include a lexical remark field or a synonym remark field in its first version; these were added later as it became clear that such fields are needed in some of the dictionaries that are extracted from the database. As a consequence of this, some of the existing fields had to be partially redefined as some of the new fields were used for some of the data that had occurred in one or more of the old fields, e.g. when the lexical remark field was added, some of the information that was previously incorporated in the meaning field was moved to the lexical remark field as this latter field contains data that is considered to be extra information to the definition provided in the meaning field. Other fields have also been redefined as they were not as useful as originally planned, e.g. at the beginning of the project, the second reference field was typically used for writing the stem of the lemma. Today, this field is instead used as an association field or "good to know" field, i.e. things that may be relevant or interesting for the user when he has made a search for the given lemma.

Other fields could be added to the list of data fields presented below, e.g. one for inserting links to images, thus narrowing down the use of the current internet link field to webpages that primarily contain textual information, a grammar field with the imperative form of the lemma or a grammar field with the passive form(s) of the lemma. It could also have been the case that fewer fields had been enough, e.g. only a grammar field, a meaning field and an internet link field. The number and types of fields will depend on their usefulness in the planned dictionaries. The types of fields to be selected could also have been other types than the ones in the list below. We could imagine a remark field for all the fields in the list below, not only for the meaning field (lexical remark), grammar field(s) (grammar/spelling remark) and synonym field (synonym remark), but also for e.g. the collocation field and the example field. We could also have included an etymology field, a pronunciation field, a style field or even a book reference field in the database. The more familiar ones of these fields will, of course, be easier for the lexicographer to work with – and also for the dictionary user to use – whereas new types of fields require detailed descriptions, i.e. describing what type of data we want in this field, and experience in using them, both for the lexicographer and the dictionary user. In addition to selecting the types and number of fields, a different number of units for each field type could be made possible. As indicated in the list below with the plural form of some of the words, fields number 10, 13–18 and 20–21 may occur more than once, thus there may e.g. be more than one collocation and more than one example connected to the same lemma field in the database. The same could have been made possible for e.g. the remark fields and the internet link field, but in this specific database example, it is only possible to have one of each of these two field types. As all of these examples demonstrate, an infinite number of data fields could be thought up and incorporated in a database. This would be no problem for the user as he will only be presented with the data occurring in the fields that have been selected for presentation in the current dictionary. This means that there is no risk of presenting the user with more information than he needs. However, even though the user would have no problem with a high number of data fields in the database, it may cause problems for the lexicographer(s) working in the project. There is quite a difference between having to get an overview of 24 or 50 fields rather than of 100, 500 or 1,000 fields. Therefore, it is important to keep in mind what use each data field will have in the planned dictionaries, and whether the data from a certain field is actually needed in one or more of these dictionaries.

The database used in the construction of the above-mentioned Danish e-dictionaries has 24 fields, but, as has already been emphasized, this list is only an example of the fields that could occur in a database – we are not saying that this is the final or the optimal result for the types and number of fields to occur in a database. In actual fact, the first version of this list had 20 fields, the current version has 24, and we are currently planning a new version with extra fields for style markers. The following list presents the current number of fields in the database:

1. Lemma
2. Sublemma
3. Homonym number
4. Polysem number
5. Meaning
6. Lexical remark
7. Lexical remark for text production
8. Grammar, word class
9. Grammar, recommended inflexion
10. Grammar, non-recommended inflexion (one or more)
11. Grammar/spelling remark
12. First reference
13. Second reference(s)
14. Collocation(s)
15. Example(s)
16. Word formation(s)
17. Synonym(s)
18. Antonym(s)
19. Synonym remark
20. Proverb(s)
21. Idiom(s)
22. Idiom meaning
23. Internet link
24. Memo field

The distribution of data not only takes place between the database fields, but also to an integrated outer text as well as to other texts outside the database (Bergenholtz/Tarp, 2005: 124–126). An outer text could be a systematic overview or it could be an integrated grammar included in the database. There could be a reference in the dictionary article to this integrated grammar instead of a long grammatical note, or the integrated grammar could be used as additional information, supporting what has been written in the article. In the case of other texts, the distribution concerns existing texts that are not integrated in the database, e.g. an external grammar. Thus, if the user is interested in learning more about the grammar of the relevant lemma, he can look it up in the external grammar to which the article refers. As for the data field distribution, there is not one fixed way of distributing the data between the selected database fields, but many ways of distributing it. Of course, for some fields it is rather clear what should occur in the field, e.g. the grammar, word class field. For other fields, e.g. the meaning field and the synonym remark field, the distribution of data can be done in sometimes a few, sometimes many different ways with different selections or consequences for the presentation in the extracted dictionaries.

Between these fields in the database, the lexicographer distributes the different types of data to be incorporated in the dictionaries. The distribution of data in these fields is made by the lexicographer and the decision behind this distribution is based on the assumed use of this data for certain user needs. This means that every time the

lexicographer adds and distributes data to the database, he has to consider in which situations this data is useful, i.e. in what fields the data should be presented (Bergenholtz/Agerbo 2014).

In the following section, we will present a number of meaning elements related to the word **blindhed** (= blindness) and illustrate what choices the lexicographer has to consider or could consider when working with the distribution of these meaning elements in the database. This example does not represent an actual dictionary article or a suggestion for a complete dictionary article, but it is an example that illustrates our argument that the distribution of meaning elements is not fixed and that there may be several useful ways of distributing the data, though some solutions are often better than others. As described in Bergenholtz/Agerbo (2014), the first step in finding the meaning elements to be applied in the database is to find different types of data and write them down in a more or less systematized way and from this set of data extract the meaning elements. The following example with **blindhed** demonstrates this type of data collection. Thereafter, the lexicographer distributes these meaning elements to the predefined and pre-chosen database fields. When performing this task, the lexicographer should keep in mind in which extracted dictionaries which fields occur, thus remembering what the dictionary users will be able to see in the dictionaries as this will not necessarily be the same as what the lexicographer has added to the database and also making sure that the distribution of the data is the most useful one for all the extracted dictionaries.

2.2 An example of data distribution

In the example below, List 1 is a list of selected data which has been found in the empirical basis, and from this list we have extracted data directly (e.g. collocations and links) as well as constructed abstractions, which we have written down in another list, List 2. The data in List 2 consists of meaning elements (for the reader who is interested in knowing how to find this data, see Bergenholtz/Agerbo 2014), hence List 2 is a worklist which the lexicographer uses in the distribution and formulation of these elements. The "titles" (grammar, etymology etc.) in the lists below are only used to categorize the data. It should not appear as though it has already been settled how to distribute the different meaning elements to the different database fields because the types and number of database fields are not fixed; they will, as described above, differ from one dictionary project to the next.

2.2.1 List 1: Data from the empirical basis

Grammar
- en blindhed (= a blindness; singular, indefinite)
- blindheden (= the blindness; singular, definite)
- blindheds (= blindness'; singular, indefinite, genitive)
- blindhedens (= the blindness'; singular, indefinite, genitive)

Etymology
The word is of Germanic origin; it is related to the Dutch and German word "blind".

Text examples (example sentences)
- Trachoma er en øjensygdom forårsaget af bakterien Chlamydia trachomatis. Hvis man ikke bliver behandlet, kan sygdommen medføre **blindhed**. (= Trachoma is an eye disease caused by the bacteria Chlamydia trachomatis. If a person is not treated for this, the disease may cause blindness.)
- **Blindhed** er en tilstand, hvor man ikke kan se, men alligevel godt kan forestille sig, hvad der er rundt omkring én. (= Blindness is a condition that entails the inability to see, but despite this you are able to imagine what occurs around you.)
- Forekomsten af **blindhed** er faldende blandt type 1 diabetikere. (= The number of occurrences of blindness is decreasing among type 1 diabetics.)

