Philipp Stöckle / Sabine Wahl (eds.)

Lexicography and Language Variation

Vienna University Press



V&R unipress



unipress

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

Wiener Arbeiten zur Linguistik

Band 8

Herausgegeben von Alexandra N. Lenz, Melanie Malzahn,
Eva-Maria Remberger und Nikolaus Ritt

Advisory Board:
Peter Auer, Universität Freiburg, Deutschland
Ina Bornkessel-Schlesewsky, Universität South-Australia, Australien
Olga Fischer, Universität Amsterdam, Niederlande
Junko Ito, UC Santa Cruz, USA
Hans Kamp, Universität Stuttgart, Deutschland
Johanna Laakso, Universität Wien, Österreich
Michele Loporcaro, Universität Zürich, Schweiz
Jim McCloskey, UC Santa Cruz, USA
John Nerbonne, Universität Groningen, Niederlande
Peter Trudgill, Universität Agder, Norwegen

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

Philipp Stöckle / Sabine Wahl (eds.)

Lexicography and Language Variation

With 63 figures

V&R unipress

Vienna University Press

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953



Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über https://dnb.de abrufbar.

Veröffentlichungen der Vienna University Press erscheinen bei V&R unipress.

Gedruckt mit freundlicher Unterstützung der Österreichischen Akademie der Wissenschaften.

© 2024 Brill | V&R unipress, Robert-Bosch-Breite 10, D-37079 Göttingen, ein Imprint der Brill-Gruppe (Koninklijke Brill BV, Leiden, Niederlande; Brill USA Inc., Boston MA, USA; Brill Asia Pte Ltd, Singapore; Brill Deutschland GmbH, Paderborn, Deutschland; Brill Österreich GmbH, Wien, Österreich)
Koninklijke Brill BV umfasst die Imprints Brill, Brill Nijhoff, Brill Schöningh, Brill Fink, Brill mentis, Brill Wageningen Academic, Vandenhoeck & Ruprecht, Böhlau und V&R unipress.
Alle Rechte vorbehalten. Das Werk und seine Teile sind urheberrechtlich geschützt.
Jede Verwertung in anderen als den gesetzlich zugelassenen Fällen bedarf der vorherigen schriftlichen Einwilligung des Verlages.

Umschlagabbildung: pixabay.com: https://pixabay.com/illustrations/system-web-network-connection-3699542/, https://pixabay.com/photos/library-old-historical-shelf-2614804/, https://pixabay.com/photos/library-tunnel-books-wallpaper-3

Vandenhoeck & Ruprecht Verlage | www.vandenhoeck-ruprecht-verlage.com

ISSN 2365-7731
ISBN 978-3-8470-1695-3

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

Contents

Philipp Stöckle / Sabine Wahl
Introduction 7

Petra Storjohann
Linguistic Variation and Diversity in a Dictionary of German Confusables 13

Andreas Nolda / Adrien Barbaresi / Alexander Geyken
A Newspaper Corpus for the Investigation of Diatopic Variation in Standard German – the *ZDL-Regionalkorpus* 31

Grit Nickel / Alexander Werth
Dialect Dictionaries and the Reconstruction of Morphological Systems: a Case Study on German Plural Allomorphy 49

Philipp Stöckle / Sabine Wahl
Exploring Diminutive Variation in the Bavarian Dialects of Austria: a Lexicographic Perspective 69

Eva Wittenberg / Almut König / Mechthild Habermann
Grammatical Iconicity in the Expression of Augmentation in East Franconian 91

Karlheinz Mörth
Refining the Structure: Applying TEI Lex-0 to the *Concise isiZulu-English Internet Dictionary* 117

Volker Harm
The *Deutsches Wörterbuch* as a Resource for the Investigation of Loan Vocabulary 139

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

Christina Katsikadeli
Samuel Krauss' *Dictionary of Greek and Latin Loanwords in the Talmud, Midrash and the Targums* (1898) Revisited: New Perspectives on Language Contact and Variation in Accordance to Recent Advances in (Digital) Lexicography (DLGenR) 155

Yunlu Wan
Some Examples of Chinese Linguistic Borrowings in the *Dicionário Português–Chinês* (1583–1588) 175

Thomas Krefeld / Stephan Lücke
Linguistic Diversity and Bidirectional Lexicography 191

Manuel Raaf
Murphy's Law in Electronic Lexicography: What Can Go Wrong, Will Go Wrong. Frequent Challenges on the Example of *Bavaria's Dialects Online* 211

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

Philipp Stöckle / Sabine Wahl

Introduction

The paradigm of language variation explores socio-cultural, ethnic, cognitive or media-specific variation as well as individual and societal multilingualism in the past and present. In order to study linguistic diversity, pre-existing (primarily) written sources are consulted or new data is collected using different methods, e. g., oral or written questionnaires, interviews, experimental tasks, etc. However, the availability of appropriate data continues to be one of the greatest challenges to empirical research.

