Automatically extracted word formation products in an online dictionary

Sabina Ulsamer
Institute for German Language (IDS)
R5, 6-13; 68161 Mannheim
E-mail: ulsamer@ids-mannheim.de

Abstract

This paper presents a method of automatically extracting compounds and derivatives relating to particular base words from the headword list of the online dictionary *elexiko*, and a layout for displaying them in the entries of the base words. The aims of this project were to illustrate word formation relations and to provide a stronger connection between the headwords. The starting point for the automatic retrieval of word formation products was the morphological analysis of the headwords using a computational morphology. The analyses were stored in a database. Having obtained the base words, all compounds and derivatives relating to them were retrieved and displayed online in the entries of the base words.

Keywords: online dictionary; word formation; automation

1. Introduction

In metalexicographic literature (cf. Ulsamer (forthcoming) for an overview of the debate), the problem of illustrating word formation in dictionaries, i.e. morphological relations between dictionary entries, has been widely discussed. Mugdan (1984:274) states that compounds and derivatives form a complex net of connections between lexemes. Word formation products enrich the vocabulary by demonstrating relations between words. In a strictly alphabetical ordering of headwords (in a printed dictionary), these connections cannot emerge. With the advantage of space and hyperlinks, electronic dictionaries offer a variety of possibilities for illustrating morphological relations between headwords. Methods from computational linguistics and software tools increasingly complement work in lexicography (cf. Klosa, 2010) and provide new approaches to dictionary writing.

The German online dictionary *elexiko*[1], already combining lexicographically edited information with automatically generated information, aims to present all compounds and derivatives relating to a particular base in the entry of that base word. As part of the project 'User-adaptive access and cross-references in *elexiko* (BZV*elexiko*)[2]' at the Institute for German Language (IDS[3]), methods were developed to extract compounds and derivatives relating to a particular base word from a database of morphologically analysed headwords. The aim is to provide a better connection between the headwords. In this paper – after a brief description of the online dictionary *elexiko* and how word formation is handled there in section two – the underlying morphological analyses and retrieval methods are described in detail in section three. Section four focuses on the online display of the automatically retrieved word formation products. Section five discusses the problems and advantages of the presented method. After a summary in section six, implications for other languages and further research are considered in section seven.

2. The online dictionary elexiko

2.1 Background

elexiko is a corpus-based dictionary for contemporary German being compiled at the IDS and integrated into the lexicographic internet portal OWID[4]. *elexiko* is intended for native and non-native speakers of German, for linguists and non-linguists alike. The 300,000 headwords were extracted based on frequency from a dynamic corpus (currently consisting of 2.8 billion tokens) specifically compiled for *elexiko*. The corpus is comprised of daily and weekly newspapers and magazines from Germany, Austria, Switzerland and the former GDR. Instead of editing the entries in alphabetical order, modules of entries, defined by specific semantic, syntactic or morphological criteria, are chosen (cf. Storjohann, 2005:55-83).

2.2 Word formation in elexiko

The dictionary aims to explain the structure and word formation process of derived and compounded words in order to fulfil Barz's (2001:89) second objective, i.e. to reconstruct the recent word formation steps of a secondary word. In order to fulfil her first objective (Barz, 2001:88), that is to demonstrate the word formation activity of a primary word, all derivations and compounds containing the primary word in question should be listed in the entry of this headword.

The adjective *jugendlich* (*juvenile*) is described as an explicit derivation consisting of the base *Jugend* (*youth*), a noun, and the suffix *-lich*. In order to illustrate the word formation activity of the primary word *Jugend* *elexiko* aims to present all derivatives and compounds from the headword list in which *Jugend* appears: two derivations – *jugendlich* and *jugendhaft* – and about 400 compounds where *Jugend* is either the first constituent

[1] www.elexiko.de
[2] www.ids-mannheim.de/lexik/BZVelexiko/
[3] www.ids-mannheim.de
[4] www.owid.de

as in *Jugendbuch* ('book for adolescents') or the second constituent as in *Dorfjugend* ('young people of the village').

The aim of this approach is to enhance connections between words. Starting from the primary word, a net of word formations expands, offering the dictionary user a way of inferring new lexicological relations. Especially for non-edited entries, *elexiko* aims to provide more information by presenting automatically retrieved word formation products under a primary word.

3. Automatic retrieval of word formation products

3.1 Morphological analysis of headwords

As a prerequisite for the retrieval of word formation products the whole *elexiko*-headword list was morphologically analysed. The headwords were split into their constituents, and these specified morphologically. The analysis and segmentation were done with the computational morphology Morphisto.

