
HERMES

Journal of Language and Communication in Business

Thematic Section (issue 66-2026)

**Phraseology and communication –
Formulaic language in communication**

Editors:

Erla Hallsteinsdóttir / Jan Engberg /

Anke Heier / Ermina Menzilovic

Published by the School of Communication and Culture, Aarhus University

ISSN: 1903-1785

**Carolina Flinz*

Collocations in Specialized Communication: Interlingual, Corpus-driven and Corpus-based Perspectives and Applications

Abstract

Starting from the claim that vocabulary learning is inherently collocation learning, this article highlights that lexical competence depends not only on knowing individual word meanings but also on understanding how words conventionally combine in context. This perspective proves particularly relevant in languages for specific purposes (LSP), where technical terms occur in fixed, domain-specific collocations essential for professional communication, such as *eine Diagnose stellen* (en. to give a diagnosis) in medical discourse or *einen Vertrag abschließen* (en. to sign a contract) in legal contexts. LSP genres are also characterized by recurrent formulaic patterns, which play a crucial role in both text comprehension and production. Collocations thus function as markers of professional competence and disciplinary affiliation, supporting not only linguistic correctness but also communicative appropriateness. Adopting a corpus-based and corpus-driven perspective (cf. Tognini-Bonelli, 2001), the article shows how collocations can be systematically extracted from LSP-corpora, capturing recurrent combinations, preferred syntagmatic environments, and genre-specific formulaicity. In the context of German as a foreign and LSP-language (DaF/DaFF), such corpus insights provide a foundation for needs-oriented vocabulary selection, lexicographic resource development, and data-driven learning (DDL), enabling learners to acquire both specialized terminology and its conventional lexical contexts. Practical applications are illustrated through lexicographic projects (TOURLEX) and teaching tools (COPLUS), showing that effective vocabulary instruction in DaFF requires attention to authentic usage patterns rather than isolated words.

Keywords

collocations; vocabulary learning; languages for specific purposes (LSP); corpus-based approach; German as a foreign and LSP-language (DaF/DaFF); formulaic patterns

1. Introduction

The¹ starting point of this article is Franz Joseph Hausmann's (1984) assumption that "Wortschatzlernen ist Kollokationslernen" (in English 'vocabulary learning is collocation learning'), which has had a lasting influence on phraseological research as well as on approaches to vocabulary acquisition. This claim highlights the fundamentally combinatorial nature of lexical knowledge and challenges views of vocabulary as a mere inventory of isolated words.

As a brief introductory exercise designed to open the discussion the guiding question "Why is vocabulary learning collocation learning?" was addressed with the support of an AI-based language model (ChatGPT 4.0, 02.05.2025). This small experiment serves merely as a starting point for illustrating and reassessing the relevance of the underlying assumption from a contemporary, usage-oriented perspective, rather than as empirical evidence. The response systematically foregrounds five interrelated arguments that are well established in linguistic and didactic research:

1. words rarely occur in isolation;
2. collocations facilitate both comprehension and production;

¹ The author would like to thank the anonymous reviewers for their valuable comments and constructive suggestions, which helped to improve this paper.

* *Carolina Flinz*
Dipartimento di Studi Umanistici
Sezione di Lingue e Letterature Straniere Moderne
Università di Pavia
carolina.flinz@unipv.it

3. many collocations are only partially semantically transparent;
4. recurrent word combinations structure the mental lexicon;
5. from a didactic point of view, learning fixed or preferred word combinations is more effective than learning isolated lexical items.

Drawing on these observations, the following considerations emerge: vocabulary knowledge does not merely consist in knowing the meanings of individual words but crucially involves knowing how these words are conventionally used in context. This usage-based dimension of lexical competence becomes particularly visible in collocational patterns, reinforcing the well-established conclusion that vocabulary learning is necessarily also collocation learning².

This insight proves to be even more pertinent when the focus is shifted from general language to languages for specific purposes (LSP). In specialized domains, lexical items are embedded in highly conventionalized, domain-specific collocations that are essential for professional communication. Technical terms typically occur in restricted combinations that cannot be freely substituted without loss of precision, acceptability, or professional credibility. In medical discourse, for example, expressions such as *eine Diagnose stellen* (en. to give a diagnosis) are standard, just as legal language relies on collocations like *einen Vertrag abschließen* (en. to sign a contract). Such combinations are not stylistic variants but constitute normative patterns within expert communities. So, although technical terms are semantically well defined, their full professional meaning frequently emerges only within typical syntagmatic environments, as illustrated by expressions such as *nach Erhalt der Buchungsbestätigung* (en. upon receipt of the booking confirmation). Accordingly, learners must acquire not only specialized terminology, but also the conventional lexical contexts in which this terminology is used. In addition, LSP genres are characterized by recurrent collocational and formulaic patterns, e.g. *Die Studie bezieht sich auf* (en. the study is based on), which are central to both text comprehension and production and therefore form an integral part of domain-specific vocabulary.

The appropriate use of collocations plays a crucial role in signalling professional competence and disciplinary affiliation. Conventionalized expressions function as markers of expertise, while deviations, although often comprehensible, may be perceived as stylistically inadequate or insufficiently professional. From this perspective, collocations contribute not only to linguistic correctness, but also to communicative appropriateness and professional identity.

Against this background, the present article wants to show how a corpus-based and corpus-driven perspective (cf. Tognini-Bonelli, 2001) is central to extract and investigate collocations as a core component of vocabulary knowledge in languages for specific purposes, with particular relevance for the field of German as a foreign language (*Deutsch als Fremdsprache*, DaF) and German as a foreign and specialist language (*Deutsch als Fremd- und Fachsprache*, DaFF) (cf. Roelcke, 2024). Corpus linguistics offers empirically grounded methods for identifying, describing, and quantifying collocational patterns as they occur in authentic professional discourse. By analyzing domain-specific corpora, it becomes possible to capture recurrent lexical combinations, preferred syntagmatic environments of technical terms, and genre-specific formulaic patterns that are characteristic of expert communication.

From a DaF/DaFF perspective, such a combined approach is particularly valuable, as learners of German in academic and/or professional contexts are confronted not only with specialized terminology but also with highly conventionalized patterns of use that are rarely made explicit in traditional teaching materials. It provides a crucial empirical foundation for needs-oriented vocabulary selection and for the development of lexicographic resources, teaching materials, and data-driven learning (DDL) scenarios that enable learners to explore authentic language use autonomously. Moreover,

² Although based on a brief exploratory IA experiment, these observations closely align with the theoretical position established in the literature (see Section 2).

collocation analysis of this type allows for a systematic distinction between general-language and domain-specific collocations (Flinz & Perkuhn, 2018), thereby making visible the ways in which lexical meaning, terminological precision, and textual conventions interact in specialized discourse. In this sense, collocations function as an interface between lexis, text, and discourse (see paragraph 2). By combining phraseological theory with corpus-linguistic methods, the article outlines how collocations can be systematically captured in lexicographic resources, as illustrated by TOURLEX (paragraph 3), and transformed into effective learning targets in DaF, as seen in COPLUS (paragraph 4). In doing so, it engages with current debates on corpus-based vocabulary teaching and highlights that effective instruction in DaFF relies on authentic usage patterns rather than isolated lexical items.

2. Towards an integrated view of collocation

The notion of collocation has been conceptualized in linguistics from different, partly complementary perspectives, most prominently along a divide between theoretical-lexicological approaches and empirically grounded, corpus-linguistic approaches. Since Firth's (1957, p. 11) dictum "you shall know a word by the company it keeps", collocations have been understood as recurrent co-occurrence patterns that contribute essentially to lexical meaning and usage. However, the criteria used to delimit and describe such patterns vary considerably depending on the underlying theoretical framework and methodological orientation.

In theoretical and phraseological traditions, collocations are often defined as conventionalized, yet semantically transparent word combinations that occupy an intermediate position between free combinations and idioms (Burger, 2015; Hausmann, 1984). Central to these approaches is the assumption of a structural asymmetry between a base and a collocator, with restricted lexical choice on the part of the collocator (Hausmann, 1984). From this perspective, collocations are lexically entrenched units that are part of speakers' linguistic competence and are particularly relevant for lexicography, phraseodidactics, and contrastive linguistics (Ďurčo, 2016; Ďurčo, Vajičková & Tomášková, 2019; Feilke, 1996).

