

Petra Storjohann

## SYNONYMS IN CONTRAST

### A Dynamic and Descriptive Resource for New Semantic Equivalents

**Abstract** This paper introduces “Synonyms in Contrast”, a new online dictionary that addresses the complexities and nuances of neologicistic (near-) synonyms in the German language. The emergence of new lexical items, often borrowings from English, has contributed to the proliferation of meaning equivalents. These share a large extent of contextual features, causing ambiguity and uncertainty among speakers regarding their appropriate usage. The dictionary, which is part of the new IDS *Neo*<sup>2020+</sup> resource, distinguishes itself by utilising corpus-driven and corpus-based approached and sophisticated lexicographic methods such as word embeddings to provide users with comprehensive, context-sensitive information on semantic similarities and differences between new synonyms. This resource aims to aid in the clear understanding of new terms, their usage, and interrelations by detailing their semantic overlaps, distinctive contexts, and collocational preferences. Through a combination of empirical data and user-specified presentation, the dictionary facilitates better linguistic decisions in both everyday and also more specialised communication. The ultimate goal is to enhance the integration of new lexical variants into the mental lexicon of German speakers by offering a reliable, dynamic, and descriptive platform for exploring and resolving lexical uncertainties with respect to meaning equivalents that have recently emerged.

**Keywords** dictionary of synonyms; discriminating meaning equivalents; contrastive lexicography; semantic similarities; new lexical variants

## 1. Introduction

Parts of the new German vocabulary are lexical items that share similar meanings (synonyms in their broadest sense)<sup>1</sup>, potentially leading to ambiguity in certain linguistic contexts or to uncertainty as to their preferences and appropriateness. This can be attributed to their unfamiliarity, semantic overlaps, conceptual proximity or it can also be a result of incomplete lexicalization processes. Some newly introduced terms are adopted from English (*Lockdown/Shutdown*). Other English terms are borrowed and serve as equivalent counterparts to existing native words (*Fake-News/Falschnachrichten/Fake-Nachrichten/Fakemeldung/Falschmeldung*, *Self-Care/Selbstfürsorge*, *Prank/Streich*). Alternatively, they enter concurrently together with their loan translations (*Wallbox/(Wand)Ladestation*, *Greenflation/grüne Inflation/Grünflation*). In situations where clusters of synonymous neologisms or a number of similar morphological variants emerge at the same time, users are presented with multiple options for expressing similar concepts and may feel the need to distinguish between them to determine their appropriate use. These challenges among German

<sup>1</sup> For a detailed discussion on classifications of synonymy see Cruse (1986), for a more pragmatic approach see Murphy (2003; 2013).

speakers become apparent in online forums, as attested by questions such as “What does “Prank” (*prank*) mean and what is the difference to a “Streich” (*hoax*)?”

This paper introduces a new dictionary for new German synonyms “Synonyme im Kontrast” (‘Synonyms in Contrast’), including corpus-driven and corpus-based methods (cf. Tognini-Bonelli, 2001) and addresses both, meta-lexicographic reflections and the practical implementation. So far, there is no corpus-related tradition in identifying and analysing synonyms, including neo-synonyms. Consequently, this gap has repercussions on the lexicographic documentation of terms with similar meanings. This absence has led to a lack of innovative reference guides that employ new approaches to detecting, describing and presenting synonyms. This is noteworthy considering because German lexicography has a well-established tradition of synonym dictionaries (cf. Hausmann, 1990; Hahn, 2004), and present-day opportunities allow for an empirically grounded online lexicography. As Murphy (2013, p. 301) points out “we can be very imaginative in how synonym information is found and presented”. As corpus-guided<sup>2</sup> reference works are inclined to follow descriptive perspectives, their aims differ. In contrast to distinctive or cumulative reference works (Hahn, 2004), a descriptive synonym resource aims to provide terms that are semantically similar or equivalent, along with comprehensive comparative explanations. The headwords are thoroughly disambiguated, highlighting context-bound similarities and differences, supported by authentic examples.