Collocations
- akut blindhed (= acute blindness)
- blindhed forårsaget af diabetes (= blindness caused by diabetes)
- blindhed forårsaget af grå stær (= blindness caused by cataract)
- blindhed hos hunde (= blindness among dogs)
- blindhed, som er en risiko for sukkersygepatienter (= blindness, which is a risk among patients with diabetes)
- blindhed, som er en tilstand, der kan være medfødt eller erhverves (= blindness, which is a condition that can be innate or acquired)
- forbigående blindhed (= temporary blindness)
- forebygge blindhed blandt diabetikere (= prevent blindness among diabetics)
- lide af blindhed (= suffer from blindness)
- medfødt blindhed (= innate blindness)
- midlertidig blindhed (= temporary blindness)
- nedsætte risikoen for blindhed (= reduce the risk of blindness)
- praktisk blindhed (= practical blindness)
- social blindhed (= social blindness)
- total blindhed (= total blindness)
- årsager til blindhed (= causes of blindness)

Related words
- blind (= blind)
- braille (= braille)
- døvhed (= deafness)
- førerhund (= guide dog)
- blindestok (= white cane)
- synshandicap (= visual disability)
- stumhed (= muteness)

Word formations
ansigtsblindhed (= face blindness)
blind (= blind)
farveblindhed (= color blindness)
flodblindhed (= river blindness)
natteblindhed (= night blindness)
ordblindhed (= dyslexia)
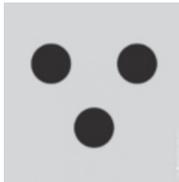sneblindhed (= snow blindness)

Synonyms
– ablepsi (= ablepsia)
– det at være blind (= the condition of being blind)
– manglende evne til at se (= lacking the ability to see)
– det at være synshandicappet (= the state of being visually handicapped)
– det at være synshæmmet (= the state of being visually impaired)

Links
– https://www.sundhed.dk/sundhedsfaglig/laegehaandbogen/oeje/tilstande-og-sygdomme/oevrige-sygdomme/svagsynethed-og-blindhed/
– http://www.danformer.com/danskblindesamfund.dk
– http://en.wikipedia.org/wiki/Blindness
– http://www.who.int/topics/blindness/en/

Symbols/illustrations

Blind man
(guide dog, sunglasses,
white cane)

Symbol of blindness (1)

Symbol of blindness (2)
 ("The man with the white
cane")

These different types of data lead to the following meaning elements that have been listed in four groups (A-D), making it easier to work with them when distributing them to the different fields in the relevant database:

2.2.2 List 2: Meaning elements

(A) The causes of blindness:
– Some people or animals are born with it; others acquire it after some time.
– In some cases it occurs suddenly, in other cases it occurs gradually.
– Some diseases cause blindness, e.g. diabetes, glaucoma and cataract.

(B) Different categories of blindness:
- Blindness can be described on a scale going from (mildly) visually impaired to completely blind.
- WHO's definition: When you cannot count fingers in a distance of three meters or more.
- Danish definition: When the ability to discriminate between things or the visual acuity is reduced to $^1/_{10}$ or less of the normal.
- Denmark: Blindness is categorized into the following numbers and terms:
  (1) normal to slightly reduced vision: > 6/18
  (2) visual impairment: ≤ 6/18 and > 6/60
  (3) social blindness: ≤ 6/60 and > 1/60
  (4) practical blindness: ≤ 1/60
  (5) total blindness: sense of light without projection or no sense of light at all

(C) Gear associated with blindness:
- Guide dog
- White cane
- Sunglasses
- A label with the pictograph or symbol for the visually impaired and blind – there are two types of symbols: In order to be recognized when moving around outdoor, blind people and visually impaired people used to wear the yellow label or armband with three black dots. In 1998, this was replaced with the traffic label "the man with the white cane", which is an international symbol for people with a visual impairment. The yellow label is still in use, though not as much as the new one. The traffic label is typically used either as a badge on one's jacket or as a wristband. The white cane, which was approved by the Danish Road Traffic Act in 1976, is also an international symbol for the blind and the visually impaired. The traffic label only serves to inform the surroundings that the person carrying it is visually impaired whereas the white cane, in addition to this purpose, functions as a personal aid that makes it possible for the blind and visually impaired to walk around without anyone escorting them.

(D) Synonyms, illustrations, word formations, links etc. transferred directly from List 1.

When the lexicographer has compiled these lists, he can begin to distribute the meaning elements. In List 2 that has been compiled for the word based on the empirical basis, there are some types of data that are more flexible than others in relation to distribution. As mentioned above, it is rather clear that that word in question (here: **blindhed**) will appear in the lemma field; that the inflexions of the word should appear in a grammar, inflexion field; and that the word class should appear in a grammar, word class field. But for many of the other meaning elements occurring in List 2, there are several distribution options in the database that the lexicographer needs to consider as these will affect the presentation of the data in the extracted dictionaries.

The meaning elements in groups (A) and (B) in List 2 show that it is relevant to mention the different categories of blindness and the causes of blindness. This could be done in the meaning field, but this may take up too much space as these meaning

elements cannot be incorporated properly in a few sentences; thus, some of this information could instead be distributed to a lexical remark field in which more detailed information could be provided to the reader who is interested in knowing this. In this field, it is not as important for the lexicographer to be concise as in the meaning field. In Denmark, the different categories of blindness are: **svagsynethed** (= visual impairment), **social blindhed** (= social blindness), **praktisk blindhed** (= practical blindness) and **total blindhed** (= total blindness). Apart from the lexical remark field, this data could (also) be presented in the collocation fields. However, if the terms are added to the collocation fields, there are two problems: One, the term **svagsynethed** only consists of one word in Danish, making it difficult to argue for its categorization as a collocation; Two, the four terms presented here are specific types of blindness with each their specific definition that a user cannot infer, thus, it may be better to place them somewhere else in the database for which it is possible to assign them their own definition – unless there is a collocation remark field connected to each collocation field in the database, in which case it is possible to provide definitions of or other types of comments for each collocation. A solution for providing the terms with each their own meaning is to present them either in a sublemma field or in the lemma field in which cases they can be assigned each their own meaning field:

Each category is a lemma:
Lemma 1: svagsynethed
Lemma 2: social blindhed
Lemma 3: praktisk blindhed
Lemma 4: total blindhed

Each category is a sublemma connected to the lemma **blindhed**:

Lemma: blindhed
   Sublemma 1: svagsynethed
   Sublemma 2: social blindhed
   Sublemma 3: praktisk blindhed
   Sublemma 4: total blindhed

These are two viable solutions, but in the former case, the terms would not appear in the dictionary article for blindness where the dictionary user would very likely search for them, and in the latter case, the word **svagsynethed** causes problems as it does not incorporate the word **blindhed**, which is a requirement for sublemmas. Therefore, a different solution could be to distribute them to the meaning field or, as mentioned, to the lexical remark field, in the former case without much elaboration as the definition should be concise, in the latter case providing as much detail as is useful for the dictionary user.

    Group (C) in List 2 includes meaning elements that are the typical items associated with blindness. These could be relevant to include in one or more of the database fields: They could be mentioned in a lexical remark field, elaborating on their function and symbolic value, but it would also or instead be possible to mention them in a

related words field, in this way referring the user to the dictionary articles containing these related words. A third option would be to mention them in one or more illustration remark fields – if each illustration field is connected to an illustration remark field – as the words are connected to the illustrations provided in List 1. In both a lexical remark field and an illustration remark field, a longer description, as provided under (C), could be given. However, the lexicographer should consider whether it is actually necessary to have both a lexical remark field as well as an illustration remark field in the database and also if the meaning items should occur in one, two or all three of the mentioned fields.

