

# Theories of the Lexicon and Decisions in Dictionary Making

*Pius ten Hacken*  
*Leopold-Franzens-Universität Innsbruck*  
*E-mail: pius.ten-hacken@uibk.ac.at*

## Abstract

There has been an on-going discussion of the role of theories in lexicography. Here, the purpose is to present a framework for applied sciences that can serve as a background for this discussion and apply this framework by comparing two theories of the lexicon. A traditional view of the lexicon in theoretical linguistics, e.g. Bloomfield (1933), is that it is an appendix to the grammar. A newer view, based on Jackendoff's (2002) Parallel Architecture, assigns it a much richer structure. Dictionaries are not theories of the lexicon, but tools. Therefore, explanations relating to dictionaries should come from applied science. In applied science, the choice, delimitation and analysis of practical problems is leading in the search for solutions that are backed up by explanatory theories. In this respect, lexicography can be compared to medicine. Some examples of relations that can be expressed in the lexicon model of Parallel Architecture suggest how such a theory can be used in lexicographic practice and theorizing.

**Keywords:** applied science, explanation, redundancy, Parallel Architecture, Relational Morphology

## 1. Introduction

In compiling a dictionary, decisions have to be taken. To the extent that these decisions can be motivated, one can say that there is a theory underlying them. What I want to explore here is how and to what extent theories of the lexicon developed in linguistics can be used as a component of the theoretical basis for lexicography. In section 2, I will outline two examples of theories of the lexicon from the field of generative linguistics that make rather different assumptions about the role of the lexicon. In section 3, I will show how the interpretation of the concept of *lexicon* in these theories differs from the nature of dictionaries as compiled in lexicography. In section 4, I turn to the position of theories as a component of a model of science and apply this model to lexicography. Section 5 proposes an analysis of lexicography as an applied science and shows how the theories of the lexicon presented earlier can play a role in it.

## 2. Two theories of the lexicon

In the traditional view of the description of a language, we need a dictionary and a grammar. The dictionary gives the words of the language and the grammar the rules for combining them into sentences. Linguists have generally preferred working on the rules, because they were thought of as more interesting. A lexicographer was seen rather as the “harmless drudge” of Samuel Johnson’s (1755) dictionary definition. In his influential overview of the field of linguistics, Bloomfield (1933: 274) characterized the lexicon as in (1).

(1) The lexicon is really an appendix to the grammar, a list of basic irregularities.

In (1) the lexicon is represented not as a central concern, but rather as a wastebasket for what cannot be described in the rules of the grammar. This perspective has been predominant also in generative linguistics. In his earliest works, Noam Chomsky did not even include a lexicon. When giving a first example of a formal grammar of English, Chomsky (1957: 26) introduces lexical items by means of rewrite rules. In their study of the position of the word in the theory of language, Di Sciullo & Williams (1987: 3) distinguish syntactic, phonological and lexical senses of *word* and characterize the latter as in (2).

(2) The lexicon is like a prison—it contains only the lawless and the only thing its inmates have in common is lawlessness.

The “lawlessness” in (2) corresponds to the “basic irregularities” in (1). Also in (2), the aim to direct attention away from the lexicon as an interesting component is obvious. In both (1) and (2), the lexicon is the result of removing all redundancies from the specification. Here a redundancy is a piece of information that can be derived from a rule. Grammar is interested in these rules, not in the remaining irregularities.

The view of the lexicon embodied in (1) and (2) has been challenged within generative linguistics. Jackendoff (1975) proposes that rules of morphology should be interpreted as redundancy rules. This means that the lexical entry specifies all information, but information that can be retrieved from a rule does not count fully to the burden of storing the entry. The information burden for a lexical entry consists of the information that is idiosyncratic in the sense of (1) plus the cost of referring to a rule for the information that is covered by this rule. Jackendoff (2002) elaborates this idea in his Parallel Architecture (PA). In PA, the information for a lexical entry is distributed over phonological, syntactic and conceptual structures. As an example, (3) gives the entry for Dutch *huis* (‘house’).

- (3) a. /hœys/  
 b. N<sub>neut</sub>  
 c. [<sub>Thing</sub> HOUSE]

In (3a), the pronunciation is given in IPA. In (3b), it is indicated that *huis* is a noun of neuter gender. In (3c), the meaning is represented. Following Jackendoff’s practice, I use a word in all-caps to represent those aspects of the meaning that are not analysed into smaller components. As this text is in English, I use English words for this. (3a) and (3c) represent prototypes. They serve to recognize what is perceived as a more or less typical instance of the stored pronunciation and meaning.