## 2. Neologistic Synonyms in Contrastive Entries

The new synonym guide “Synonyme im Kontrast” is a fundamental integral component of the recently designed neologism resource IDS *Neo*<sup>2020+</sup> which is currently being implemented.<sup>3</sup> A central feature of the synonym guide is its elaborate contrastive entries which allow for context-bound comparison of lexical use. Following the main principles developed for the German paronym dictionary for easily confused words (“Paronyme – Dynamisch im Kontrast”), the new synonym resource continues to establish a lexicography of lexical-semantic doubt. In order to retrieve and accommodate for necessary data, complementary corpus-related approaches are employed. On the one hand, potential synonyms are being extracted automatically through the comparison of similar collocation profiles, thus using a corpus-driven method. On the other hand, both near-synonym terms are subsequently searched for in close contextual vicinity in the underlying corpus. Then, the lexicographer can verify their semantic relation using the corpus-based method.

### 2.1 Lexicographic Corpus Methods

New methods for retrieving synonyms include word embedding tools, which use vector-based analysis to identify paradigmatic commonalities. These tools can identify

<sup>2</sup> In this paper, the term ‘corpus-guided’ refers to any method related to corpus-linguistic work.

<sup>3</sup> IDS *Neo*<sup>2020+</sup> will comprise a neologism dictionary (“Neo Dictionary”) launching in 2024; a dashboard featuring various formats and details on neologism research (“Neo Dashboard”) and a neologistic synonym guide (“Synonyms in Contrast”) launching in 2025.

terms with similar collocation profiles, often pinpointing synonyms. As these tools provide a visualisation with different levels of proximity indicating varying degrees of semantic closeness, lexicographers can quickly detect such closeness through the positioning of neighbouring terms. With regard to German, the tool DERKOVECS, which analyses data based on the German Reference Corpus (DeReKo), is employed to obtain and verify potential synonym candidates (Figure 1, see also <https://corpora.ids-mannheim.de/derekovecs?word=Lockdown&cutoff=500000&n=100&N=2000>). Among the words most similar to *Lockdown* are *Shutdown* as a semantic alternative and compounds specifying the type of Covid restriction such as *Corona-Lockdown*/*Corona-Shutdown* or *Teil-Lockdown* ('partial lockdown').