As group (D) in list 2 shows, there are several identified synonyms that could be applied in the synonym field in the database. Some of these synonyms are clearly synonyms while others may be regarded as definitions. Many dictionaries provide the simple definition of **blindness** as 'lacking the ability to see', which may be enough information for certain dictionary users in certain situations, but for others this may not be enough. Thus, the lexicographer may have to add more meaning elements to the meaning field. If the lexicographer applies definition 6 in the meaning field in the database, he could apply the simple definitions 2 and 3 as synonyms in each their synonym field. If there is no synonym field, but a meaning field in the database, the word **ablepsi** may instead be applied in the meaning field, though of course the lexicographer should consider whether or not such a short definition is sufficient.

Suggested synonyms → synonym field
1. ablepsi (= ablepsia)
2. det at være blind (= the condition of being blind)
3. manglende evne til at se (= lacking the ability to see)
4. det at være synshandicappet (= the state of being visually handicapped)
5. det at være synshæmmet (= the state of being visually impaired)

Suggested definitions → meaning field
1. ablepsi (= ablepsia)
2. det at være blind (= the condition of being blind)
3. manglende evne til at se (= lacking the ability to see)
4. manglende evne til at kunne se, hvilket kan være medfødt eller opstået senere i livet (= lacking the ability to see, which can be innate or can occur later in life)
5. manglende evne til at kunne se, hvilket kan være medfødt eller opstået senere i livet, og hvilket inddeles i forskellige sværhedsgrader gående fra en mindre synshæmning til fuldstændig blind (= lacking the ability to see, which can be innate or can occur later in life, and which can be categorized based on the severity of the impairment, a scale going from a minor impairment to total blindness)
6. manglende evne til at kunne se med sine øjne, hvilket kan være medfødt eller opstået senere i livet på grund af skade eller sygdom såsom diabetes eller grå stær, og hvilket inddeles i forskellige sværhedsgrader gående fra en mindre synshæmning til fuldstændig blind (= lacking the ability to see, which can be innate or can occur later in life due to an injury or a disease such as diabetes or cataract, and which can be categorized based on the severity of the impairment, a scale going from a minor impairment to total blindness)

Group (D) in List 2 also includes the meaning elements that are formed from the lemma, i.e. word formations (compounds) such as **natteblindhed** (= night blindness) and **farveblindhed** (= color blindness). Contrary to the terms above representing the four categories of blindness, these compounds are not different categorizations or degrees of blindness, but specific types of blindness, which means that adding these meaning elements to a lexical remark field would probably not be as good a solution as distributing them either to a word formation field or to a related words field. It may be relevant to add some of the words to both of these database fields, but only if the fields occur in different dictionaries or else the two fields just repeat the same data. Of course, it is also possible to lemmatize these compounds, which makes it possible to provide them with each their definition. If there were a word formation remark field in the database, comments about the meaning of these compounds could also be made in the dictionary article for **blindhed** and not in separate dictionary articles.

Group (D) also includes a number of links. The obvious place to include these would be in a link field in which each link is inserted. Each link could be presented in the extracted dictionary either with its exact address or with a given title, in the latter case telling the user directly what the link contains. However, if the lexicographer for example added a comment in the lexical remark field about WHO and this organization's definition of blindness, it may be of use to insert the link to a webpage containing this definition directly in the text in the lexical remark field rather than in a separate link field.

These examples of data distribution for the Danish lemma **blindhed** demonstrate how many options a lexicographer has at hand when distributing data to the different fields in the database. The decision about which of the solutions to incorporate in the dictionary article should be based on the function of the dictionary.

3 Searching, navigating and data presentation

We follow the clear distinction described by Bothma (2011, 81) between searching and navigating. When navigating, the user follows predefined links integrated in the search process. In printed dictionaries, the user can go to different parts of the outer text or to different lemma lists, or he can follow a cross reference hoping to get the wanted information. In e-dictionaries, navigating is similar as the users follow different kinds of links or go to other parts of the dictionary. Searching implies that the user has to define and use a search string – this could consist of a single word, a phrase, or a number of words/phrases combined by means of Boolean operators. Both options, searching and navigating, can be used by the lexicographers in their use of the database and by the dictionary users in their use of the dictionaries. In addition, the user can select a certain filter with predefined search options for a certain type of user situation (Bothma, 2011, 86f). From a functional viewpoint, this filtering leads to different dictionaries. The articles that are derived as a selection of such predefined searches

can also be shown in distinct printed dictionaries (Bergenholtz 2012). In the working language of lexicography, we normally say, "I work with the database". This does not mean that the lexicographer makes changes in the database itself. He puts data in or takes data out of single fields in the user interface that has been specifically made for the lexicographer, not in the database itself. In some cases, it is possible to make use of different user interfaces for different members of the group of lexicographers working in the same project. Another user interface is made for the dictionary user. In the case of a polyfunctional dictionary, we often have only one user interface or one printed dictionary extracted from the database, but we can have many different user interfaces, that is, a vast number of monofunctional dictionaries derived from one database.

3.1 Lexicographers searching in the database

As mentioned above, the lexicographer makes use of the search facilities in the user interface designed for the lexicographer during his work with the database. The lexicographer can search in one or more fields to find one or more cards during the working process. When we open our database, we get all the cards in the database. The first of the cards shown is the one that was made as the very first card and all other cards appear in the sequence corresponding to when they were made. If a lexicographer in the running work works with a certain letter, e.g. all cards beginning with the letter D, he first searches for all cards with the first letter D and then asks the program to sort these cards alphabetically. The lexicographer can also make a search among all the cards in the database using Boolean operators, e.g. in the lemma field for all cards with a certain last stem or with a certain meaning element in the main meaning field and at the same time only for nouns. You can sort the cards alphabetically, or you can leave them unsorted, which means that you will be given the first created card as the first one and the last created card as the last card. Defining the different ways of searching in the database is a part of structuring: each result gives a certain selection with a certain structure.

3.2 Lemma presentation in dictionaries

With a printed dictionary, you can either search for the lemma in the lemma list or you can use an index, e.g. with the page number. The last possibility is the one normally needed for a systematic macrostructure, the first one for an alphabetical macrostructure (Wiegand/Gouws 2013). In both cases, when looking at printed dictionaries, you have the whole lemma list shown and have to find the wanted lemma at its place in the lemma list. It is in many ways easier and quicker to find a certain lemma in an e-dictionary, but you cannot talk about macrostructure in the same way as in a printed diction-

ary – this is only possible for primitive e-dictionaries that have been made like printed dictionaries where you see the whole lemma list in one long list or in a separate list after the first letter. In real e-dictionaries, you will use a search string and in most cases get one and only one lemma and see the whole dictionary article with this lemma. If you search with a search string that results in more than one dictionary article, the selection can be shown as a list with the lemmas that have been found. In addition to this, some dictionaries show an extract of the whole lemma list, but the user can scroll up or down in this list and in principle see the whole lemma list and click on one of the lemmas in the list in order to navigate to the dictionary article with this lemma. You could call this a macrostructure, but such a use of the term would be quite different from the use of macrostructure for printed dictionaries where you start with the whole list. In an e-dictionary like The Danish dictionary (2014), you start with a certain lemma and get additional links to the rest of the lemma list. There are many other possibilities to show parts of the lemma list. In The Danish Internet dictionary (2014), you get four lists. The first one is a list of the last five lemmas the specific user has used as a search string. The lemmas in this and in the following lists can be used as links to navigate to the single lemmas. The second list contains the search string as the first stem in other lemmas, e.g. the search string **tillæg** (= addition):