In PA, information specified in the lexicon is used to build up the phonological, syntactic and conceptual structures of an expression and link them to each other. As such, (3) is a linking rule. In (4), we have an example of an expression in which it is used.

- (4) Ze spelen in eigen huis tegen RKC.  
 lit. 'They play in own house against RKC'  
 i.e. They play a home match against RKC

The expression *in eigen huis* is an idiomatic expression used in the context of football and other sports. It cannot be translated literally, but the gloss and the translation in (4) indicate its structure and meaning. In PA, this means that *in eigen huis* has a lexicon entry of its own specifying its pronunciation, syntactic structure and meaning. The fact that *in eigen huis* consists of three words is represented by references to the entries for *in*, *eigen* and *huis*, not by removing the specifications that can be found in these entries. This is the central difference from the model in (1) and (2).

Another point to be noted in (4) is that not all entries need a specification for all three structures. In the entry for *ze* ('they'), the information that it is a third person plural pronoun is specified in syntax. The actual reference needs to be determined in the context. This means that the entry for *ze* does not have any information for conceptual structure. We can represent it as in (5).

- (5) a. /zə/  
 b. Pron<sub>3pl</sub>  
 c. ∅

When we combine the ideas that information in a lexical entry can be structured, as is required for *in eigen huis*, and that the information for a particular representation can be empty, as in (5c), we can derive that syntactic rules can also be represented as lexical entries. A typical syntactic rule will combine zeros in phonology and conceptual structure with a structural specification in syntactic structure. Jackendoff (2002: 152-182) makes this point in detail and concludes that there is no categorical difference between words and rules. As a consequence, all information that is involved in knowing a language is encoded in lexical entries. Jackendoff & Audring (2020) elaborate this idea into a theory of the lexicon they call *Relational Morphology*. This theory is in a sense the exact opposite of the theory advocated in (1) and (2). Whereas Bloomfield (1933) aims to minimize the content of lexical entries in favour of the grammatical rule component, Jackendoff & Audring (2020) include the grammatical rules as entries in the lexicon.

### 3. Lexicon and dictionary

Dictionaries are generally thought of as descriptions of the vocabulary of a language. This is often reflected in their title. Van Dale (2022), the latest edition of the major dictionary of Dutch, published in three volumes, is entitled *Groot woordenboek van de Nederlandse taal* ('Large dictionary of the Dutch language'). In the preface to this dictionary, Parqui (2022) describes the task of such a dictionary as in (6).

(6) [h]et is de taak van een woordenboek om de woordenschat zo volledig mogelijk te beschrijven.<sup>1</sup>

Here *woordenschat* is a compound of *woord* ('word') en *schat* ('treasure'). Van Dale (2022: *woordenschat*) gives as the first sense the vocabulary of a specific language and as the second sense the vocabulary of a particular person. Parqui (2022) makes it clear in the context that he intends to refer to the first sense. This raises the question how realistic the task in (6) is.

In order to investigate this question, we can consider how it is determined that a word is part of the *woordenschat* in the sense of (6). We are not surprised to find that *boek* ('book') is in the dictionary, but *book*, the English equivalent is not. It is important to see, however, that someone has to decide this. Van Dale (2022: 1769) describes the procedure as follows. Before the availability of large electronic corpora, dictionary users signalled new words and the dictionary editor had to decide on the basis of intuition whether the word was sufficiently common to deserve an entry. Since the 1990s, the use of corpora has made it possible to check the use in recent years. This description suggests that it is the corpus that determines whether a word is included in the dictionary.

One problem with this reasoning is that *book* actually occurs in Dutch corpora. CHN (2021), a large corpus of written and spoken Dutch, gives 13,370 occurrences (4.72 per million words). This is much more than, for instance, *wolfram* ('tungsten'), which has 213 occurrences, and more than twice the number for *gebarentaal* ('sign language'), which has 5,329 occurrences. These observations do not imply that the decision not to include *book* in van Dale (2022) is arbitrary or wrong, but they exclude the reasoning that the decision is a direct consequence of the frequency as found in the corpus. Instead, they demonstrate that the lexicographer's intuition remains an essential factor in the decision, even when corpora can be used as supportive evidence.

