

1 | Lexicography

I know of no more enjoyable intellectual activity than working on a dictionary. Unlike most research, lexicography rarely sends one in fruitless quests; one does not devote days, months, or even years to testing an hypothesis only to decide that it is not tenable, or to attempting to collect evidence to prove a theory only to have to conclude that sufficient facts are no longer in existence to clinch it. It does not make one's life anxious, nor build up hopes only to have them collapse. Every day one is confronted by new problems, usually small but absorbingly interesting; at the end of the day one feels healthily tired, but content in the thought that one has accomplished something and advanced the whole work towards its completion. JAMES R. HULBERT

1.1 Lexicography and society

Dictionaries are a cultural phenomenon. It is a commonplace to say that a dictionary is a product of the culture in which it has come into being; it is less so to say that it plays an important part in the development of that culture.

Different dictionaries have different purposes. They are produced in order to meet either the individual's needs for information (**UTILITY LEXICOGRAPHY**) or the needs of a community – national, political, scientific, etc. – to preserve information for the future (**DOCUMENTARY LEXICOGRAPHY**).

Utility lexicography serves two main purposes. One of them is to support communication (in the widest sense), either in the user's native language or in a foreign language. The other purpose is to support the learning of language, either one's native language or a foreign language.

There are, and there have always been, several driving forces behind the emergence of lexicographic activity. Among the areas where such forces have been active and where, consequently, dictionaries have played an important part, the following may be mentioned (Hausmann 1989c).

In religion, special lexicographic aids have been required for the study of holy scriptures and for contacts with foreign cultures and languages made necessary by missionary work and other activities for the dissemination of religious faith.

In literature, the reading of ancient or foreign-language works has called for a particular kind of lexicographic product, while individuals' own literary activities

have required others. Education, where there is a great need for lexicographic aids in order to improve native-language proficiency and facilitate the learning of foreign languages, has been a major driving force behind the emergence of the types of dictionary that are most common nowadays: the monolingual and bilingual general-purpose dictionaries. In earlier times, there was also a need for dictionaries when reading or writing in the international language of the learned – a function which, for instance, Latin once fulfilled in Europe.

The role of lexicography in politics is multifaceted. Governments have always needed linguistic and, hence, lexicographic support for their diplomatic activities, and this need has been particularly obvious in countries cherishing colonial aspirations or having already acquired full-fledged colonial empires. The activities of present-day international organizations, such as the United Nations and the European Union, make similar demands for linguistic aids. Governments may also find it necessary to influence their citizens ideologically, and in such cases the linguistic aspect and, consequently, dictionaries will be of great importance. Dictionaries may also be of interest to political actors other than governments, for instance those engaged in promoting nationalist or regionalist efforts.

In today's information and communication society, developments in science and technology create a greater demand for lexicographic efforts than before, not least in connection with the study, development and use of technical language. Trade and industry need linguistic aids when communicating with foreign business connections and immigrated labour, and increasing tourism has brought about a demand for dictionaries as communicative support when travelling abroad.

1.2 Lexicography and related disciplines

1.2.1 Lexicography, lexicology and linguistics

Taking the description in the preceding section as a basis, it would seem natural to define lexicography as 'the compilation of dictionaries'. However, such a definition is too narrow; it is perfectly possible to work on lexicography without actually being engaged in a dictionary project (for instance, one can write a book such as this one). It would therefore be more correct to define lexicography in the following way:

- (a) **LEXICOGRAPHY** is an activity which consists in observing, collecting, selecting, analysing and describing, in a dictionary, a number of lexical items (words, word elements and word combinations) belonging to one or more languages. In cases where two or more languages are involved simultaneously, the description takes on the nature of a comparison between the items that

have been selected from the vocabularies of the languages in question. This part of the subject, the compilation of dictionaries, is called **PRACTICAL LEXICOGRAPHY**, or simply **DICTIONARY-MAKING**.