> tillægsplade (= extra leaf)
> tillægsform (= participle)
> tillægsagtig (= attributive)
> tillægsprøve (= supplementary test)
> tillægsspørgsmål (= supplementary question)
> tillægsaftale (= additional agreement)

This list and the following third list are not identical to the word formation items in the database. They are extra lists provided for the interested user. As in the case of the list with all the lemmas in The Danish dictionary (2014), these lists are lemma lists without a clear function. The third list is formed when the search string is the last stem in other lemmas, here again exemplified with the stem **tillæg** (= addition):

> særtillæg (= supplement)
> udetillæg (= expatriation allowance)
> nattillæg (= night supplement)
> løntillæg (= bonus)
> genetillæg (= nuisance compensation)
> gagetillæg (= salary increment)

The fourth list is a list with semantically related lemmas:

> udover (= in addition to)
> udvidelse (= expansion)

yderligere (= further)
tilskud (= supplement)
ændringer (= changes)
ud over (= in addition to)

This list is not identical to the synonym items in the database because in the case of the synonyms items, there are more and other synonyms to the lemma **tillæg** than shown in this list, and these synonyms have been divided into three groups belonging to each their polysem (the lemma **tillæg** has three polysems in THE DANISH INTERNET DICTIONARY). It is a list made automatically by extracting the words from the meaning items, i.e. it is not a list with synonyms.

Dictionaries could offer many other lemma lists, and IT specialists are indeed sometimes too fond of these lists. It is not quite clear if such additional lists are really helpful for the user in specific user situations, apart from the function called lexicographical entertainment. However, the central point of our discussion is another: We can call the construction of such lemma lists for lexicographical structuring, but these do not form a macrostructure in the sense of the word as it is applied in printed dictionaries.

3.3 Predefined searching and dictionary article presentation

If the lexicographer is preparing a polyfunctional dictionary, he only has to decide on the sequence of data types from the single fields in the database and to add field markers to help the user to quicker access. This is the normal procedure in the development of printed polyfunctional dictionaries. The resulting dictionary articles can be described in a static way as a microstructure, cf. Wiegand/Smit (2013). Of course, this procedure could also be applied in the production of monofunctional printed dictionaries.

In this section, we will focus on e-dictionaries and the search and presentation options that exist for electronic information tools. However, it should be mentioned that the three e-dictionaries described also exist as printed dictionaries based on the 2012 versions of the e-dictionaries.

In our experience, the predefined searching and the selected sequences of article presentation, also called **filtering**, is changed several times every year – this happens as a consequence of comments received by our dictionary users as well as of our own discoveries of possible or needed changes in the preselected searches or presentations. We are therefore much more interested in the process of lexicographical decisions that we call **lexicographical structuring**.

In the case of the selection of fields for a database, the lexicographer defines as many fields for different data types as he expects to use in order to fulfil the needs for the planned dictionary or dictionaries. In the case of the database for the Danish general language dictionaries mentioned earlier, we had planned at least three mono-

functional and three polyfunctional dictionaries: a dictionary that included all items from the database, a grammar dictionary, a text reception dictionary, a synonym dictionary, a text production dictionary and a special text production dictionary for when the user does not know the word he needs to use. For those functions, we prepared 24 fields as shown above in chapter 2.

For all types of searches in e-tools, there are a number of search criteria to choose between. You could let the programme search in one or more fields in the following ways (XXX represents the applied search string):

1. *XXX*
2. *XXX
3. XXX*
4. =XXX
5. ==XXX
6. Fuzzy search

Each type of search has its own advantages and disadvantages:

(1) The search string has to correspond exactly to the word or to a part of one of the words in the search field. If the search field, for example the lemma field, contains a compound, this word will also be found and shown. This method will often lead to information overload as almost always a high number of articles will be found.

(2) The search string has to correspond exactly to the word or be the last part of one of the words in the search field. This method will also lead to information overload as a high number of articles will almost always be found, though not as many as when applying method (1).

(3) The search string has to correspond exactly to the word or be the first part of one of the words in the search field. This method will often lead to information overload as a high number of articles will almost always be found, though not as many as when applying method (1).

(4) The search string has to correspond exactly to one of the words in the search field. This method entails much fewer selections. In addition, it has the advantage that when searching the lemma field, it also shows articles containing multiword lemmas.

(5) The search string has to correspond exactly to the word(s) in the search field. This method has the advantage that typically there will only be one dictionary article corresponding to this string, but it also has the disadvantage that for example it does not include articles with multiword lemmas. Neither is this method suitable for searching in other fields containing more than one word, e.g. the collocation and example fields. However, this is the method that we apply in many of our dictionaries because it normally only provides one search result, i.e. there is no risk of information overload.

(6) In the case of a fuzzy search, a word corresponding to the search criterion will not be found, e.g. due to a spelling mistake. Instead, the programme suggests a number of lemmas (the maximum number is 10), which are orthographically close to the applied search string.

It should be added that for all six methods, it is the case that a search in the lemma field is not only a search in the lemma name, but also in every inflexion provided in the inflexion field for the given lemma.

If the programme is supposed to search in more than one field, it is in the programme not enough only to determine in what sequence the search string is to be used, but also in what way. Here, there are two basic options: a minimizing search and a maximizing search. In the case of a minimizing search, the programme stops if the search in a field type has given one or more selections. In the case of a maximizing search, the programme searches all the field types that the lexicographer has decided should be searched. The minimizing method has the advantage that often it only results in one dictionary article, thus avoiding information overload. The maximizing search instead has the advantage that many dictionary articles with much data are found. This could lead to information overload, but also to exactly the kind of information that the user wants to get and which would not be found if applying a minimizing search.

When the programme finds a dictionary article that matches the search string exactly, the whole article will appear on the screen. Most e-dictionaries also show this when a small number of articles, for example 10, are found. But when there are many articles, the lexicographer may choose only to show the lemmas that have been found, or maybe some more data belonging to these lemmas, but less than the whole articles. This is what we have chosen to do in for example THE DANISH MEANING DICTIONARY. In this dictionary, when there are more than 10 dictionary articles, the lemma and the first line of the definition are provided to make it easier for the user to choose which one of the articles may contain the data he is looking for.

Finally, an explanation text may be provided if the dictionary presents an article that does not correspond to his search string. This is what we have chosen to do in for example THE DANISH SYNONYM DICTIONARY, in which case we use a minimizing search to show a dictionary article for which the search string does not correspond to a word in the lemma field, but instead to a string in the synonym field that belongs to a different lemma. Therefore, this search does not result in a dictionary article with the search string as the lemma, but it results in another relevant article. This will then be explained to the user in an explication text. The presentation of data is formed by the choice of which fields to show in a certain dictionary and in what order to show them. This is similar for both printed dictionaries and e-dictionaries.

We will only discuss the structuring of three of the six mentioned dictionaries. The first one is called THE DANISH SYNONYM DICTIONARY, which is a communicative dictionary, primarily with the subfunction text production, secondarily with the subfunction text reception. The target group of this dictionary are users with Danish as their mother tongue. The first column of the following table shows the sequence of searches in the different fields. It involves a minimizing search, which means that the program stops if one or more articles have been found after search number 1. Searches 2 and 3 are then incorporated in case search 1 did not provide any selections because

the search string was not a lemma or one of the inflexion forms of this lemma. In the latter case, the user will be shown another article where the search string is given as a synonym or antonym item. In the second column, you see the selection and sequence of the item types in the dictionary article(s) that have been found. It is for example our decision that the dictionary contains the word class, but not the inflexion. If a lemma has not been found, but instead a synonym or an antonym has been found based on the current search string, an explication text is inserted at the beginning of the dictionary article(s), se column 5. In column 4, there is a notation for the case that the program should find more than 10 dictionary articles. In this case, the user will get a lemma list instead of 1–10 full dictionary articles, and he can use each lemma as a link to the article.