For the decision whether a word belongs to the vocabulary of Dutch, we can use two basic types of information. One is based on language use. This includes electronic corpora, but also individual observations as signalled by dictionary users or indeed lexicographers. The other one is linguistic intuition. In practice this is primarily the lexicographer's intuition whether a word is part of the vocabulary of a large enough number of speakers of Dutch. What is not possible is to observe Dutch as an entity that is realized somewhere. Many speakers of Dutch may refer to van Dale (2022) or an earlier edition as the authority on whether a word exists in Dutch, but the lexicographers compiling the dictionary cannot use it in this sense, because they take the ultimate decisions about inclusion. We should therefore reject (6), as it does not specify a realistic task.

The rejection of (6) gives a new urgency to the question of what a dictionary actually is. Ten Hacken (2009) proposes that a dictionary should be understood as a tool, not as a description. The purpose of a dictionary is to solve lexical problems identified by the users. In the same way as a hammer does not solve the problem of fixing a clock to the wall, the dictionary does not offer solutions, but constitutes a tool that the user can work with in order to solve the problem. In fixing a clock to the wall, the user has to determine the course

---

1 'it is the task of a dictionary to describe the vocabulary as comprehensively as possible' [my translation, PtH]

of action, involving, for instance, attaching it with a nail. The hammer plays a crucial part in getting the nail into the wall, but it is the user who has to think of this and take up the hammer and their skill in using it determines the success of the action. This is not unlike the use of a dictionary. The dictionary user decides what to look up and their skill in selecting where to look and interpreting what they find there determines the success of the action.

As ten Hacken (2009) emphasizes, the interpretation of a dictionary as a tool for problem solving does not reduce the value of dictionaries or of the lexicographer's expertise. It is broadly compatible with Tarp's (2008) functional theory of lexicography. It requires that the lexicographer assesses the user's needs when compiling a dictionary. This involves both the macrostructure and the microstructure of the dictionary and both the question of what information to include and the question of how to represent it.

When a dictionary is a tool, it is of a quite different nature from the lexicon as intended in section 2. In the generative framework adopted by Di Sciullo & Williams (1987) in (2), the lexicon is a component of a theory of language. In Jackendoff's (2002) PA, the lexicon has a different relationship to the theory of language, but it is still a component of it. The purpose of a theory of language in the generative framework is to explain certain aspects of language as it is observed in a speaker's competence. Chomsky (1986) presents this purpose as explaining how language acquisition is possible and how knowledge of language must be represented in a speaker in order to make it possible. His idea is that language acquisition depends on innate facilities present in the genetic specification of the human species. Hauser et al. (2002) shift the central question one level up and ask how language could emerge in the evolution of the human species. As explained in ten Hacken (2007a: 258-266), it is in their approach to this question that Hauser et al. (2002) and Jackendoff (2002) propose different paths. Whereas Hauser et al. (2002) propose that there was a single catastrophic change that triggered the emergence of language, Jackendoff (2002) proposes a gradualist emergence with different intermediate stages that each produced evolutionary advantages. As we have seen in section 2, the latter approach leads to Jackendoff & Audring's (2020) Relational Morphology. Reboul (2017) gives an elaborate argumentation for the emergence of language in a single step and outlines how this influences the view of the lexicon.

In conclusion, lexicography and generative linguistics approach vocabulary in a very different way. Whereas lexicographers produce a tool for speakers of a language to answer questions they have about vocabulary items, generative linguists propose theories about the possibility of language acquisition and the emergence of language as a property of the human species in its evolution. Nevertheless, they are both concerned with vocabulary items as knowledge elements in speakers.

#### **4. Models of applied and empirical science**

In discussing the models of science underlying generative linguistics and lexicography, it is first of all necessary to distinguish empirical science and applied science. A prime example of an empirical science is astronomy. Astronomy has often been used as the basis for building up a model of how a science works. Astronomists, and scientists more generally, tend to operate in the understanding that they make observations, often as a result of experiments, then generalize over these observations and propose a theory as a hypothesis

about why the generalization holds. The theory is then used as a source of predictions. Testing these predictions gives rise to new observations that may lead to modifications or extensions of the theory. In the case of astronomy, positions of the planets at particular times are observations, orbits are generalizations and gravity is a central notion in the theory explaining why planets remain in their orbits. Nagel (1961) is a classical overview of this view of science. Kuhn (1970) demonstrates that in practice, science is more complex than this, but scientists are protected from these complications by the paradigm they work in. Andersen et al. (2006) elaborate this view. Ten Hacken (2007a) applies it to generative linguistics.