By contrast, corpus-linguistic and usage-based approaches adopt a primarily empirical view of collocations, defining them as statistically salient co-occurrences of lexical items in large corpora (Belica & Perkuhn, 2015; Evert, 2009; Steyer, 2013, 2023). Here, collocational status is not determined a priori by categorical linguistic criteria but emerges from distributional patterns observable in authentic language data. Measures of association strength and frequency serve as key diagnostics, allowing for a gradient and probabilistic understanding of collocativity (Bubenhofer, 2009, 2017). In this view, collocations are part of a broader spectrum of recurrent multi-word patterns, ranging from weak co-occurrences to strongly fixed constructions.

Recent research increasingly emphasizes the complementarity of these perspectives, also for lexicographic and didactic perspectives (Flinz, 2024; Steyer, 2023). Corpus analyses provide empirical evidence of usage preferences and lexical attraction and enable a nuanced understanding of collocational phenomena by distinguishing combinations that reflect not only statistical significance but also linguistically motivated compatibility—whether syntagmatic or paradigmatic. In this way, they go beyond mere frequency effects, capturing collocations as higher-level units that emerge through their embedding in specific discourses and communicative practices (cf. Steyer, 2023). Conversely, theoretical and phraseological models offer interpretative frameworks that account for their functional, semantic, and didactic relevance (Mollica & Steyer, 2024; Schafroth, Mollica & Mellado Blanco, 2021). As a result, collocations are now commonly viewed as empirically observable yet cognitively and lexicographically motivated units, whose description benefits from the systematic integration of quantitative corpus methods and qualitative linguistic analysis.

2.1. Collocations in Specialized Communication

In specialized communication, collocations are of particular relevance because linguistic economy and unambiguous reference are central requirements. Technical and professional discourse aims at precision, efficiency, and a high degree of conventionalization; in this context, collocations function as condensed realizations of domain-specific conventions and functional patterns of meaning (cf. Flinz, 2019; Steyer, 2013). At the same time, they pose a significant challenge: while contributing to clarity and communicative economy, their form-meaning relations are often opaque, discipline-specific, and not freely compositional. Consequently, collocations constitute both a structural and a didactic problem in LSP contexts.

The complexity of collocations in LSP is further increased by the fact that there is no single “German specialized language” (Roelcke, 2020, p. 41). Instead, each discourse and subject area develops its own specialized language with distinct norms, conventions, and collocational patterns. Individual specialized languages are characterized by different levels of abstraction on both horizontal and vertical planes, varying degrees of specialization, and a high diversity of text types and communication genres (Swales, 1990; Buhlmann & Fearn, 2018; Flinz, 2019; Flinz & Pietzuch, 2024; Hoffmann, 1976; Hoffmann et al., 1998/99; Roelcke, 2020).

This perspective aligns with a broader strand of LSP research beyond German, where studies on English, French, Italian, and other languages have similarly highlighted the patterned and recurrent nature of specialized discourse (a.o. Swales, 1990; Biber, 2006; Coxhead, 2000; Tribble, 2000; Gledhill, 2000; Tutin, 2007; Pecman, 2007). In particular, research on academic and scientific English has shown that collocational behaviour is a key component of disciplinary discourse organization, while French and multilingual studies have described comparable phenomena as “scientific” or “generic” collocations, emphasizing recurrent phraseological patterns across domains.

Moreover, specialized communication frequently involves the simultaneous use of multiple specialized languages, whether intrafield, interfield, or extrafield. In addition, specialized language often co-occurs with professional language and everyday language within the same communicative setting (Efing, 2014, p. 420; Flinz & Pietzuch, 2024; Roelcke, 2020). This overlap creates hybrid registers in which collocations may shift in meaning, form, or pragmatic function. Cultural variation adds another layer of complexity, as specialized communication increasingly takes place in intercultural and transcultural contexts, requiring sensitivity to discourse conventions across cultures (Heier & Hallsteinsdóttir, 2024; Roelcke, 2024).

From a linguistic point of view, the interaction between polysemy and valency plays also a crucial role in the formation and interpretation of collocations (cf. Konecny, 2018). This applies not only within specialized domains (intrafield) but also across domains (interfield) and in the interface with educational or academic language, where collocational patterns often mediate between everyday and specialized usage.

Collocations in specialized communication can therefore be approached from two complementary perspectives (Flinz, 2019; Flinz & Pietzuch, 2024):

- From a theoretical perspective, the focus lies on the identification and analysis of collocations, their lexicographic elaboration, and their adequate didactic presentation. This includes questions of how collocations are defined, classified, and represented in dictionaries, as well as how their discourse-specific functions can be captured.
- From a practical perspective, attention shifts to acquisition and learning processes. This concerns learners of DaF/DaFF, for whom specialized collocations often pose significant difficulties, as they are closely tied to domain-specific knowledge that laypersons or non-experts do not typically master. At the same time, collocational competence is also relevant for native speakers of German (L1), particularly when entering unfamiliar professional domains, as well as for lay-

persons without subject-specific expertise. In all these cases, mastering collocations is a prerequisite for successful participation in specialized communication.

2.2. Collocations and DaFF

From a DaF and DaFF perspective³, collocations in specialized communication constitute a core interface between linguistic structure, usage, and didactic mediation. Rather than being treated as a purely descriptive phenomenon, collocations are approached here as pedagogically relevant units whose acquisition is crucial for the development of both general and domain-specific communicative competence in German. Insights from corpus linguistics, phraseology, and LSP research therefore converge in a framework explicitly oriented towards teaching, learning, and professional language use.

In the field of German as a Foreign Language (DaF), collocations increasingly become accessible through processes of didacticization, diversification, and digitalization (cf. Roelcke, 2024). Corpus-based resources and digital learning environments allow for the systematic identification, contextualized presentation, and guided exploration of collocational patterns, thereby fostering data-driven, learner-centered approaches to vocabulary development.

In German as a Language for Specific Purposes (DaFF), the focus shifts towards dynamization, differentiation, and decentralization (cf. Roelcke, 2024) of collocational knowledge. Collocations are analyzed as integral components of evolving specialized discourses, where they index domain-specific norms, terminological conventions, and professional practices. From a didactic perspective, this implies moving beyond static inventories towards usage-based models that reflect disciplinary variation and communicative needs.

Building on these considerations the present contribution adopts a corpus-driven and corpus-based approach to the systematic analysis of collocations, with a particular focus on German in professional and academic contexts. The empirical examples discussed in the following sections are drawn primarily from specialized corpora and are intended to illustrate how collocational patterns vary across subject domains, text types, and communicative settings. This perspective enables both intralingual and interlingual comparisons in LSP communication and highlights the relevance of collocational analysis for linguistic research as well as for DaF and DaFF teaching practice.

Examples of practical application include:

- Tourism discourse (treated as a specialized domain) from a lexicographic and LSP-oriented perspective, focusing on the identification of conventionalized collocational patterns, their frequency profiles, and their discourse-functional roles.
- Scientific and academic language (*Wissenschaftssprache*) in DaF contexts – here more appropriately understood as a discourse type encompassing multiple specialized domains rather than a single homogeneous field (see also *Allgemeine/Alltägliche Wissenschaftssprache*). Collocational patterns show both shared and discipline-specific tendencies, reflecting differences in epistemologies and communicative practices. For learners, this may pose a challenge, as they are required to navigate both shared academic collocational patterns and discipline-specific variations.

Across both domains, collocations are characterized with respect to their statistical salience, understood as recurring co-occurrence patterns measurable in corpora; their syntagmatic or paradigmatic motivation; and their corpus-semantic and discourse-pragmatic relevance.

Drawing on German-Italian example material and insights from ongoing projects such as Tourlex and COPLUS, this article illustrates the complexity of collocational behaviour (cf. Steyer, 2023) and

³ For a discussion of the characteristics and interrelations between DaF and DaFF, see Roelcke (2017).

demonstrates how corpus-driven and corpus-based approaches can inform both lexicographic and didactic practice. By synthesizing disciplinary perspectives, the contribution underscores the central role of collocations in specialized language communication and highlights their methodological and pedagogical relevance for DaF and DaFF.

