Cognitive Features in a Corpus-based Dictionary of Commonly Confused Words

Petra Storjohann

Institut für Deutsche Sprache, Mannheim (Germany)
E-mail: storjohann@ids-mannheim.de

Abstract

This paper discusses how cognitive aspects can be incorporated into lexicographic meaning descriptions based on corpus-driven analysis. The new German Online dictionary "Paronyme – Dynamisch im Kontrast" is concerned with easily confused words such as *effektiv/ effizient*, *sensibel/ sensitiv*. It is currently in the process of being developed and it aims at adopting a more conceptual and encyclopaedic approach to meaning. Contrastive entries emphasise usage, comparing conceptual categories and indicating the mapping of knowledge. Adaptable access to lexicographic details offers different perspectives on information, and authentic examples reflect prototypical structures.

Some of the cognitive features are demonstrated with the help of examples. Firstly, I will outline how patterns of usage imply conceptual categories as central ideas instead of sufficiently logical criteria of semantic distinction. In this way, linguistic findings correlate better with how users conceptualise language. Secondly, it is pointed out how collocates are family members and fillers in contexts. Thirdly, I will demonstrate how contextual structure and function are included by summarising referential information. Details are drawn from corpus data; they are usage-based patterns illustrating conversational interaction and semantic negotiation in contemporary public discourse. Finally, I will show flexible consultation routines where the focus on structural knowledge changes.

Keywords: cognitive lexicography; corpus semantics; paronyms, easily confused words;

encyclopaedic-conceptual approach

1. Introduction

Lexicography has undergone dramatic changes over the past two decades. These mainly concern approaches to lexical analysis, the editorial process and the digitation/presentation of data. The relationship between semantic theory and practical lexicography has always been a difficult one (cf. Rundell, 2012). When it comes to employing semantic foundations, it is above all the field of corpus linguistics that has made its mark on dictionary writing. Corpora and their tools have turned lexicography into a more objective and empirical trade which makes use of authentic language data. Lexicographers have also continually taken advantage of hypertextual opportunities to present lexical information in innovative ways, although their full potential has not been exploited, nor have users' needs been extensively considered (Müller-Spitzer, 2014).

Cognitive linguistics, however, has had no major impact on general dictionaries. In particular, the structuring of entries and the definition of senses are two areas where cognitive principles could be used to implement descriptions of conceptual structures and to show how meaning is construed or represented. As Ostermann (2015) points out, novel cognitive theories have been neither recognised nor successfully integrated into general English dictionaries. A few specialised frame-based English systems such as Pattern Dictionary of English Verbs (PDEV) or FrameNet[1] facilitate their meaning descriptions with cognitive foundations (e.g. Fillmore, 1976; 1977). As far as general monolingual dictionaries are concerned, both in English and in German, there is a complete lack of guiding cognitive paradigms (e.g. conceptualisation, construction, categorisation, representation) being incorporated into semantic descriptions with a theoretical foundation.

In this paper, it is argued that cognitive ideas can be successfully implemented in descriptions of meaning and the structuring of entries, and that these provide relevant information which primarily benefits users. In the following, the new German dictionary of commonly confused words "Paronyme-Dynamisch im Kontrast" (Storjohann, 2016) is taken as an example that breaks with tradition by including central conceptual information and by representing both linguistic and encyclopaedic knowledge. Within the German context, it is a first attempt at a more cognitively infused lexicography calling for more realistic documentations of language and the way speakers perceive, conceptualise and linguistically represent the world. For the purpose of illustration, some cognitive features will be demonstrated, particularly those emphasising the interaction of details for more adequate depictions of flexible usage and contextual categorical implications.

2. Where to look for information on commonly confused words

Paronyms are easily confused words which regularly cause problems for both native speakers and language learners. As these lexical items often share morphological roots, they are similar with respect to sound, spelling and/or meaning, e.g. *effektiv/effizient (effective/efficient) sensibel/sensitiv (sensitive/delicate)*, *formell/formal/förmlich (formal/official), Method/Methodologie/Methodik (method/methodology)*, *Elektrik/Elektronik (electrics/electronics)*.[2] Generally, such pairs/sets are not regarded as synonymous (cf. Làzàrescu, 1995; 1999) although corpus analyses suggest that some items undergo meaning change due to the rivalry between the words. Sometimes, they can develop synonymous notions and simply become lexical alternatives (cf. Storjohann, 2015). In other cases, they remain similar in meaning but show subtle differences and restrictions in usage. Inevitably, situations of confusion arise when

[1] A related project in German is the German Frame-Semantic Online Lexicon GFOL (http://coerll.utexas.edu/frames/).

[2] For more examples see Schnörch (2015).

speakers' intuitions contradict information in existing reference works.

The importance of paronyms is based on the assumption that these items play a vital role for users in the process of second language acquisition and foreign language communication in order to avoid misunderstandings. Confusing paronyms is sometimes regarded as a violation of semantic correctness. Prescriptive analysts favour semantic correction and the avoidance of such mishaps (Bolshakov & Gelbukh, 2003). Indeed, the alleged misuse of morphologically and semantically similar words also leads to linguistic uncertainties for native speakers, as numerous language-related Internet blogs show. However, corpus-guided investigations of paronyms partly reveal recent semantic changes, conventionalised overlappings and newly established contexts. Therefore, empirically sound, descriptive documentation is necessary to capture the current use of paronyms. Corpus-assisted investigations of easily confused words and their usage over recent decades can provide valuable insight into principles of semantic shift. It is argued here that such analyses might enable semanticists to integrate the phenomenon into a wider theoretical framework on the one hand and into appropriate lexicographic descriptions on the other hand.

