Svensén (2009) Handbook of Lexicography

21 | Macrostructure

> L'ordre alphabétique [...], symbole même du dictionnaire pour l'immense majorité des lecteurs, ne peut devoir son succès qu'à son efficacité pratique, puisque son absurdité conceptuelle et linguistique est universellement reconnue. ALAIN REY

The term **MACROSTRUCTURE** is used to denote the relationships and order between the lemmas included in a certain lemma list. Thus defined, the macrostructure is an element of the **OUTER ACCESS STRUCTURE**, whose purpose is to guide the users to the lemma they seek. The topic is not so trivial as it might at first appear since the dictionaries under discussion here can differ quite considerably in this respect.

21.1 Two main types of macrostructure

A dictionary has either of two main types of macrostructure. One of them is based on the expression side, i.e. the spelling of the lemmas, which means that these are arranged according to an alphabetical principle (**ALPHABETICAL MACROSTRUCTURE**). The other one is based on the content side, i.e. the meaning of the lemmas, which means that these are arranged according to a systematic or thematic principle (**SYSTEMATIC MACROSTRUCTURE**).

In general-purpose dictionaries, alphabetical macrostructure is almost universally prevalent, and the present chapter will thus be primarily concerned with that type and its different variants. Systematic macrostructure occurs mainly in technical dictionaries but also in other types of specialized dictionary, for instance thesauri; a short survey of different variants of systematic macrostructure is given at the end of the chapter.

21.2 Alphabetization

The term **ALPHABETIZATION** denotes the sorting of lexical items according to a certain alphabetical system. Alphabetization is based on the set of characters (**GRAPHEMES**) forming the alphabet of the entry language of the dictionary. The relationships and order between the graphemes of the alphabet (the **CANONICAL**

ORDER OF LETTERS) represent a convention that varies between languages, although the differences between the western Latin alphabets are rather small. For instance, in Swedish the last three letters of the alphabet are *å, ä, ö*, while Danish and Norwegian have *æ, ø, å*. In German, *ü* is sorted as *u*, *ä* as *a* and *ö* as *o*. In Spanish, *ñ* is regarded as a letter in its own right, being sorted after *n*.[1]

However, a distinction must be made between the 'ordinary' alphabet of the entry language and the alphabetical principle forming the basis of the macrostructure of the dictionary (its **ACCESS ALPHABET**). The access alphabet contains a considerably greater number of units, and sorting is done according to strictly determined rules. One such important rule is concerned with the sorting of allographs, for instance upper-case and lower-case letters. Other rules determine the treatment of ligatures (*æ, œ*), letters with diacritics (*ç, é, ñ*), special letters such as Icelandic *ð* and German *ß*, numerals, certain punctuation marks (primarily hyphens), etc. In this context, it is also important to determine the order in which these factors are to be taken into account in the course of the sorting process.

The sorting of lemmas containing numerals (**NUMERICAL LEMMAS** and **ALPHANUMERICAL LEMMAS**) will often cause special problems. Usually, such lemmas are placed in the lemma list as if the numerals were written in letters (Bergenholtz & Tarp 1995: 191). Some L2 dictionaries have tried to solve the problem by providing a separate list of 'Numbers that are entries' or 'Numbers that are used as words'.

Another important matter that must be taken into consideration is whether alphabetization is to be done according to the **WORD-BY-WORD PRINCIPLE** or the **LETTER-BY-LETTER PRINCIPLE**. The difference between these two principles is apparent when the lemma list includes multi-word lexical items (multi-word lemmas). Word-by-word alphabetization takes into account the spaces (which are always ranked before letters, according to the 'nothing precedes something' principle), while the letter-by-letter method disregards the spaces, treating multi-word items as if they were written solid (Table 21.1, p. 370).

For Germanic languages, word-by-word alphabetization is usually recommended, not least because it will bring phrasal verbs together immediately after the simplex verb, as can be seen from the example. In a language such as English, however, there may be difficulties since the varying way of writing compounds (separately, hyphenated or solid) rather speaks in favour of letter-by-letter sorting, which has the advantage that the user who is unsure about the spelling will not have to search in several places in the dictionary.