Lexicography – one of the oldest linguistic sub-disciplines – began to compile extensive corpora early on as the basis for dictionary work. Surprisingly, these corpora and the resulting dictionary articles have not been used very frequently for the study of language variation, although most dictionaries do not only contain information about word senses and grammar, but also deal with variation in many ways, such as regional distribution, style level, language contact, language change or pragmatics.

This volume explores the value of dictionary articles and lexicographical data in the study of language variation. The contributions focus on different types of dictionaries for different languages and target groups as well as on various linguistic research questions ranging from the dictionaries' approach to loan words or morphology to practical issues regarding digital frameworks for lexicographic work.

Some of the superordinate research questions addressed in this volume are:

- How is linguistic variation treated in lexicography?
- How is linguistic variation represented in dictionaries or in different dictionary types?
- Which levels of a language system (in addition to lexis and semantics) can be studied by using lexicographical data?
- Which methods can be used to access lexicographical material beyond the dictionaries?
- Which tools are available?

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

The idea for this thematic volume goes back to and was inspired by the workshop "Dictionary Articles and Corpora – a Research Laboratory for Linguistic Diversity" organized by the editors at the 42nd Annual Conference of the German Linguistic Society ('Deutsche Gesellschaft für Sprachwissenschaft', DGfS) "Sprachliche Diversität: Theorien, Methoden, Ressourcen / Linguistic Diversity: Theories, Methods, Resources" in 2020 in Hamburg. Some of the participants wrote full papers for this volume, but it also includes contributions from other dictionary projects.

The volume is grouped into four thematic sections. The first section includes dictionaries dealing with standard language. Petra Storjohann's paper "Linguistic Variation and Diversity in a Dictionary of German Confusables" discusses pairs of competing words in German, known as confusables or paronyms, which share roots and are easily confused due to similarities in form and concept (e. g., *autoritär/autoritativ* 'authoritarian/authoritative'). The dictionary *Paronyme – Dynamisch im Kontrast* documents these patterns and variations in usage across different contexts, genres, and registers, providing insights into shifts in usage over time and between different types of communication.

The contribution "A Newspaper Corpus for the Investigation of Diatopic Variation in Standard German – the *ZDL-Regionalkorpus*" by Andreas Nolda, Adrien Barbaresi & Alexander Geyken introduces the *ZDL-Regionalkorpus*, a corpus aimed at investigating diatopic variation in Standard German. It contains over 31 million articles with 9 billion tokens from local and regional sections of German daily newspapers. The *ZDL-Regionalkorpus* is integrated into the online dictionary *Digitales Wörterbuch der deutschen Sprache* (DWDS), allowing registered users to access and query it for geographic labels. The paper describes in detail its design, user interface, and a typical use case.

Within the second thematic section of our volume, the emphasis is placed on dialect dictionaries. The paper "Dialect Dictionaries and the Reconstruction of Morphological Systems: a Case Study on German Plural Allomorphy" by Grit Nickel & Alexander Werth explores the potential of using dictionary data from seven dictionaries covering all major German dialect areas for research on dialect inflection, particularly focusing on noun plural marking strategies. The study not only contributes insights into areal variations in plural marking strategies and declension classes, but also evaluates the methodological considerations and usefulness of using dictionary data for such research. The authors conclude that while there are limitations, dictionary data complement other data types and existing research effectively.

In the paper "Exploring Diminutive Variation in the Bavarian Dialects of Austria: a Lexicographic Perspective", Philipp Stöckle & Sabine Wahl explore the various diminutive forms (*-erl*, *-(e)l*, *-elein*, *-i*) in historical Bavarian dialects in Austria and South Tyrol from a lexicographic perspective, drawing upon the

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

dictionary articles and the underlying database of the *Wörterbuch der bairischen Mundarten in Österreich* (WBÖ) 'Dictionary of Bavarian Dialects in Austria'. The study describes how diminutives are presented in the dictionary articles and analyzes the morphological variation across the entire database as well as the semantic and, wherever possible, pragmatic classification of diminutives with the *-i* suffix.

Using corpus data from the *Franconian Dictionary* (WBF), the contribution "Grammatical Iconicity in the Expression of Augmentation in East Franconian" by Eva Wittenberg, Almut König & Mechthild Habermann examines the hypothesis that in East Franconian, an enhanced referent concept is primarily conveyed through the grammatical structure [noun1 + noun2] (e. g., *ein Trumm Kerl* 'a (big) piece (of a) bloke'), while notions of smallness or weakness are expressed through the [adj + noun] construction (e. g., *ein dürrer Hecht* 'a scrawny pike'). By comparing East Franconian data with Standard German constructions, the study concludes that nouns like *Brocken* 'chunk' function as augmentative modifiers, supported by a picture-matching task, bridging dialect lexicography and psycholinguistics.