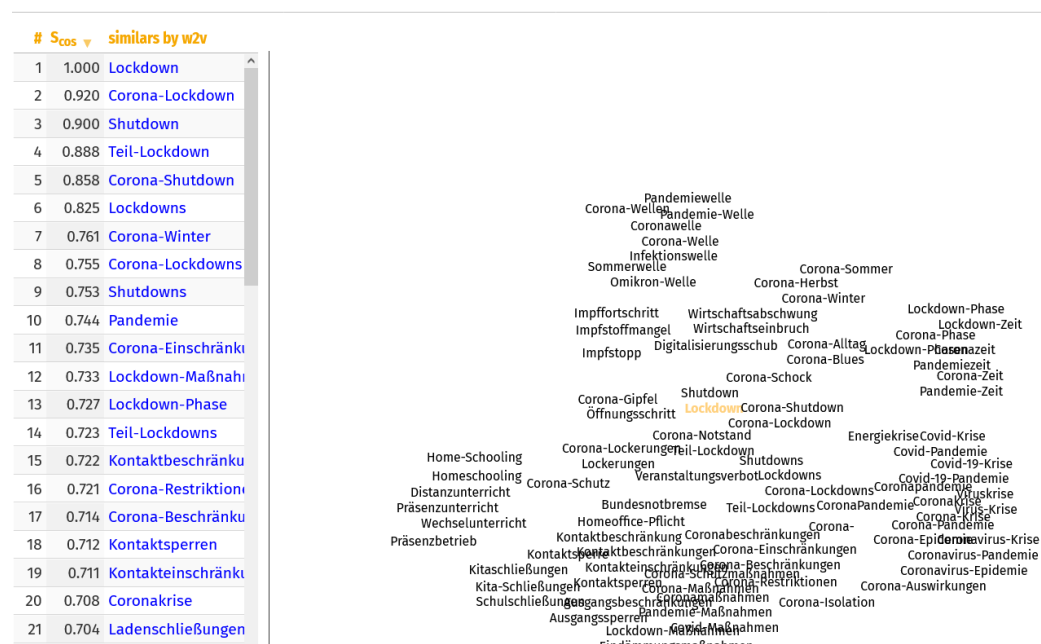


Fig. 1: Synonyms of *Lockdown* detected through DERKOVECS

Each word listed provides an interactive link to generate the corresponding word embedding accordingly. To answer user questions about the exact differences between two terms, one must employ a complementary method and conduct a collocation analysis. Collocations and examples that demonstrate contextual usage serve as more than just means of lexicographic analysis and interpretation; even more so they constitute a vital component within each entry, requiring users to consult them for a comprehensive understanding. Since the early stages of corpus-guided lexicography, as advocated by Partington (1998), the importance of this approach has been recognized and continues to be emphasised.

Comparing the collocational behaviour of so-called synonyms in concordance data may supplement dictionary information and help learners decide in what circumstances substitution of one item for the other is possible in a text. (Partington, 1998, p. 29)

The clear demonstration of the precise circumstances in which a new lexical item can be substituted for a similar item, by including typical syntagmatic patterns,

empowers users. This approach helps them to appropriately connect new words with similar meanings to the unknown concepts they represent and to understand the relationships between the lexical items (see Table 1). The question then arises: How can such pairs or groups be semantically registered and documented in a dictionary? Providing a parallel description of patterns in an entry enables users to discern where differences lie, or whether both terms are interchangeable.

**Table 1:** Typical collocates allocated to context, word class and semantic group

	Lockdown	Shutdown
adjectives		
How is X being characterised?	<i>hart</i> ('hard'), <i>light</i> ('light'), <i>erneut</i> ('anew/again'), <i>total</i> ('total'), <i>partiell</i> ('partial'), <i>monatelang</i> ('for months'), <i>flächendeckend</i> ('nationwide'), <i>coronabedingt</i> ('covid-related'), <i>strikt</i> ('strict'), <i>neuerlich</i> ('repeated')	<i>hart</i> ('hard'), <i>light</i> ('light'), <i>erneut</i> ('anew'), <i>komplett</i> ('complete'), <i>teilweise</i> ('partial'), <i>erschöpfend</i> ('tiring'), <i>coronabedingt</i> ('Covid-related'), <i>wochenlang</i> ('for weeks'), <i>verschärft</i> ('tightened'), <i>zielführend</i> ('effective')
verbs		
What does someone do in X?	<i>überstehen</i> ('endure'), <i>leiden</i> ('suffer'), <i>nutzen</i> ('use'), <i>vermissen</i> ('miss'), <i>entschleunigen</i> ('slow down'), <i>weitemachen</i> ('continue'), <i>verlängern</i> ('extend'), <i>verschärfen</i> ('intensify'), <i>verhindern</i> ('prevent'), <i>lockern</i> ('loosen')	<i>überstehen</i> ('endure'), <i>überleben</i> ('survive'), <i>aushalten</i> ('endure'), <i>abgewinnen</i> ('gain'), <i>beschließen</i> ('agree'), <i>verhindern</i> ('prevent'), <i>fordern</i> ('demand'), <i>verlängern</i> ('extend'), <i>lockern</i> ('loosen'), <i>verschärfen</i> ('intensify')
What does X do ?	<i>bremsen</i> ('slow down'), <i>(an-)dauern</i> ('last'), <i>gelten</i> ('apply'), <i>enden</i> ('end'), <i>herrschen</i> ('dominate'), <i>kommen</i> ('come'), <i>beginnen</i> ('begin'), <i>verhindern</i> ('prevent'), <i>drohen</i> ('threaten')	<i>(an-)dauern</i> ('last'), <i>zwingen</i> ('force'), <i>bremsen</i> ('slow down'), <i>drohen</i> ('threaten'), <i>treffen</i> ('affect'), <i>kommen</i> ('come'), <i>verhindern</i> ('prevent'), <i>enden</i> ('end'), <i>kosten</i> ('cost'), <i>eintreten</i> ('happen')
nouns		
Who is affected by X?	<i>Ungeimpfte</i> ('unvaccinated'), <i>Gastronomie</i> ('gastronomy'), <i>Geimpfte</i> ('vaccinated'), <i>Schulen</i> ('schools'), <i>Einzelhandel</i> ('retail'), <i>Wirtschaft</i> ('economy'), <i>Geschäfte</i> ('shops'), <i>Restaurant</i> ('restaurants'), <i>Friseure</i> ('hairdressers')	<i>Wirtschaft</i> ('economy'), <i>Möbelindustrie</i> ('furniture industry'), <i>Einzelhandel</i> ('retail'), <i>Selbstständige</i> ('self-employed'), <i>Restaurants</i> ('restaurants'), <i>Onlinehändler</i> ('online retailer'), <i>Kulturinstitutionen</i> ('cultural institutions'), <i>Bürger</i> ('citizens'), <i>Ungeimpfte</i> ('unvaccinated'), <i>Gastronomie</i> ('gastronomy')
What is X being associated with X?	<i>Coronapandemie</i> ('Covid pandemic'), <i>Homeoffice</i> ('homeoffice'), <i>Isolation</i> ('isolation'), <i>Ausgangs- und Kontaktbeschränkungen</i> ('curfews and contact restrictions'), <i>(Corona)-Infektionszahlen</i> ('Covid infection numbers'), <i>Öffnungsperspektive</i> ('opening perspective'), <i>Impfstrategie</i> ('vaccination strategy'), <i>Hamsterkäufe</i> ('panic buying'), <i>Impfpflicht</i> ('compulsory vaccination'), <i>Homeschooling</i> ('home schooling')	<i>Reisesperren</i> ('travel bans'), <i>Lockerungen</i> ('easing'), <i>Corona-Pandemie</i> ('Covid pandemic'), <i>Homeoffice</i> ('home office'), <i>Infektionszahlen</i> ('infection numbers'), <i>Einschränkungen</i> ('restrictions'), <i>Grundrechte</i> ('basic rights'), <i>Homeschooling</i> ('home schooling'), <i>Social Distancing</i> ('social distancing'), <i>Eindämmung</i> ('containment')