As most general German reference guides still favour a traditional style and structure, recent or new phenomena are hardly captured nor adequately described. Taking a closer look at resources such as Duden online, their lexicographic deficiencies become apparent. Users interested in the differences between *Elektrik/Elektronik* or *effektiv/effizient* find the following facts:

Elektrik	*Elektronik*
Gesamtheit einer elektrischen Ausstattung	Gesamtheit einer elektronischen Anlage oder Ausstattung
effektiv	*effizient*
wirksam, wirkungsvoll lohnend, nutzbringend sich tatsächlich feststellen lassend, wirklich	wirksam und wirtschaftlich

Table 1: Definitions taken from Duden online

The entries of *Elektrik/Elektronik* are circular and "married with content from antiquated dictionaries – the type that define pedantic as 'of, pertaining to, or characteristic of a pedant'" (Rundell, 2012: 74). The entries of *effektiv/effizient* mainly summarise synonyms. Users do not obtain sufficient details concerning their conceptual potential and contextual usage. Quite likely, users will miss information, for example, on semantic reference, relevant conceptual domains or categories, discourse structures and contextual situations. Who/what is specifically characterised

as *effektiv/effizient* and in what kind of contextual circumstances? This question remains open. Similarly, German Wiktionary describes the meaning of *effektiv* as follows:

[1] die Fähigkeit besitzend, eine Aufgabe erfolgreich zu erledigen,
(to have the ability to complete a task successfully),
[2] *ohne Steigerung:* sich tatsächlich feststellen lassend, wirklich,
(*without comparison*: in fact, real).

The adjective *effizient* is described as 'to be able to be productive relative to the invested effort':

[1] fähig, viel Leistung in Relation zum Aufwand zu erbringen.

Again, conceptual details, preferred discourse situations and further encyclopaedic knowledge are not documented. Also, both descriptions suggest a small semantic spectrum for both adjectives.

Today, speakers face a range of consultation options, from traditional print dictionaries to free online resources. As most German e-dictionaries are copied or digitised versions of conventional reference books, unfortunately these often do not offer satisfactory answers to questions about paronym behaviour.

> bezüglich *sensitiv/sensibel*: „Ich hab zwar überall nach einer Definition dieser beiden Wörter gesucht, aber je mehr ich finde, desto irritierender ist es" (aus: http://depriforum.phpbb8.de/diskussionen-f16/sensitiv-sensibel-t1258.html).[3]

Consequently, online forums have turned into widely used social media sources where users consult the community for their linguistic problems (see Figure 1[4]).

Figure 1: Typical Language Question in Internet

[3] Translation: I've looked for a definition of these two words everywhere, but the more I find, the more irritating it becomes.

[4] Example taken from:
http://www.gutefrage.net/frage/was-ist-der-genaue-unterschied-zwischen-effektiv-und-effizient.

In some cases, they explain whole contextual situations in which their uncertainties occur. They seek information on lexical use, prototypical contexts, possible constructions, and conceptual as well as encyclopaedic issues. The answers from the language community are impressively diverse and revealing. As a matter of fact, speakers have good intuitions as to what linguistic and extra-linguistic information is required to form essential parts of authentic communication. In online forums, people share their concerns about easily confused words. It is here, through the study of blogs, that detailed insights into the specific linguistic problems of users, their consultation behaviour and their needs, can be gained. However, it is also here where we see that users do not always obtain satisfactory answers (see Figure 2).

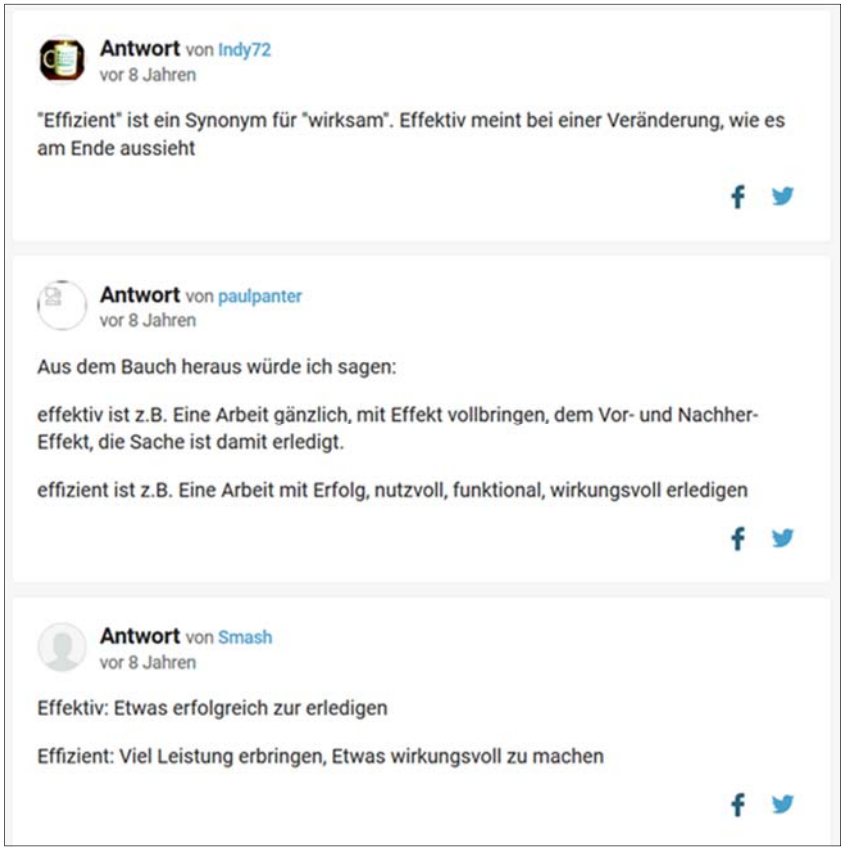

Figure 2: Exemplary Answers in Internet Forum[5]

Undoubtedly, Internet forums are not a reliable source of information. Consultations can be helpful but they are not guaranteed sources of reliable information.

[5] Translation answer 1 “efficient” is a synonym for “productive/effective”. Effective refers to a change of state and what it looks like in the end. Translation answer 2: Intuitively, I would say: effective is, for example, to finish some work with effect, the pre and after effect, the job is then done. Efficient is, for example, to do a work in a useful, functional, effective way. Translation answer 3: effective: to so something successfully. efficient: to work productively, to do something effectively.