When speaking of alphabetical ordering, one normally thinks of a procedure where the graphemes are considered in turn starting from the beginning of the

[1] Not very long ago, Spanish *ch* and *ll* were treated as independent graphemes as well; nowadays, they are sorted as *c+h* and *l+l*, respectively.

Table 21.1 *Two sorting principles*

Word by word	Letter by letter
pass	**pass**
pass away	**passage**
pass for	**pass away**
pass off	**passenger**
pass out	**pass for**
pass up	**passion**
passage	**passive**
passenger	**pass off**
passion	**pass out**
passive	**passport**
passport	**pass up**
password	**password**

Source: Gavare 1988: 76.

Table 21.2 *Reverse-order (final-alphabetical) macrostructure*

beak
leak
bleak
sneak
peak
speak
break
creak
freak
streak
wreak
teak
steak
squeak
weak
tweak

lemma, i.e. by proceeding from left to right. A macrostructure alphabetized according to this principle is called an **INITIAL-ALPHABETICAL MACROSTRUCTURE**. However, the graphemes may also be considered in turn starting from the end of the lemma, i.e. by proceeding from right to left, and the result of the alphabetization will then be a **REVERSE-ORDER** (or **FINAL-ALPHABETICAL**) **MACROSTRUCTURE** (Table 21.2).

It should be emphasized that reverse-order sorting is exclusively concerned with the order in which the individual graphemes are considered. It does not involve sorting from Z to A; the sort order is still 'forwards', as can be seen from the example.

21.3 Types of alphabetical macrostructure

There are several types of alphabetical macrostructure.

A distinction is made between macrostructures where all the lemmas are sorted according to the access alphabet and macrostructures where deviations from the access alphabet occur.

A further distinction is made between macrostructures where each entry constitutes an independent typographical unit (an **ENTRY BLOCK**) and macrostructures where an entry block may include several entries. In the latter case, the lemmas are said to have been subjected to **GROUPING**; the first lemma in such an entry block is called **ENTRANCE LEMMA**, and the other lemmas are called **SUBLEMMAS**. Grouping as such has nothing to do with alphabetical order but is exclusively concerned with the position of the lemmas on the dictionary page.

21.3.1 Strict-alphabetical macrostructure

A macrostructure where all the lemmas are sorted according to the access alphabet is called a **STRICT-ALPHABETICAL MACROSTRUCTURE**. There are two kinds of strict-alphabetical macrostructure: without grouping and with grouping.

21.3.1.1 Strict-alphabetical macrostructure without grouping

A strict-alphabetical macrostructure without grouping is called a **STRAIGHT-ALPHABETICAL MACROSTRUCTURE**. In such a macrostructure, all the lemmas are equal, and each lemma appears at the extreme left of the column:[2]

Schau *f* show, view
Schaubild *n* diagram
Schauder *m* shudder, shiver
schauderhaft *adj.* horrible, dreadful
schaudern *vb.* shudder, shiver
schauen *vb.* look, see

[2] In the examples shown in this chapter, indications given with a lemma have an identifying function only.

Schauer *m* **1** shower (of rain) **2** shudder, shiver
schauerlich *adj.* horrible, dreadful
schauern *vb.* shudder, shiver
Schaufel *f* shovel, scoop
schaufeln *vb.* shovel, scoop
Schaufenster *n* shop window
Schaukel *f* swing, see-saw
schaukeln *vb.* swing, see-saw, rock, sway
Schaukelpferd *n* rocking horse
schaulustig *adj.* curious
Schaum *m* foam, froth
Schaumbad *n* bubble (foam) bath
schäumen *vb.* foam, froth
Schauplatz *m* scene
schaurig *adj.* creepy; horrible
Schaurigkeit *f* creepiness; horribleness
Schauspiel *n* play, drama
Schauspieler *m* actor

Straight-alphabetical macrostructure is gaining more and more ground in general-purpose dictionaries, being easier to handle and normally not causing any search problems. However, in many monolingual dictionaries – particularly L2 dictionaries, where the aspect of vocabulary learning is important – different kinds of macrostructure with grouping have been preferred.