3. Collocations in tourism language: Lexicographic perspective

Collocations have long been a central topic in lexicographic research, both in theory and practice. This concept has been widely discussed in the literature, with foundational contributions addressing their definition, classification, and treatment in dictionaries (cf. Wiegand, 1989; Mellado Blanco, 2008; Schafroth, 2011). Beyond theoretical considerations, a substantial body of research has focused on how collocations are represented in lexicographic resources (cf. a. o. for the language pair German-Italian Nied Curcio, 2006, 2012; Giacomini, 2014, 2015; Konecny, 2010, 2011; Mollica, 2023; Flinz, 2015, 2018). These studies examine different dictionary types and languages, addressing issues such as the selection of collocational data, the structuring of entries, the use of labels and examples, and the degree to which collocational information is made explicit and accessible to users. A recurring concern in this line of research is the discrepancy between the richness of collocational patterns observed in corpora and the often limited or inconsistent way in which they are encoded in dictionaries. Several shortcomings are identified⁴, suggesting that collocations are overall inadequately represented and that key insights from metalexigraphic discussions are frequently overlooked.

By contrast, over the past decades and particularly with the rise of digital lexicography, a range of high-quality online lexicographic and corpus-based resources for general language (German) has emerged, including *Duden*, DWDS, and *Wortschatz Leipzig*. These resources offer various ways of accessing collocational information. Collocations can be explored through word clouds (as provided by *Duden*, DWDS, and *Wortschatz Leipzig*), static collocation tables (*Wortschatz Leipzig*), and dynamically sortable tables that allow rankings based on different statistical association measures (DWDS). In addition, visual representations of synchronic and diachronic preferred combinations are available, for example through *diacollo* (DWDS). Even if these tools can also be used for general language, they do not always allow for the clear identification of preferred collocational patterns within a specific domain. In specialized contexts, general-language resources may obscure domain-specific usage, as illustrated by cases such as *Rücktritt*, where collocational preferences differ significantly depending on the discourse domain. Typical collocations highlighted there include for example: *vom Rücktritt zurücktreten* (en. to withdraw a resignation), *den Rücktritt erklären* (en. to announce one's resignation), *seinen Rücktritt ankündigen* (en. to announce one's resignation), *nach dem Rücktritt* (en. after the resignation) and *Rücktritt vom Amt* (resignation from office). These typically reflect political and public discourse. Domain-specific usage (LSP examples) is drawn almost exclusively from the law discourse (*vom Vertrag zurücktreten*, *wirksamer Rücktritt*, *Rücktritt wegen Mängeln*⁵), sports discourse (*Rücktritt vom Profisport*, *Rücktritt vom aktiven Leistungssport*, *seinen Rücktritt vom Wettbewerb erklären*⁶), economics context (*Rücktritt des Vorstandsvorsitzenden*, *Rücktritt aus persönlichen Gründen*, *Rücktritt infolge wirtschaftlicher Verluste*⁷). In general-language resources, political collocations are statistically dominant and therefore highly visible,

⁴ A.o.: Collocational information is often organized in an unsystematic manner, which makes it difficult for users to identify relevant patterns; the distinction between base and collocate is frequently left implicit, and the positioning of collocational information within the entry remains unclear or arbitrary; labelling practices and visual highlighting are often inconsistent or absent, reducing the salience of collocations within the dictionary microstructure; meaning explanations and usage examples tend to be conflated without clear functional differentiation, and information on the underlying data basis or sources is frequently omitted.

⁵ In English: to withdraw from a contract; effective withdrawal; withdrawal due to defects.

⁶ In English: retirement from professional sport; retirement from active competitive sport; to announce one's retirement from competition.

⁷ In English: resignation of the CEO; resignation for personal reasons; resignation due to financial losses.

while other domain-specific preferred combinations (e.g. legal *Rücktritt vom Vertrag*⁸) are either underrepresented or not explicitly marked as LSP-specific. As a result, users, especially learners, may struggle to identify which collocations are appropriate in a given professional or disciplinary context.

Specific resources dedicated to collocations are also available for general language, both in print and online formats. These include *Festwortverbindungen des Deutschen* (Francke), the *Wörterbuch der Kollokationen im Deutschen* (de Gruyter), the *Stilwörterbuch* (Duden), the online *Kollokationenwörterbuch* and *Feste Wortverbindungen* (OWID). These resources provide more targeted access to collocational patterns than general-purpose dictionaries and represent an important advancement in the lexicographic treatment of collocations. Nevertheless, the situation largely mirrors that observed in general-language dictionaries and corpus portals. Collocational information is predominantly based on general corpora, with a strong focus on journalistic and public discourse. Using the entry *Rücktritt* as an example, these resources typically highlight collocations such as *den Rücktritt erklären* (en. to announce one's resignation), *seinen Rücktritt ankündigen* (en. to announce one's resignation) or *Rücktritt vom Amt* (en. resignation from office), which are characteristic of political contexts. In contrast, other domain-specific collocations, such as legal expressions like *vom Vertrag zurücktreten* (en. to withdraw from a contract), or *wirksamer Rücktritt* (en. effective rescission) are often underrepresented or not explicitly marked as belonging to a specialized domain. As a result, users cannot easily identify preferred collocational patterns for specific professional or disciplinary contexts. This limitation reduces the usefulness of these otherwise valuable resources for languages for special purposes and poses particular challenges for learners who rely on dictionaries to guide appropriate usage.

With regard to the German-Italian language pair, the situation proves to be more problematic. Although there are high-quality bilingual lexicographic resources that include collocational information, such as *Il nuovo dizionario di tedesco* (from the publisher Zanichelli) and various PONS online dictionaries, collocations are generally not treated in a systematic or consistent manner. In most cases, collocational information is scattered across entries, often mixed with examples or translation equivalents, and rarely made explicit as such. Moreover, domain-specific collocations are only partially represented and are seldom clearly marked, making it difficult for users to identify preferred combinations in specific contexts. This becomes evident when considering *Rücktritt*. In bilingual dictionaries, *Rücktritt* is typically translated as *dimissioni* in political or institutional contexts and as *recesso* in legal contexts. However, the collocational patterns associated with these meanings are rarely made explicit. For instance, German collocations such as *den Rücktritt erklären* (en. to announce one's resignation) or *seinen Rücktritt ankündigen* (en. to announce one's resignation) correspond to Italian expressions like *annunciare le dimissioni* or *rassegnare le dimissioni*, while legal collocations such as *vom Vertrag zurücktreten* (en. to withdraw from a contract) correspond to *recedere dal contratto*. In most bilingual resources, these distinctions are not systematically represented, and collocational information is often embedded in example sentences without clear labelling or domain indication. As a result, users, especially learners and translators, may struggle to identify the appropriate collocations for a given discourse domain. This limitation significantly reduces the usefulness of existing German-Italian dictionaries for languages for special purposes and highlights the need for more structured and domain-sensitive collocational descriptions.

3.1. Tourlex

This situation is particularly challenging for many languages for specific purposes (LSP), such as tourism (Flinz, 2018; Flinz & Perkuhn, 2018; Flinz, 2021). For this domain, specialized dictionaries are largely lacking. In the field of tourism, the only notable resources are the *Langenscheidt*

⁸ In English: withdrawal from the contract.

Praxiswörterbuch Tourismus Italienisch and the *Langenscheidt Online-Fachwörterbuch Tourismus Italienisch*.