Karlheinz Mörth's paper "Refining the Structure: Applying TEI Lex-0 to the *Concise isiZulu-English Internet Dictionary*" discusses the application of TEI Lex-0, a new format aimed at refining the structure of dictionaries for better interoperability, using the *Concise isiZulu-English Internet Dictionary* as an example. The investigation addresses challenges in modelling lexical data, particularly in languages like isiZulu, which lacks examples and features unique linguistic phenomena. It also explores issues of lexical variation, including morphological doublets and graphematic variation in foreign loans.

The third thematic section includes contributions on loan vocabulary and language contact. Volker Harm's contribution "The *Deutsches Wörterbuch* as a Resource for the Investigation of Loan Vocabulary" explores the potential of historical dictionaries, specifically focusing on the *Neubearbeitung des Deutschen Wörterbuchs von Jakob und Wilhelm Grimm* (^{2}DWB), to investigate loan vocabulary in German. Through two case studies, the paper examines the dictionary's usefulness in analyzing German loan-prefixes and tracing larger trends of loan activities over centuries. The study compares its findings with data from the *Deutsches Fremdwörterbuch* (DFWB), another significant resource for understanding linguistic diversity in German.

In the paper "Samuel Krauss' *Dictionary of Greek and Latin Loanwords in the Talmud, Midrash and the Targums* (1898) Revisited: New Perspectives on Language Contact and Variation in Accordance to Recent Advances in (Digital) Lexicography (DLGenR)", Christina Katsikadeli utilizes findings from the *Digital Dictionary of Loanwords of the Midrash Genesis Rabbah* (DLGenR) to showcase how advancements in theoretical and historical linguistics, corpus-based ap-

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

proaches, and digital lexicography encoding (TEI Lex-0) can enhance the revision of Samuel Krauss's dictionary. By leveraging digital tools, the revision process holds potential for yielding significant insights not only within Judaic studies but also in broader research areas such as the study of diachronic variation in Greek and other historical languages.

Yunlu Wan's paper "Some Examples of Chinese Linguistic Borrowings in the *Dicionário Português–Chinês* (1583–1588)" explores examples of linguistic borrowing in the *Dicionário Português–Chinês*, compiled by Jesuit missionaries during the Ming dynasty. It delves into neologisms introduced through loan words and semantic loans, shedding light on their phonological and semantic factors through comparative linguistics methods. By analyzing the functional equivalency of Portuguese-Chinese borrowings, it contributes to understanding the impact of cultural contact on Chinese vocabulary development in historical linguistics and lexicology.

The fourth and last thematic section contains contributions dealing with methodological questions. The contribution "Linguistic Diversity and Bidirectional Lexicography" by Thomas Krefeld & Stephan Lücke discusses the transformative impact of digitization and the internet on lexicography, emphasizing the need for both old and new dictionary projects to adapt to these changes. It highlights the potential of persistent identifiers like Wikidata to serve as a central reference, enabling the creation of a complex lexicographic network that facilitates research across dictionaries, languages, and epochs. The authors envision a future of bidirectional lexicography that not only explores semantic meanings, but also considers linguistic and cultural-historical connections in a more interconnected manner.

Manuel Raaf's paper "Murphy's Law in Electronic Lexicography: What Can Go Wrong, Will Go Wrong. Frequent Challenges on the Example of *Bavaria's Dialects Online*" examines the challenges faced in publishing dictionary content online, using *Bavaria's Dialects Online* (BDO), a language information system for the *Bavarian Dictionary* (BWB), the *Franconian Dictionary* (WBF), and the *Dialectological Information System of Swabian in Bavaria* (DIBS), as a case study. It discusses issues such as geo-referencing, unique identifiers of entries, and XML-related languages, aiming to raise awareness and provide insights to prevent similar problems in future projects. Additionally, it questions the continued usefulness of XML and other languages within the electronic lexicography field, considering the encountered challenges and potential alternatives.

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

Acknowledgements

We would like to thank the publishing house Vandenhoeck & Ruprecht GmbH & Co. / V&R unipress and the series editors Alexandra N. Lenz, Melanie Malzahn, Eva-Maria Remberger and Nikolaus Ritt for accepting this volume in the series "Vienna Studies in Linguistics" ('Wiener Arbeiten zur Linguistik'). Needless to say, we would like to thank all contributors for their interesting papers. We would also like to thank all reviewers who have contributed to the quality of the volume with their suggestions and comments. Furthermore, we would like to thank our student assistants Marlene Haslinger-Fenzl and Susanne Sophie Schmalwieser for their help with the formatting. Finally, we would like to thank the Austrian Centre for Digital Humanities and Cultural Heritage (ACDH-CH) at the Austrian Academy of Sciences for the financial support that made this volume possible.