3. Dictionaries and the Cognitive Perspective

The subject of paronymy has not been revisited with empirical, data-driven methods either in terms of semantic theory or practical lexicography. Lexicographically, some German paronyms have been documented in printed dictionaries (Müller, 1973; Pollmann & Wolk, 2010), although not systematically. However, there is no corpus-guided reference guide empirically describing paronym sets enabling readers to find the correct usage of such lexical items.

Placing the user in focus, it is essential to strive for conceptual approaches and to document the interplay of lexical, structural and encyclopaedic knowledge in meaning descriptions. While analysing the needs and various interests of users we have come across two prerequisites. On the one hand, it is necessary to implement a semantic structure and network that is closer to actual usage and this requires information on patterns of conceptualisation, on categories, reference and concrete lexical prototypes. For quite some time, there are endeavours to reconcile the branch of lexicography with cognitive semantic theories. As Geeraerts (2007: 1168) has pointed out:

> […] what Cognitive Linguistics seems to offer to lexicography is a conception of semantic structure that is perhaps in a number of respects more realistic than what many other semantic theories (in particular, theories of a structuralist persuasion) can provide.

On the other hand, we need to overcome a rigid, linear ordering of information and strive for a realistic representation of multi-dimensional facets of semantic configurations in language use to be closer to the structure of the mental lexicon (cf. Ostermann, 2015).

3.1 The New German e-Dictionary "Paronyme – Dynamisch im Kontrast"

"Paronyme-Dynamisch im Kontrast" is an electronic dictionary that breaks new ground by adopting a more conceptual and encyclopaedic approach to meaning by incorporating cognitive features. It will be published in the dictionary portal OWID (Online-Wortschatz-Informationssystem Deutsch, www.owid.de) in 2017. It is currently in the process of being developed and includes conceptual, prototypical, and referential categorisation and a flexible structural access to knowledge. This dictionary does not follow sufficiently logical criteria of semantic distinction for its sense disambiguation. Instead, different patterns of usage and their underlying conceptual categories and prototypical realisations function as parameters of contextual distinction. These are then accessed flexibly via menu navigation. As a quick guide, short paraphrases define characteristics of conceptual referential categories. Concerning the adjectival pair *effektiv/effizient*, relevant topic areas (or frame

presentations) are given for each adjective. These are coded as "guide words"[6] together with a synonym (see Figure 3).

