

# Developing a German-English Dictionary of the Common Language of Academia

Markus Rheindorf  
University of Vienna, Austria

Birgit Huemer  
University of Vienna, Austria

## Abstract

The factual dominance of English as *Lingua Franca* of the international scientific community continues to increase in most disciplines. University students, lecturers and researchers are more and more commonly required to acquire English for academic purposes. Meanwhile, linguistic research into the similarities and differences between academic English and its German, Italian or other counterparts has been conducted. However, no comprehensive study has yet dealt with the specific problems faced by German-speaking academics and students acquiring English for academic purposes. It is in this context that this article presents an ongoing research project designed to address that gap. It concerns the lack or partial lack of equivalent linguistic means between the German and the English common language of academia. The common language of academia is, by definition, used across the disciplines and contains or realises fixed, language- and culture-specific conceptions of what actually constitutes 'science' or 'academia'. It is largely based on but differs from language use in everyday contexts, thus posing significant challenges for students even in their first language. In this article we address the challenge of achieving not only a functional comparison of the common language of academia in English and German, but of complementing the underlying inventory and analysis of each. Following a discussion of previous work, we detail our research design and criteria for the comparative analysis of the common language of academia in English and German, as well as some preliminary results of the project. The findings of our project will ultimately lead to a German-English dictionary for the common language of academia designed to support both students and academics with understanding and participating in international scientific communities with different linguistic backgrounds.

## Introduction

As the debate over English as *Lingua Franca* of the international scientific community rages on, its factual dominance continues to rise in most disciplines. This development has been noted and variously decried or hailed for several decades (e.g. Phillipson 1992). National languages of science, such as German, Italian or Spanish, meanwhile have been pushed more and more into the niches of specialized forms of publication. For over a decade, a similar development has extended into the domain of teaching at university: undergraduate students are more and more commonly required to acquire English for academic purposes in their Bachelor studies (Ammon and McConnell 2002).

Positions on this development and suggested responses vary greatly. They range from warnings – for instance of the loss of national research traditions or the possibilities to popularize research – to calls for multilingual publication practices and language policies on a European level. Meanwhile, linguistic research into the similarities and differences between academic English and its German, Italian or other counterparts has been conducted. However, no comprehensive study has yet dealt with the specific problems faced by German-

speaking academics and students acquiring English for academic purposes. The latter often face the challenge of learning to use both languages simultaneously. We know that this entails more than the common problems of language acquisition, including the specific language- and culture-based differences in text structure and organisation, paragraphing, verb focus and nominalisation, compounding, lexis and register, metonymy etc. (Swales 1990, Bazerman 1991, Ventola and Mauranen 1996, Nesi et al. 2004 and Nesi and Gardner 2006).

Such differences between English and German academic texts have been studied largely from a contrastive perspective (Clyne 1987 and Clyne and Kreuz 1987). Unlike a comparative approach, such a perspective is ill-suited for didactic applications – for instance within Writing-Across-the-Curriculum programmes aimed at German-speaking students. The necessary shift in perspective has been slow to take hold since the late 1990s. As noted by Thielmann (1999) with respect to the comparative analysis of academic languages, a contrastive perspective relies on ethnocentrically determined parameters that must be overcome by an analytic shift to the functions of linguistic means.

It is in this context that this article identifies and delineates a gap in such comparative, application-oriented research that has been noted before (Ehlich 2002, Fandrych 2002 and Steinhoff 2009) and presents an ongoing research project designed to address that gap. It concerns the lack or partial lack of equivalent linguistic means between German and English academic language, henceforth referred to as 'equivalence problems'. This is, above all, a lexical issue in the domain of the common language of academia (or German 'Allgemeine Wissenschaftssprache'), henceforth referred to as CLA. This domain is crucial in the acquisition of academic writing competence and hence in the didactics of teaching academic writing, particularly in a multilingual context. CLA is, by definition, used across the disciplines and contains or realises fixed, language- and culture-specific conceptions of what actually constitutes 'science' or 'academia' (Ehlich 1994: 334). It is largely based on, but differs from language use in everyday contexts, thus posing significant challenges for students even in their first language (Kruse and Battaglia 1998: 193 and Thielmann 2009: 49). Acquiring academic writing competence in a second language is thus made all the more challenging by equivalence problems between the two languages (Vossen 1985, Hyland 2003, 2004 and Pearson-Casanave 2004). While some studies (Belcher 1995 and Leki 2007) have analysed the challenges faced by non-native students acquiring English for academic purposes, the significance of the students' respective first language has been largely ignored. Interference and, specifically, equivalence problems in CLA thus remain to be studied.

While research into the register of academic English and its teaching has been extensive but remains conceptually fragmented and without any comparative interest, research into the register of academic German has produced rather isolated analyses of limited scope but yielded conceptually promising, comparative approaches (Kretzenbacher and Weinrich 1995). We thus face the task of not only achieving a functional comparison of CLA in English and German, but of complementing the underlying inventory and analysis of each. Following a discussion of previous work, we will present our research design and criteria for the comparative analysis of CLA in English and German, as well as some preliminary results of the project.