The model of empirical science cannot be used for medicine. In medicine it is not enough to explain observations, e.g. how patients suffer and recover or die, but it is expected that medicine develops cures, i.e. procedures that change the natural course of events for the better. Whereas astronomy solves only problems of understanding, the problems that medicine solves are practical problems. It is sometimes argued, e.g. by Cosmacini (2008), that medicine is not a (real) science. However, medicine is traditionally thought of as being scientific. In fact, it can be seen as the prime example of applied science. It is applied, because of its concern with solving practical problems. At the same time, there is a crucial difference between applied science and technology. Both can use the results of empirical science, but the nature of their use is different. When I go to my general practitioner, I expect that she can explain why she prescribes a certain cure and that this explanation is based on scientific evidence. When I bring my mobile phone for repair, I am not interested in such an explanation. Although a mobile phone is based on the application of knowledge that originates in sciences, someone repairing it only needs the practical knowledge to make it work again, not the explanatory knowledge based on physics and mathematics. Applied science is not just the application of science, but it is a science itself. Medicine shares with astronomy the concern for a proper scientific explanation.

Compared to the large body of studies of empirical science, there are only few studies of applied science aiming to analyse the process of doing science epistemologically. For the analysis of computational linguistics as an applied science, ten Hacken (2007b: 261) proposes four guiding questions. In (7), the headings for these questions are maintained, but the first two questions have been reformulated so as to be applicable to applied science more generally.

- (7)
  - a. *Problem identification*: What are the boundaries of the problem?
  - b. *Evaluation*: How can it be determined that the problem has been solved?
  - c. *Problem decomposition*: How can the problem be analysed into components?
  - d. *Knowledge selection*: Which types of knowledge should be used to solve the component problems?

An example from the field of medicine can be taken from the reaction to a novel disease. In this context, we can compare the plague with Covid-19. In medicine as an applied science, a new disease is considered as a problem to be solved. In (7a), the question is how to delimit this problem, i.e. what counts as an instance of the particular disease. The starting point for a disease is usually a set of symptoms that tend to occur together. For the plague, it took centuries to develop an understanding (cf. Orent 2004). In our current understanding, it is caused by a particular bacterium, *Yersinia pestis*. Its discovery by Alexandre Yersin

in 1894 provided the basis for an answer to (7a). In (7b), criteria for success have to be specified. In the case of plague, this might be based on the amount of *Yersinia pestis* found in lymph nodes. If the value drops below a threshold, the patient is declared *cured*. In (7c), the aim is to understand the process that leads from infection to disease. Finally, (7d) uses theories of what Cosmacini (2008) calls the “*scienze di base*” to identify where and how intervention in the process of the disease is possible. For Covid-19, the process of delimitation of the disease, discovery of its cause and development of a cure took a much shorter time. This can be attributed to progress in medical knowledge. However, the emergence of long Covid shows that our current understanding is still partial (cf. Codignola 2022).

It is the relevance of theories and of their explanations that distinguishes applied science from technology. Using knowledge from empirical science to solve practical problems is not sufficient to qualify as science. Rossi (1997: 41-45) makes this point in the context of the work of Leonardo da Vinci (1452-1519), which he qualifies as technology because of Leonardo’s lack of interest in explanation. Explanation is the goal of a theory. Theories of empirical science are addressed quite explicitly in (7d). Theories of applied science address all questions in (7). In the case of medicine, the theories of anatomy, bacteriology and virology involved in the explanation of what happens in a case of plague or of Covid-19 are theories of empirical science. The explanation of how interventions are possible belongs to the theory of medicine as an applied science. At a more abstract level, the explanation of what is a good approach to each of the questions in (7) is also the subject of theories of applied science.

## 5. Dictionary making as applied science

In order to consider the role of the theories of the lexicon presented in section 2 in dictionary making, we have to model dictionary making in terms of the questions in (7). The question in (7a), applied to lexicography, requires a specification of the range of problems that a particular dictionary should solve. As we saw earlier, (6) is not an adequate problem definition. Nevertheless, views such as this one are widespread and have a long tradition. Thus, for bilingual dictionaries, Zgusta (1971: 294) makes the claim in (8).