The selection of the order of the listed fields could have been different. We could, for example, have added inflexions or left out the part of speech classifications. In the most recent edition of this dictionary, we have chosen to leave out inflexions because the user is not interested in information about the lemma, but in being provided with a number of synonyms from which he can choose the one(s) he needs. We have, however, chosen to show the part of speech in the case of homonymy. In an earlier edition, we also provided the user of this dictionary with a lexical remark, which is a field with elaborate lexical information of a more encyclopaedic nature. This piece of data was only provided in the case of polysemy or homonymy, as was the case with the definition, but as the lexical remark is often rather long and detailed, the most important piece of information was almost hidden away. Therefore, the lexical remark was removed from this dictionary. The sequence of data types could hardly be any different, though in the case of polysemy, we could have chosen to place the definition after the synonym group to which it belongs.

Search in fields + search sequence	Field	Sequence of fields in the dictionary article	Sequence when more than 10 cards are found	Explication text
1 == lemma or inflexion 4 fuzzy search	1. Lemma	1	1	
	2. Sublemma	5		
	3. Homonym number	2		
	4. Polysem number	4		
	5. Meaning	6 But only if there is polysemy or homonymy.		

Search in fields + search sequence	Field	Sequence of fields in the dictionary article	Sequence when more than 10 cards are found	Explication text
	6. Lexical remark			
	7. Lexical remark for text production			
	8. Grammar, word class	3		
	9. Grammar, recommended inflexion			
	10. Grammar, not recommended inflexion			
	11. Grammar/ spelling remark			
	12. First reference			
	13. Second reference(s)			
	14. Collocation(s)			
	15. Example(s)			
	16. Word formation(s)			
2	17. Synonym(s)	7		No article with *** was found, instead we show article(s) with *** as a synonym to other lemma(s).
3	18. Antonym(s)	8		No article with *** was found, instead we show article(s) with *** as an antonym to other lemma(s).
	19. Synonym remark	9		

Search in fields + search sequence	Field	Sequence of fields in the dictionary article	Sequence when more than 10 cards are found	Explication text
	20. Proverb(s)			
	21. Idiom(s)			
	22. Idiom meaning			
	23. Internet link			
	24. Memo field			

Using the search string **jomfru** (= virgin), you get the following dictionary article with 3 polysems in THE DANISH SYNONYM DICTIONARY. The third polysem has a so-called sublemma, which is a certain type of word combination that includes the lemma:

> **jomfru** substantiv
> **1. Betydning**
> kvinde eller mand, der endnu ikke har haft samleje for første gang
> **Synonymer**
> virgo
>
> **2. Betydning**
> ung, ugift pige
> **Synonymer**
> frøken
> mø
> skønjomfru
> teenager
> ung pige
> ungmø
>
> **3. kold jomfru**
> **Betydning**
> kok eller kokkepige, som er særligt uddannet til at tilberede og anrette smørrebrød
> **Synonymer**
> smørrebrødsjomfru
>
> **virgin** noun
> **1. Meaning**
> woman or man who has never had intercourse
> **Synonyms**
> virgo
>
> **2. Meaning**
> young, unmarried girl
> **Synonyms**

miss
maiden
maiden
teenager
young girl
maiden

3. (literal translation) **cold virgin**
Meaning
chef who has specialized in making and arranging open-faced sandwiches
Synonyms
~ woman who makes open-faced sandwiches

The second dictionary is called WHAT IS THE WORD I AM LOOKING FOR?, which is a communicative dictionary with the function text production for a user who either knows something about the meaning of the word he wants to use, but has forgotten the word, or does not know the word at all. The target group of this dictionary consists of speakers with Danish as their mother tongue. In this dictionary, a Boolean search with one or more meaning elements can be used. It is, of course, an e-dictionary, but it could also be printed. In this case, we would need a very long index with all the words derived from their meaning. A maximizing search is used. This means that each step from 1 to 5 is used with the result that the user often gets a lot of dictionary articles, which are then shown in a list (see below).

Our structuring considerations are similar to the continuous microstructure discussion, see Wiegand/Smit (2013): Which fields are shown, how and in which sequence. The difference is that our main object of the process of microstructuring is related to certain functions. Here, we have made the choice not to show idioms and proverbs because they can better be explained for the function text production in a separate dictionary. We have also decided only to show the recommended inflexion, not other inflexion variants. Such possible, but not recommended variants are only shown in THE DANISH GRAMMAR AND SPELLING DICTIONARY and in THE DANISH INTERNET DICTIONARY. The sequence of data from the different data fields is shown in column 3. Here, we have already made several changes, and we are currently discussing whether synonyms and antonyms should be placed directly after the meaning item instead of after the word combination items. We use a minimizing search following the principle *XXX*, i.e. the search string has to correspond to the words or a part of the words in the search field.

Search in fields + search sequence	Field	Sequence of fields in dictionary article	Sequence when more than 10 cards are found	Explication text
	1. Lemma	1	1	
	2. Sublemma	7		
	3. Homonym number	2		
	4. Polysem number	6		
1 *XXX* 6 fuzzy search	5. Meaning	9	2 Only the first line.	
	6. Lexical remark			
2	7. Lexical remark for text production			
	8. Grammar, word class	3		
	9. Grammar, recommended inflexion	4		
	10. Grammar, not recommended inflexion			
	11. Grammar/ spelling remark	5		
	12. First reference	8 If a word is not recommended, the items from steps 10–15 are not shown.		This variant is not recommended, instead use → ***.
	13. Second reference(s)			
	14. Collocation(s)	10		
	15. Example(s)	11		
4	16. Word formation(s)	12		
3	17. Synonym(s)	13		

Search in fields + search sequence	Field	Sequence of fields in dictionary article	Sequence when more than 10 cards are found	Explication text
5	18. Antonym(s)	14		
	19. Synonym remark	15		
	20. Proverb(s)			
	21. Idiom(s)			
	22. Idiom meaning			
	23. Internet link			
	24. Memo field			

If we make a search with the following search string: seksuel - mand + kvinde (= sexual - man + woman) we get a list with 33 lemmas, each provided with its definition. Below, we only show a few of them. The formulation "Se artikel" (= see article) functions as a link, which the dictionary user can click on, and which leads him to the article for the relevant lemma.

allemandspige (= **whore** See article)
Betydning (= Meaning)
kvinde, som ernærer sig ved at have seksuelle forhold til mange mænd (= woman who makes a living by having sexual relationships with many men)

escortpige (= **escort** See article)
Betydning (= Meaning)
kvinde, som tilbyder seksuelle ydelser til gengæld for betaling (= woman who offers sexual services for money)

fingersex Se artikel (= **finger sex** See article)
Betydning (= Meaning)
seksuel stimulering, hvor partneren stikker en eller flere fingre op i kvindens skede (= sexual stimulation where the partner sticks one or more fingers up the woman's vagina)

frigid (= **frigid** See article)
Betydning (= Meaning)
som ikke reagerer på seksuelle tilnærmelser eller stimuli; især om kvinder (= who does not respond to sexual approaches or stimuli; especially about women)

frigiditet (= **frigidity** See article)
Betydning (= Meaning)
tilstand, hvor man ikke reagerer på seksuelle tilnærmelser eller stimuli; især om kvinder (= state characterized by a lack of response to sexual approaches and stimuli; especially about women)
...

klinikdame Se artikel (= **brothel receptionist** See article)
Betydning (= Meaning)
kvinde, som varetager forskellige opgaver af rutinemæssig administrativ karakter i en mindre, privat virksomhed, der tilbyder seksuelle ydelser mod betaling; om medarbejder

i et bordel (= woman who is in charge of different administrative tasks in a small private business that offers sexual services for money; about an employee in a brothel)

...