- (b) Lexicography also includes the examination and development of theories concerning the compilation, characteristics, purposes and use of dictionaries. This part of the subject is generally called **THEORETICAL LEXICOGRAPHY** or **METALEXICOGRAPHY** (= 'lexicography which deals with lexicography'). A handbook of lexicography is a typical metalexigraphic product.

Another concept which often occurs in connection with lexicography is **LEXICOLOGY**. Opinions differ as to the relationship between these two. There are those who argue that, on the whole, lexicography and lexicology are the same thing, while others regard lexicography as a branch of lexicology; there is also a third opinion maintaining that lexicology is equivalent to metalexigraphic work. In this book, lexicology is regarded as a science concerned with the study of vocabulary, its structure and other characteristics. This refers first of all to the study of the meanings of words and the relationships between meanings (semantics), but also to the study of the formation and structure of individual words, i.e. morphology. Thus defined, lexicology is not the same as lexicography or metalexigraphic work, nor does the term represent a wider concept of which lexicography constitutes only a part. However, lexicology obviously forms an important basis for lexicographic work.

Similarly, there are differing opinions as to the relationship between lexicography and **LINGUISTICS**. There are those who argue that lexicography is a kind of applied linguistics, while others (among them the author of this book) regard it as an independent discipline. However, there is almost general agreement that linguistics forms an important basis for lexicographic work, and, conversely, that lexicography and its products have made considerable contributions to linguistic research.

1.2.2 General-language lexicography, technical lexicography and terminography

It is usual to make a distinction between **GENERAL-LANGUAGE LEXICOGRAPHY**, which is primarily concerned with general vocabulary, and **TECHNICAL LEXICOGRAPHY**, the object of which is the terminology of various specialist fields. General-language dictionaries usually include a considerable number of technical terms, particularly those encountered by everyone in everyday life. The converse is not the case, however: technical dictionaries normally do not include words belonging to general language only.

Work on technical vocabulary is also done within a discipline usually named **TERMINOLOGY** or (by analogy with 'lexicography') **TERMINOGRAPHY**. The object and

purpose of this work is in many respects the same as that of technical lexicography, but theories and methods are partly different.

The relationship between technical lexicography and terminography is a controversial issue. Some people (particularly terminographers) argue that the differences between the two are so fundamental that they should be regarded as different disciplines, while others (particularly technical lexicographers) emphasize the similarities, regarding terminography and technical lexicography as being largely the same thing. The disagreement is about, for instance, the semiotic foundations of the work, the use of subject classification, the normative aspect and the presentation format. The matter is not of very great importance to the present book, which deals primarily with general-language lexicography and only sporadically touches upon problems concerning technical lexicography.

1.2.3 Lexicography and language technology

Lexicographic methods and products are also important within a discipline named **LANGUAGE TECHNOLOGY**. The purpose of language technology is to make it possible for humans and machines to interact by means of natural language, and to develop digital tools for text processing in a wide sense, primarily the production, editing and translation of both spoken and written text. In this context, certain types of electronic dictionaries (see Ch. 26) play an important part.

1.3 The terminology of lexicography

Lexicographic theories, methods and products developed considerably during the last decades of the twentieth century. However, the terminology of lexicography did not develop with equal speed and was initially insufficient and unsystematic. There were hardly any specialized dictionaries of lexicographic terminology available, and in linguistic reference works lexicographic terms were included only in so far as they were relevant from a general-linguistics point of view. This state of affairs caused problems of communication for both the theoretical and the practical side of the subject, and this applied not only to term formation and term selection but also to concept formation, conceptual relationships and conceptual systems.

The publication of the comprehensive three-volume encyclopedia *Wörterbücher/Dictionaries/Dictionnaires* (Hausmann et al. (eds.) 1989–91) represented a considerable advance, not least where concept formation and terminology were concerned, although the choice of languages was restricted to English, German and French. In 1991 the Nordic Association for Lexicography declared at its foundation that

one of its main tasks would be the development of a Scandinavian lexicographic terminology. A joint working group was established and six years later the Association was able to publish *Nordisk leksikografisk ordbok* (1997; 'Nordic Dictionary of Lexicography'), with definitions and explanations in Norwegian *Bokmål* and terms in all the major Scandinavian languages as well as in English, German and French. The following year saw the publication of *Dictionary of Lexicography* by Hartmann and James, with definitions and terms in English.