Morphisto[5] is a computational morphology for German developed at the IDS within the TextGrid-project[6]. The tool is based on SMOR (Schmid et al., 2004). Morphisto is able to generate inflection paradigms as well as to analyse words. Additionally, Morphisto assigns a binary branching hierarchical structure to complex word formations (Zielinski et al., 2009). Morphisto was used to decompose every word of the *elexiko*-headword list into complement and head. The head denotes the grammatical and morphosyntactic category of the whole word. According to the Righthand Head Rule (Williams, 1981), the right-most constituent is the head of a German word. The complement is the accompanying constituent that specifies the head. In compounds, the second constituent is the head, while the first constituent denotes the complement. In derivatives with suffixes, the suffix functions as head, whereas the base is the complement. In prefixed words, the base is the head and the prefix the complement (cf. fig. 1).

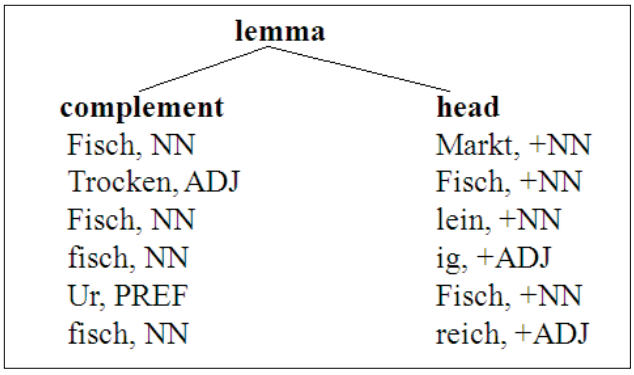

Figure 1: Tree structure of various words

[5] www.ids-mannheim.de/lexik/TextGrid/morphisto.html
[6] www.textgrid.de

Furthermore, every complement and head was marked with a part-of-speech-tag, and the head POS-tag labelled with an additional plus sign. Suffixes are tagged with the part of speech they create. The diminutive suffix *-lein* (*-let*), which creates nouns, is labelled +NN. Prefixes were tagged PREF.

Moreover, Morphisto generated the word formation rule for every headword (cf. table 1). Word formation rules for compounds have the general form <+TAG>→<TAG> <+TAG>. A POS-tag surrounded by angle brackets is followed by an arrow which is itself followed by two further POS-tags, each in angle brackets. A nominal compound such as *Fischmarkt* (*fish market*) is described as <+NN>→<NN> <+NN>, which reads 'a noun is comprised of two nouns, where the second noun is the head of the whole construction'. In compounds, the second constituent always functions as the head which is marked by the plus sign before the second POS-tag after the arrow. The adjectival compound *fischreich* ('fish rich' – *rich in fish* [waters]) consists of a noun and an adjective: <+ADJ>→<NN> <+ADJ>. Word formation rules for derivatives can have two forms. Prefixed words are almost identical to compounds in their rule format with the general form being <+TAG>→<PREF> <+TAG>. A word like *Urfisch* (*ancient fish*) is therefore written as <+NN>→<PREF> <+NN>. Suffixed words, however, are designated with rules in the form of <+TAG>→<TAG> <SUFF>.

lemma	word formation rule
Fischmarkt	<+NN>→<NN> <+NN>
Trockenfisch	<+NN>→<ADJ> <+NN>
Fischlein	<+NN>→<NN> <SUFF>
fischig	<+ADJ>→<NN> <SUFF>
Urfisch	<+NN>→<PREF> <+NN>
fischreich	<+ADJ>→<NN> <+ADJ>

Table 1: A few words and their word formation rule

Simple words of any part of speech are annotated <+TAG>→lemma, e.g. <+NN>→*Fisch* (*fish*), <+ADJ>→*grün* (*green*), and <+V>→*schlafen* (*to sleep*).

The results of the morphological analysis for the entire *elexiko*-headword list were written into a relational database table (called the base table from now on) in the database management system Oracle. For each entry, the complement and the head, as well as their corresponding tags and their word formation rule, were inserted into different columns (cf. table 2). Storing the analyses in a database made it possible to search for specific patterns and retrieve the various word formation products accordingly.

lemma	complement	compl _tag	head	head_tag	rule
Fisch			Fisch	+NN	<+NN>→Fisch
Flussfisch	Fluss	NN	Fisch	+NN	<+NN>→<NN> <+NN>
Fischfabrik	Fisch	NN	Fabrik	+NN	<+NN>→<NN> <+NN>
Fischfabrikschiff	Fischfabrik	NN	Schiff	+NN	<+NN>→<NN> <+NN>
Trockenfisch	trocken	ADJ	Fisch	+NN	<+NN>→<ADJ> <+NN>
Urfisch	ur	PREF	Fisch	+NN	<+NN>→<PREF> <+NN>
fischig	Fisch	NN	ig	+ADJ	<+ADJ>→<NN> <SUFF>
fischen	Fisch	NN	en	+V	<+V>→<NN> <SUFF>
fischreich	Fisch	NN	reich	+ADJ	<+ADJ>→<NN> <+ADJ>

Table 2: Simplified diagram of the base table

3.2 Retrieval of the word formation products

Starting point for the retrieval of word formation products was the extraction of simple words for each part of speech. In order to retrieve compounds and derivatives containing simple nouns, these nouns had to be extracted beforehand. The aim was to create a database table in which every simple word is associated with its compounds and derivatives, so that *Fischmarkt* (*fish market*), *Flussfisch* (*river fish*), *fischen* (*to fish*), and *fischig* (*fishy*) are all related to *Fisch* (*fish*).