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

Petra Storjohann

Linguistic Variation and Diversity in a Dictionary of German Confusables

Abstract: German has pairs of competing words which are easily confused. These often share a root and are hence similar both in form and in the concepts they designate (e. g., *autoritär/autoritativ* 'authoritarian/authoritative'). These confusables or paronyms can be subject to semantic variation. The dictionary "Paronyme – Dynamisch im Kontrast" documents conventionalised patterns and use. It provides information on preferences and tendencies as observed in real communication, thereby also accounting for variation in use. Shifts are observed diachronically and also in terms of genre, discourse and register. This paper explores how confusables are used in different situations and how this variation unfolds in the reference guide. It demonstrates how differences in usage between technical language and everyday communication or between oral and written conversation are captured.
Keywords: confusables, paronyms, easily confused words, dictionary of paronyms, semantic variation

1 Introduction

Two or more lexical expressions of the same part of speech that are similar in form and often refer to similar semantic concepts may, in certain circumstances, lead to linguistic uncertainty; they can be confused with one another or create communicative misunderstandings. Examples of such instances of monolexemic uncertainty (cf. Làzàrescu 1999; Klein 2018), also known as paronyms, include: *fachkundig/fachkundlich* ('*competent/with regard to a specific topic*'); *Methode/Methodik/Methodologie* ('*method/methodology*'); *original/originell/originär* ('*original/orginative/orginary*'). These words often differ on account of particular morphological elements. Also considered in more detail here as cases of linguistic uncertainty are phonetically similar expressions that have no semantic overlap at all, such as *ethisch/ethnisch* ('ethical/ethnical'), as well as alternative plural forms (*Wörter/Worte,* '*words*') and antonyms (*konkav/konvex,* '*concave/convex*').[1]

1 There are approx. 2,000 pairs/groups of German paronyms (cf. Schnörch 2015).

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

The many discussion threads on online language forums dealing with uncertainty about particular words demonstrate that both foreign language learners and native speakers have a keen interest in the appropriate use of these expressions. Language speakers also turn to these forums because they have encountered difficulty when consulting established reference works (cf. Murphy 2013). Their questions remain unanswered because the information there is vague, is absent altogether, or runs counter to their own intuition. Until 2018, there was no dedicated paronym dictionary for German. This omission was made good by the online dictionary "Paronyme – Dynamisch im Kontrast", which presents the first corpus-based analysis of paronyms and a descriptive documentation of their current usage.[2] It contains 370 entries, which, under a number of headings, attest to the heterogeneity in the behaviour of individual pairs or groups of lexemes according to the type of text, medium, style, situation, function, or chronological period.[3] The dictionary also describes linguistic variation that emerges in the form of conventionalised linguistic patterns and, as such, as a constituent part of the semantic spectrum. By incorporating variational linguistic behaviour, the entries illustrate usage in different language situations, provide linguistic and extralinguistic knowledge, and raise users' awareness of semantic alternatives. Using examples from the dictionary, this contribution will outline how extensive the context-dependent differences are that exist between paronyms and how they are represented in the new reference work. All the linguistic examples mentioned below are described in detail in the dictionary, and corpus evidence is provided for all the descriptions in the relevant dictionary entries.

2 The Paronym Corpus – A Mirror of Semantic Variation and Diversity

2.1 Variation and Diversity

Specific linguistic usage in the same language that differs according to particular parameters or dimensions (medium, speaker, context, time, place, function) is referred to as variation (cf. Berruto 1987: 263). There is a consensus that variation is a fundamental characteristic of language (cf. Sinner 2014; Felder 2016; Elspaß 2018). As far as paronyms are concerned, this begins with the object of enquiry itself. Many paronyms are morphologically motivated lexical variants. This ap-

2 The dictionary was compiled at the Leibniz-Institut für Deutsche Sprache as part of the "Paronymwörterbuch" project. Dictionary access: <https://www.owid.de/parowb>.

3 Regional differences have not been considered as part of the "Paronymwörterbuch" project.

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

plies above all to pairs that are alternative plural forms (*Denkmäler*/*Denkmale* ('*monuments*')) or that are competing lexemes that refer to the same conceptual entity (*kristallin*/*kristallen*/*kristallisch*/*kristallinisch* ('*crystal*/*crystalline*')). In cases of monolexemic uncertainty, linguistic heterogeneity exists particularly at a semantic level and is rooted in a diverse spectrum of meaning. For this reason, variation will be understood in what follows in the sense of standard or near-standard semantic uses, considering "language at a particular moment in time, that is, the simultaneous coexistence of different meanings or functions and usages, and change in language variation from a diachronic perspective, that is, the chronological succession of different meanings or functions and usages" (Elspaß 2018: 87; translation: PS)[4]. Two forms of variation are discussed in relation to pairs or groups of paronyms. First, pairs or groups demonstrate types of usage that vary according to the form of communication or function (technical variety), according to medium (situational and medial variety), or according to a particular historical period (diachronic variety). These variations affect a lexical expression or a paronym pair and are described in the keyword entry (see section 3). Second, we encounter semantic alternatives across pairs. This form of heterogeneity corresponds to semantic diversity or variation in the general form of paronyms and can only be established across different entries in the dictionary (see section 4).