Starting from this situation, the Tourlex project aimed to fill this gap. In the planning of the dictionary, the user was considered a central factor in the design of the lexicographical processes (Klosa & Tiberius, 2016, p. 75). To achieve the project's objectives, a combined approach was adopted, employing different types of corpora for different purposes (Flinz, 2021, p. 66). First, a comparable German-Italian specialized corpus was created in *Sketch Engine*, based on a sub-genre of the tourism domain: the *Allgemeinen Reisebedingungen* and *Condizioni generali di contratto di vendita di pacchetto turistico*. From this corpus, a provisional list of lemma candidates was generated, taking into account absolute and relative frequencies from integrated corpora (German Web and Italian Web). Equivalents and collocations were then identified. In a second step, additional reference corpora for German, such as DeReKo and the DWDS corpus, were integrated to extract further examples and contextual information for the target entries. This multi-corpus approach supported several analytical objectives:

- meaning description;
- identification of equivalents;
- detection of domain-specific fixed expressions within tourism language;
- identification of deviating meanings that occur in general language and/or in other specialized languages.

The identification of some domain-specific meanings and collocations was not always straightforward (Flinz & Perkuhn, 2018). For instance, the above discussed German noun *Rücktritt* appears in collocations such as *den Rücktritt erklären*, whose precise meaning must be carefully described and differentiated depending on the context.

The project relied not only on multiple corpora but also on a range of tools and analytical methods to ensure comprehensive coverage and reliable descriptions. The tools employed included *Sketch Engine*, *LexPan*, COSMAS II, and the CCDB DeReKoll platform⁹ to compare the general language and the LSP use (see Figure 1). The methods were adapted to the specific requirements of multilingual LSP lexicography (Flinz, 2021).

⁹ *LexPan* is a tool for querying multilingual parallel corpora. COSMAS II is a corpus analysis system for large German corpora developed by IDS Mannheim. CCDB DeReKoll provides access to the DeReKo corpus, allowing analysis of word collocations and distributional patterns in contemporary German.

Reisende Reiseplan Reisepreis Reiserücktrittsversicherung Reiseteilnehmer Reisetermin Reiseveranstalter Reiseversicherung Reisevertrag Reiseziel Reservierung Restzahlung Rückbeförderung Rücktritt Rücktrittserklärung Rücktrittsrecht Sachschaden Schaden Schadenersatz Schadenersatzanspruch Sicherungsschein Sonderwunsch Stornierung Storno Stornogeühr	Rücktritt Frequenz Ratio¹ syntagmatisches Muster 120 93% hat <u>überraschend</u> seinen Rücktritt [...] erklärt 35 97% hat seinen ihren Rücktritt vom <u>Leistungssport</u> erklärt 57 82% seinen Rücktritt aus der <u>Nationalmannschaft</u> [...] erklärt 3252 87% seinen Rücktritt [...] erklärt 1 100% <u>indirekt</u> ... Rücktritt aufgefördert worden 32 100% <u>indirekt</u> zum Rücktritt aufgefördert 150 99% zum Rücktritt [...] aufgefördert worden 1443 97% zum Rücktritt [...] aufgefördert 88 92% <u>forderte</u> ... den sofortigen Rücktritt des 14 78% <u>verlangte</u> den sofortigen Rücktritt des ... 17 76% <u>verlangen</u> den seinen sofortigen Rücktritt der 921 97% seinen den sofortigen [...] Rücktritt 328 80% hat [...] seinen Rücktritt [...] <u>eingereicht</u> 69 73% auf Ende Jahr seinen Rücktritt [...] <u>eingereicht</u> 1813 97% den Rücktritt [von ...] <u>gefordert</u> 526 78% hat [...] seinen Rücktritt [...] <u>angekündigt</u> 1892 87% <u>forderte</u> [...] den zum Rücktritt auf ... 1269 84% und <u>forderten</u> [...] den Rücktritt des von ... 148 89% seinen Rücktritt [...] als [...] <u>Parteichef</u> 62 96% seinen Rücktritt [...] als [...] <u>Parteivorsitzender</u> 150 88% seinen Rücktritt [...] als [...] <u>Vorsitzender</u> der des ...
--	---

Figure 1. Screenshot of CCDB DeReKoll¹⁰

With DeReKoll (Figure 1), for example, it was possible to compare the use of a lexeme in general language corpora versus specialized corpora, highlighting differences in terminology, phraseology, and usage conventions.

To address the identification of equivalents, the use of ad hoc LSP-comparable corpora has been shown to be particularly useful, as it enables lexicographers to infer domain-specific equivalents from authentic texts in both languages (Varantola, 2003).

These findings were then integrated into the microstructure of the dictionary entries, ensuring that collocations¹¹ are clearly highlighted to support users in recognizing typical word combinations and their correct usage. In the case of *Restzahlung*¹² (en. final payment), the collocations reflect recurrent syntactic and semantic patterns typical of tourism and travel-contract discourse, such as payment modalities, due dates, and institutional procedures. They are organized in a systematic, usage-oriented way, rather than on a purely alphabetical order.

¹⁰ For each syntagmatic pattern, the tool provides: the frequency (the number of occurrences in the corpus), (e.g., *seinen Rücktritt [...] erklärt* appears 120 times), the ratio (the percentage of occurrences following the specific pattern), the syntagmatic pattern (common word combinations with the lexeme, including a.o. verbs, syntactic structures and context-specific phrases).

¹¹ The choice of the label *combinazioni tipiche* (en. typical combinations) instead of collocations, was motivated by user studies conducted with dictionary users (see Flinz, 2018).

¹² See the entry *Restzahlung*: <https://wiki.uni-mannheim.de/tourlex/index.php?title=Restzahlung> (Accessed: 20 January 2026).

Combinazioni tipiche

• Sostantivo - Verbo:

Restzahlung (AKK) abbuchen: addebitare il saldo

Esempi:

Anzahlung und Restzahlung werden zu den in den Ziffern 2.1 und 2.2 angegebenen Fälligkeitsdaten von Ihrem Konto abgebucht. (Olimar 2018)

die Restzahlung (AKK) leisten: corrispondere il saldo, effettuare il saldo

Esempi:

Im Ausnahmefall kann sowohl die Anzahlung als auch die Restzahlung bei Entgegennahme der Reiseunterlagen in bar im Reisebüro geleistet werden. (Olimar 2018)

La misura dell'acconto, fino ad un massimo del 25% del prezzo del pacchetto turistico, da versare all'atto della prenotazione ovvero all'atto della richiesta impegnativa e la data entro cui, prima della partenza, dovrà essere effettuato il saldo, risultano dal catalogo, dall'opuscolo o da quanto altro. (Columbia 2018)

die Restzahlung liegt (nicht) vor: il saldo non è stato eseguito

Esempi:

Liegen Anzahlung und/oder Restzahlung nicht entsprechend den vereinbarten Fälligkeiten vor, so sind wir berechtigt, nach Mahnung mit Fristsetzung vom Reisevertrag zurückzutreten und Sie mit den Rücktrittskosten laut diesen Reisebedingungen zu belasten. (Interchalet 2018)

Il saldo del prezzo si considera avvenuto nel momento in cui perviene materialmente all'organizzatore. (I Viaggi del Turchese 2018)

die Restzahlung ist fällig: dover versare il saldo; il saldo è da versarsi

Esempi:

Die Restzahlung wird 4 Wochen vor Reisebeginn zur Zahlung fällig, sofern der Versicherungsschein übergeben ist und die Reise nicht mehr aus dem in Ziffer 5 genannten Grund abgesagt werden kann. (Piccolonia 2018)

Il saldo da versarsi entro 20 giorni dalla data di partenza, potrà essere corrisposto con carta di credito o con bonifico. (Alpitour 2018)

die Restzahlung ist zahlungsfällig: dover versare il saldo; il saldo è da versarsi

Esempi:

Soweit der Versicherungsschein übergeben ist, und im Einzelfall keine anderweitigen Vereinbarungen getroffen wurden, ist die Restzahlung 3 Wochen vor Reisebeginn zahlungsfällig. (Urscher 2018)

die Restzahlung geht bei jm (DAT) ein: il saldo viene accreditato a qualcuno

Esempi:

Die Anzahlung muss in jedem Fall so rechtzeitig unter Angabe der Rechnungsnummer auf das angegebene Bankkonto überwiesen werden, dass sie innerhalb von 10 Tagen nach Rechnungsdatum, die Restzahlung 28 Tage vor dem Reiseternin bei uns eingeht. (Dertour 2018)

Figure 2. Screenshot of the entry *Restzahlung* illustrating its typical combinations

As can be seen from this example, different types of combinations are identified:

- recurrent syntactic patterns, especially noun + verb combinations (e.g. *eine Restzahlung leisten, abbuchen, eingehen*¹³), which show the typical actions associated with the concept of balance payment;
- copular and stative constructions (e.g. *die Restzahlung ist fällig/zahlungsfällig, die Restzahlung liegt vor*¹⁴), which express the legal or temporal status of the payment;
- fixed terminological pairings, such as *Anzahlung* and *Restzahlung* or *An- und Restzahlung*¹⁵, reflecting conceptually linked payment stages in travel contracts;
- Pragmatic and procedural contexts, including due dates, receipt of payment, and the dispatch of travel documents, which situate the lemma within authentic institutional and contractual discourse.