For nominal simplexes, a simplified definition of simplex was adopted, according to which all nouns are regarded as simple nouns if they cannot be segmented. Therefore, conversions such as *Tanz* (*dance* [noun]) or *Schlaf* (*sleep* [noun]) are simple nouns as well. In this paper, the retrieval of nominal simplexes and their compounds and derivatives is explained through the use of examples. The extraction of simple adjectives and verbs with their corresponding word formation products was done by analogy.

Simple words are characterised by the fact that they do not have a complement, i.e. this column is empty in the base table. Furthermore, the head equals the whole word and therefore the entries in the columns lemma and head in the base table are identical. The head POS-tag of simple nouns is +NN, and their word formation rule is always <+NN>→lemma. That is, the arrow is not followed by an angle bracket because that would indicate a derived or compounded word. The base table was queried for these conditions, and every line that fulfilled the conditions was extracted. The query resulted in about 5,600 simple nouns which were saved in a separate table, the simplex table.

For the following extraction of noun compounds that contain these simple nouns, two cases had to be considered. The simplex is either the first constituent or the second constituent. In the first case, the simple nouns from the simplex table are in the complement-column of the base table, as in the word *Fischmarkt* pertaining to the simplex *Fisch*. In the second case, the simple nouns from the simplex table are in the head column of the base table, as in *Flussfisch* to *Fisch* (cf. fig. 2).

To retrieve the compounds from the base table, a query was run in which the complement-column of the base table was compared to the lemma-column in the simplex table, and the head-column of the base table was compared to the lemma in the simplex table. Identity of the columns resulted in compounds with the simple nouns either as complement or as head. In the case of head compounds, it further had to be ensured that the entries in the columns head and lemma of the base table were not identical, because this would have resulted in simple nouns as in the query mentioned above. Additionally, the word formation rule had to be <+NN>→<NN> <+NN>.

The queries led to roughly 88,000 noun-noun compounds with one of the simple nouns as complement and about 106,000 compounds with the respective simplexes as head. Again, the results were stored in separate tables.

simplex table

lemma
Aal
…
Brot
…
Ei
…
Fisch
…
Haus
…

base table

lemma	compl	head	rule
Aal		Aal	<+NN>→Aal
atmen		atmen	<+V>→atmen
Beutefisch	Beute	Fisch	<+NN>→<NN> <+NN>
bremsen		bremsen	<+V>→bremsen
Brotfisch	Brot	Fisch	<+NN>→<NN> <+NN>
dunkel		dunkel	<+ADJ>→dunkel
Fischfabrik	Fisch	Fabrik	<+NN>→<NN> <+NN>
Fischmarkt	Fisch	Markt	<+NN>→<NN> <+NN>
Floß		Floß	<+NN>→Floß
Flussfisch	Fluss	Fisch	<+NN>→<NN> <+NN>

Figure 2: Comparison between the simplex table and the base table for retrieving the compounds relating to *Fisch*

Noun-adjective compounds (*fischreich*), adjective-noun compounds (*Trockenfisch*; 'dry fish'), verb-noun compounds (*Backfisch*; 'fried fish in batter') and the various derivatives were extracted following the same pattern. Each time, the simplex table and the base table were compared in terms of whether the simple noun was head or complement, and whether a specific word formation rule applied. Word formation products with simple adjectives or verbs were obtained likewise.

Furthermore, the corpus frequency for every product was retrieved and implemented as a column in the separate tables. All compounds and derivatives were labelled with a letter combination denoting the type of word formation and the role the simplex plays in that word formation. The nominal compound *Fischmarkt* is assigned the label 'compound-c-nn' when associated with the simplex *Fisch* (*Fisch* is the **c**omplement) but 'compound-h-nn' when associated with the simplex *Markt* (*Markt* is the **h**ead). The adjective *fischig* derived from the noun *Fisch* and is therefore labelled 'deriv-c-nadj'.