Using the link "Se artikel", you can read the whole dictionary article, which is structured as if it occurred in a text production dictionary for the word **klinikdame** (1 = medical receptionist; 2 = brothel receptionist):

klinikdame substantiv <en; klinikdamen, klinikdamer, klinikdamerne>

1. Betydning
kvinde, som er ansat på et mindre behandlingssted, fx en lægepraksis eller tandlæge, og som varetager opgaver af rutinemæssig administrativ eller teknisk art

Kollokationer
arbejde som hans sekretær og klinikdame
blive ansat som klinikdame hos tandlæge
klinikdame og lægesekretær på forskellige speciallægeklinikker
klinikdamen, der hjælper lægen
være klinikdame hos en læge

Eksempler
– Folkene oppe ved tandlægen var vældig søde og hjælpsomme – især den omsorgsfulde klinikdame, der gav mig en beroligende pille, fordi jeg så nervøs ud.

Synonymer
klinikassistent
lægesekretær

2. Betydning
kvinde, som varetager forskellige opgaver af rutinemæssig administrativ karakter i en mindre, privat virksomhed, der tilbyder seksuelle ydelser mod betaling; om medarbejder i et bordel

Kollokationer
ansætte en udsmider og en klinikdame, der tager sig af bestillinger og betalinger
en klinikdame, der tager telefonen
trafficking, narkoprostitution, hobbydamer og klinikdamer

Eksempler
– Dog er der vel en årsag til, at det er nødvendigt også for escortpiger og klinikdamer at have en chauffør/dørmand? Hvorfor betale en hyre til sådan en fyr, hvis der ikke var noget arbejde at gøre?

medical receptionist noun <the medical receptionist, medical receptionists, the medical receptionists>

1. Meaning
woman who is employed at a small place for treatment, e.g. a medical practice or a dental clinic, and who is in charge of administrative and technical tasks

Collocations
work as his secretary and medical receptionist
be hired as a medical receptionist in a dental clinic
medical receptionist and medical secretary in different specialized medical clinics
the medical receptionist, who helps the doctor
work as a medical receptionist in a medical clinic

Examples
– The people working at the dental clinic were very nice and helpful – especially the caring medical receptionist, who gave me a tranquillizer because I looked nervous.
Synonyms
clinical assistant
medical secretary

2. Meaning
woman who is in charge of different administrative tasks in a small private business that offers sexual services for money; about an employee in a brothel
Collocations
hire a bouncer and a brothel receptionist to take care of orders and payments
a brothel receptionist, who answers the phone
trafficking, drug prostitution, massage parlor prostitutes and brothel receptionists
Examples
But there must be a reason behind the necessity for escorts and brothel receptionists to have a chauffeur/bouncer? Why pay wages to him if there is nothing for him to do?

The structuring of another dictionary based on the same database is made for The Danish Grammar and Spelling Dictionary. It is primarily a communicative dictionary for text production, secondarily a cognitive dictionary for grammar learning. The dictionary is intended for users with Danish as their mother tongue. This dictionary uses a minimizing search following the method ==XXX, i.e. the search string has to be identical to the lemma or one of the inflexions provided. The choice of fields to be shown in the dictionary and the order to present these in are based on the assumption that the users speak Danish as their mother tongue. When these users have a problem related to text production, they will know and understand the word for which they are uncertain about the spelling or the grammatical application. Therefore, in most cases, it is not necessary to include definitions. In the 2001 version of The Danish Spelling Dictionary, this principle is applied consistently, also when there are different spellings or grammatical forms of each definition. The approach is the complete opposite of the one applied for Duden's spelling dictionaries, which are practically speaking general language dictionaries with all types of data: grammar, meaning, collocations, synonyms, word formations etc. (a criticism of this approach can be found in Bergenholtz 1991). Instead of this approach, we have decided to provide definitions only when there is homonymy, which in our definition of the term occurs when the same orthographical word belongs to different parts of speech or when different inflexions of the word within the same part of speech are connected to different meanings. The chosen order of presenting the data follows tradition, i.e. the lemma is followed by the word class, which is followed by the inflexion. A possible remark concerning the inflexion or orthography of the word occurs after and not before the inflexions.

Search in fields + search sequence	Field	Sequence of fields in the dictionary article	Sequence when more than 10 cards are found	Explication text
1 == lemma or inflexion 2 fuzzy search	1. Lemma	1	1	
	2. Sublemma	7 Only if there is a grammatical remark to a card with a sublemma.		
	3. Homonym number	2		
	4. Polysem number	8 Only if there is homonymy or a grammatical remark.		
	5. Meaning	9 Only if there is homonymy or a grammatical remark to a polysem.		
	6. Lexical remark			
	7. Lexical remark for text production			
	8. Grammar, word class	3		
	9. Grammar, recommended inflexion	4		
	10. Grammar, not recommended inflexion	5		The use of this inflexion is not recommended.
	11. Grammar/ spelling remark	6		

Search in fields + search sequence	Field	Sequence of fields in the dictionary article	Sequence when more than 10 cards are found	Explication text
	12. First reference	10 An arrow is placed in front of the word to which the article refers, and this word functions as a link to the relevant article.		This variant is not recommended, instead use → ***.
	13. Second reference(s)			
	14. Collocation(s)			
	15. Example(s)			
	16. Word formation(s)			
	17. Synonym(s)			
	18. Antonym(s)			
	19. Synonym remark			
	20. Proverb(s)			
	21. Idiom(s)			
	22. Idiom meaning			
	23. Internet link			
	24. Memo field			

In Danish, a word often has inflexional variants of which the dictionary only recommends one, which is presented as the first one; the non-recommended variant(s) is/are mentioned afterwards with an accompanying production recommendation to the dictionary users. The following is an example with the word **plaster** (= band-aid):

> **plaster** *substantiv* <et; plasteret, plastre, plastrene>
> også, men anbefales ikke <plastret>
>
> **band-aid** *noun* <the band-aid, band-aids, the band-aids>)
> also, but not recommended <the band-aid[2]>

---

2 In Danish, for the inflexion of the singular, definite form of the noun, we have two variants: plasteret and plastret. Both variants can be used, but only one is recommended.

This principle is also applied to orthographical variants. A remark, which is placed after the inflexions, explains why one variant is recommended or not recommended:

> **AIDS** *forkortelse* <en; AIDS'en, ikke flertal>
> Denne skrivemåde anbefales ikke, brug i stedet → aids
> **Grammatisk anmærkning**
> Dansk Sprognævn godtager ikke denne skrivemåde, men kun aids med små bogstaver. Der henvises der til brev fra Landsforeningen for bøsser og lesbiske, der mener, at skrivemåden med store bogstaver kan virke unødigt dramatisk på de mennesker, der har aids. Mange fagfolk og også de fleste aviser bruger dog kun den ikke-tilladte skrivemåde.