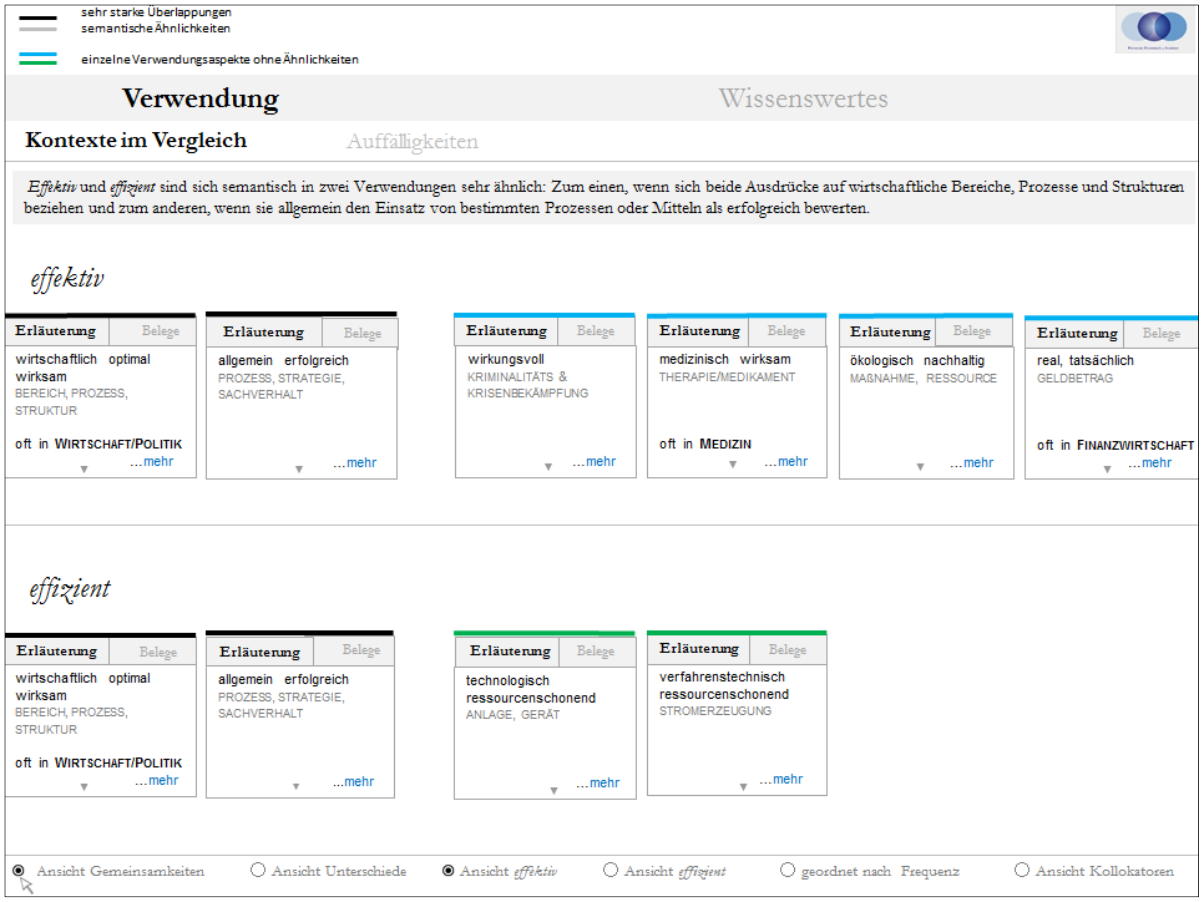

Figure 3: Default Conceptual Navigation Structure

A large amount of knowledge about words, meanings and concepts is derived from experience and from the categories we construct, i.e. mentally represented frames or schemas. It is these categories (e.g. AREA/PROCESS, STRUCTURE, PROCESS/STRATEGY/STATE OF AFFAIRS, CRIME/CRISIS, MEDICINE, MEASURE/RESOURCE, MONEY, TECHNOLOGICAL DEVICE, ENERGY) that justify a distinction of patterns and help to correlate situations of language use to different contexts. Similar ideas of how to use guide words to exemplify contextual frames in which the words are prototypically embedded can be found in Ostermann (2015). In the dictionary, these categories build up a quick contrastive guide and a concept-driven navigation structure (see Figure 1). They are also able to activate corresponding concepts of polysemous words. These also help a user to encode contexts and to identify metonymic and metaphoric mappings (cf. Fillmore & Atkins). Users can more easily relate the adjectives to their meanings and relate these then to the preferred contextual reference (here nouns), e.g.:

[6] Guide words are also used in Cambridge International Dictionary of English.

- *Effektiv* means 'economically optimal' with respect to an AREA, PROCESSES or STRUCTURE and it often occurs in economy or politics.
- *Effektiv* means 'generally successful' with regard to a PROCESS, a STRATEGY or a SITUATION/SPECIFIC MATTERS.
- *Effektiv* means 'working' in terms of FIGHTING CRIME/CRISES.
- *Effektiv* means 'working' in terms of MEDICATION or a THERAPY and it often occurs in medical contexts or contexts describing health issues.
- *Effektiv* means 'ecologically sustainable' in terms of MEASURES or RESOURCES.
- *Effektiv* means 'real' with regard to AMOUNTS OF MONEY and it is often used in contexts describing financial issues.

Compared to traditional dictionaries (see table 1) much more information is provided which can be consulted and then mentally stored together.

Through the more visual explanations, it is possible to answer questions such as *Can German effektiv be used synonymously with effizient in contexts of business to characterise economic methods or structures? Can a motor be described as effektiv or effizient?* or *Is a powerful production of electricity better referred to as being effektiv or effizient? Do I use effizient or effektiv when I want to say that a medical treatment is working well?* With the help of the given synonyms and guidewords in the short paraphrase it is also possible to compare individual contexts of the two paronyms and quickly identify similarities and differences.

3.2 Contextual Fillers as Prototypical Lexical Realisations

Users also have the option of consulting more detailed information on demand. Conceptual reference and encyclopaedic ideas are then explicitly integrated into the longer paraphrase. The relevant ontological category or domain is then specifically illustrated using lexical preferences, i.e. collocates. With a dynamic electronic display at hand, these are shown optionally, as a list of frequent and conventionalised contextual partners, introduced by *such as* underneath the definition (see Figure 4). In this approach, collocates are concrete lexical realisations (or fillers[7]) in specific contexts illustrating the referential category given in the definition.[8]

[7] For verbs, which only make up only a small section of the dictionary, collocates serve as fillers in frame-like constructions. Collocates are then grouped into different sets (argument roles).

[8] The linguistic analysis of corpus-driven collocates is also indicative evidence of distinct usage and senses. They are a primary source for lexicographers for deriving definitions and disambiguating meaning.