In devising a terminology suitable for this handbook, the three works mentioned above formed the basis; however, deviations and additions proved to be necessary in quite a number of cases.

1.4 A simple descriptive model for lexicography

1.4.1 The dictionary and its content

A description of the various types of information provided in a dictionary may well be based on the concept of the **LINGUISTIC SIGN**. This can be said to consist of an **EXPRESSION** (a certain form) and a **CONTENT** (a certain meaning). Furthermore, the sign has a **FUNCTION** (a certain way of behaving in combination with other linguistic signs).

Starting from the three components of the linguistic sign, we can thus group the different characteristics of words and word combinations, as they are described in the most common type of dictionary, in the following way:

1. formal characteristics (expression-side characteristics), i.e. their spelling, pronunciation and morphology (inflection and word formation);
2. semantic characteristics (content-side characteristics), i.e. their meaning;
3. syntagmatic characteristics (function), i.e. how they combine with other words and word combinations;
4. pragmatic characteristics, which, among other things, include factors involved in their use.

The following gives an overall picture of the basis for, and the work on, the most common type of dictionary and of its structure and content. It also provides a simple descriptive model for practical lexicography and simultaneously serves as a quick survey of the contents of this book.

1.4.2 Dictionary users and dictionary types

The form of a dictionary is largely determined by the types of linguistic activities it is meant to be used for, by the needs of the users in this context and by the capabilities of the users. Since these factors can vary greatly, there are, in response

to this, many different types of dictionaries. This range of topics is discussed in Chapter 2.

1.4.3 Data collection and data selection

The work of description cannot start until suitable materials to form the basis of the description have been collected and selected. A general survey of the collection and selection of data is given in Chapter 3.

1.4.4 The dictionary as text

A dictionary is not just any kind of text but has certain special characteristics. In order to describe the components and structures of a dictionary text and their interrelationships, there is a need for a uniform set of concepts and a suitable terminology. The most important tools for such a description are presented in Chapter 4.

1.4.5 The lemma

The **LEMMA** ('headword')¹ has a very special position in the dictionary, and its establishment, form and function is the topic of a special section (Chapter 5).

1.4.6 Information types

Several **INFORMATION TYPES** can occur in a single dictionary entry.² The division of lexicographic data into information types can be made in many ways, and the division made in this book is only one among several possibilities. A survey of the information types, their linguistic background and their treatment in dictionaries takes up a large part of the book (Chapters 6–19).

1.4.6.1 *Formal information*

As the term suggests, the different types of **FORMAL INFORMATION** are concerned with the form of the lemma.

One of these categories is the orthographic form, i.e. **SPELLING**, which may be regarded as also including **WORD DIVISION** (or **HYPHENATION**, or **SYLLABIFICATION**). These matters are discussed in Chapter 6.

¹ On the choice of the term 'lemma' instead of 'headword', see p. 93.

² Despite its ambiguity, the term 'entry', being firmly established in English metalexigraphic literature, will be used throughout this book instead of its competitor 'article'.

The second formal category is concerned with the phonetic form of the lemma, i.e. **PRONUNCIATION** (Chapter 7).

A third formal information type is **MORPHOLOGY**, comprising on the one hand **INFLECTION** and on the other **WORD FORMATION**. These matters are dealt with in Chapter 8.

1.4.6.2 *Syntagmatic information*

The **SYNTAGMATIC INFORMATION** provided in a dictionary is concerned with the behaviour of the lemma in combination with other words, both grammatically and lexically. Different information types specify this with different degrees of precision.

The least precise information simply states the **PART-OF-SPEECH MEMBERSHIP** of the lemma. This provides merely a rough description of how it functions grammatically along with other words (Chapter 9).

More precise grammatical information is given by the next category, which deals with the **CONSTRUCTIONS** (or **SYNTACTIC VALENCY**) of the lemma. Information on constructions is discussed in Chapter 10.