2.2 Corpus, Variation, and Lexicography

Constructing a descriptive, corpus-based paronym dictionary represents a challenge, since users are primarily accustomed to traditional, and often prescriptive, reference works and treat these works as sources of linguistic authority (cf. Haß-Zumkehr 2001). When it comes to uncertainty between similar expressions, speakers are interested in finding the 'correct' word in the 'correct' context or in finding simple and distinctive features. However, corpus-based analysis shows that any objective description of paronym usage must include what is usual, prototypical, and prevalent across different linguistic levels and variants. The "Paronymwörterbuch" project compiled its own corpus,[5] which ensured analysis, above all, of general public and technical communication (cf.

4 German original: "Sprache zu einem bestimmten Zeitpunkt, also das gleichzeitige Nebeneinander von verschiedenen Ausdrucksformen und Bedeutungen bzw. Funktionen, Wandel auf die Veränderlichkeit von Sprache in diachroner Perspektive, also das zeitliche Nacheinander von verschiedenen Ausdrucksformen und Bedeutungen bzw. Funktionen."

5 The corpus extends to 2.3 billion words and includes different textual sources with a relative balance between regions. The corpus is publicly accessible in order to ensure lexicographical transparency. It can be accessed via COSMAS II.

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

Storjohann 2019a). Historical corpora (from DeReKo) were employed for (micro)diachronic analyses and the Datenbank für Gesprochenes Deutsch (DGD) was used to analyse spontaneous and oral language use. The scope, composition, and range of text-types made it possible to evaluate both general and technical usage on a supraregional level, as well as thematic and discursive contextualisations, spoken peculiarities, and discernible differences in more recent language history.[6] The findings from specific pieces of corpus-linguistic research found their way into different rubrics in the dictionary. The meaning of paronyms in contemporary public communication, as manifested in written language, is presented contrastively under the rubric "Vergleich" ('Comparison') (as seen in Figure 2). Information about spoken and historical usage, together with variations between everyday and specialist communication can be consulted under the rubrics "Sprachwandel" ('Language Change'), "Mündliches" ('Spoken Language Use'), and "Fachsprache" ('Technical Language Use') (as seen in Figure 1), with the proviso that some normalised deviations exist.



Figure 1: Dictionary sections for the heading "Wissenswertes" ('Nice to Know')

3 Semantic Variation within a Paronym Pair

An online dictionary benefits from the space, functionality, and means of representation that allow extensive or diverse spectrums of meaning to be described without any restrictions.[7] Describing general meanings in pairs occupies the central position in the paronym dictionary (as seen in Figure 2, "Vergleich" ('Comparison') rubric). The key element is the dynamic lexical-semantic comparison between keywords or individual contextual readings. Different descriptive levels of meaning or functions are available in contrastive entries, in both a summary view and a detailed view.

6 On different corpus methods, see Schnörch (2019) and Teichmann (2019).
7 Müller (1973) and Pollmann / Wolk (2001) often list just older or individual readings that demonstrate only semantic distance to the contrasting partner word.