## **Groundwork: A Common Language of Academia**

The term 'common language of academia' is based on the observation that terminology does not equal 'the language of academia', prominent as it may be. Unlike discipline-specific language, the common register of academia is closely connected to the 'elevated vocabulary' used and taught in secondary education (Schepping 1976). Unlike the specific terminology of individual disciplines, it is clearly bounded and much less productive, generally not open to topical innovation, not tied to a specific subject and displays a high rate of recurrence. CLA thus provides the lexical foundation of the entire range of academic communication (Strauß and Zifonun 1985: 84).

Referring to the rudimentary inventory established by Erk (1972), Schepping concluded that German CLA consists of a limited number of lexical items and recurrent phrases that function as the necessary basis for more specialized terminology in all disciplines. It represents a repertoire of linguistic means to communicate about academic work, processes and texts, relating to actions such as comparing, studying, analysing, concluding, assuming, naming, reasoning and explaining (Schepping 1976: 21). Similarly, Ehlich (1993: 33) defines CLA as the fundamental linguistic means used by most disciplines in identical or similar fashion, with identical or similar meaning and frequency. It relates academic fields of knowledge to each other, to experiential reality, but also to academia as a field of practices (Ehlich 1998: 856-858; 1999: 9-11). In the case of German, therefore, the register of CLA developed from and retained close ties with common language (rather than Latin).

This development, however, depended on a shift of meaning in the usage of common, everyday expressions in the academic context, often creating challenges even for German-speaking students, but especially for non-native speakers of German (Johnen 2008: 4). Such challenges are not easily overcome with the help of either monolingual or bilingual dictionaries, because they do not explicate and distinguish the specific meanings of CLA. Discipline-specific dictionaries are equally inadequate in this respect as they are, by definition, limited to the specific terminology of the respective field (Johnen 2008: 5). Acquiring discipline-specific terminology is generally less challenging, as it is clearly defined. A second language's CLA, however, cannot simply be learned by extension. It comprises many 'false friends' (Johnen 2008: 8), metaphorical and figurative borrowings from everyday language (Graefen 2009: 150), idiomatic phrases (Feilke 1998), diverging and specialised register and specific conceptions of what constitutes academic work (Fandrych 2002). It is these factors that cause equivalence problems, leading in turn to infidelities, imprecision and unintended meanings in academic writing.

Few equivalence problems between German and English CLA are as trivial as the false friend *actual* for German *aktuell*. German CLA contains many items that are derived from everyday language and now possess extended, figurative meanings – e.g. *heranziehen*, *herausarbeiten* or *beleuchten* – that are difficult to translate or even approximate because English CLA has developed other figurative fields. The verb *herausarbeiten* is an illustrative example; the English translation *work out* given by most dictionaries demonstrates that even when parallel phrases exist, they need not have developed the same or any academic meaning at all. Fandrych (2002: 2) suggests that such divergences stem from the fact that the respective CLA of a national language will encode central conceptualisations of what it means to do academic work, i.e. conduct research, interpret data and achieve understanding. Ehlich (1995: 343) also suggests that present-day CLA contains 'sediments' of distinct eras of scientific pursuit, such as the Hellenic or Scholastic traditions, in the form of recurrent words and phrases. Interlingual equivalence problems on the level of CLA are thus not simply issues of stylistics or collocation, but relate to diverging traditions. In many instances, there may not be a direct equivalence; our focus must therefore be to find functional equivalents or approximations, discuss limitations and point out irreducible discrepancies. In terms of application, this requires a special kind of dictionary outlined at the end of this article.

CLA as previously defined and studied is largely understood as manifest on the lexical level, even though 'fundamental linguistic means' can be taken to include tenses, aspect or thematic progression (Ehlich 1995, 1999 and Graefen 1999, 2000). Even more narrowly, most studies of German CLA or comparisons of German and English CLA have focused on verbs or verb phrases, arguably because these present the most dissimilar and thus most challenging aspect. It may also be due to the fact that German CLA is dominated by verbs, i.e. performative verbs in figurative use (Graefen 1997, Hund 1999 and Fandrych 2004). Nonetheless, a comprehensive comparative study must include nouns, adjectives, adverbs and adverbials (Graefen 2009) as well as phrasal constructions.

Previous work has shown that the development of German CLA is closely tied to the shift from Latin to national languages in academia and began in the 17<sup>th</sup> century and extended well into the 18<sup>th</sup> century, with protagonists such as Giordano Bruno, Martin Luther and Leibniz drawing on but also systematically extending and adapting everyday German vocabulary.

While they also used existing Latin and Greek terms, their usage was strongly limited (Ehlich 2000: 15). On a purely lexical level, English CLA is strikingly different inasmuch as common English at that point in history had already incorporated substantial Latin-Romance vocabulary (Fandrych 2002: 25). Studies of the lexical dimension of English CLA have so far – particularly within Writing Across the Curriculum – focused on a distinct range or inventory of items (e.g. Coxhead 2000). Frequency across the disciplines alone, we would argue, is not a sufficient criterion for studying CLA, nor for compiling an inventory. Conceptually and analytically, a comparative study must rely on robust operators defining criteria for what is ‘common’ across the disciplines and of what is ‘distinct’ from everyday language, as well as on identifying the functional-semantic fields that CLA has evolved to cover.