Table 1 shows that both lexical items overlap semantically to a large degree. Context-bound semantic commonalities as well as differences discernible through the interpretation of collocates, thereby establish the analytical foundation for frame-based referential use, as registered in the dictionary. A semantically equivalent relationship is also confirmed by a number of corpus examples (see example (1)).

Die Wissenschaftler hatten mehrere Szenarien entwickelt. Sie erwarten demnach, dass bei einem zweimonatigen **Shutdown** die Emissionen bundesweit um etwa 40 Millionen Tonnen CO<sub>2</sub> gegenüber 2019 sinken. Bei einem zweiten **Lockdown** aufgrund einer zweiten Infektionswelle wären Einsparungen von weiteren rund 71 Millionen Tonnen CO<sub>2</sub> möglich. (DeReKo, Welt Online, 25.05.2020) (Scientists had developed several scenarios. Accordingly, they expect that a two-month shutdown would reduce emissions nationwide by about 40 million tons of CO<sub>2</sub> compared to 2019. With a second lockdown due to a second wave of infections, additional savings of about 71 million tons of CO<sub>2</sub> could be possible.)

## 2.2 Two-level Presentation of Similarities and Differences

The typical co-occurrences in Table 1 and the corpus examples suggest a synonymous relationship between *Lockdown* and *Shutdown*. However, this relationship is strictly context-bound, as *Shutdown* is a polysemous term with two older senses not attested for *Lockdown*. Consequently, these neologisms also exhibit differences in their usages. It is therefore important to clearly indicate both semantic overlaps and differences, as identifiable through the use of collocates in a dictionary entry (Figure 2). In an entry overview, all senses of the terms in question are listed, with identical contextual uses appearing first by default. However, other listing options are available via a top menu. Each sense (or tile) contains a sense identifier in the form of a synonym or hyperonym, a domain and exemplary collocates.