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

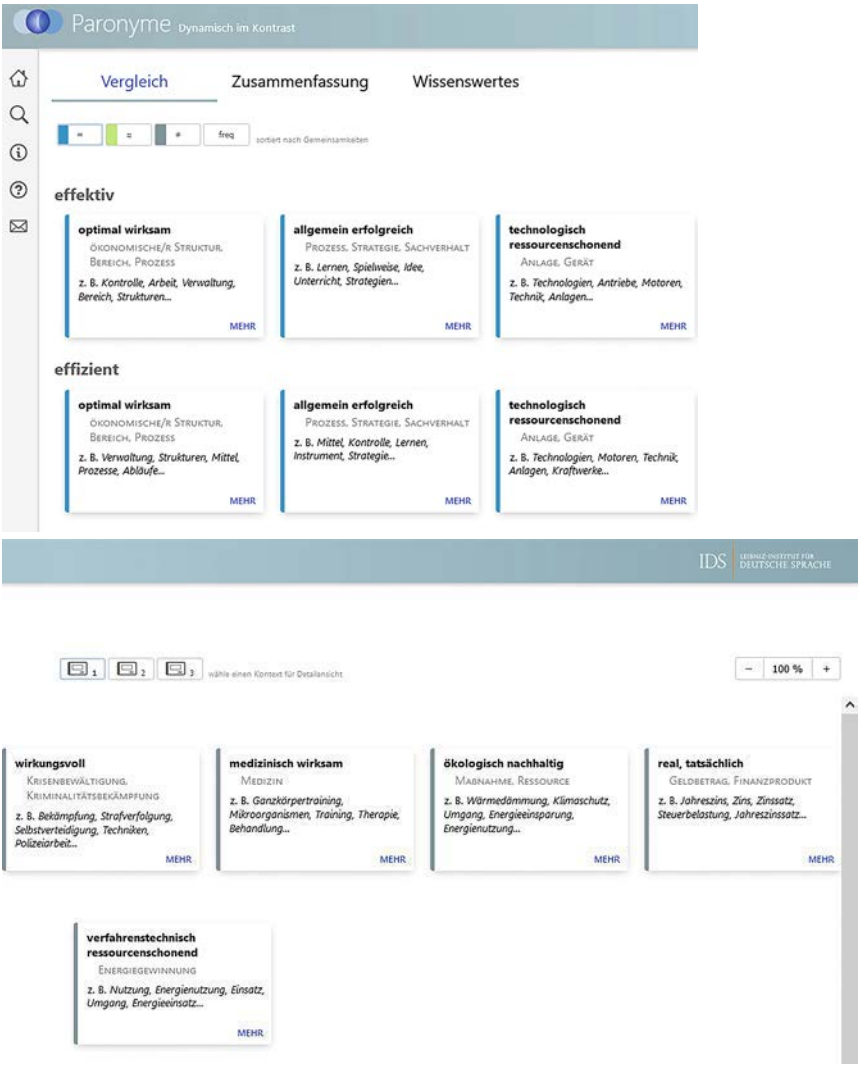


Figure 2: Extract of contrastive overview of the semantic spectrum for *effektiv/effizient*

Each individual tile on the page represents an established contextual usage, which designates a particular domain of meaning with reference words that can be clearly delimited from one another. Every context of use is identified by a short paraphrase (e.g., *effektiv* ('*effective*'): 'optimal wirksam' ('optimal, successful in producing a desired result')); the contextual references beneath illustrate ontological categories (e.g., PROZESS, GERÄT, MEDIZIN ('PROCESS, DEVICE, MEDICINE')), which are exemplified with collocates (e.g., *Kontrolle*, *Motoren*, *The-*

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

rapie (‘*control*, *engine*, *therapy*’)). The summary serves as an overview of the meaning for the two expressions in the pair and as a presentation of context-dependent relationships at the same time. Coloured highlighting and positioning on the page (either directly underneath one another or offset) distinguish between identical, similar, or divergent relationships. This view also acts as a navigating structure through which users access additional information in the detailed view.[8] Finally, it also provides a summary of which usages are more commonly or more rarely employed, since it can be sorted, via a menu, according to frequency. In addition to these central descriptions of meaning from everyday communication, the dictionary supplies further documentation of variational behaviour in additional rubrics, which are explained below.

3.1 Variation in Usage between Registers

Paronyms can often be technical and academic expressions that are part of general speech and widely understood. These include lexemes from medicine (*Arthritis/Arthrose,* (‘*arthritis/arthosis*’)), education (*autoritär/autoritativ,* (‘*authoritarian/authoritative*’)), psychology (*Egozentrik/Egozentrismus* (‘*egocentrism*’)), history (*historisch/historistisch/historizistisch* (‘*historic/historical/historicist*’)), and geology (*Stalaktit/Stalagmit* (‘*stalactite/stalagmite*’)).[9] Here, the specialist usage may diverge from that found in everyday language. For example, *Parodontitis/Parodontose* (‘*periodontitis/periodontosis*’) are synonymous in general speech, but in dentistry they refer to different forms of illness. In computing, *authentisieren* (‘*authenticate*’) denotes the action of a ‘user typing in logins/and password to get authorisation’, while *authentifizieren* (‘*authenticate*’) designates an ‘IT-process where the computer checks back if the user is permitted to be authorised’. By contrast, in everyday language the two verbs are used in the sense of ‘to authenticate/authorise accounts’. There are similar differences for the keywords *Methode/Methodologie/Methodik* (‘*method/methodology*’) between academic language and everyday conversation, while *lebenslang/lebenslänglich* (‘*lifelong/life sentence*’) have different interpretations in legal discourse and colloquial language. These kinds of particularities are presented in narrative form in the “Fachsprache” (‘Technical Language Use’) rubric with examples, quotations, or references to literature.

8 One or more contexts can be studied in the detailed (comparative) view. Additional information includes a more extensive definition, further collocates, attestations, linguistic constructions, and synonyms/antonyms.

9 For a systematic list, see Storjohann (2019b).

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

How do these kinds of difference come about? Technical language is a "constitutive feature of language in science, technology, and in institutions" (Roelcke 2015: 372; translation PS)[10] and is used by particular groups of speakers in order to ensure that technical information can be exchanged in a specific and precise way. As a result, meanings are strictly defined. Expressions and concepts that are frequently a matter of discussion outside of the specialist field and that, as such, occur regularly in public discussion may cross over into the general language. A range of influences, such as controversial public debates or advertising, ensure that specialist pairs (e. g., *Parodontose*/*Parodontitis* ('*periodontosis*/*periodontitis*') from medicine) can also find a place in the standard language.