- (8) The basic purpose of a bilingual dictionary is to coordinate with the lexical units of one language those lexical units of another language which are equivalent in their meaning.

What (8) shares with (6) is a complete neglect of the user’s purpose. *Equivalence* is presented as if it were a theoretically neutral, objective property of the relationship between lexical units in two languages. In section 3, I showed that such assumptions are problematic and argued that a dictionary should be viewed as a tool. In their discussion of bilingual dictionaries, Atkins & Rundell (2008: 40) start with the distinction between decoding and encoding dictionaries. This offers a better starting point for answering (7a), but it is only a starting point. Bilingual dictionaries are used by translators and by learners, but they have very different requirements on the information they want to find in a decoding dictionary. Adamska-Sałaciak (2016) discusses this matter in more detail.

Identifying the scope of a dictionary type extends also to the mention of the language. As shown in section 3, van Dale’s (2022) use of *Nederlandse taal* (‘Dutch language’) in its

title cannot be interpreted as a reference to an entity that is realized in the outside world. Instead, it indicates the range of entries and the type of information to be found in the dictionary. The dictionary presents itself as a major general dictionary in the sense of van Sterkenburg (2003), which means that it aims for what can be called a ‘usage-based authority’. Intended users of van Dale (2022) want to solve problems of what is correct in Dutch and expect the information they find in it to be well-founded. *Well-foundedness* is interpreted in different ways depending on expertise. Lexicographers will see it as ‘corpus-based’, but linguists will know, as indicated above, that the lexicographer’s competence plays a crucial role in interpreting corpus data and formulating the entry. Many users will not ask the question how the authority is founded, but assume there is a proper foundation.

The evaluation question in (7b) is the basis for user studies. The question in such studies is whether users can solve the problems they consult a dictionary for with the information they find in it. An example of a user study with a specific question is Ptasznik (2023). Here the question is whether giving dictionary users more examples leads to better solutions or to more confusion. Ptasznik set up a study where users had to solve pre-defined problems with a smaller or a larger set of examples. His study is a typical example of a range of lexicographic user studies. In order to address (7b), it needs to be determined to what extent the problems in the scope of (7a) are in fact solved. Experimental user studies aim to find more precise information by controlling the conditions, but their findings are meant to be typical for actual dictionary use.

In (7c), a strategy for breaking down the general problem in (7a) into smaller problems is asked for. In the case of lexicography, we have to distinguish the problem of dictionary making and the problems solved through dictionary use. The process of dictionary making can be decomposed into the stages in (9).

- (9) a. Collection of texts to be included in a corpus
- b. Selection of headwords for the dictionary
- c. Selection and representation of information about each headword

The analysis lexicographic work into the stages in (9) has a long tradition. Although the use of computers changes the way a corpus is compiled and can be queried, the idea of using a corpus of texts as a basis for a dictionary is found in the 17<sup>th</sup> century, for instance in the 1612 Italian dictionary of the Accademia della Crusca and the 1680 French dictionary by Pierre Richelet. For each stage in (9), practical problems need to be solved. Van Sterkenburg (1992) documents this process for the Dutch *Woordenboek der Nederlandsche Taal* and Giliver (2016) for the *Oxford English Dictionary*. Atkins & Rundell (2008) give a more general overview of the issues and considerations.

Finally, (7d) addresses the question of which theories can be used. Here, the distinction between applied science and technology is reflected in the degree to which theories are used not only as a source of practical knowledge, but also as a source of explanation.

In order to see how the analysis of lexicography as an applied science works out, we can draw a parallel between lexicography and medicine. Table 1 lists elements that can be seen as corresponding.



<i>Medicine</i>	<i>Dictionary use</i>	<i>Dictionary making</i>
Analysis of a disease	Analysis of user problems	Analysis of decision problems
Development of a cure	Compilation of a dictionary	Development of guidelines
Contact with patients	Consultation of a dictionary	Writing dictionary entries

**Table 1:** Correspondence between medicine and lexicography as applied sciences.

In medicine, the starting point is the observation of conditions that patients present with. A generalization of which conditions belong together as symptoms of a disease is crucial for the further process, but it is not applied science. At this stage, the focus is on the explanation of the patient’s conditions, so that it is a matter of empirical science. Recognizing a disease is like discovering the orbit of a planet in the sense that it does not change anything in the outside world. It is only in the second stage that we turn to interventions in this natural process. The explanatory connection between the intervention and the analysis of the disease determines the extent to which medicine is an applied science rather than technology. The contact with patients uses the results of applied science, while monitoring the effects of the cure yields further data for the refinement of the theory.