21.3.1.2 Strict-alphabetical macrostructure with grouping

A strict-alphabetical macrostructure where grouping may occur is called a **NICHING MACROSTRUCTURE**. In such a macrostructure, an entry block with several lemmas is called a **NICHE**, and the formation of niches is termed **NICHING**:

Schau *f* show, view
  Schaubild *n* diagram
Schauder *m* shudder, shiver
  schauderhaft *adj.* horrible, dreadful
  schaudern *vb.* shudder, shiver
schauen *vb.* look, see
Schauer *m* **1** shower (of rain) **2** shudder, shiver
  schauerlich *adj.* horrible, dreadful
  schauern *vb.* shudder, shiver
Schaufel *f* shovel, scoop
  schaufeln *vb.* shovel, scoop

Schaufenster *n* shop window
Schaukel *f* swing, see-saw
  schaukeln *vb.* swing, see-saw, rock, sway
  Schaukelpferd *n* rocking horse
schaulustig *adj.* curious
Schaum *m* foam, froth
  Schaumbad *n* bubble (foam) bath
  schäumen *vb.* foam, froth
Schauplatz *m* scene
schaurig *adj.* creepy; horrible
  Schaurigkeit *f* creepiness; horribleness
Schauspiel *n* play, drama
  Schauspieler *m* actor

Typographically, the grouping of a niche can be realized in two different ways. Either all the sublemmas are arranged vertically (**LISTING**, as in the above example), or they are run on to each other (**CLUSTERING**):

Schau *f* show, view **Schaubild** *n* diagram
Schauder *m* shudder, shiver **schauderhaft** *adj.* horrible, dreadful
  schaudern *vb.* shudder, shiver
schauen *vb.* look, see
Schauer *m* **1** shower (of rain) **2** shudder, shiver **schauerlich** *adj.* horrible, dreadful **schauern** *vb.* shudder, shiver
Schaufel *f* shovel, scoop **schaufeln** *vb.* shovel, scoop
Schaufenster *n* shop window
Schaukel *f* swing, see-saw **schaukeln** *vb.* swing, see-saw, rock, sway
  Schaukelpferd *n* rocking horse
schaulustig *adj.* curious
Schaum *m* foam, froth **Schaumbad** *n* bubble (foam) bath **schäumen** *vb.* foam, froth
Schauplatz *m* scene
schaurig *adj.* creepy; horrible **Schaurigkeit** *f* creepiness; horribleness
Schauspiel *m* play, drama **Schauspieler** *m* actor

Sublemmas are sometimes subjected to textual condensation:

Schauder *m* shudder, shiver ~**haft** *adj.* horrible, dreadful ~**n** *vb.* shudder, shiver

As a rule, grouped lemmas are morphologically related, i.e. formed from the same base word. However, morphology often comes into conflict with the alphabetical

order and has to give way, which is illustrated in the example by the sequence starting with **Schau**, which has to be interrupted before **Schauder**, starts again after **schaudern**, stops before **Schauer**, starts again after **schaufeln**, stops before **Schaukel**, starts again after **Schaukelpferd**, etc.

In some dictionaries, morphology is not taken into account at all, and the formation of niches is based on spelling only, mostly in combination with condensation of the sublemmas. This method is hardly to be recommended (Rettig 1985: 110):

gig|antesque [...] **-antisme** [...] **-olo** [...] **-ot** [...]

The main advantage of a niching macrostructure is that it saves space. On the other hand, the pedagogical advantage of grouping is substantially reduced in a niching macrostructure by the fact that the sequence of related words must often be interrupted. Another disadvantage from a pedagogical point of view is that the entrance lemma is not always identical with the base word from which the other words have been formed:

isol|able [...] **-ant** [...] **-ateur** [...] **-ation** [...] **-é** [...] **-ement** [...] **-ément** [...] **-er** [...]