The variation in meaning can be set out as follows. In medicine, terms with the ending *-itis* denote inflammatory diseases and those with the ending *-ose* denote atrophic diseases, i. e., those involving tissue degeneration. Hence, *Parodontitis* designates inflammation of the gums, while *Parodontose* refers to a degenerative disease of the gums originally thought to lack accompanying inflammation.[11] As such, it would be technically incorrect to refer to a case of periodontitis as periodontosis. However, in general speech *Parodontitis* and *Parodontose* are used as synonyms. Both lexemes describe an inflammation of the gums without any further differentiation. Overlaps in meaning also become clear through common surrounding contexts: both collocate with *Zähne* ('*teeth*'), *Zahnarzt* ('*dentist*'), *Hauptursache* ('*main cause*'), *Ursachen* ('*reasons*') and with *Patient* ('*patient*'), *Entzündungen* ('*inflammation*'), *behandeln*, *vorbeugen* ('*treat, prevent*'). Using *Parodontose* in the sense of *Parodontitis*, something which happens regularly in everyday communication, is judged to be scientifically incorrect from the perspective of specialist language. This is demonstrated by constructions such as the following: … *die* ***Parodontose*** (*wissenschaftlich korrekt:* ***Parodontitis***), … ***Parodontitis****, fälschlich oft* ***Parodontose*** *genannt.* This mistake is not commented on evaluatively in general language, but the commonly held equivalence is often marked metalinguistically …. ***Parodontitis****, im Volksmund auch/oft "****Parodontose****" genannt, …. **Parodontitis*** (*umgangssprachlich* ***Parodontose*** *genannt*). The trigger for this everyday equivalence might lie in the use of *Parodontose* in the sense of *Parodontitis* in advertising for toothpaste.[12] This resulted in a continuing reception of the terms as equivalent and spread quickly in general speech. This kind of semantic process has not taken place with other

10 German original: "konstitutives Merkmal der Sprache in Wissenschaft, Technik oder Institutionen."

11 Cf. <http://www.gesundheit.de/lexika/medizin-lexikon>.

12 A German commercial by the brand "blend-a-med" in 1970 advertised a toothpaste suitable to prevent *periodontosis*, using the term in the sense of *periodontitis* (see also Storjohann 2020).

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

pairs of words following the same *-itis*/*-ose* pattern (*Arthritis*/*Arthrose*, *Dermatitis*/*Dermatose*, *Gastritis*/*Gastrose*, *Neuritis*/*Neurose*, *Spondylitis*/*Spondylose*).

3.2 Media-Specific Variation

Linguistic uncertainty appears more frequently in sources of spoken language than in written sources and can encourage new dimensions of meaning to be established. In order to improve spoken competence in relevant cases of linguistic uncertainty, the specific aspects of spoken communication are documented in the dictionary under the rubric "Mündliches" ('Spoken Language Use'). A range of additional details can be consulted there.[13]

First, there is information about contexts of usage that occur exclusively in the spoken language and that, as such, are not included in conventional reference works as, for example, with the word *effektiv* ('*effective*'). In this semantic variant, *effektiv* is used as an adverb and can be paraphrased as 'for the most part' or 'actually'. This usage is not attested regularly in written language.

Second, discourse-specific constraints are noted. The term *virtual* appears frequently in the spontaneous discussion formats of newsgroups and forums and usually in relation to questions relevant to IT. Particularly frequent are the following combinations of words that denote specialist concepts from computing and that do not occur in general written language: *virtuale Maschine*, *virtuale Diskette*, *virtualer Rechner*, *virtuale Umgebung* ('*virtual machine*, *virtual disc*, *virtual computer*, *virtual environment*').

Third, users can find information about the usage that is conveyed exclusively in spoken form. The article for *erfreut*/*erfreulich* ('*pleased*/*pleasing*') records the fact that *erfreut* is used as a polite greeting in the expression *sehr erfreut* ('*very pleased*'), as a polite response to the personal introduction of someone unknown, or also as a polite answer.

Fourth, lexicographers have documented when the confusion of particular expressions can be observed more frequently in spoken interactions. The expression *Worte* (usually referring to an entire utterance or speech) is used more frequently in the DGD in the sense of *Wörter* ('several individual expressions') than in written communication. This usage may not lead to contextual misunderstandings, but this kind of situation does accelerate changes in meaning. This also applies to *scheinbar* ('*seemingly*'), which is already a partial synonym for *anscheinend* ('*apparent*') and which, in particular, regularly elicits the same meaning as the adverbial usage 'as it seems, apparently' in spoken texts. By

13 All following senses given here are taken from the corresponding entries of the paronym dictionary.

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

contrast, the meaning 'supposedly, not really' has been displaced, a process that is more advanced in the spoken language than in the written language.