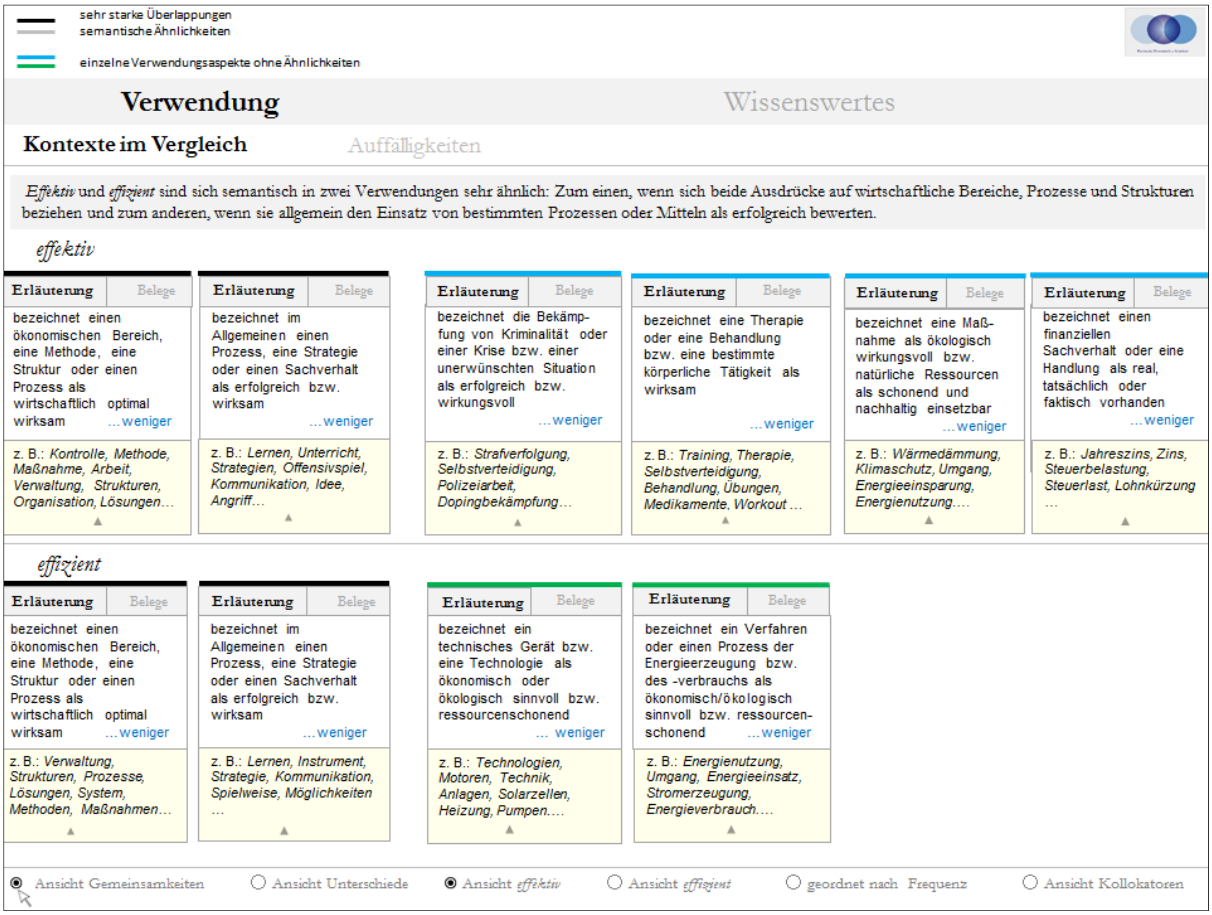

Figure 4: Long Definition and Prototypical Realisation (Fillers)

For example, polysemous *effektiv* prototypically means something like 'economically effective' or 'efficient'. It is the conceptual background where the adjective refers to nouns functioning as non-human subjects or objects and denoting ECONOMIC AREAS, PROCESSES, STRUCTURES or MATTERS OR AFFAIRS such as *control, method, measure, work, administration, structures, organization* or *solutions*. Similarly, German *effizient* also refers to nouns expressing the concepts of ECONOMIC AREAS, PROCESSES, STRUCTURES or MATTERS OR AFFAIRS illustrated by *administration*, *structures*, *processes*, *solutions*, *system*, *methods* and *measures*. Alternatively, both items can refer to STRATEGIES or PROCESSES as 'generally being successful': for *effektiv* these are typically *learning*, *teaching*, *strategy*, *offense*, *communication*, *idea* and *attacks*.

For *effizient* these are *learning, instrument, strategy, ways of playing, communication* or *possibilities*. In other contexts, they differ in terms of their conceptual referents. For example, *effektiv* can be used to describe the successful fight against crime or crises (as exemplified by *self-defense*, *police work*). It refers to the positive results of a therapy or medicine (illustrated by *training*, *therapy*, *exercise*) and the adjective describes measures and natural resources (demonstrated by *insulation*, *climate protection*, *energy saving*) as successful. An adverbial usage is also attested for *effektiv* with

referenced to money or interests, meaning 'real, actual' (indicated *annual interest*, *tax burden*).

Effizient exhibits two further contexts which refer to technological equipment or instruments such as *motors*, *solar cells*, *heating* and *pumps*. It also occurs in contexts were the adjective describes procedures of generation or consumption of power/energy as ecologically sustainable (illustrated by co-occurring *electricity consumption, electricity supply*, *power generation*).

In essence, the lexical representations are prototypical domain elements and structured mental representations of human experience. They shed light on strong affinities to constructions and contextual preferences, and they point to properties correlating with aspects of meaning structure. With prototypical details, we have the possibility of handling polysemous contexts in a way that "more faithfully reflects what corpus data tells us" (Rundell, 2012: 82). For polysemous items, metonymous and metaphorical contexts are listed. These show cognitive processes in which conceptual elements motivate the configuration of another semantically related conceptual entity (cf. Kövecses & Csabi, 2014).

The lexical representations are not intuition-based examples but statistically significant occurrences provided by corpus instances (see Section 4). They are lexicographically analysed, interpreted and classified manually, once automatically retrieved collocation analyses have provided the necessary access to typical contextual structures. Each paraphrased context is illustrated by up to three citations editorially picked from the corpus. The entry as such is not automatically retrieved, corpus tools pre-analyse complex data sets and provide systematic access to significant patterns. These then undergo editorial scrutiny where corpus findings are essential evidence of cognitive entities and categories. In the entries, corpus lexicography meets cognitive lexicography.

3.3 The Organisation of Knowledge

Hypertext dictionaries can break up conventional sequential ordering of information. A granular XML-architecture allows for different data structures and therefore flexible access routes, adaptable presentations and complex searches. As digital data systems can represent their content in a structure that is not dependent on its presentation, it is possible to generate adaptable displays. Tailor-made user-adaptivity is technologically feasible but will only become a realistic option once we know more about the users. Content can be arranged dynamically, changing linguistic focus to "allow users to recreate and re-represent their own dictionary data" (Fuertes-Olivera, 2013: 330).

By focusing on the needs of the user, we have learned that these vary considerably (cf. Storjohann, 2016). Given this, a system of various options has been developed which

enables us to configure different perspectives on the organisation of knowledge. In essence, this dictionary is an XML-based hypermedia resource. Its system is customisable and can adaptively generate and prioritise information for specific user groups. Apart from regular search options, with multi-functional specifications at hand, dictionary data can be individually "reshuffled" by setting different parameters during the consultation process. Consequently, focal points on conceptual structures change and different linguistic aspects are emphasised.

Firstly, as a default, the different instances of usage of each lexical item are established in relation to the individual contexts of the corresponding paronym item, with identical contexts first, followed by similar and dissimilar contexts. Through this, an instant overview of overlapping uses and differences is provided (Figure 1 or 2). Secondly, depending on personal interest, users can also choose parameters for listing the different contexts first. Thirdly, as neither ordering necessarily corresponds to the frequency of occurrence in actual usage, all contexts can be shown according to their distribution in the corpus, so that the predominance or centrality of certain contexts can be seen. Fourthly, it is the user's decision to choose the ordering of paronym items and determine which one appears at the top of the entry. Finally, the menu options also include a visualisation of collocation profiles with behavioural networks and interactive functions (see Figure 5).