Obviously, an arrangement such as this one will provide an unsatisfactory picture of the morphology of the word group concerned.

21.3.2 Non-strict-alphabetical macrostructure

A macrostructure where there is deviation from the strict-alphabetical order is called a **NON-STRICT-ALPHABETICAL MACROSTRUCTURE**.

Non-strict-alphabetical macrostructure occurs only in combination with grouping.[3] In such a macrostructure (**NESTING MACROSTRUCTURE**), an entry block with several lemmas is called a **NEST**, and the formation of nests is termed **NESTING**. It is only the sublemmas that may deviate from the strict-alphabetical order; the order of the entrance lemmas is still a strict-alphabetical one:

Schau *f* show, view **Schaubild** *n* diagram **schauen** *vb.* look, see **Schaufenster** *n* shop window **schaulustig** *adj.* curious **Schauplatz** *m* scene **Schauspiel** *m* play, drama **Schauspieler** *m* actor
Schauder *m* shudder, shiver **schauderhaft** *adj.* horrible, dreadful **schaudern** *vb.* shudder, shiver
Schauer *m* 1 shower (of rain) 2 shudder, shiver **schauerlich** *adj.* horrible, dreadful **schauern** *vb.* shudder, shiver **schaurig** *adj.* creepy; horrible **Schaurigkeit** *f* creepiness; horribleness

[3] Obviously, a non-strict-alphabetical macrostructure without grouping would be fairly chaotic.

Schaufel *f* shovel, scoop **schaufeln** *vb.* shovel, scoop
Schaukel *f* swing, see-saw **schaukeln** *vb.* swing, see-saw, rock, sway **Schaukelpferd** *n* rocking horse
Schaum *m* foam, froth **Schaumbad** *n* bubble (foam) bath **schäumen** *vb.* foam, froth

As in a niching macrostructure, the grouped lemmas are as a rule morphologically related. Also, the same conflict between morphology and alphabetical order will often occur, although in a nesting macrostructure it is the alphabet that has to give way. As can be seen from the example, all the lemmas related to **Schau** and **Schauer**, respectively, have now been brought together in their respective nests.[4]

In the example, the nest has been realized by way of clustering, but listing is equally conceivable.

As is the case with the niching macrostructure, the sublemmas are sometimes subjected to textual condensation:

Schau *f* show, view **~bild** *n* diagram **~̈en** *vb.* look, see **~fenster** *n* shop window **~̈lustig** *adj.* curious **~platz** *m* scene **~spiel** *n* play, drama **~spieler** *m* actor

Within the nest, sorting need not always be strict-alphabetical. In some dictionaries, for instance, derivatives have been separated from compounds, and sometimes further subdividing has been done. An example:

Natter *f*
Natur *f* • **Naturanlage** *f* – ... – **Naturwissenschaft** *f* **~ler** *m* **~̈lich** *adj.* • **Naturalien** *pl* **~kabinett** *n* **~sammlung** *f* – ... – **natürlich** *adj.* **~erweise** *adv.* **~̈keit** *f*
Nautik *f*

As usual, the entrance lemmas (**Natter, Natur, Nautik**) are sorted alphabetically. After the entrance lemma (**Natur**) the sublemmas come in: first the compounds having the entrance lemma as first element (**Naturanlage** ... **Naturwissenschaft**), then the derivatives of the entrance lemma (**Naturalien** ... **natürlich**). This means that there are two nests within the nest. However, within each one of these two nests, there are further nests containing derivatives of compounds (**Naturwissenschaftler, naturwissenschaftlich**), compounds with derivatives as

[4] It should be pointed out that a nesting macrostructure may very well include niches too. This will occur when all the members of a group of related words happen to form an alphabetical sequence where no other words 'intervene'. However, should the lemma list contain one single grouping that deviates from the strict-alphabetical order, this will suffice for the macrostructure to be considered as a nesting one.

first elements (**Naturalienkabinett, Naturaliensammlung**) and derivatives of derivatives (**natürlicherweise, Natürlichkeit**), an arrangement which could be described as SECOND-LEVEL NESTING (Gouws 2005b).