The screenshot shows the IDS Neo 2020+ interface for 'Synonyme im Kontrast'. It features a search bar, navigation icons, and a top menu with options like '=', '≈', '≠', and 'freq'. The main content area displays three tiles:

- Lockdown (Neulexem)**:
  - Stillegung** (GESUNDHEIT, GESELLSCHAFT, POLITIK): z. B. Ungeimpfte, Gastronomie, Geimpfte, Schulen, Einzelhandel ...
- Shutdown (Neubedeutung)**:
  - Stillegung** (GESUNDHEIT, GESELLSCHAFT, POLITIK): z. B. Wirtschaft, Einzelhandel, Selbstständige, Restaurants, Ungeimpfte, ...
  - Haushaltssperre** (POLITIK): z. B. US-Kongress, US-Wirtschaft, Regierungsgeschäfte, US-Regierung, Bundesbehörden, ...
  - Herunterfahren** (INFORMATIK): z. B. Computer, PC, Server, Netzwerk, Prozesse, ...

Fig. 2: Contrastive overview of *Lockdown/Shutdown* with all their senses

Senses that are not shared are also marked by different colours and positions. Each use can be further explored by clicking on it individually or by selecting up to three tiles for comparison (Figure 3). In Figure 3, differing senses such as *Lockdown* ‘Stilllegung’ (‘closing down’) and *Shutdown* ‘Haushaltssperre’ (‘budget freeze’) were chosen for the detailed view to illustrate context-bound collocational differences. Here, an explicit definition is followed by details on domain and subject, a list of grouped collocates and background information. All details are presented side by side for direct comparison.

Verwendung	Grammatik	Verwendung	Grammatik
<p><b>Lockdown, ‘Stilllegung’</b> (Neulexem: seit 2020)</p> <p>bezeichnet einen Zeitraum, in dem fast alle wirtschaftlichen und gesellschaftlichen Aktivitäten auf politische Anordnung hin stillgelegt sind (z.B. zum Infektionsschutz)</p> <p><b>Domäne:</b> meist in GESUNDHEIT, GESELLSCHAFT, POLITIK <b>Thematischer Bezug:</b> BESONDERS WÄHREND CORONAPANDEMIE</p> <p><b>Kollokationen</b> <b>Wer oder was ist von einem Lockdown betroffen?</b> <i>Ungeimpfte, Gastronomie, Geimpfte, Schulen, ...</i></p> <p><b>Was macht man in bzw. mit einem Lockdown?</b> <i>überstehen, leiden, nutzen, vermissen, entschleunigen, ...</i></p> <p><b>Was macht ein Lockdown?</b> <i>bremsen, dauern, gelten, enden, losgehen, ...</i></p> <p><b>Wie ist ein Lockdown?</b> <i>hart, light, erneut, total, partiell, monatelang, ...</i></p> <p><b>Was wird im Zusammenhang mit Lockdown thematisiert?</b> <i>Coronapandemie, Homeoffice, Isolation, Ausgangssperren, ...</i></p> <p><b>Hintergrund:</b> Der Ausdruck <i>Lockdown</i> bezieht sich im Zuge der Corona-Pandemie auf einen Ausnahmezustand und Zeitraum, in dem zahlreiche Bereiche des öffentlichen Lebens stillgelegt waren. Der Ausdruck gehört damit zum Coronawortschatz und war besonders zwischen 2020-2022 besonders präsent.</p>		<p><b>Shutdown, ‘Haushaltssperre’</b></p> <p>bezeichnet eine Haushaltssperre, die zu einem Teil-Stillstand der Regierungsgeschäfte in den USA führt und daraufhin die Tätigkeiten des Staats- und Verwaltungsapparates einstellt.</p> <p><b>Domäne:</b> meist in POLITIK <b>Thematischer Bezug:</b> US-Regierung</p> <p><b>Kollokationen</b> <b>Wer oder was ist von einem Shutdown betroffen?</b> <i>US-Kongress, US-Wirtschaft, Regierungsgeschäfte, US-Regierung, ...</i></p> <p><b>Was macht man in bzw. mit einem Shutdown?</b> <i>abwenden, verhindern, beenden, vermeiden, erzwingen, ...</i></p> <p><b>Was macht ein Shutdown?</b> <i>andauern, anhalten, treffen, lähmen, kommen, verhindern, ...</i></p> <p><b>Wie ist ein Shutdown?</b> <i>erneut, teilweise, amerikanisch, neuerlich, wochenlang, ...</i></p> <p><b>Was wird im Zusammenhang mit Shutdown thematisiert?</b> <i>Fristablauf, Stilllegung, Trump, US-Geschichte, Demokraten, Republikaner, ...</i></p> <p><b>Hintergrund:</b> Bei <i>Shutdown</i> handelt es sich möglicherweise um eine Kurzform von Government-Shutdown und der Ausdruck bezieht sich auf US-amerikanische politische Zustände, die erfolgen, wenn sich Senat, Repräsentantenhaus und Präsident nicht rechtzeitig über die Bewilligung von Haushaltsmitteln einigen. <i>Shutdowns</i> ereignen sich regelmäßig.</p>	