For lexicography, Table 1 gives two ways of modelling it as an applied science. One takes the cure in medicine as correlating with the dictionary. Here the disease corresponds to a problem the user experiences, for which the dictionary might be used in the solution. The other one takes the guidelines for lexicographic work as the correlate of the cure. In this perspective, the disease corresponds to the feeling of being overwhelmed by the task of compiling a dictionary, as described by Gilliver (2016) at various points in the genesis of the OED, especially in the early stages. Documentation of a working method with details about the tasks in (9) can then be used in the writing of entries. The final column in Table 1 can be seen as an elaboration of the second component of the middle column.

In Table 1, the applied component of the work is in the bottom field of each column. The references to empirical science occur in the first two stages. In medicine, anatomy and microbiology are central in the analysis of the disease and pharmacy is central in the development of a cure. We need to explain the way the disease evolves by reference to biological knowledge. Pharmacy explains the working of substances used in the cure. Chemistry can also contribute to the explanation. In lexicography, various branches of linguistics can be used. Both the user problems and the decision problems referred to in Table 1 have to be analysed in terms of the language system as it is realized in a speaker in order to work towards a solution. The effectiveness of the dictionary and of the guidelines depends on the way the dictionary user and the lexicographer process the relevant information in order to work with their language competence.

It is interesting to note that in lexicography, some of the tasks evoked in Table 1 are often labelled as *metalexigraphic* (cf. Atkins & Rundell, 2008: 1). No correlating term is used in the medical context. A GP does medicine as well as a researcher analysing a disease. If we want to consider lexicography as an applied science, it seems reasonable to label the

work of writing dictionary entries as well as the work of analysing the decision problems involved as *lexicography*.

## 6. Applying theories of the lexicon

At this point, we have all the ingredients for a well-founded application of theories of the lexicon in dictionary making. The point here is to compare the benefits of adopting a Bloomfieldian view of the lexicon as an appendix, as formulated in (1) and (2), with the advantages of the model of PA in which the lexicon is the encoding of all linguistic knowledge. As a first example, I will use the Dutch expression *in eigen huis* as illustrated in (4).

At first sight, the Bloomfieldian view of the lexicon is closer to the one generally adopted in lexicography. After all, we make a distinction between dictionaries and grammars. In PA, this distinction is not maintained, because grammar rules are also formulated as lexicon entries. When we consider *in eigen huis*, however, we are dealing with a multi-word expression for which it is difficult to specify what exactly is the additional information that we should specify in a special entry. In a dictionary of Dutch, there are already entries for *in* ('in'), *eigen* ('own') and *huis* ('house'). Excluding all redundant information leaves very little room for a separate entry for *in eigen huis*.

Corpus data can be used to give a more detailed picture of the use of the expression. When we look up *in eigen huis* in CHN (2021), we get 10,646 occurrences, i.e. 5.41 per million. The most common following words are the prepositions *met* (853 cases) and *tegen* (640 cases). An example of each is given in (10).

- (10) a. Heerenveen ging gisteravond in eigen huis met 1-3 onderuit tegen Steaua.  
lit. 'Heerenveen went last-night in own house with 1-3 down against Steaua.  
i.e. Last night, Heerenveen lost at home against Steaua with 1-3.
- b. Eind augustus leed Beerschot in eigen huis tegen Standard nog een eer-  
volle nederlaag  
lit. 'late August, suffered Beerschot in own house against Standard still  
an honourable defeat'  
i.e. In late August, Beerschot still lost honourably at home against Stan-  
dard

As shown in (10), *met* ('with') is usually followed by a result and *tegen* ('against') by a club. Both prepositions introduce modifiers of a verb indicating playing, often specified as winning or losing. These modifiers follow *in eigen huis*, which is also such a modifier. Examples such as (10) are typical of reports of football matches. It can be assumed that information about the meaning and use of *in eigen huis* is stored in the knowledge of language of speakers who often read reports of football matches. As a result, authors of such reports take the relevant information to be well-known. This makes the use less transparent for readers that do not have the expression in their mental lexicon. Therefore, it is the kind of expression they are likely to look up in a dictionary.