Lexico-syntactic regularities and communicative functions are made transparent. Each typical combination is highlighted in bold and followed by possible Italian equivalents and by authentic usage examples in both languages, with explicit source references. This organization supports not only comprehension but also accurate production in specialized communication.

4. Collocations in AWS. Didactic perspectives

Allgemeine or *Alltägliche Wissenschaftssprache* (AWS), has been the subject of extensive discussion (cf. among others in German linguistics Ehlich, 1999; Graefen & Moll, 2011; Meißner 2014; Meißner & Wallner, 2018; Rheindorf, 2016). Scientific work is in fact fundamentally connected to everyday language (Ehlich, 1999, p. 7): Beyond the technical terminology of individual disciplines, scientific communication relies on the specific use of elements of everyday language that are adapted to scientific purposes. These linguistic means constitute what Ehlich describes as that which “zwischen den Fachtermini steht” (en. stands between the technical terms, see Ehlich, 1999, p. 8), and they occur both in spoken and written forms of scientific communication.

While Ehlich refers to a traditional conception of technical vocabulary characterized by well-definedness, unambiguity, and context-independence, this view has been relativized and redefined in

¹³ In English: to make the final payment; to debit a final payment; to receive a final payment.

¹⁴ In English: the remaining payment is due; the remaining payment is due; the remaining payment has been received.

¹⁵ In English: deposit and final payment.

more recent research on languages for special purposes (Meißner, 2014). Graefen and Moll (2011, p. 17) describe AWS as: “Wörter und Ausdrucksweisen, die auf das wissenschaftliche Handeln im weiteren Sinne bezogen sind: Forschen, Nachdenken, und Analysieren, Austausch mit anderen über wissenschaftliche Themen. Man spricht hier von der Allgemeinen Wissenschaftssprache oder – noch deutlicher – von der Alltäglichen Wissenschaftssprache“¹⁶.

A key issue associated with AWS is its relative inaccessibility for learners of German as a foreign or second language (DaF/DaZ) (Meißner, 2014). Expressions that are familiar from everyday language are often used in scientific contexts with different or more specialized meanings. Their interpretation depends strongly on co-text and discourse, resulting in polysemy and semantic variability. Moreover, depending on communicative needs, these expressions may appear in different syntactic structures and exhibit varying valency patterns.

An even greater challenge lies in the collocational and partially idiomatically fixed nature of AWS, which becomes apparent in recurring phrases and conventionalized formulations (Graefen, 1999), as exemplified by expressions such as *einen Hinweis geben* (Arnone & Flinz, 2025). These features pose difficulties not only for learners of German as a foreign language, but also for native speakers. Phenomena such as the overlap of subject and object roles, possessive constructions, or semantic shifts of lexemes that have different meanings in general language further contribute to the complexity of AWS (see below *Rahmen*). Although first steps toward addressing AWS in the DaF context have been taken (Graefen & Moll, 2011), a systematic description – especially from a didactic perspective – has largely been lacking so far (Meißner, 2014, p. 55).

In response to this gap, the project COPLUS¹⁷ emphasizes the development of an accessible application for the acquisition of AWS, the promotion of metalinguistic competence (using English as a bridge language), open-access materials tailored to different needs. These include didactic units that combine theory and practice, offering theoretical introductions, examples, bibliographic references, and exercises, as well as curated bibliographic databases, for instance based on Zotero.

The app creation process¹⁸, analogous to the lexicographic process, involved several phases. In the preparation phase, relevant linguistic phenomena for didactic learning units were identified using literature and corpora. Data collection involved compiling primary resources, including ad hoc corpora, while computerization relied on tools such as *Sketch Engine* to structure and manage the data. Data processing produced a provisional entry list (macrostructure), and analysis focused on individual entries (microstructure) and their interconnections (mediostructure). Finally, the app was prepared for online release.

For the selection and description of the central single and multiple lexemes in the domain of AWS, a combined inductive-deductive approach was adopted, based on the systematic evaluation of the *ad hoc* COPLUS corpus, which is a comparable corpus in three languages (DE-EN-IT)¹⁹. Between 2017 and 2024, 30 PhD theses were collected in connection with the project, covering a broad range of disciplines. These include research funded by the European Research Council (ERC) across the Social Sciences and Humanities (SH), Life Sciences (LS), and Physical, Engineering, and Mathematical Sciences (PE), which encompasses mathematics, physical sciences, information and communication

¹⁶ In English: consisting of words and expressions related to scientific activity in a broader sense, such as researching, reflecting, analyzing, and exchanging ideas about scientific topics. Accordingly, it is referred to as General Scientific Language or, more explicitly, Everyday Scientific Language.

¹⁷ See COPLUS: <https://sites.google.com/uniroma1.it/coplus/home> (Accessed: 20 January 2026). The project, led by Sabine E. Koesters Gensini (Sapienza University of Rome), was funded by the Italian Ministry of University and Research (MUR) and the European Union under the NextGenerationEU initiative (PNRR – PRIN: Projects of Significant National Interest, Call 2022, Project No. 2022FXCBM3).

¹⁸ The app has been developed by Sapienza University of Rome and will initially be made available to partner universities, followed by a wider release to other institutions.

¹⁹ The German corpus (COPLUS_DE_2025) contains 2,773,177 tokens, the Italian corpus (COPLUS_IT_2025) 2,286,625 tokens, and the English corpus (COPLUS_EN_2025) 2,861,199 tokens.

technologies, engineering, and universe and earth sciences. This multidisciplinary engagement highlights the diverse applications and impact of the project's methodologies, ensuring both empirical grounding and theoretical alignment with existing specialist literature.

The analysis of multi-word expressions (see Figure 3) also followed this dual approach. Inductively, frequency lists, keyword analyses, and N-gram profiles were used to identify recurring patterns in scientific discourse. Complementing this, collocations described in the research literature were examined deductively and compared with the structures observable in the corpora.



Figure 3. Word cloud of the multiwords entries

These expressions were categorized according to their communicative purpose (see Koesters Ginsini, Costa & Flinz, 2026) for example:

- Observing and describing: expressions such as *einen Blick auf etwas werfen*; *einen Überblick liefern*; *einen Einblick in etwas geben*, used to guide the reader through content;
- Investigating and interpreting: e.g. *Unterschiede herausarbeiten*; *der Unterschied liegt in*; *eine Ausnahme darstellen*, which help clarify distinctions;
- Arguing, providing evidence, and proving: formulations for posing questions, hypotheses, and theories e.g. *die Frage aufwerfen*; *eine Hypothese aufstellen*; *eine These aufstellen*;
- Positioning: e.g. *den Eindruck vermitteln*; *von Bedeutung sein*; *in den Fokus rücken*; *den Schwerpunkt auf etwas legen*; *der Schwerpunkt liegt in*, which enable authors to highlight important points;
- Assuming and deducing: phrases like *in Frage kommen*; *in Betracht ziehen*; *zugrunde legen*, supporting reasoning;
- Attributing and referring: expressions such *im Bezug auf*; *im Folgenden*; *in Hinblick auf*; *im Rahmen der/des*, which allow authors to link statements to texts, contexts, or factual matters.

These categories were organized not only according to their onomasiological characteristics but, where relevant, also according to learner proficiency levels, providing a systematic framework for both corpus analysis and didactic application.