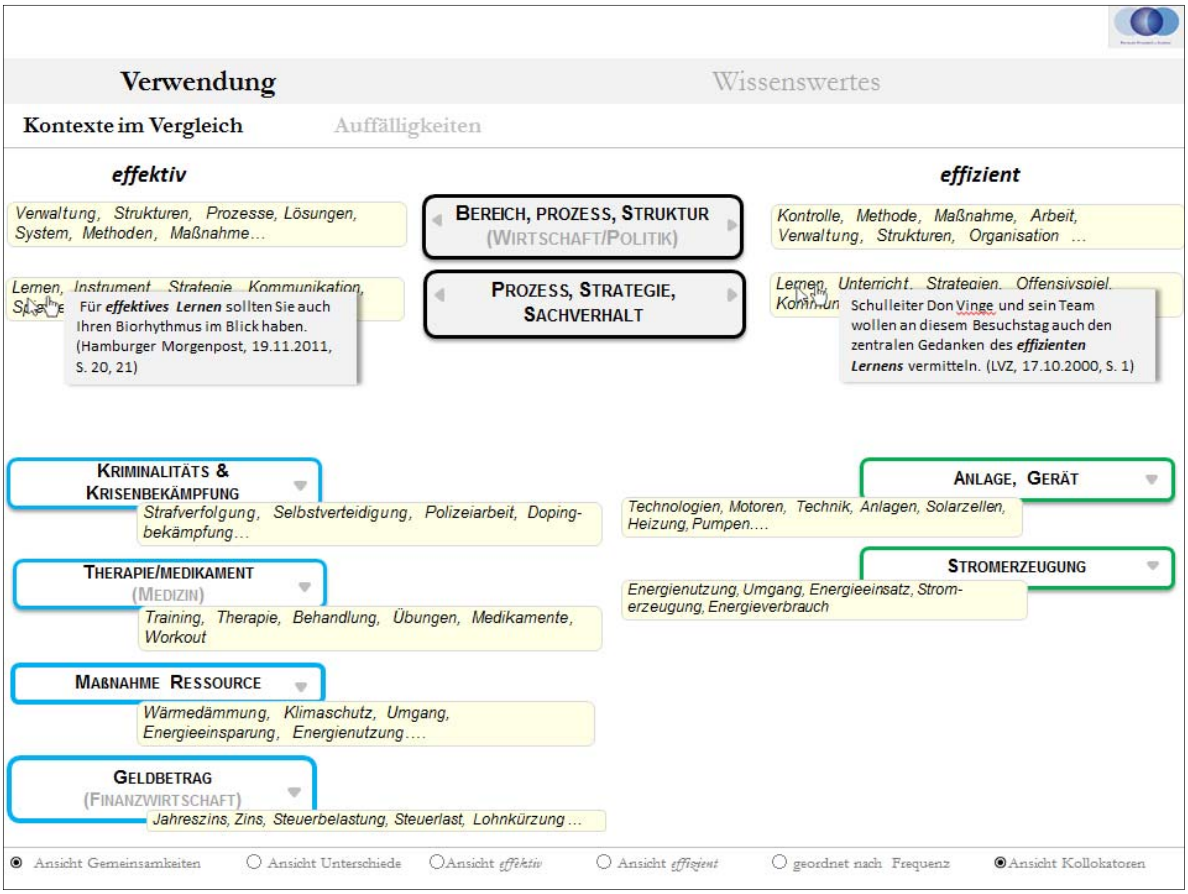

Figure 5: Visualisation of Collocational Profiles and Interactive Functions

Taking the conceptual categories as a starting point, their corresponding collocational representations can be studied contrastively. The denoted concepts that are commonly shared are in the centre, followed by dissimilar concepts arranged separately to the left and right below. Each category, together with its individual lexical realisations, is exemplified by corpus instances. It is a simplified diagram with abstract concepts directly representable in a contrastive conceptual organisation. Overall, this e-dictionary exploits text- and hypertechnological possibilities and offers consultation routines by optionally generating different facets of structural knowledge.

3.4 Corpus-guided dictionaries vs traditional dictionaries

As we can see from Figures 3–5, corpus data strongly suggests that *effektiv* and *effizient* are used synonymously with respect to two contexts. The underlying corpus provides numerous attestations (see Examples 1–3).

(1) Arbeit **effizienter** machen: Mit einem guten Computer-Netz kann jedes Unternehmen **effektiver** arbeiten - gerade, wenn seine Büros auf viele Orte verteilt liegen. (Rhein-Zeitung, 15.03.2002, Das Dekanat soll "Computer-fit" werden.)

(2) Derzeit gilt der Vertrag von Nizza aus dem Jahr 2000. Doch die Strukturen sind nicht mehr **effizient**. Eine Kommission mit 27 Kommissaren kann ebenso wenig **effektiv** arbeiten wie ein Parlament mit fast 800 Abgeordneten. (Braunschweiger Zeitung, 21.06.2007, Fragen und Antworten zum Gipfel.)

(3) Der sture Ablauf, der fast immer eingehalten wird, sei vielmehr das Ergebnis **effektiver** Arbeitsteilung von Spitälern, Bestattern und den Friedhofsbetreibern. Fast alle größeren Bestattungsunternehmen pflegen in Deutschland eine **effiziente** Arbeitsteilung. (Die Zeit, 15.04.2004, Wie man in Deutschland begraben wird)

This entry, demonstrated in Figures 3–5, is a good example to show differences to other existing prescriptive reference books such as Pollmann & Wolk (2010). Its documentation aims at guiding users to the allegedly correct usage and suggests a clear distinction between the items in question (see Figure 6).


effektiv ··· effizient

- Ist etwas gemessen an den Mitteln, die zur Erreichung eines bestimmten Ziels ausgerichtet sind verhältnismäßig, dann ist es effektiv.
- Effizient ist etwas, das besonders wirtschaftlich ist, also die Kosten mit dem Nutzen vergleicht. (Eselsbrücke: -zient / Cent).

Ein Insektenspray vertreibt oder tötet beispielsweise Insekten, somit ist es effektiv. Wesentlich effizienter ist jedoch eine Fliegenklatsche.