Finally, it is worth mentioning in this context that differences in frequency of occurrence between medial variants are also described in the dictionary. The expression *scheinbar* is attested around three times more frequently than *anscheinend.* However, *anscheinend* is more common in spoken sources where it occurs about one and a half times more frequently than *scheinbar.*

4 Variation between Paronym Pairs

This section is devoted to the forms of variation or heterogeneity that exist between multiple paronym pairs. To that end, a number of entries are compared with one another in relation to particular semantic aspects.

4.1 Semantic Heterogeneity among Morphologically Analogous Pairs

To identify suitable keywords, word pairs were extracted systematically from the corpora. These pairs were identical in form except for particular tokens, mostly morpheme chains such as *-isch/-istisch*, *-bar/-lich.* Nine thousand potential paronym pairs were extracted automatically and then analysed and edited by the lexicographers (cf. Schnörch 2015). Some of the morphological groups are very extensive, such as the adjective pairs involving *-frei/-los* (*-free/-less*), fourteen of which are included in the paronym dictionary, including: *autofrei/autolos* ('*car free/carless*'), *fehlerfrei/fehlerlos* ('*flawless*'), *geruchsfrei/geruchslos* ('*odour free/odourless*'), *kinderfrei/kinderlos* ('*child free/childless*'), *kostenfrei/kostenlos* ('*free of charge/for free*'), *schmerzfrei/schmerzlos* ('*pain free/painless*'), *steuerfrei/steuerlos* ('*tax free/rudderless, leaderless*'), *wertfrei/wertlos* ('*unbiased/worthless*'). Pairs with identical patterns are listed in the rubric "Analoge Muster" ('Analogous Patterns'). Reckenthäler (2019) investigates whether pairs belonging to the same group in terms of their form also exhibit homogenous behaviour semantically. This corpus-linguistic analysis points to some differences to existing theories of word formation (cf. Fleischer / Barz 2012, Eichinger 2000) and also paints a very diverse semantic picture. Generally, individual adjective pairs cannot be organised into a single semantic category with fixed aspects of meaning on the basis of their wordfinal morphological segments (e. g., suffixes or suffixoids). There are pairs that are absolutely identical in meaning (*kostenfrei/kostenlos* ('*free of charge/for free*'), *zinsfrei/zinslos* ('*interest free*'), but there are also adjectives that overlap in some contexts while at the same time exhibiting slight nuances in their references or discursive field (*geruchsfrei/geruchslos*

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953

('*odour free/odourless*'), *parteifrei/parteilos* ('*non-party/independent*'). Equally, there are other pairs that are polysemous, but which have the same meaning in some usages while diverging in others (*fleischfrei/fleischlos* ('*meet free/meatless*'), *schmerzfrei/schmerzlos* ('*pain free/painless*')). A final group comprises pairs that demonstrate no semantic overlaps at all, but exhibit only distinct usages because of their considerable discourse-dependent differences (*steuerlos/steuerfrei* ('*tax free/rudderless, leaderless*'), *wertlos/wertfrei* ('*unbiased/worthless*')).[14] Comparing the corresponding dictionary entries testifies to the semantic heterogeneity identified in the behaviour of some paronyms in cases where theories of word formation would predict greater uniformity.[15]

4.2 Types, Reasons, and Effects of Lexical Confusion

Not in every case do both keywords in a pair actually cause linguistic problems. Among the pairs that generate uncertainty for speakers, there are some cases where only one of the two lexemes is used in the same sense as the other one, while in other cases both paronyms are confused with one another. These two types of cases occur with differing degrees of regularity. The causes of linguistic errors are equally varied, as are their consequences.[16] Underlying reasons for paronym uncertainty can be conceptual-semantic proximity and mutual association, phonetic similarities, regular contextual co-occurrence, a mismatch between the concept and the lexical item, insufficient extralinguistic knowledge, and new socio-political (critical) linguistic influence (cf. Storjohann 2020).

One prototypical case of uncertainty is that which arises from two expressions that are very similar semantically-conceptually, such as *sensibel/sensitiv* ('*sensible/sensitive*')[17]. The two adjectives are frequently used with the same meaning, something which results from a considerable extension of the meaning of *sensitiv* and its shift towards *sensibel.*[18] The greatest overlaps in meaning are found in contexts in which information, dates, or domains are characterised as controversial, awkward, connected to safety, or delicate/difficult. Numerous common collocates occur in the usage of both terms. These overlaps in meaning can

14 These lexemes are derivations of a homonym (*die Steuer/das Steuer* ('*tax/wheel*') or of a polysemous noun (*Wert* ('*value, worth, usefulness*')). As a result, they refer to different meanings.

15 As Reckenthäler (2019) points out, this assumption, however, needs more empirical validation.

16 These observations are described under "Zusammenfassung" ('Summary').

17 Please note, the English formal equivalents are false friends with their German terms.

18 The extension of meaning may also create a lack of semantic-referential precision and increased polysemy (cf. Blank 1997).

© 2024 V&R unipress | Brill Deutschland GmbH
ISBN Print: 9783847116950 – ISBN E-Book: 9783847016953