As an example, consider the German noun *Rahmen* (in English framework) which occurs both in general language and within AWS with a different meaning (see figure 4):

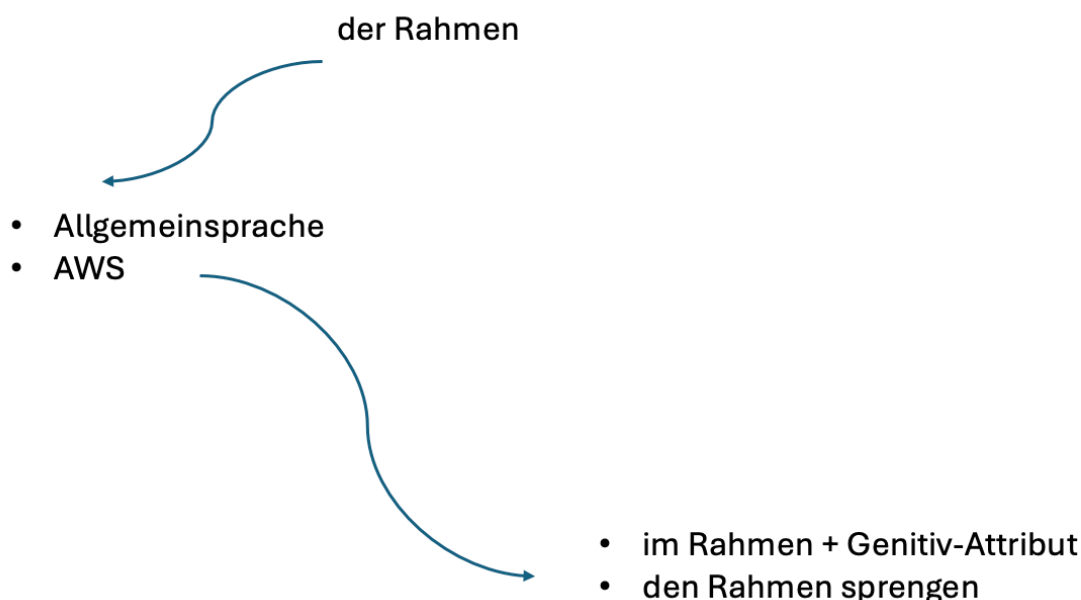


Figure 4. The item *Rahmen*²⁰

In everyday language, *Rahmen* (en. frame) most commonly refers to a physical frame, such as the border of a picture, a door or window frame, or the structural frame of a vehicle. In these uses, the meaning is concrete and easily visualized: for instance (1-2):

- (1) Die Breite des Rahmens muss mindestens 6,35 cm betragen (deTenTen23) (en. The width of the frame must be at least 6.35 cm);
- (2) Dies könnte zu Problemen am Motor und/oder Rahmen führen (deTenTen23) (en. This could cause problems with the engine and/or the frame).

In scientific or specialized contexts, however, *Rahmen* (en. context, framework) is used figuratively to denote a conceptual framework or context that defines the scope, boundaries, or structure of a study, discussion, or particular field. This figurative use extends the general meaning of frame into an abstract, organizational sense, providing a framework for thought, research, or argumentation.

Examples from the COPLUS 2025 corpora illustrate this usage:

- (3) Auch dieser Frage soll im Rahmen der Untersuchung nachgegangen werden (COPLUS_DE_2025) (en. This question will also be examined in the context of the study);
- (4) Diese sollen im Rahmen dieser Arbeit auf ihre Praktikabilität untersucht werden (COPLUS_DE_2025) (en. these are to be analysed for their practicability in the context of this work).

²⁰ Cf. Ponzi (2027).

Morphosyntactically, *Rahmen* is a masculine noun following strong declension. While it appears in both singular and plural forms in general language, in AWS it is predominantly used in the singular. Related lexical items include partial synonyms such as *Kontext*, *Zusammenhang*, and *Hintergrund* (en. context), and partial antonyms such as *Fokus* and *Vordergrund* (en. focus, foreground).

Rahmen exemplifies how lexical meaning can shift from a concrete, physical sense in everyday language to a figurative, abstract sense in scientific discourse. Its study inside the COPLUS-Project demonstrates the importance of understanding both syntagmatic combinations and context-specific usage when compiling lexicographic resources or teaching specialized language.

Typical constructions include *im Rahmen + genitive* as *im Rahmen der Untersuchung* (en. in the context of the study) and *im Rahmen von* (en. in the context of). More figurative examples include *den Rahmen sprengen* (en. to exceed the scope, to go beyond the scope), demonstrating idiomatic flexibility. Building on this figurative meaning, the idiomatic expression *den Rahmen sprengen* combines *Rahmen* with the transitive verb *sprengen* (en. to blow up, break) to indicate exceeding conceptual, thematic, or procedural limits. Literally, it would mean ‘to blow up the frame’, but in context it conveys going beyond the intended scope of a project, study, or discussion. In general language, it often refers to exceeding financial, temporal, or other practical limits, e.g., *Die finanzielle Belastung sprengt den Rahmen* (en. The financial burden exceeds the limits).

In AWS, *den Rahmen sprengen* (en. to exceed the scope, to go beyond the scope)²¹ is used metaphorically to describe exceeding the scope of academic work or analysis. Corpus examples include:

- (5) Daher entfällt die Diskussion der Literaturdaten, da diese den Rahmen dieser Arbeit sprengen würden (COPLUS_DE_2025) (en. Therefore, the discussion of the literature data is omitted, as it would exceed the scope of this work)
- (6) Ihre umfassende Ausführung würde jedoch den Rahmen dieser Dissertationsschrift sprengen. (COPLUS_DE_2025) (en. A comprehensive discussion of it, however, would go beyond the scope of this dissertation)
- (7) Dem Anspruch nachzukommen, das soziale Profil aller Diskursteilnehmer detailliert in den Deutungshorizont miteinfließen zu lassen, würde allerdings den Rahmen dieser Arbeit sprengen. (COPLUS_DE_2025) (en. However, meeting the requirement to incorporate in detail the social profile of all discourse participants into the interpretive framework would exceed the scope of this work.)

Syntactically, *den Rahmen sprengen* is usually singular and frequently followed by a genitive noun phrase specifying the work or project (*den Rahmen dieser Arbeit*, *den Rahmen dieser Studie*, *den Rahmen dieser Dissertationsschrift*²²). Although the passive is possible, it is rare; usage in the conditional mood (Konjunktiv II) with *würden* or *hätte* is most common for hypothetical or potential scenarios. The expression occurs in both main and subordinate clauses and may include modifiers or demonstratives (*den vorgegebenen Rahmen*, *diesen Rahmen*²³).

Partial synonyms include *den Rahmen überschreiten* (en. to exceed the frame/limits), reinforcing the notion of going beyond established boundaries. Importantly, the figurative AWS meaning of *Rahmen* and *den Rahmen sprengen* derives from the literal sense of physical frames: just as a frame defines spatial boundaries, the conceptual *Rahmen* defines the intellectual or thematic limits of a discussion, study, or project.

Overall, the study of *Rahmen* and *den Rahmen sprengen* highlights the interplay between literal and figurative meaning, the importance of syntagmatic combinations, and the role of context in understanding and teaching specialized language. This entry illustrates how corpus-based analysis, as

²¹ Cf. Flinz (2027).

²² In English: the framework of this paper, the framework of this study; the framework of this dissertation.

²³ In English: the given framework, this framework.

in the COPLUS 2025 datasets, provides empirical grounding for lexicographic description and supports learners in navigating idiomatically and syntactically complex expressions.

These examples demonstrate how fixed syntagmatic patterns combine with grammatical structures to convey precise contextual and functional meanings in scientific and formal discourse.

5. Summary

As we have seen, collocations represent a highly complex and multifaceted aspect of language (cf. Steyer, 2023), posing specific challenges for learners of German as a foreign language, particularly in specialized domains. These challenges arise from differences in meaning across contexts, as well as syntactic divergences between general and specialized language. The systematic analysis of corpora, such as in the TOURLEX and COPLUS projects, combined with a dual inductive-deductive approach, has enabled to identify recurring patterns, verify established collocations, and provide empirically grounded descriptions for both general and specialized language.