However, within a word combination involving a particular word, there may also be a need for semantic and lexical identification of the other components. For example, if one of the other components has three synonyms, perhaps only one or two of these generally occur with the word in question. Such regularly occurring word combinations are called **COLLOCATIONS**, and the specification of these provides even more precise information about the syntagmatic properties of the lemma (Chapter 11).

There are also word combinations that are so firmly established that it is hardly possible to alter the construction or replace any of the components without forfeiting the meaning. Such a combination (an **IDIOM**) should not actually be regarded as an instance of the syntagmatic properties of one of its components but as an independent lexical item. The treatment of idioms is discussed in Chapter 12.

1.4.6.3 *Semantic information*

The **SEMANTIC INFORMATION** is concerned with the meaning of the lemma and of the word combinations of which it may form a part.

In monolingual dictionaries, meaning is specified by **DEFINITIONS**. The semantic description also includes information about semantically related words, in particular synonyms but also antonyms, hyponyms and hyperonyms (**CONTENT-PARADIGMATIC INFORMATION**). Semantic information in monolingual dictionaries is dealt with in Chapter 13.

In bilingual dictionaries, the meaning of words and word combinations in the **ENTRY LANGUAGE** (the language of the lemmas) is specified by means of **EQUIVALENTS** in the other language. This type of information is the topic of Chapter 14.

1.4.6.4 *Examples*

Syntagmatic and semantic information, but also other kinds of information, is often provided by means of **EXAMPLES**. A survey of the form and function of examples is presented in Chapter 15.

1.4.6.5 *Encyclopedic and pragmatic information*

ENCYCLOPEDIAIC (or **FACTUAL**) **INFORMATION** is of two kinds, verbal and visual. Verbal encyclopedic information may sometimes be needed in order to clarify the information given about the language, but it can also be provided for its own sake, independently of the linguistic description (Chapter 16). Visual encyclopedic information, which has the same purpose as its verbal counterpart, is provided by **ILLUSTRATIONS** (Chapter 17).

PRAGMATIC INFORMATION may be provided by various means (examples, explanations, comments, etc.) and at various levels (word, phrase, clause, sentence, etc.). A more formalized way of conveying pragmatic information about lexical items is to specify their occurrence and use in different dimensions (temporal, geographic, stylistic, situational, etc.), i.e. with a generic term, their **MARKING**. Marking information is dealt with in Chapter 18.

1.4.6.6 *The diachronic perspective*

Certain monolingual dictionaries also specify the **ETYMOLOGY** of the words, i.e. their origin and cognates. This information type differs from any previously discussed through its historical viewpoint; it lies along the time axis and cuts across most of the other categories, incorporating elements from several of them (Chapter 19).

1.4.7 Dictionary structure

After the exposition of the information types, the next few chapters are devoted to **DICTIONARY STRUCTURE**. This topic will have been touched upon very briefly in Chapter 4, but now the description will be systematized and more exhaustive.

The information given within an individual dictionary entry must be arranged in a suitable way. The internal structure of the entries, called the **MICROSTRUCTURE** of the dictionary, is dealt with in Chapter 20.

Furthermore, the entries must themselves be arranged in a particular order in accordance with the purpose and intended use of the dictionary. This aspect, the **MACROSTRUCTURE** of the dictionary, is the topic of Chapter 21.

The main component of a dictionary is obviously the ordered set of entries (the **LEMMA LIST**; see p. 379). However, it is often the case that the dictionary includes a number of other, more or less obligatory, components of different kinds. The order

and relationships between the components of the dictionary (including the lemma list), the **MEGASTRUCTURE** of the dictionary, is discussed in Chapter 22.

Dictionary structure also includes the use of **CROSS-REFERENCES**, which can occur at several levels. A special section (Chapter 23) is devoted to their form and function.

1.4.8 Dictionary work and dictionary publishing

Working on a dictionary differs in some respects from working on books of other kinds. General aspects of **DICTIONARY PROJECTS** – their planning, implementation and follow-up – are discussed in Chapter 24.