It is important to keep in mind, however, that *in eigen huis* is not necessarily interpreted in the opposition between home and away matches. CHN (2021) also gives examples such as (11).

- (11) De provincies moeten vooral in eigen huis orde op zaken stellen.  
lit. 'The provinces have above-all in own house order on business to-put.'  
i.e. Above all, the provinces have to put things right within their own organization

In (11), *huis* is used metaphorically for an organization and the expression *in eigen huis* is to be interpreted compositionally on the basis of the three lexical entries it consists of. In Relational Morphology (RM), the theory of the lexicon based on PA described by Jackendoff & Audring (2020), modelling rules for complex entries as redundancy rules accounts for such constellations very naturally. In language processing, the context will determine the prominence of the special reading in (10). In a lexicographic context, this may be mimicked in electronic dictionaries by entering a user profile as a background for searches. User profiles are the implementation of the concept of priming in Jackendoff's (2002) account of language processing. They have no obvious correlate in the Bloomfieldian view of the lexicon.

As a second example, let us consider the verb *samenwerken* ('collaborate'). This is a separable verb composed of *samen* ('together') and *werken* ('work'). The examples from CHN (2021) in (12) illustrate its use.

- (12) a. Het is essentieel dat we in vertrouwen samen kunnen werken met andere inlichtingendiensten  
lit. 'it is essential that we in trust together can work with other intelligence services'  
i.e. It is essential that we can trustfully collaborate with other intelligence services
- b. [een plaats] waar mensen met en zonder beperking samen kunnen werken en ontspannen  
lit. '[a place] where people with and without handicap together can work and relax'  
i.e. [a place] where people with and without handicap can work and relax together

In (12a), the two components of *samenwerken* are separated by *kunnen* ('can'), showing that it is a separable verb. In (12b), the same two components are used, but they are not part of the separable verb. This shows that we need an entry for *samenwerken*, but also entries for *samen* and *werken*. What redundancy rules in PA do is to express the relationships between related entries. These associations concern not only *samenwerken* and its component parts, but also other separable verbs with *werken* and with *samen*, e.g. *inwerken* ('train, introduce to a new job'), *samenvatten* ('summarize'). In an electronic dictionary, such information can be presented to support vocabulary learning and show connections between words. In a Bloomfieldian dictionary there are no provisions for representing connections between entries.

## 7. Conclusions

In this paper, I discussed the question of how theories of the lexicon can be used in lexicography. I presented the Bloomfieldian model of a lexicon as a minimal set of information

that cannot be produced by rules and contrasted it with the model adopted in Jackendoff's PA, which relies on redundancy rules and was elaborated as RM, a theory of the structure of the lexicon, by Jackendoff & Audring (2020). In order to find how these models can be used in lexicography, I presented an analysis of applied science, contrasting it on one hand with empirical science, on the other with technology. On the basis of this model, I compared lexicography to medicine and indicated the place of theories of the lexicon in the analysis of lexicography.

A first set of conclusions can be drawn from the comparison of medicine and lexicography as instantiations of applied science. It is obvious that medicine can draw on a much larger body of theories providing a scientific basis. For lexicography, the fast growth of studies of dictionary use is promising, but it will be crucial that such studies are theoretically informed. In the same way as knowledge of anatomy does not cure patients, but no explanatory approach is possible without it, linguistics is essential for theoretical explanations of lexicographic insights, even though it does not directly guide the writing of dictionary entries. Separating theoretical considerations from practical work by the opposition of lexicography and metalexicography is not helpful in this respect. In the same way as in medicine, lexicography operates best if it integrates the practical and theoretical orientations. *Medicine* includes both the work of general practitioners and advanced research into safe vaccines. Similarly, writing dictionary entries should not be separated from research on explaining their effect.

A second set of conclusions pertains to the comparison of RM with the Bloomfieldian approach to the lexicon. Whereas for Bloomfield, the lexicon is no more than an appendix to the grammar, RM aims to offer a theoretical foundation for the lexicon on the basis of its use in language processing and its origin and extension in language acquisition. As the examples in section 6 suggest, it is in particular the encoding of connections between lexical entries in RM that can serve as an inspiration for electronic dictionaries, especially for multi-word units and morphologically complex words.