Ensuring that these findings are effectively conveyed to users requires careful design and presentation of lexicographic resources, including digital applications, interactive tools, and structured didactic materials. Moreover, ongoing contact with target users is essential to align the resources with their practical needs, support language acquisition, and facilitate the comprehension and production of idiomatically and syntactically complex expressions. Through the integrated use of different types of corpora and the methodologies developed in the two projects, this approach demonstrates how corpus-based research and applied lexicography can jointly address the challenges of collocation learning in specialized language contexts.

Literature

- Arnone, C., & Flinz, C. (2025). Risorsse lessicografiche e corpora per la ricerca e la didattica della AWS tedesca: il caso di *Einen Hinweis auf etwas geben*. *Italiano LinguaDue*, 17(2), 229–243. <https://doi.org/10.54103/2037-3597/30546>
- Belica, C., & Perkuhn, R. (2015). Feste Wortgruppen / Phraseologie I: Kollokationen und syntagmatische Muster. In U. Haß & P. Storjohann (Hrsg.), *Handbuch Wort und Wortschatz* (pp. 201–225). De Gruyter.
- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. John Benjamins.
- Bubenhof, N. (2009). *Sprachgebrauchsmuster: Korpuslinguistik als Methode der Diskurs- und Kulturanalyse* (Sprache und Wissen, 4). De Gruyter.
- Bubenhof, N. (2017). Kollokationen, n-Gramme, Mehrworteinheiten. In K. S. Roth, M. Wengeler, & A. Ziem (Hrsg.), *Handbuch Sprache in Politik und Gesellschaft* (pp. 69–93). De Gruyter.
- Buhlmann, R., & Fearn, A. (2018). *Handbuch des fach- und berufsbezogenen Deutschunterrichts: DaF, DaZ, CLIL* (Forum für Fachsprachen-Forschung, 141). Frank & Timme.
- Burger, H. (2015). *Phraseologie: Eine Einführung am Beispiel des Deutschen* (5th ed.). Erich Schmidt.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238.
- Đurčo, P. (2016). Zum Konzept der Kollokationsdidaktik und des Kollokationslernens bei Germanistikstudenten. In P. Đurčo (Ed.), *Kollokationsforschung und Kollokationsdidaktik* (pp. 147–172). LIT.
- Đurčo, P., Vajičková, M., & Tomášková, S. (2019). *Kollokationen im Sprachsystem und Sprachgebrauch: Ein Lehrbuch*. Kirsch-Verlag.
- Efing, C. (2014). Berufssprache & Co.: Berufsrelevante Register in der Fremdsprache. *Informationen Deutsch als Fremdsprache*, 41(4), 415–441. <https://doi.org/10.1515/infodaf-2014-0403>
- Ehlich, K. (1999). Alltägliche Wissenschaftssprache. *Informationen Deutsch als Fremdsprache*, 26(1), 3–24. <https://doi.org/10.1515/infodaf-1999-0102>
- Evert, S. (2009). Corpora and collocations. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics: An International Handbook* (Vol. 2, pp. 1212–1248). Mouton de Gruyter.
- Feilke, H. (1996). *Sprache als soziale Gestalt*. Suhrkamp.
- Festwortverbindungen des Deutschen. (n.d.). *Festwortverbindungen des Deutschen*. Francke.
- Firth, J. R. (1957). A synopsis of linguistic theory, 1930–1955. *Studies in Linguistic Analysis (Special volume of the Philological Society)*, 1–32.
- Flinz, C. (2015). *Projekt zur Erstellung eines Online-Fachwörterbuches der Linguistik (Deutsch-Italienisch)* (European University Studies, 384). Peter Lang.

- Flinz, C. (2018). Tourlex: Ein deutsch-italienisches Fachwörterbuch zur Tourismussprache. *Lexicographica*, 34(1), 9–36.
- Flinz, C. (2019). Fachsprachen – aktuelle Fragen zu Forschung und Lehre. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 24(1), 1–20.
- Flinz, C. (2021). Korpora als primäre Quellen von Tourlex. In J. Taborek, M. Piosok & M. Woznicka (Hrsg.), *Korpora in der Lexikographie* (Lexicographica. Series Maior, pp. 57–83). De Gruyter.
- Flinz, C. (2024). *Linguistica dei corpora: Una nuova prospettiva teorico-metodologica per lo studio di morfologia e sintassi in ambito DaF*. Milano University Press. <https://doi.org/10.54103/milanoup.165>
- Flinz, C. (2027). Den Rahmen sprengen. Portale COPLUS. La comunicazione tecnico-specialistica plurilingue. Ricerca linguistica e progettazione di un percorso didattico digitale per il Tedesco (L3) attraverso l'Inglese (L2) <https://sites.google.com/uniroma1.it/coplus/home>
- Flinz, C., & Perkuhn, R. (2018). Wortschatz und Kollokationen in „Allgemeine Reisebedingungen“. In J. Čibej et al. (Eds.), *Proceedings of the XVIII EURALEX International Congress* (pp. 959–966). Ljubljana University Press.
- Flinz, C., & Pietzuch, J. P. (2024). Fach- und berufssprachliches Handeln: Zwischen Theorie und Praxis. *Muttersprache* 134(2), 105–111. (Sonderheft, C. Flinz & J. P. Pietzuch (Hrsg.), *Fach- und berufssprachliches Handeln*). <https://doi.org/10.53371/61154>
- Giacoma, L. (2014). Quello che i dizionari possono fare. In A. Abel, C. Vettori, & N. Ralli (Eds.), *Proceedings of the XVI EURALEX International Congress: The User in Focus* (pp. 551–562). Institute for Specialised Communication and Multilingualism.
- Giacoma, L. (2015). Die Entwicklung der Mikrostruktur in zweisprachigen Wörterbüchern. In J. Calañas & F. Robles Sabater (Hrsg.), *Wörterbücher des Deutschen* (pp. 175–198). Peter Lang.
- Gledhill, C. (2000). *Collocations in science writing*. Narr.
- Graefen, G. (1999). Wie formuliert man wissenschaftlich? In H. Barkowski & A. Wolff (Hrsg.), *Alternative Vermittlungsmethoden* (pp. 222–239). FaDaF.
- Graefen, G., & Moll, M. (2011). *Wissenschaftssprache Deutsch: lesen – verstehen – schreiben. Ein Lehr- und Arbeitsbuch*. Frankfurt am Main: Lang.
- Hausmann, F. J. (1984). Wortschatzlernen ist Kollokationslernen. *Praxis des neusprachlichen Unterrichts*, 31(4), 395–406.
- Heier, A., & Hallsteinsdóttir, E. (2024). Interkulturelle Kompetenz in der beruflichen Kommunikation. *Muttersprache*, 134(2), 119–131. (Sonderheft, C. Flinz & J. P. Pietzuch (Hrsg.), *Fach- und berufssprachliches Handeln*). <https://doi.org/10.53371/61156>
- Hoffmann, L. (1976). *Kommunikationsmittel Fachsprache*. De Gruyter.
- Hoffmann, L., Kalverkämper, H., & Wiegand, H. E. (Hrsg.). (1998/1999). *Fachsprachen / Languages for Special Purposes*. De Gruyter.
- Il nuovo dizionario di tedesco. (n.d.). *Il nuovo dizionario di tedesco: Italiano-Tedesco / Tedesco-Italiano*. Zanichelli / Klett.
- Klosa, A., & Tiberius, C. (2016). Der lexikographische Prozess. In A. Klosa & C. Müller-Spitzer (Hrsg.), *Internetlexikografie. Ein Kompendium* (pp. 65–110). De Gruyter.
- Koesters Gensini, S., Costa, M. & Flinz, C. (2026). Von Korpora zur Lernapp: Das COPLUS-Projekt zur Vermittlung der allgemeinen Wissenschaftssprache des Deutschen als Drittsprache. In A. Ziegler et al. (Hrsg.), *Sprache und Literatur in Krisenzeiten. Herausforderungen, Aufgaben und Chancen der internationalen Germanistik*. Beiträge des XV. IVG-Kongresses in Graz 2025 (Publikationen der Internationalen Vereinigung für Germanistik (IVG)). Graz University Library Publishing.
- Konecny, C. (2010). Von ‘hinkenden’ Stühlen, ‘tanzenden’ Zähnen und ‘verlorenen’ Verkehrsmitteln. Erfassung und Darstellung italienischer lexikalischer Kollokationen für deutschsprachige L2-Lerner (auf der Grundlage des *Dizionario di base della lingua italiana – DIB*). In A. Dykstra & T. Schoonheim (Eds.), *Proceedings of the XIV EURALEX Congress* (pp. 1207–1221). Fryske Akademy – Afûk.
- Konecny, C. (2011). Divergenze e convergenze in collocazioni lessicali italiane e tedesche. In E. Lavric, W. Pöckl & F. Schallhart (Eds.), *Comparatio delectat. Akten der VI. Internationalen Arbeitstagung zum romanisch-deutschen und innerromanischen Sprachvergleich, Innsbruck, 3.-5. September 2008* (Teil 1, pp. 295–309). Peter Lang.
- Konecny, C. (2018). Valenza e coesione collocazionale. In S. Dallabrida & P. Cordin (Eds.), *La grammatica delle valenze* (pp. 143–161). Cesati.
- Langenscheidt Praxiswörterbuch Tourismus Italienisch. (n.d.). *Praxiswörterbuch Tourismus Italienisch*. Munich: Langenscheidt.
- Meißner, C. (2014). *Figurative Verben in der allgemeinen Wissenschaftssprache*. Stauffenburg.
- Meißner, C., & Wallner, F. (2018). GeSIG. In W. P. Klein et al. (Hrsg.), *Vernakuläre Wissenschaftskommunikation* (pp. 457–471). De Gruyter.
- Mellado Blanco, C. (Ed.). (2008). *Colocaciones y fraseología en los diccionarios*. Peter Lang.
- Mollica, F. (2023). Collocazioni nella lessicografia bilingue. In E. Autelli et al. (Hrsg.), *Dialektale und zweisprachige Phraseographie* (pp. 241–268). Stauffenburg.