Finally, data from the various tables containing the different word formation products were put together and inserted into an overall table (cf. table 3). In this table, every simple noun, adjective and verb is associated with the compounds and derivatives in which the simple lemma appears. In this way, the aim of establishing a table containing all of *elexiko*'s word formation products relating to a particular simple headword was achieved.

lemma	product	product_frequency	word formation role
Fisch	Anglerfisch	6	compound-h-nn
Fisch	Aquarienfisch	30	compound-h-nn
Fisch	Beutefisch	23	compound-h-nn
Fisch	…	…	…
Fisch	Fischabfall	32	compound-c-nn
Fisch	Fischadler	93	compound-c-nn
Fisch	…	…	…
Fisch	fischen	3670	deriv-c-nv
Fisch	fischig	49	deriv-c-nadj
Fisch	Fischindustrie	68	compound-c-nn
Fisch	…	…	…

Table 3: An extract from the overall word formations' table exhibiting the word formations associated with *Fisch*

4. Presenting the automatically extracted word formation products online

An *elexiko* article is divided into two parts: sense-independent information and sense-related information. The sense-independent information concerns the lexeme itself and focuses on information applying to the entire entry (Storjohann, 2005:62), i.e. details on spelling, spelling variation, and syllabification, among other things. Sense-related information provides information on a specific sense of the headword.

Due to the fact that automatic retrieval does not allow the association of the various word formation products with the specific sense of their base (e.g. the simple noun *Zug* has about nine senses), the compounds and derivatives have to be arranged under the sense-independent information. A link was placed reading 'Wortbildungsprodukte (automatisch ermittelt) weiter »' ('word formation products (automatically retrieved) more »') on the initial page of a headword under the sense-independent information (cf. fig. 3).

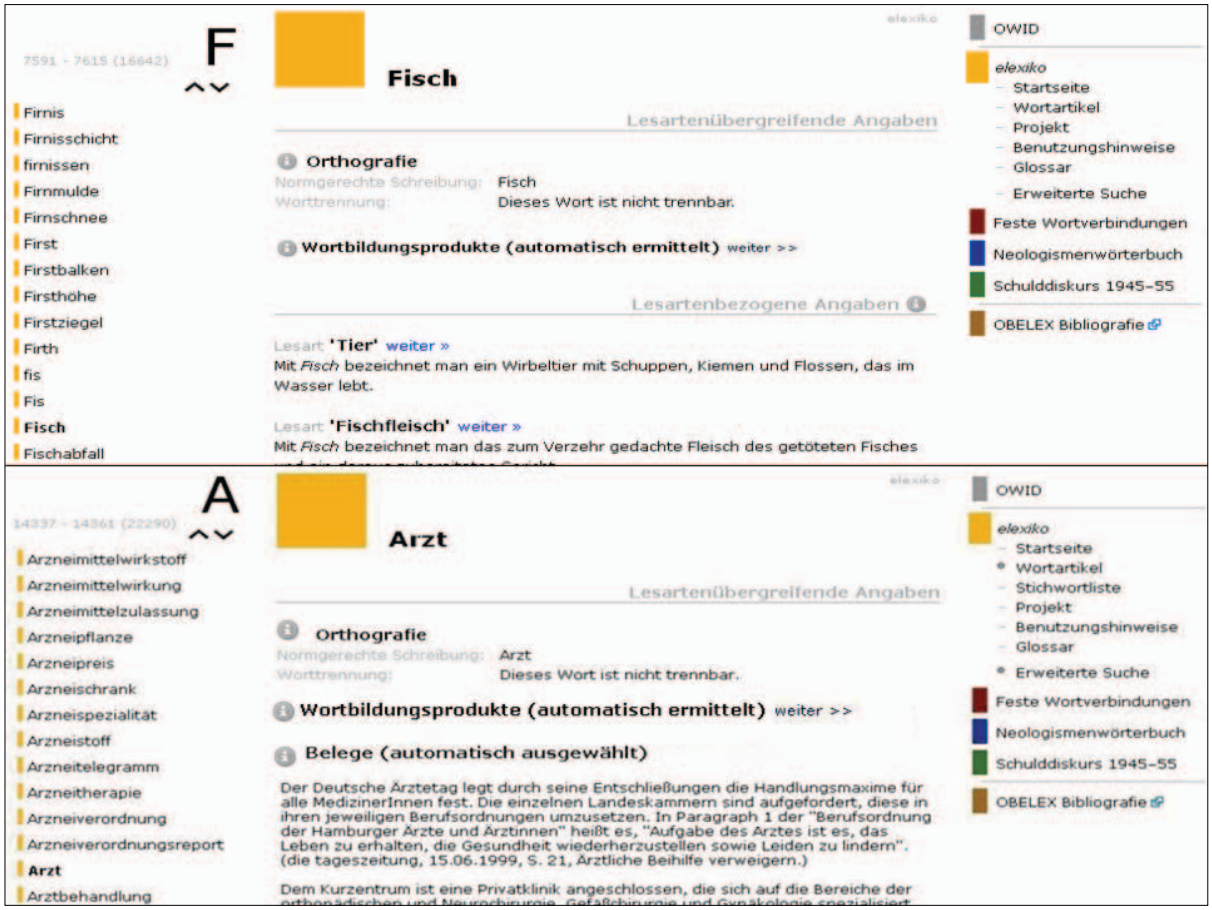

Figure 3: Start pages of the edited entry *Fisch* (above) and of the unedited entry *Arzt* (*doctor*; below) with the link to the automatically retrieved word formation products