> **AIDS** *abbreviation* <the AIDS, no plural form>
> This spelling is not recommended, instead, use → aids
> **Grammatical remark**
> The Danish Language Council does not accept this spelling, only aids written with small letters. A reference is made to The Danish National Organisation for Gay Men and Women, who think that the spelling with capital letters may have an unnecessary dramatic effect on people who have aids. However, many experts and also most newspapers use the non-recommended spelling.

As already mentioned, in case of homonymy, definitions are also included in the article:

> **lam** I *substantiv* <et; lammet, lam, lammene>
> endnu ikke udvokset afkom af får
> **lam II** *adjektiv* <lamt, lamme>
> som ikke kan føle og bevæge visse dele af kroppen

> **lamb** I *noun* <the lamb, lambs, the lambs>
> a sheep's offspring which is not yet full-grown
> **paralysed II** *adjective* <paralysed, paralysed>
> who cannot feel nor move certain parts of the body

3.4 User-defined searching and dictionary article presentation

The user could also be given the same search possibilities as the lexicographer. By controlling if all cards are arranged in the same systematic way, the lexicographer makes e.g. the following searches:

– Find all verbs belonging to more than one paradigm (in our database, we found 154 cards)
– Find all nouns with the following last three letters: *cer (in our database, we found 53 cards)
– Find all adjective cards which the meaning 'colour' (in our database, we found 253 cards)
– Find all interjections (in our database, we found 294 cards)

There are thousands of possibilities for lexicographers in the use of their user interface, indeed many millions as Fuertes-Olivera/Tarp (2014, 95f) show. The user would then not only search for certain types of data in one or more fields, but could also choose to present some of the data types in a certain sequence using a table like the one we used in chapter 3.2 with predefined searches and presentations for single dictionaries. It will give the user millions of different types of presentations. In contrast to Fuertes-Olivera/Tarp (2014, 95f), we would not call each possibility a dictionary. Only the presentations selected by the lexicographer should be called dictionaries. But it is still lexicographical structuring as it is structured by the single user for his personal need. The personal structuring will – similar to the known use of advanced searches in other IT sources – perhaps only be used by about 5% of all dictionary users. But those users are normally the most active users, and they will not confine themselves to only one personal selection of searches and presentations, but make a new search with a new kind of presentation if they are not happy with the given dictionary article or if they get another personal need different from the predefined possibilities.

4 Concluding remarks

When the lexicographer produces a polyfunctional dictionary, he will have to explain the structure in a preface or in a user instruction, both in the case of a printed dictionary and an e-dictionary. This is beneficial when you have articles with many abbreviations, a complicated structure, or if the lexicographer has to present much data within little space. On this basis, some lexicographers conclude that only the dictionary user who has read the user instruction thoroughly will be able to use the dictionary with great or full benefit instead of searching haphazardly:

> The knowledge of non-alphabetical inner access structures which is particularly related to partially and completely condensed dictionary articles has to be gained by means of an intensive reading of the guidelines for use. Only after doing so, the user is able to execute internal access procedures without looking for something in longer dictionary articles haphazardly. In the guidelines for use, non-alphabetical inner access structures are not explained to the user explicitly, but implicitly by means of the introductions to the so-called "composition of article". By reading these explanations, the user can only grasp (different) structure excerpts … (Wiegand/Beer 2013, 140)

It is true that familiarity with a dictionary's structure makes it easier to use the dictionary and that this knowledge about the structure makes is easier for the user to find the data he is searching for. But demanding this of the dictionary user is unrealistic. In a two-year period, only few users had read the user instructions, both a short and a more detailed user instruction, for THE DANISH DICTIONARY OF FIXED EXPRESSIONS and THE DANISH MUSIC DICTIONARY. Bergenholtz/Johnsen (2013, 563) provide data

from the logfiles of these two dictionaries. Over a long period, there were 3,038,932 searches in total in The Danish Dictionary of Fixed Expressions, but only a fraction of these (0.19%) were made in one of the user instructions:

short user instructions	3,089
detailed user instructions	2,717
	5,806

In The Danish Music Dictionary, there were 160,157 searches in total, and of these the number of searches in the user instructions corresponded to 3.18% of all searches:

short user instructions	1,822
detailed user instructions	3,267
	5,089

This data is very discouraging for those lexicographers who emphasize the importance of the users' familiarity with and understanding of the structure of the specific dictionary. The users' lack of knowledge about and interest in user instructions can partially be explained by the fact that within a three-year period, which included 18 million searches, 80% of all users (or, to be more precise, each specific IP address) had only used the dictionary three times, two times or one time (Bergenholtz/Norddahl 2012, 209).

This data speaks in favour of developing and offering monofunctional dictionaries with only a few, but necessary pieces of data, which should be presented without any use of abbreviations and with terms used in everyday language to refer to the different types of data, e.g. "meaning" for definitions. In other words, a dictionary should be produced in a way that makes it easy and unproblematic for the user to use it without first having read the user instruction. However, this does not mean that the lexicographer should not write a user instruction for those dictionary users who actually read it.

This idea about monofunctional dictionaries with data presented in a simple format applies to most users, but it cannot be applied to a certain type of user: The user who will define his search methods and presentations according to his own personal wishes in future e-dictionaries. This type of user will be a well-informed dictionary user. This may be the type who uses the dictionary almost every day. But first and foremost, this will be the user who carries out a part of the lexicographical structuring himself.

5 Bibliography

Almind 2005a = Almind, Richard: First Impressions Last: Layout in Dictionaries. In: Schreiben, Verstehen, Übersetzen und Lernen: Zu ein- und zweisprachigen Wörterbüchern mit Deutsch, hrsg. von Irmhild Barz / Henning Bergenholtz / Jarmo Korhonen. Frankfurt a.M./Bern/New York/Paris: Peter Lang 2005, 93–102.

Almind 2005b = Almind, Richard: Designing internet dictionaries. In: Schreiben, Verstehen, Übersetzen und Lernen: Zu ein- und zweisprachigen Wörterbüchern mit Deutsch, hrsg. von Irmhild Barz / Henning Bergenholtz / Jarmo Korhonen. Frankfurt a.M./Bern/New York/Paris: Peter Lang 2005, 103–118.

Almind/Bergenholtz 2000 = Almind, Richard/Bergenholtz, Henning: Die ästhetische Dimension der Lexikographie. In: Ulla Fix/Hans Wellmann (Hrsg.): Bild im Text – Text und Bild. Heidelberg: Winter 2000, 259–288.

Bergenholtz 1991 = Bergenholtz, Henning: Rechtschreibewörterbuch, bleib bei deinen Leisten! In: Rechtschreibwörterbücher in der Diskussion, hrsg. von Augst, Gerhard und Schaeder, Burkhard. Frankfurt a.M./Bern/New York/Paris: Peter Lang, 403–417.

Bergenholtz 2012 = Bergenholtz, Henning: What is a Dictionary? In: Lexikos 22, 20–30.

Bergenholtz/Agerbo 2014 = Bergenholtz, Henning/Agerbo, Heidi: Extraction, selection and distribution of meaning elements for monolingual information tools. In: Lexicographica 30, 2014, 488–512.

Bergenholtz/Bergenholtz 2011 = Bergenholtz, Henning/Inger Bergenholtz: A dictionary is a tool, a good dictionary is a monofunctional tool. In: Pedro Fuertes-Olivera/Henning Bergenholtz (eds.): e-Lexicography: The Internet, Digital Initiatives and Lexicography. London & New York: Continuum 187–207.