Figure 6: Dictionary Entry *effektiv*/*effizient* in Pollmann & Wolk (2010).

Strictly normative language use is also propagated in the German Wiktionary[9], a popular electronic resource which under an explicit headline points out that confusion over the two words *effektiv* and *effizient* should be avoided. Conventional reference guides have so far focussed on the differences between commonly confused words. They entirely fail to explain existing similarities. The usage restrictions that are documented in these reference books cannot be confirmed through corpus data. As is the case for *effektiv/effizient*, strict usage lines cannot be sharply drawn which might have been expected intuitively. The meanings of typically confused words are more freely exposed to semantic negotiation. Following a descriptive empirical view, the semantics of some paronymic lexical items have adopted new semantic aspects and undergone meaning changes that are observable as regular patterns in a corpus and not as single misused occurrences.

Overall, all reference guides mentioned are neither based on semantic examinations of current natural language in use nor on investigations of large data. It is empirical corpus explorations that open up the discrepancies to traditional descriptions. Corpus studies allow for the description of similarities which, on the one hand, might offer a deeper understanding why two words are regularly being confused and, on the other hand, it might indicate ongoing linguistic change worth documenting. Consequently, corpus-driven research on paronymy demands a more differentiated look at the phenomenon than has previously been offered.

4. Corpus Lexicography meets Cognitive Lexicography

The paronym dictionary bases its information on a comprehensive purpose-built corpus comprising 2.3 billion words.[10] The underlying corpus is publicly accessible and provides for transparent lexicographic practices. As the subject of paronymy has not been revisited with empirical, data-driven methods, either in terms of semantic theory or in terms of practical lexicography, suitable corpus methods for contrastive investigation needed to be tested (cf. Storjohann & Schnörch 2014). Currently, complementary software-driven resources facilitating the search for similarity and difference are being exploited, each of which is based on the analysis and interpretation of contextual profiles, collocations and colligations, corresponding semantic roles and syntactic functions.[11] Corpus data reveals how meaning is constantly being negotiated in usage events and how communicative acts can create semantic rivalry or increase vagueness of easily confused words. Accordingly, variation and uncertainties arise from lexical similarity, sometimes leading to the adoption of new conceptual-semantic nuances. It is corpus-guided investigations that uncover discrepancies between conventionalised language use, speakers' intuition and

[9] See Wiktionary entry https://de.wiktionary.org/wiki/effektiv.

[10] See: http://www1.ids-mannheim.de/lexik/paronymwoerterbuch/dasparonymkorpus.html.

[11] For verbs, these would be based on the extraction of complementation patterns.

traditional dictionary entries. They are essential in the tracing of regular, conventionalised or new semantic components. The analysis and interpretation of patterns shows that meaning is conceptualisation, constantly negotiated in usage. Aspects of discourse, domain, reference and ontological categorisation are mentally processed and stored as information on lexical use and meaning.

In the case of the paronym dictionary, linguistic and encyclopaedic details are drawn from corpus data and are included in usage-based linguistic patterns, illustrating conversational interaction and semantic negotiations in contemporary public discourse. Cognitive elements play an essential role when users confuse lexical items. This confusion is often not only related to formal similarities but also to conceptual closeness. Corpus-derived data allow for the search of minimal semantic differences and the integration of necessary encyclopaedic knowledge, information that is complementary to linguistic information and needed by users. While this is not news to cognitivists, lexicographers still have to learn how to integrate this insight into usable tools. Bridging the gap between corpus lexicography and cognitive paradigms is a slow but steady process (Gries, 2006; Rundell, 2012; Hanks, 2013). Writing dictionaries should be informed by theoretical grounding and lexicographers should be linguistically aware corpus analysts. As Lew (2007: 221) points out "let us hope that lexicographers will keep an open mind to developments in linguistics […]".

5. Summary

So far, there is no corpus-assisted German reference guide empirically describing commonly confused words and enabling readers to find the correct contemporary usage.[12] The paronym dictionary is committed to overcoming the discrepancy between traditional practice and insights from language use. This necessarily means finding a way of educating users by showing how linguistic knowledge, encyclopaedic knowledge and human experience are inextricably linked. Given these goals, the dictionary breaks down the binary distinction of dictionary vs. encyclopaedia. Solutions to a number of lexicographical challenges were required. One aim was to bridge the gap between cognitive semantics and corpus lexicography by simultaneously considering user needs. It was argued that cognitive aspects can successfully be incorporated into meaning descriptions based on corpus-driven analysis. Insights into collocational use and the interpretation of contexts can lead to the implementation of more abstract encyclopaedic or conceptual categories as central ideas. Together with concrete prototypical contextual realisation these replace circular definitions and uncommented lists of synonyms. Authentic examples reflect prototypical structures as manifested in discourse and in the mental lexicon.

In contrastive entries, the interaction between lexemes is emphasised. The dictionary

[12] Intuition-based dictionaries include (Müller) 1973 and Pollmann & Wolk (2010).

strives to adequately reflect ideas such as conceptual structure, categorisation and knowledge. While Kövecses & Csábi (2014) argue that employing cognitive linguistics is a profitable theoretical underpinning for lexicographers, we favour the description in terms of cognitive principles as it predominantly embraces user needs.

Only a digital resource is able to solve problems of strict macrostructural ordering. Indeed, "an online dictionary can be adapted to the needs of each dictionary user" (Kwary, 2012: 35). Dynamic look-up options replace rigid structures. An adaptable access to lexicographical information has been suggested, where variable search options enable different foci and perspectives on linguistic information. In addition, the implementation of interactive collocation networks is a more onomasiological approach which offers an alternative access to language and knowledge structures relevant in actual usage events. The Paronymwörterbuch is a dynamic source of information where the interests of different users will hopefully be met.

6. References

Bolshakov, Igor A. & Gelbukh, Alexander (2003). Paronyms for accellerated correction of semantic errors. In *International Journal Information Theories & Applications* 10, pp. 198–204.