If the user clicks on 'more', a new view opens where the automatically extracted compounds and derivatives are presented in a tabbed pane with one tab for compounds, one for derivatives and one for further word formation products. Inside each tab, the word formation products are grouped into the resulting word class or according to the words they consist of. For simple nouns, the compounds are sorted into the categories noun-noun compounds, noun-adjective compounds, adjective-noun compounds and verb-noun compounds (cf. fig. 4). The noun-noun compounds are presented in two columns. In the left-hand column, the compounds with the simple noun as complement are listed, and in the right-hand column are the compounds with the simple noun as head.

In the derivations tab, the various derivatives are grouped according to their parts of speech, as illustrated in fig. 5. The nouns *Kinderei*, *Kindheit*, *Kindschaft* as nominal derivatives to the simple noun *Kind* (*child*) are classified under nouns. As with compounds, they are divided into prefixed and suffixed words. The same applies for adjectives deriving from adjectives (*ungut* – 'not good' from *gut* – *good*; *dümmlich* – 'slightly silly' from *dumm* – *silly*).

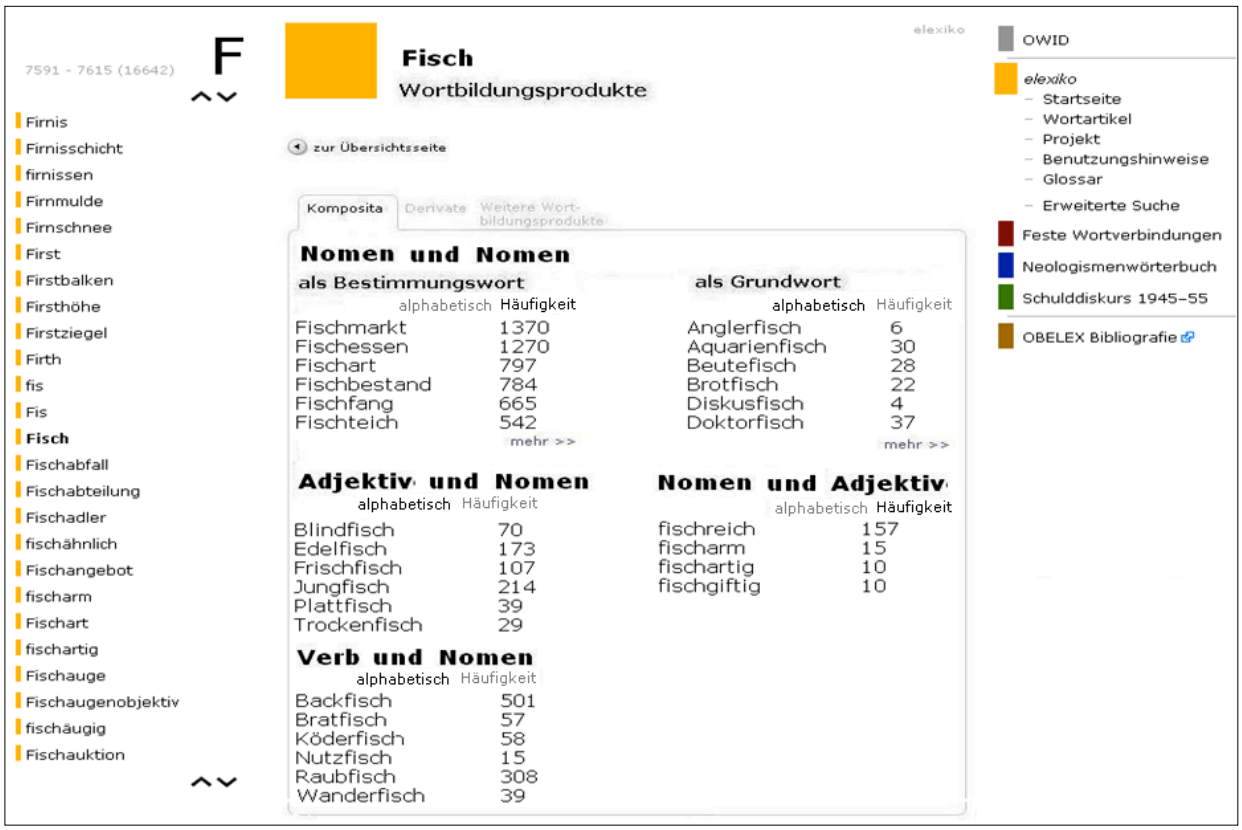

Figure 4: View of the tabbed pane for word formation products relating to the lemma *Fisch* with the tab 'compounds' opened