- Mollica, F., & Steyer, K. (2024). Korpusdaten und Phraseodidaktik. *Zeitschrift Korpora Deutsch als Fremdsprache*, 4(1), 1–7. <https://doi.org/10.48694/kordaf.4000>
- Nied Curcio, M. L. (2006). La lessicografia tedesco-italiana. In F. San Vicente Félix (Ed.), *Lessicografia bilingue e traduzione* (pp. 57–70). Polimettrica.
- Nied Curcio, M. L. (2012). Die Valenz in der zweisprachigen Lexikographie. *Studi Germanici*, 1(12), 175–191.
- Pecman, M. (2007). Approche onomasiologique de la langue scientifique générale. *Revue française de linguistique appliquée*, 12(2), 79–96.
- Ponzi, M. F. (2027). Der Rahmen. Portale COPLUS. La comunicazione tecnico-specialistica plurilingue. Ricerca linguistica e progettazione di un percorso didattico digitale per il Tedesco (L3) attraverso l'Inglese (L2) <https://sites.google.com/uniroma1.it/coplus/home>
- Rheindorf, M. (2016). Die Figurativität der allgemeinen Wissenschaftssprache des Deutschen. *Linguistik Online*, 76(2), 177–195. <https://doi.org/10.13092/lo.76.2820>
- Roelcke, T. (2017). Dynamisierung – Differenzierung – Dezentralisierung. Tendenzen beruflicher Kommunikation im Deutschen am Beispiel der Verordnung über die Berufsausbildung zum Kraftfahrzeugmechatroniker und zur Kraftfahrzeugmechatronikerin. *Glottology* 8(2), 155–170.
- Roelcke, T. (2020). *Fachsprachen* (4th ed.). Erich Schmidt.
- Roelcke, T. (2024). Deutsch als Fremd- und Fachsprache. *Muttersprache* 134(2), 112–118 (Sonderheft, C. Flinz & J. P. Pietzuch (Hrsg.), *Fach- und berufssprachliches Handeln*). <https://doi.org/10.53371/61155>
- Schafroth, E., Mollica, F., & Mellado Blanco, C. (2021). Kollokationen – kontrastiv, didaktisch, theoretisch. In E. Schafroth et al. (Hrsg.), *Kollokationen* (pp. 7–24). Peter Lang.
- Schafroth, E. (2011). Caratteristiche fondamentali di un learner's dictionary italiano. *Italiano LinguaDue*, 3(1), 23–52.
- Steyer, K. (2013). *Usuelle Wortverbindungen*. Narr Francke Attempto.
- Steyer, K. (2023). Dynamik, Adaptation, Unschärfe. Zur Vielschichtigkeit usueller Wortverbindungen. *Deutsch als Fremdsprache*, 60(4), 201–212.
- Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. John Benjamins.
- Tribble, C. (2000). Genres, keywords, teaching: Towards a pedagogic account of the language of academic writing. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective* (75–90). Peter Lang.
- Tutin, A. (2007). Modélisation linguistique et annotation des collocations : application au lexique transdisciplinaire des écrits scientifiques. In S. Koeva et al. (Eds.), *Formaliser les langues avec l'ordinateur : de INTEX à Nooj* (189–215). Presses universitaires de Franche-Comté.
- Varantola, K. (2003). Translators and disposable corpora. In F. Zanettin, S. Bernardini, & D. Stewart (Eds.), *Corpora in translator education* (pp. 55–70). St. Jerome.
- Wiegand, H. E. (1989). Die lexikographische Definition. In F. J. Hausmann et al. (Hrsg.), *Wörterbücher / Dictionaries / Dictionnaires* (Vol. 1, pp. 409–462). De Gruyter.
- Wörterbuch der Kollokationen im Deutschen*. De Gruyter.

Online-Resources

- Deutsches Referenzkorpus (DeReKo)*, Leibniz-Institut für Deutsche Sprache (IDS). <https://www.ids-mannheim.de/digspra/kl/projekte/korpora/> (Accessed: 20 January 2026).
- Duden. *Wörterbuch der deutschen Sprache*. Duden Online. <https://www.duden.de/woerterbuch> (Accessed: 20 January 2026).
- Digitales Wörterbuch der deutschen Sprache (DWDS)*. <https://www.dwds.de/> (Accessed: 20 January 2026).
- Feste Wortverbindungen. Online-Wortschatz-Informationssystem Deutsch (OWID)*. <https://www.owid.de/> (Accessed: 20 January 2026).
- IDS Mannheim (2023a). *LexPan* [Corpus tool]. Available at: <https://www.ids-mannheim.de/lexpan> (Accessed: 20 January 2026).
- IDS Mannheim (2023b). *COSMAS II* [Corpus tool]. Available at: <https://www.ids-mannheim.de/cosmas-ii> (Accessed: 20 January 2026).
- IDS Mannheim (2023c). *CCDB DeReKoll* [Corpus tool]. Available at: <https://www.ids-mannheim.de/derekoll> (Accessed: 20 January 2026).
- Kollokationenwörterbuch*. <https://kollokationenwoerterbuch.ch/> (Accessed: 20 January 2026).
- Online-Fachwörterbuch Tourismus: Italienisch-Deutsch*. Munich: Langenscheidt. <https://it.langenscheidt.com/> (Accessed: 20 January 2026).
- OpenAI. (2024). *ChatGPT* (version GPT-4). Large language model. <https://chat.openai.com/> (Accessed: 20 January 2026).

PONS Online. <https://it.pons.com/traduzione> (Accessed: 20 January 2026).

Lexical Computing Ltd. (2023). *Sketch Engine* [Corpus tool]. Available at: <https://www.sketchengine.eu> (Accessed: 20 January 2026).

Stilwörterbuch (Duden). Duden Online. <https://www.duden.de/woerterbuch> (Accessed: 20 January 2026).

Tourlex. <https://wiki.uni-mannheim.de/tourlex/index.php?title=Hauptseite> (Accessed: 20 January 2026).

Wortschatz Leipzig. Universität Leipzig. <https://wortschatz.uni-leipzig.de/> (Accessed: 20 January 2026).