Fillmore, Charles J. (1976). Frame semantics and the nature of language. *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech* 280, pp. 20–32.

Fillmore, Charles, J. & Atkins, B. T. Sue (1992). Toward a Frame-based Lexicon: The Semantics of RISK and its Neighbors. In Adrienne Lehrer & Eva Feder Kittay (eds.) *Frames, Fields and Contrast. New Essays in Semantic and Lexical Organization*. Hillsdale & London: Erlbaum, pp. 75–102.

Fillmore, Charles, J. (1977). The need for a frame semantics in linguistics. In Hans Karlgren (ed.) *Statistical Methods in Linguistics*. Stockholm: Scriptor, pp. 5–29.

Fuertes-Olivera, Pedro A. (2013). e-lexicography: The Continuing Challenge of Applying New Technology to Dictionary-Making. In Howard Jackson (ed.) *The Bloomsbury Companion to Lexicography*. London: Bloomsbury, pp. 323–340.

Geeraerts, Dirk (2007). Lexicography. In Dirk Geeraerts & Hubert Cuyckens (eds.) *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford university Press, pp. 1160–1174.

Gries, Stefan Th. (2006). Corpus-based methods and cognitive semantics: The many senses of to run. In Stefan Th. Gries & Anatol Stefanowitsch (eds.) *Corpora in Cognitive Linguistics Corpus-Based Approaches to Syntax and Lexis*, Berlin & New York: de Gruyter, pp. 57–100.

Hanks, Patrick (2013). *Lexical Analysis: Norms and Exploitations*. Cambridge (MA): MIT.

Kövecses, Zoltán & Csábi, Szilvia (2014). Lexicography and cognitive linguistics. *Revista Española de Lingüística Aplicada* 27(1), pp. 118–139.

Kwary, Deny Arnos (2012). Adaptive hypermedia and user-oriented data for online

dictionaries: A case study on an English dictionary of finance for Indonesian students. *International Journal of Lexicography* 25(1), pp. 30–49.

Làzàrescu, Ioan. (1995). Deutsche Paronyme. In *Grazer Linguistische Studien* 43, pp. 85–93.

Làzàrescu, Ioan. (1999). *Die Paronymie als lexikalisches Phänomen und die Paronomasie als Stilfigur im Deutschen*. Bukarest: Editura Anima.

Lew, Robert. 2007. Linguistic semantics and lexicography: A troubled relationship. In Małgorzata Fabiszak (ed.) *Language and Meaning. Cognitive and Functional Perpectives*. Frankfurt: Lang, pp. 217–224.

Müller-Spitzer, Carolin (ed.). (2014). *Using Online Dictionaries*. Berlin & New York: de Gruyter.

Ostermann, Carolin (2015). *Cognitive lexicography. A New Approach to Lexicography Making Use of Cognitive Semantics*. Berlin & Boston: de Gruyter

Rundell, Michael (2012). It works in practice but will it work in theory? The unseasy relationship between lexicography and matters theoretical. In Ruth Vatvedt Fjeld, & Julie Matilde Torjusen (eds.) *Proceedings of the 15th EURALEX International Congress,* EURALEX 2012. Oslo: University of Oslo, pp. 47–92. Available at: www.euralex.org/elx_proceedings/.../pp47-92%20Rundell.pdf (10 May 2017).

Schnörch, Ulrich (2015). Wie viele Paronympaare gibt es eigentlich? Das Zusammenspiel aus korpuslinguistischen und redaktionellen Verfahren zur Ermittlung einer Paronymstichwortliste. *Sprachreport* 4 (2015), pp. 16–26.

Storjohann, Petra & Schnörch, Ulrich (2014). Empirical approaches to German Paronyms. In Abel, Andrea & Chiara Vettori & Natascia Ralli (eds.) *Proceedings of the 16th EURALEX International Congress: The User in Focus*. EURALEX 2014, Bolzano/Bozen: EURAC, pp. 463-476. Available at: http://euralex2014.eurac.edu/en/callforpapers/Pages/default.aspx (10 May 2017).

Storjohann, Petra (2015). Was ist der Unterschied zwischen sensitiv und sensibel? *Zeitschrift für Angewandte Linguistik* 62(1), pp. 99–122.

Storjohann, Petra (2016). Vom Interesse am Gebrauch von Paronymen zur Notwendigkeit eines dynamischen Wörterbuchs. *Sprachreport* 4 (2016), pp. 32–43.

Dictionaries & Websites:

Cambridge International Dictionary of English. 1995. Cambridge: Cambridge University Press.

DWDS: *Digitales Wörterbuch der Deutschen Sprache*. Accessed at: http://www.dwds.de. (10-14 March 2015).

Duden Online. Accessed at: http://www.dude-online.de. (10 May 2017).

FrameNet. Accessed at: framenet.icsi.berkeley.edu/fndrupal/. (10 May 2017).

German Frame-Semantic Online Lexicon GFOL. Accessed at: http://coerll.utexas.edu/frames/. (10 May 2017).

Müller, Wolfgang. 1973. *Leicht verwechselbare Wörter.* Duden Taschenwörterbücher Vol. 17. Mannheim: Bibliographisches Institut.

OWID - *Online Wortschatzinformationssystem des Deutschen.* Accessed at: http://www.owid.de. (10 May 2017).

Paronyme - Dynamisch im Kontrast. URL: http://www1.ids-mannheim.de/lexik/paronymwoerterbuch.html. (10 May 2017).

Pollmann, Christoph & Wolk, Ulrike. 2010. *Wörterbuch der verwechselten Wörter. 1000 Zweifelsfälle verständlich erklärt.* Stuttgart: Pons.

PDE - *Pattern Dictionary of English Verbs.* Accessed at: http://www.pdev.org.uk. (10 May 2017).

Wiktionary-*Das freie Wörterbuch*: Accessed at: www.wiktionary.de. (10 May 2017).