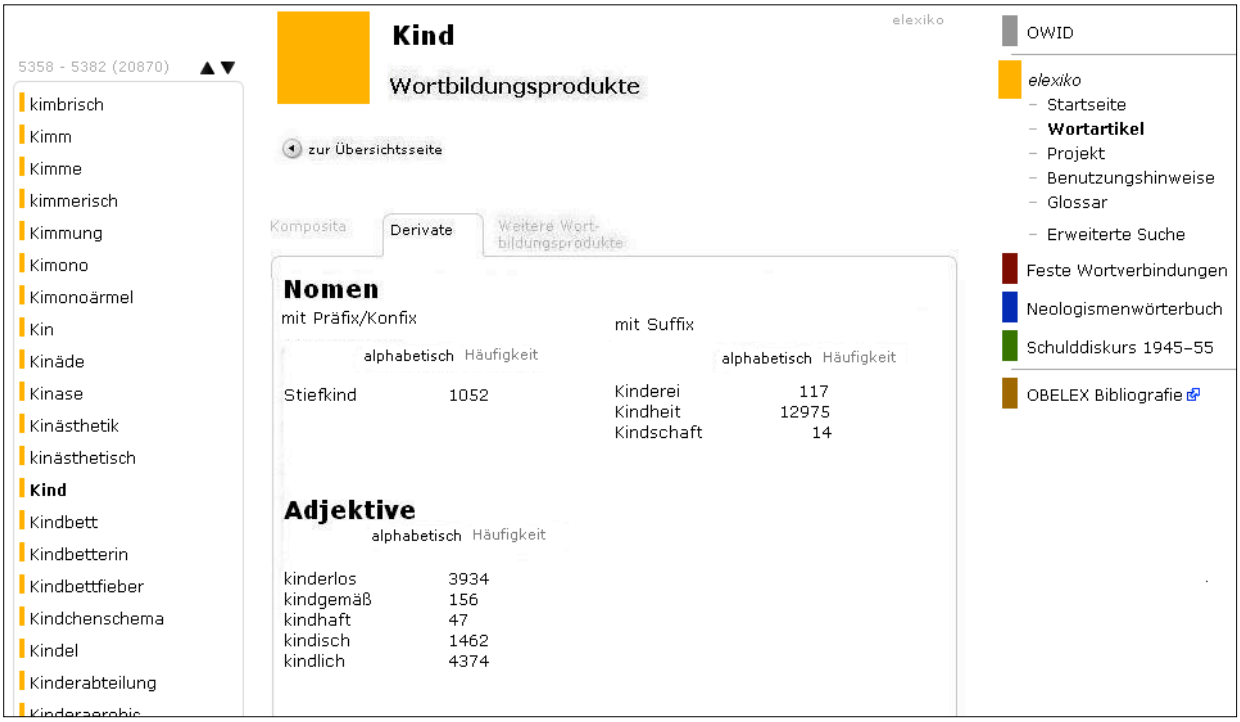

Figure 5: View of the tabbed pane for word formation products relating to the lemma *Kind* with the tab 'derivatives' opened

The frequency of every word formation product in the *elexiko*-corpus is given. The compounds and derivatives can then be ordered either alphabetically or by frequency. The default presentation is alphabetical order because a usage study indicated a preference for this ordering (cf. Klosa/Koplenig/Töpel, forthcoming).

It must be pointed out that the listed compounds and derivatives are not 'nested' under the entry of the simple word. They do have their own unique address to which each of the word formation products presented in the simplex entry are linked and where further information is given. In the presentation chosen for *elexiko* and presented here, the hyperlinked word formation products are grouped according to morphological criteria. The compounds and derivatives are not nested, but, according to Engelberg & Lemnitzer (2009:150), are groups of reference lemmas organized by word formation.

5. Problems and Advantages of automatically retrieved word formation products

5.1 Problems

In order to get an idea of the percentage of errors, a sample of 88 simple nouns was taken and the number of their relating compounds counted. The simple nouns, ranging from very low to very high corpus frequency, led to 6,652 noun-noun compounds containing one of the simple nouns as complement or as head. On average a simple noun is either complement or head in 76 noun-noun compounds. About 792 of the retrieved 6,652 compounds were erroneous, i.e. almost 12% in total or 9 falsely analysed compounds per simple noun.

It became clear that several problems arise from false or missing morphological analyses. Particularly foreign words such as *Toxoplasmose* (*toxoplasmosis*) or very complex words such as *Staatssekretärsrunde* ('meeting of the secretaries of state') were not analysed by Morphisto and therefore not split into complement and head and assigned a word formation rule. Structurally ambiguous words led to false segmentations. A word like *Konzertsaal* (*concert hall*) is a compound consisting of the complement *Konzert* (*concert*) and the head *Saal* (*hall*). However, the word was decomposed into *Konzert*, linking element *-s* and head *Aal* (*eel*); from a structural point of view a possible analysis.