A system such as this one is very difficult for inexperienced dictionary users. Actually, it is even more difficult to handle than shown by the example since the compounds and derivatives here are easily recognizable and the spelling would allow all of them, if sorted in strict-alphabetical order, to fall between **Natter** and **Nautik**.

Things are considerably harder when the entrance lemmas of several nests are similar to each other. In the following example, the numbers show the strict-alphabetical sequence:

> **Näh|arbeit** *f* (1) **-faden** *m* (20) **-garn** *n* (21)
> **Nah|aufnahme** *f* (2) **-einstellung** *f* (6) **-erholungsgebiet** *n* (14) **-kampf** *m* (22)
> **nahe** *adj., prep.* (3) – **nahe|bringen** *vb.* (4) **-gehen** *vb.* (5) **-kommen** *vb.* (7) **-legen** *vb.* (8) **-liegen** *vb.* (9) **-stehen** *vb.* (18) **-zu** *adv.* (19)
> **nahen** *vb.* (10)
> **nähen** *vb.* (11) – **Näherei** *f* (13)
> **näher** *adj.* (12) – **näherkommen** *vb.* (15)
> **nähern** *vb.* (16) – **Näherung** *f* (17)

However, sublemmas need not be so near to each other alphabetically as has been shown hitherto (ALPHABET-RELATED GROUPING). It also occurs, for instance in certain French dictionaries, that derivatives by prefixation are included in the nest, and as a consequence some of the sublemmas may belong to quite other parts of the alphabet (ALPHABET-INDEPENDENT GROUPING):

> 1. FORMER […] • **formé, e** […] • **formage** […] • **formation** […] • **formateur, trice** […] • **formatif, ive** […] • **malformation** […] • **néo-formation** […] • **préformer** […] • **reformer** […] • **reformation** […]
> (*Lexis*)

Obviously, a dictionary having a macrostructure of this kind must have a very extensive cross-reference system (see further pp. 395f.).

The purpose of nesting is usually pedagogical: bringing together lemmas that are morphologically related will promote word comprehension and vocabulary learning. As a consequence, a nesting macrostructure is primarily suitable for monolingual learners' dictionaries, for instance L2 dictionaries. In bilingual dictionaries, on the other hand, where quick and easy access normally has higher priority than the pedagogical aspect, nesting is hardly to be recommended, particularly in L2→L1 dictionaries, since this will presuppose knowledge of the morphology of a foreign language.

21.4 Systematic macrostructure

Systematic macrostructure occurs primarily in technical dictionaries and certain other kinds of specialized dictionary, for instance thesauri. A typical characteristic of dictionaries having a systematic macrostructure is that, at a certain level, the systematic classification must be supplemented by an alphabetical access facility. This combination of systematic and alphabetical access can be realized in various ways.

As a rule, the concepts belonging to a certain subject field are assigned to different thematic groups or main areas. Within each area, the concepts are arranged in CONCEPTUAL SYSTEMS describing their interrelationships (superordinate, subordinate or co-ordinate). These conceptual systems should as far as possible be reflected in the macrostructure of the dictionary. Therefore, each lemma must be provided with a unique label showing the position of the corresponding concept within the conceptual system. This label functions as a GUIDING-ELEMENT CARRIER, i.e. it determines the position of the lemma in the macrostructure (see p. 000). A German example involving the conceptual system 'handsaws for wood' (adapted from Wüster 1985: 139):

> **Handsäge für Holz**
> **1. Handsäge ohne Vorspannung**
> **1.1. Heftsäge**
> **1.1.1. Heftsäge für Rundholz**
> **1.1.2. Heftsäge für Schnittholz**
> **1.2. Freispannsäge**
> **1.2.1. Zweimann-Blattsäge**
> **1.2.2. Zugkettensäge**
> **2. Handvorspannsäge**
> **2.1. Bogensäge**
> **2.1.1. Baumsäge**
> **2.1.2. Laubsäge**
> **2.2. Strecksäge**
> **2.2.1. Strecksäge für Rundholz**
> **2.2.2. Strecksäge für Kantholz**
> **2.2.3. Strecksäge für Bretter**