Fig. 3: Contrastive detail view of *Lockdown/Shutdown* and two differing senses

The dictionary includes details specifically pertinent to new words, such as the type of neologism (e.g., loan word/translation, newly coined through word formation, see Figure 2), time of introduction and extent of its integration (e.g., as seen in grammar paradigms, see Figure 3).

### 3. Summary

The new German dictionary “Synonyme im Kontrast” will include neologistic near-synonyms of various kinds. Firstly, it will comprise pairs in contrastive entries consisting of one neologistic term and one older, often native term. Secondly, one will find entries with two new borrowed words that have entered the German language simultaneously. Thirdly, there will also be cases where a new loanword emerged and was given a German loan translation at the same time. In some cases, synonyms act as lexical rivals, displaying clear preferences in usage. In other instances, they coexist, leading to situations of lexical uncertainty where users are unsure about which term to use. The new reference guide aims at answering questions commonly found in online forums regarding the use of two new, semantically similar words.

Successful storage of information in the mental lexicon can only be achieved when linguistic and encyclopaedic knowledge are intricately intertwined within an entry. This can only be guaranteed when information related to two headwords is studied together, addressing both semantic and syntactic questions, including collocational and colligational issues. The explanations are supported by examples extracted from corpus data, with entries that aim to advance lexicographic knowledge in the realm of synonymy.

## References

- Cruse, A. (1986). *Lexical Semantics*. Cambridge University Press.
- DeReKo: *Deutsches Referenzkorpus*. Retrieved May 1, 2024, from <https://cosmas2.ids-mannheim.de/cosmas2-web/>.
- DEREKOVECS. Retrieved May 1, 2024, from <https://corpora.ids-mannheim.de/derekovecs>
- gutefrage.net*. Retrieved May 1, 2024, from <https://www.gutefrage.net/>
- Hahn, M. (2004). Synonyme und Quasisynonyme in lexikografischer Hinsicht. *Neuphilologische Mitteilungen*, 105(1), 3–20.
- Hausmann, F. J. (1990). The Dictionary of Synonyms: Discriminating Synonymy. In F. J. Hausmann, O. Reichmann, H. E. Wiegand & L. Zgusta, L. (Eds.), *Wörterbücher. Dictionaries. Dictionnaires*. HSK 5.2. (pp. 1067–1075). De Gruyter.
- Murphy, M. L. (2003). *Semantic Relations and the Lexicon*. Cambridge University Press.
- Murphy, M. L. (2013). What we talk about when we talk about synonyms. *International Journal of Lexicography*, 26(3), 279–304. <https://doi.org/10.1093/ijl/ect023>
- Paronyme – Dynamisch im Kontrast*. Retrieved May 1, 2024, from <https://www.owid.de/parowb/>
- Partington, A. (1998). *Patterns and Meanings*. Benjamins.
- Tognini-Bonelli, E. (2002). *Corpus Linguistics at Work*. Benjamins.

## Contact information

### Petra Storjohann

Leibniz-Institut für Deutsche Sprache, Mannheim  
storjohann@ids-mannheim.de