Wrong or even missing segmentations will lead to false results in the retrieval of word formation products and then later to problems in the presentation as well. If *Konzertsaal* is split into *Konzert*, *-s*, and *Aal*, it cannot be retrieved as a compound to *Saal* and therefore it will not be listed under the noun-noun compounds of the simplex *Saal*. Instead, it will falsely occur under *Aal*. In order to prevent false compounds and derivatives being presented in the entry of the simple headword, several corrections had to be performed. As far as possible, groups of words containing similar errors were extracted and their segmentations and word formation rules corrected manually over a period of 14 months by two student assistants. Only a small amount of words could be corrected by executing an overall update statement on the base table.

Almost 15% of 2,762 noun-adjective compounds were analysed wrongly, mainly due to false segmentations. After correction, the number of noun-adjective compounds was reduced to almost 2,550. From 5,866 adjective-noun compounds, 19% were decomposed into false constituents or provided with a false word formation rule. False segmentations mostly occurred with prefixes such as *haupt-* (*main*) or *vorder-* (*front*), incorrectly analysed as adjectives. Another source of errors was homographic constituents. After correction, approximately 4,700 adjective-noun compounds remained. 22% of 8,594 compounds of a verb and noun exhibited incorrect segmentations. Homographs were responsible for most of the errors here as well. Fortunately, almost 1,000 of those could have been corrected automatically. However, at the time of writing approximately 7,600 remained to be checked manually. With 38 false segmentations per 100 words, the 1,262 simple adjectives were the most erroneous class of words. The main source of errors was complex adjectives being falsely analysed as simple adjectives. The number of simple adjectives was reduced to 940 after being corrected and excluding participles. Still, a certain amount of errors remains.

A further problem concerns sense disambiguation. Due to the fact that Morphisto merely decomposes compounds and derivatives into their constituents regardless of their respective senses, it is not possible to relate compounds and derivatives to a specific sense of their base word. The highly polysemous noun *Zug* (*drag*, *draught*, *draw*, *flue*, *move*, *puff*, *stroke* [sports], *platoon*, *train*) appears in compounds such as *Zugbrücke* (*draw bridge*), *Schwimmzug* (*swim stroke*) or *Zugabteil* (*train compartment*). However, the compounds will all be listed under *Zug*, leaving the dictionary user with the task of relating them to the different senses.

5.2 Advantages

Despite the difficulties outlined above, the advantages of an automatic morphological analysis and an automatic retrieval of word formation products are clear. The tool Morphisto made it possible to have a list of 300,000 headwords morphologically analysed, a task that could not have been done manually in a reasonable amount of time, considering that it took two student assistants 14 months to correct only about 2,000 words by hand.

Storing all word formation products in a relational database table as explicated above (cf section 3.2) proved suitable for the extraction of the various

compounds and derivatives. With a single query it is possible to retrieve all word formation products relating to a particular headword. Even lemmas that are subject to umlaut during word formation are associated with their respective bases, e.g. the adjective *ärztlich* (*medical*) to its base noun *Arzt* (*doctor*) from which it derives. A simple string-based query would not have found the umlauted case.

Presenting the compounds and derivatives that are associated with a particular simple headword in the entry of the simple word puts the simplex at the centre of a net of morphologically related words, illustrating the simple word's word formation activity. Corpus frequencies show how often the simple word forms the base of particular compounds and derivatives. In the case of compounds frequency information additionally offers an insight into the discourses in which the simple word mainly appears.

Links to the compounds' and derivatives' own entries provide the dictionary user with elaborate information on meaning, typical usage, paradigmatic relations and grammar (in the case of edited entries). On the one hand, the linkage serves to elucidate the morphological relations between the base word and its word formation products. On the other hand, the linkage ensures that the various word formation products presented in the entry of the base word are not left without explanation. Dictionary users, especially learners, are led directly to the answers to their specific questions regarding gender, inflection class and morphological make-up of words.

6. Conclusion

This paper gave an example of how automatic methods can accompany and complement lexicographic work and enhance dictionary writing.

In order to illustrate word formation relations in the online dictionary *elexiko*, each word from the headword list was analysed with a computational morphology. Every word was decomposed into its constituents, and its individual word formation rule given. The results of the analyses were stored in a relational database. The aim was to extract all compounds and derivatives relating to a particular (simple) base word, and to integrate the retrieved word formation products online into the entries of the simple word in question.

In spite of problems resulting from the automatic morphological analysis and automatic retrieval, the presented method and online presentation give the opportunity to interconnect the dictionary data. The display of all compounds and derivatives in which a particular lemma appears not only reflects the lemma's word formation potential, but also reveals relations between the lemma in question and other headwords that do not and cannot emerge in a printed dictionary. An online dictionary such as *elexiko* is able to connect the headwords on the basis of word formation so that a net of morphologically related words spans the dictionary.