In dictionary-making it is not only linguistic and lexicographic matters that are important. There are also **LEGAL AND ETHICAL ASPECTS** that must be taken into account, and these are briefly dealt with in Chapter 25.

Lexicographic data can be distributed in various ways and a dictionary can be published in various forms. For a very long time we have been accustomed to using dictionaries published in printed form on paper. But now the distribution of lexicographic data has to a great extent been transferred to digital media, and Chapter 26 is therefore devoted to a presentation of the **ELECTRONIC DICTIONARY** and its general characteristics.

1.4.9 Dictionary use and dictionary criticism

Once the dictionary has been published, it will hopefully reach its intended users. The branch of metalexigraphy that is concerned with investigating how different kinds of dictionaries are used by different kinds of users in their search for different kinds of information in different kinds of situations is called **USER RESEARCH**. A survey of this field is given in Chapter 27.

Furthermore, a recently published dictionary will at best be subjected to systematic and competent assessment, the results of which are published in specialist journals and (in some cases, and then in easily accessible form) in daily papers. Some aspects of this branch of metalexigraphy, **DICTIONARY CRITICISM**, are discussed in Chapter 28.

Literature

General introductions: Hausmann 1985b (concise and systematic); Béjoint 2000 (restricted to monolingual English dictionaries and not quite as modern as indicated by the year of publication); Landau 2001 (comprehensive and very readable

but dealing exclusively with monolingual British and American dictionaries); Hartmann 2001 (not a complete introduction but useful through its survey of current problems, projects and resources); Jackson 2002 (an introductory book for students, also including the elements of lexicology, but restricted to monolingual English lexicography); Herbst & Klotz 2003 (an introductory book in German dealing with both monolingual and bilingual dictionaries); Atkins & Rundell 2008 (a 'how to do it' textbook on dictionary-making).

Other general works: Zgusta 1971 (the 'classic', now partly out of date but still useful in many contexts); Hausmann et al. (eds.) 1989–91 (a monumental reference work, comprising 3,355 pages in all, now somewhat out of date as regards methods and forms of distribution);³ van Sterkenburg (ed.) 2003 (a collection of texts by some thirty writers on various aspects of lexicography); Hartmann (ed.) 2003 (a three-volume anthology of metalexigraphic texts, both 'classics' and contributions of more recent date); Fontenelle (ed.) 2008 (a collection of essays on lexicology and lexicography).

Bibliography: Wiegand 2006 (the first two volumes out of four planned).

Lexicography and society: Hausmann et al. (eds.) 1989–91, articles 1–25.

The history of lexicography: Hausmann et al. (eds.) 1989–91, articles 174–284 (lexicography of individual languages) and 308–28 (bilingual lexicography); Cowie 1998a, 1999 (English learners' dictionaries); McArthur 1998: 133–48 (lexicography in general, English learners' dictionaries); Béjoint 2000: 92–106, 169–208 (monolingual general-purpose dictionaries); Hartmann 2000 (European lexicography during the latest decades); Pruvost 2002 (French monolingual learners' dictionaries 1800–2000).

The history of metalexigraphy: Hausmann 1989d, 1989f.

The future of lexicography: Grefenstette 1998; Sinclair 2001; Rundell 2002; Teubert 2004.

The relationship between theoretical and practical lexicography: Atkins 1992/3.

The relationship between dictionary and encyclopedia: Lara 1989; Stark 1999: 6–14; Bauer 2005; Considine 2005.

Encyclopedic dictionaries in general: Hupka 1989b.

Encyclopedic learners' dictionaries: Crowther 1999; Stark 1999.

Technical (specialized) lexicography: Bergenholtz & Tarp 1995 (a general introduction); Schaefer & Bergenholtz (eds.) 1994 (topics and problems in technical lexicography); Wiegand 1988, Schaefer 1994 (the theory of technical lexicography); Kalverkämper 1990 (technical dictionaries for laypersons); Opitz 1990b

³ At the time of writing (autumn 2008), work is in progress aiming at the publication of a comprehensive supplementary volume.