## References

- Adamska-Sałaciak, Arleta (2016). 'Explaining meaning in bilingual dictionaries'. In Durkin, Philip (ed.), *The Oxford Handbook of Lexicography*. Oxford: Oxford University Press, pp. 144-160.
- Andersen, Hanne; Barker Peter & Chen, Xiang (2006). *The Cognitive Structure of Scientific Revolutions*. Cambridge: Cambridge University Press.
- Atkins, B.T. Sue & Rundell, Michael (2008). *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Bloomfield, Leonard (1933). *Language*. London: Allen & Unwin.
- CHN (2021). *Corpus Hedendaags Nederlands*. Leiden: Instituut voor de Nederlandse Taal, <http://hdl.handle.net/10032/tm-a2-s8>.
- Chomsky, Noam (1957). *Syntactic Structures*. Den Haag: Mouton.
- Chomsky, Noam (1986). *Knowledge of Language: Its Nature, Origin, and Use*. Westport (Conn.): Praeger.
- Codignola, Agnese (2022). *Il lungo Covid: La prima indagine sulle conseguenze a lungo termine del virus*. Milano: UTET.

- Cosmacini, Giorgio (2008). *La medicina non è una scienza: Breve Storia delle sue scienze di base*. Milano: Cortina.
- van Dale (2022). *Van Dale Groot Woordenboek van de Nederlandse Taal*. 16th ed., den Boon, Ton & Hendrickx, Ruud (eds.), Utrecht/Antwerpen: Van Dale Lexicografie.
- Di Sciullo, Anna Maria & Williams, Edwin (1987). *On the Definition of Word*. Cambridge (Mass.): MIT Press.
- Gilliver, Peter (2016). *The Making of the Oxford English Dictionary*. Oxford: Oxford University Press.
- ten Hacken, Pius (2007a). *Chomskyan Linguistics and its Competitors*. London: Equinox.
- ten Hacken, Pius (2007b). 'Computational Linguistics as an Applied Science'. In Dodig Crnkovic, Gordana & Stuart, Susan (eds.), *Computation, Information, Cognition: The Nexus and the Liminal*, Cambridge: Cambridge Scholars Press, pp. 260-269.
- ten Hacken, Pius (2009). 'What is a Dictionary? A View from Chomskyan Linguistics'. *International Journal of Lexicography* 22: 399-421.
- Hauser, Marc D.; Chomsky, Noam & Fitch, W. Tecumseh (2002). 'The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?'. *Science* 298: 1569-1579.
- Jackendoff, Ray (1975). 'Morphological and Semantic Regularities in the Lexicon'. *Language* 51: 639-671.
- Jackendoff, Ray (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Jackendoff, Ray & Audring, Jenny (2020). *The Texture of the Lexicon: Relational Morphology and the Parallel Architecture*. Oxford: Oxford University Press.
- Johnson, Samuel (1755). *A Dictionary of the English Language*. London: Knapton, Longman, etc.
- Kuhn, Thomas S. (1970). *The Structure of Scientific Revolutions, Second Edition, Enlarged*. Chicago: University of Chicago Press.
- Nagel, Ernest (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Indianapolis: Hackett.
- Orent, Wendy (2004). *Plague: The mysterious past and terrifying future of the world's most dangerous disease*. New York: Free Press.
- Parqui, Jaap (2022). 'Woord vooraf'. In van Dale (2022), p. ix.
- Ptasznik, Bartosz (2023). 'More examples may benefit dictionary users'. *International Journal of Lexicography* 36: 29-55.
- Reboul, Anne (2017). *Cognition and Communication in the Evolution of Language*. Oxford: Oxford University Press.
- Rossi, Paolo (1997). *La nascita della scienza moderna in Europa*. Roma / Bari: Laterza.
- van Sterkenburg, Piet (1992). *Het Woordenboek der Nederlandsche Taal: Portret van een Taalmonument*. 's-Gravenhage: Sdu.
- van Sterkenburg, Piet (2003). 'The' dictionary: Definition and history'. In van Sterkenburg, Piet (ed.), *A Practical Guide to Lexicography*. Amsterdam: Benjamins, pp. 3-17.
- Tarp, Sven (2008). *Lexicography in the Borderland between Knowledge and Non-Knowledge: General Lexicographical Theory with Particular Focus on Learner's Lexicography*. Tübingen: Niemeyer.
- Zgusta, Ladislav (1971), *Manual of Lexicography*. Praha: Academia & Den Haag: Mouton.