7. Looking ahead

This final section aims to look ahead and to consider implications for other languages or suggest research ideas.

Although the proposed model is intended for dictionary users, researchers will find the number and variety of word formation products displayed along with the information on their corpus frequency helpful for studying the productivity of individual constituents. One might ask whether there is a correlation between frequency of base word and number of word formation products in total, or between polysemy of base word and number of word formation products. Further research questions might include to what extent a particular base word appears in compounds or derivatives, and in what word classes these compounds and derivatives are. Additionally, one might ask whether the base word has a propensity to combine with a certain word class, or whether it combines with words of a particular semantic field. The automatically retrieved word formation products offer initial answers to these questions.

The extraction of word formation products depends primarily on the morphological analysis with a computational morphology. The method presented in this paper was based on a computational morphology for German. Given a language-specific computational morphology the procedure outlined here should also work for languages in which compounds and/or derivatives not only exist but can also be decomposed into a binary complement-head-structure, e.g. Turkish *elma suyu* ('apple juice-suffix' *apple juice*). For languages that use different strategies to express determinatives such as *apple juice*, the method presented here cannot simply be adopted as it is. The whole model depends on the computational morphology which has to be language-specific. In order to extract the French *jus de pommes* ('juice of apples') as a kind of word formation related to either *jus* (*juice*) or *pomme* (*apple*), a computational morphology cannot simply deconstruct the phrase into a first constituent and a second constituent, which is not even the head. As a result of the structural difference, the database design might have to vary, and retrieval methods be adjusted.

8. Acknowledgements

This project is financed by the Gottfried Wilhelm Leibniz Scientific Community ('Joint Initiative for Research and Innovation').

9. References

Barz, I. (2001). Wortbildungsbeziehungen im einsprachigen Bedeutungswörterbuch. In J. Korhonen

(ed.) *Von der mono- zur bilingualen Lexikografie für das Deutsche*. Frankfurt/M.: Peter Lang, pp. 85-100.

Engelberg, S., Lemnitzer, L. (2009). Lexikographie und Wörterbuchbenutzung. 4th revised and enlarged edition. Tübingen: Stauffenburg.

Klosa, A., Koplenig, A. & Töpel, A. (forthcoming). Benutzerwünsche und Benutzermeinungen zu dem monolingualen deutschen Onlinewörterbuch *elexiko*. In C. Müller-Spitzer (ed.) *Using Online Dictionaries*.

Klosa, A. (2010). On the combination of automated information and lexicographically interpreted information in two German online dictionaries. In S. Granger, M. Paquot (ed.) *eLexicography in the 21st century. New challenges, new applications. Proceedings of eLex 2009*. Louvain-la-Neuve, pp. 157-163.

Mugdan, J. (1984). Grammatik im Wörterbuch: Wortbildung. In H.E. Wiegand (ed.) *Studien zur neuhochdeutschen Lexikographie IV*. (Germanistische Linguistik 1-3/83). Hildesheim, New York: Georg Olms. pp. 237-308.

Schmid, H., Fitschen, A. & Heid, U. (2004). SMOR: A German Computational Morphology Covering Derivation, Composition, and Inflection. *Proceedings of the IVth International Conference on Language Resources and Evaluation (LREC 2004)*, Lisbon, pp. 1263-1266.

Storjohann, P. (2005). *elexiko*: A Corpus-Based Monolingual German Dictionary. *Journal of Linguistics* 34, pp. 55-82.

Ulsamer, S. (forthcoming). Wortbildung in Wörterbüchern – Zwischen Anspruch und Wirklichkeit. In A. Klosa (ed.) *Wortbildung im elektronischen Wörterbuch*. Tübingen: Narr.

Williams, E. (1981). On the Notions 'Lexically Related' and 'Head of a Word'. *Linguistic Inquiry* 12(2), pp. 245-274.

Zielinski, A., Simon, C. & Wittl, T. (2009). Morphisto: Service-oriented open source morphology for German. In C. Mahlow, M. Piotrowski (ed.) *State of the Art in Computational Morphology Workshop on Systems and Frameworks for Computational Morphology SFCM 2009, Zurich, Switzerland, September 4, 2009. Proceedings* Berlin, Heidelberg: Springer, pp. 64-75.

BZV*elexiko* – Benutzeradaptive Zugänge und Vernetzungen in *elexiko*. Accessed at: www.ids-mannheim.de/lexik/BZVelexiko/.

elexiko (2003 ff.). In Institut für Deutsche Sprache (ed.) OWID – Online-Wortschatz-Informationssystem Deutsch. Mannheim. Accessed at: www.elexiko.de.

OWID – Online-Wortschatz-Informationssystem Deutsch. Accessed at: www.owid.de.

Morphisto. Accessed at: www.ids-mannheim.de/lexik/TextGrid/morphisto.html.

TextGrid. Accessed at: www.textgrid.de.