However, it may not always be meaningful or possible to let the conceptual relationships at all levels be reflected in the macrostructure. In such a case, one has to restrict oneself to establishing a hierarchy of CONCEPTUAL FIELDS, providing each one of these with a heading under which the lemmas belonging to it are brought together and sorted alphabetically. We will now imagine a situation where

the German handsaws cannot be accounted for in a conceptual system; instead, alphabetical sorting has been chosen at this level:

02.13 Handsäge für Holz
Baumsäge
Bogensäge
Freispannsäge
Handsäge ohne Vorspannung
Handvorspannsäge
Heftsäge
Heftsäge für Rundholz
Heftsäge für Schnittholz
Laubsäge
Strecksäge
Strecksäge für Bretter
Strecksäge für Kantholz
Strecksäge für Rundholz
Zugkettensäge
Zweimann-Blattsäge

The heading of the conceptual field has been provided with a designation ('02.13') showing the position of the field 'Handsäge für Holz' within the hierarchy of fields and functioning as a guiding-element carrier for the field. In an alphabetical index involving all the terms of all the conceptual fields dealt with in the dictionary, the terms belonging to this particular field are marked with the designation '02.13', thus directing the user to the right place in the lemma list.

Literature

Alphabetization: Gavare 1988.

Monolingual dictionaries: Hausmann & Wiegand 1989: 336–7; Wiegand 1989.

Bilingual dictionaries: Hausmann & Werner 1991: 2746–8 (general); Martin & Gouws 2000 (closely related languages).

Multilingual dictionaries: Haensch 1991a: 2922–4.

Technical dictionaries: Bergenholtz & Tarp 1995: 190–9.

Learners' dictionaries: Rundell 1998: 327f., Rundell 1999: 48f., Cowie 1999: 145–50, Heuberger 2000: 115–22 (English dictionaries); Bornäs 1986 (French dictionaries).

Homonymization, grouping, etc.: Zöfgen 1994: 84–104 (learners' dictionaries); Gouws 2005b (special types of niching).

22 | Megastructure

> While a dictionary without headwords is not a dictionary, one without outside matter does not lose its status as a dictionary. MARGARET COP

The **MEGASTRUCTURE** of a dictionary is the relationships and order between its main components. Except for the lemma list, these have only sporadically been touched upon earlier in the book; therefore, this chapter will also provide a survey of the most important types of components and their functions.

22.1 Main components of the dictionary

The central component of a dictionary is the **LEMMA LIST**. It could even be asserted, with some simplification, that the lemma list is the only obligatory component of the dictionary: without a lemma list, no dictionary. In a sense, it is misleading to speak of 'the lemma list', i.e. with the definite article and in the singular, since there is nothing to prevent a dictionary from having several lemma lists, for instance a **CENTRAL LEMMA LIST** and one or more **SUBSIDIARY LEMMA LISTS**. Here, however, the account will be restricted to what is usual in the types of dictionary dealt with in this book, assuming that we are concerned with a dictionary that has just one lemma list.

Around the lemma list, a number of other components of different kinds are grouped. Those preceding the lemma list are called, with a generic term, **FRONT MATTER** and those following it **BACK MATTER**. In addition, there is a third component, **MIDDLE MATTER**, consisting of items inserted in the lemma list without actually being part of it, for instance certain types of illustrations and survey entries. A generic term for front, back and middle matter is **OUTSIDE MATTER**.

This classification of different kinds of outside matter is based on the purely physical position of the components in the dictionary. Viewed in another, and more fruitful, perspective, they may also be classified according to their function: whether they (a) provide information about the object language(s), (b) have a metafunction, i.e. provide information about the dictionary itself, (c) are elements of the access structure of the dictionary, or (d) have some other function.

The outside-matter components of a dictionary often pass unnoticed by the users. In order to improve accessibility, front and back matter should, in one way