Bergenholtz/Johnsen 2013 = Henning Bergenholtz/Mia Johnsen: User research in the field of electronic dictionaries: Methods, first results. In: Dictionaries. An International Encyclopedia of Lexicography: Supplementary volume: Recent Developments with Special Focus on Computational Lexicography. Ed. by Rufus H. Gouws, Ulrich Heid, Wolfgang Schweickard, Herbert Ernst Wiegand. Berlin: Mouton de Gruyter 2013, 556–568.

Bergenholtz/Norddahl 2012 = Henning Bergenholtz/Bjarni Norddahl: Ordbogsartikler, som ingen læser. In: LexicoNordica 19, 2012, 206–222.

Bergenholtz/Skovgård Nielsen 2013 = Bergenholtz, Henning / Skovgård Nielsen, Jesper: What is a lexicographical database? In: Lexikos 23, 2013, 77–87.

Bothma 2011 = Bothma, Theo J.D.: Filtering and adapting data and information in the online environment in response to user needs. In: Fuertes-Olivera, Pedro and Bergenholtz, Henning (Eds.). e-Lexicography: The Internet, Digital Initiatives and Lexicography. London & New York: Continuum 69–102.

Fuertes-Olivera/Tarp 2014 = Fuertes-Olivera, Pedro/Tarp, Sven: Theory and practice of specialised online dictionaries. Lexicography versus terminography. Berlin/Boston: de Gruyter.

Tarp 2002 = Tarp, Sven: Translation Dictionaries and Bilingual Dictionaries – Two Different Concepts. In: Journal of Translation Studies 7, 59–84.

Tarp 2008 = Tarp, Sven. Lexicography in the borderland between knowledge and non-knowledge. General lexicographical theory with particular focus on learner's lexicography. Tübingen: Niemeyer.

The Danish dictionary 2014 = Den Danske Ordbog = Den Danske Ordbog. København: Det Danske Sprog- og Litteraturselskab/Gyldendal. http://ordnet.dk/ddo (accessed August 2014)

The Danish Dictionary of Fixed Expressions 2007 = Henning Bergenholtz and Vibeke Vrang. Ordbogen over Faste Vendinger. Århus: Handelshøjskolen i Århus. Database and design: Richard Almind. http://www.idiomordbogen.dk. (accessed November 2009)

The Danish Grammar and Spelling Dictionary 2014 = Bergenholtz, Henning in cooperation with Agerbo, Heidi and Bock Michelsen, Andreas / Brosbøl Eriksen, Kathrine / Bodilsen, Andreas / Gudmann, Helene R. / Koed, Aleksander / Nørgaard Poulsen, Jon / Lybkær Kronborg Nielsen, Mia / Nguyen, Jane / Thers, Henrik. Database: Almind, Richard / Carlsen, Martin: Den Danske Grammatik- og Staveordbog. Odense: Ordbogen.com. http://www.ordbogen.com/ (accessed August 2014)

The Danish Internet dictionary 2014 = Bergenholtz, Henning in cooperation with Agerbo, Heidi and Bock Michelsen, Andreas / Brosbøl Eriksen, Kathrine / Bodilsen, Andreas / Gudmann, Helene R. / Koed, Aleksander / Nørgaard Poulsen, Jon / Lybkær Kronborg Nielsen, Mia / Nguyen, Jane / Thers, Henrik. Database: Almind, Richard / Carlsen, Martin: Den Danske Netordbog. Odense: Ordbogen.com. http://www.ordbogen.com/ (accessed August 2014)

The Danish Meaning Dictionary 2014 = Henning Bergenholtz in cooperation with Halkjær, Helene and Agerbo, Heidi / Bock Michelsen, Andreas / Brosbøl Eriksen, Kathrine / Bodilsen, Andreas / Gudmann, Helene R. / Koed, Aleksander / Nørgaard Poulsen, Jon / Lybkær Kronborg Nielsen, Mia / Nguyen, Jane / Thers, Henrik. Database: Almind, Richard / Carlsen, Martin : Den Danske Betydningsordbog. Odense: Ordbogen.com. http://www.ordbogen.com/ (accessed August 2014)

The Danish Music Dictionary 2009 = Inger Bergenholtz in cooperation with Richard Almind and Henning Bergenholtz: Musikordbogen. Århus: Handelshøjskolen i Århus. http://www.musikordbogen.dk. (accessed August 2009)

The Danish Spelling Dictionary 2001 = Dansk Sprognævn: Retskrivningsordbogen. 3. udgave. København: Alinea-Aschehoug.

The Danish Synonym Dictionary 2014 = Bergenholtz, Henning in cooperation with Odgaard Bodilsen, Filip and Agerbo, Heidi / Bock Michelsen, Andreas / Brosbøl Eriksen, Kathrine / Bodilsen, Andreas / Gudmann, Helene R. / Koed, Aleksander / Nørgaard Poulsen, Jon / Lybkær Kronborg Nielsen, Mia / Nguyen, Jane / Thers, Henrik. Database: Almind, Richard / Carlsen, Martin: Den Danske Synonymordbog. Odense: Ordbogen.com. http://www.ordbogen.com/ (accessed August 2014)

What is the word I am looking for? 2012 = Bergenholtz, Henning in cooperation with Busk Hedegaard, Stine and Bodilsen, Filip / Brosbøl Eriksen, Kathrine / Gudmann Helene R. / Halkjær Jensen, Helene / Halkjær Jensen, Jesper / Kjølhede, Christian / Kroman Liin, Anders / Nguyen, Jane / Agerbo, Heidi. Database: Almind, Richard / Theodor Styrk, Rasmus / Christensen, Peter: Hvad er det nu, det hedder? In: Verdens største ordbog, ed. by Norddahl, Bjarni and Bergenholtz, Henning. Odense: Ordbogen A/S, 79131–85956.

What is the word I am looking for? 2014 = Bergenholtz, Henning in cooperation with Busk Hedegaard, Stine and Agerbo, Heidi / Bock Michelsen, Andreas / Brosbøl Eriksen, Kathrine / Bodilsen, Andreas / Gudmann, Helene R. / Koed, Aleksander / Nørgaard Poulsen, Jon / Lybkær Kronborg Nielsen, Mia / Nguyen, Jane / Thers, Henrik. Database: Almind, Richard / Carlsen, Martin: Hvad er det nu, det hedder?. Odense: Ordbogen.com. http://www.ordbogen.com/ (accessed August 2014)

Wiegand/Beer 2013 = Wiegand, Herbert Ernst/Beer, Sandra: Access structures in printed dictionaries. In: Dictionaries. An International Encyclopedia of Lexicography: Supplementary volume: Recent Developments with Special Focus on Computational Lexicography. Ed. by Gouws, Rufus H. / Heid, Ulrich / Schweickard, Wolfgang / Wiegand, Herbert Ernst. Berlin: Mouton de Gruyter, 110–148.

Wiegand/Gouws 2013 = Wiegand, Herbert Ernst/Gouws, Rufus H.: Macrostructures in printed dictionaries. In: Dictionaries. An International Encyclopedia of Lexicography: Supplementary volume: Recent Developments with Special Focus on Computational Lexicography. Ed. by Gouws, Rufus H. / Heid, Ulrich / Schweickard, Wolfgang / Wiegand, Herbert Ernst. Berlin: Mouton de Gruyter 2013, 73–109.

Wiegand/Smit 2013 = Wiegand, Herbert Ernst/Smit, Maria: Microstructures in printed dictionaries. In: Dictionaries. An International Encyclopedia of Lexicography: Supplementary volume: Recent Developments with Special Focus on Computational Lexicography. Ed. by Gouws, Rufus H. / Heid, Ulrich / Schweickard, Wolfgang / Wiegand, Herbert Ernst. Berlin: Mouton de Gruyter 2013, 149–213.