### **Approaches to the Inventory: Corpora**

Previous research into German CLA is limited by its fragmentation in terms of methods, categories and data, as well as by disparate foci and corpus size<sup>1</sup>. While linguistic analyses such as Ehlich (1995, 1998, 1999, 2000, 2002), Fandrych (2002, 2004), Graefen (1999, 2000, 2009) and many others (e.g. Meißner 2009) must be considered valuable groundwork, they are limited in scope, focusing on a small number of verbs or groups such as reporting verbs. We thus have studies on segments of German CLA rather than an integrative, comprehensive picture as called for by Ehlich (2000: 18). Taken together, these limitations indicate a substantial gap in our understanding of CLA as well as the lack of foundation for establishing a bilingual dictionary of the common language of academia.

Unlike previous work, we use large corpora and an integrative analytical approach based on previous work to compile an inventory of German and English CLA (see below). The fundamental question of what to include in such an inventory has so far not been answered; rather, researchers have relied on their experience as academics and their intuition as native speakers, using the corpora to study the use of items already selected. The first criterion has to be the usage of a given item across the disciplines (Ehlich 1993: 33). From a corpus-linguistic point of view, this characteristic needs to be verified in two ways: (1) occurrence in all or most disciplines distinguished and (2) similar frequency of use across the disciplines.

### **Comparative Analysis: Issues of Categorisation**

Previous work has used widely differing categories and parameters to study CLA. Because of the wide range of processes, actions and textual aspects covered by CLA, categorisation is crucial and has been a preoccupation of earlier work (Steinhoff 2009: 101). For us, that meant reviewing and adapting existing suggestions to find a categorisation of functional-semantic fields for a comprehensive and comparative analysis.

Schepping (1976: 24ff) distinguished two groups of verbs: (1) logical-methodological vocabulary (logical deduction, linguistic-terminological representation) and (2) strategic-rhetorical text elements with predominantly meta-communicative and meta-discursive function (indicating progression: introduction, transition, conclusion; modi of argumentation: evidence, comparison etc.; emphasis, focus and reference). In a rudimentary fashion, Schepping's categories already anticipate Thielmann's (2009) call for a comparative analysis of CLA using categories relating linguistic means to communicative purposes.

A substantial part of the words analysed by Schepping (1976) and his successors are verbs, complemented by a handful of adjectives, adverbs and nouns – the latter clearly have to be dealt with differently and systems of categorisation developed to describe only verbs need to be revised or complemented. Regarding verbs, CLA comprises mostly reporting verbs and performative verbs denoting actions characteristic of academic practice. While many of these

---

<sup>1</sup> In Fandrych (2002), e.g., the corpus consisted of 17 English- and 19 German-language papers.

also occur in everyday language with related meanings, their specific meaning and usage in academic contexts is not easily identified or translated (see *herausarbeiten* or *ausarbeiten* vis-à-vis *work out*, which has not developed an academic usage). Many of these verbs, e.g. *zeigen*, *offenlegen*, *deuten* or *bearbeiten*, do not denote a communicative practice in common German. In other words, we are dealing with figurative usage that is seen as encoding and manifesting crucial conceptions of scientific and academic practice (Hund 1999: 308ff). Such verbs, in particular, have been the subject of attempts at categorisation, e.g. in terms of types of action: 'commenting on text' and 'discussing literature' (Fandrych 2002: 4) or 'movement', 'positioning', 'transferring', 'connecting', 'grasping', 'showing' and 'perception' (Meißner 2009: 101). These categories work well within the analyses of limited scope, but cannot capture the full range of meanings encoded by CLA.

Such narrow categorisation has sparked criticism and suggestions for a more comprehensive view (Fandrych 2002 and Steinhoff 2009). An integrative set of categories to be used for all verbs of CLA (and their nominalisations) would need to consider:

- Scientific and academic practices<sup>2</sup>
  - Perception and attention
  - Insight, recognition, understanding
  - Giving evidence or proof
  - Conceptual practices<sup>3</sup>
  - Logical and practical action in research
- Scientific and academic communication
  - Describing and presenting
  - Mentioning
  - Express and define
  - Report and address
  - Suggest
  - Pose questions
  - Argue, explain, justify
  - Summarise
  - Emphasise
  - Illustration, Figurative Representation<sup>4</sup>
    - Broach or sketch
    - Exemplify or Demonstrate
    - Visual representation
  - Indicate or point out
  - Refer to text topic
    - Text as space
    - Text as chronology
    - Highlight
  - Critical discussion<sup>5</sup>
  - Reception and integration<sup>6</sup>

---

<sup>2</sup> The verbs in this category are ambiguous inasmuch as they relate to mental or cognitive as well as communicative practices. To write that *Smith views* or *focuses on something* is to tell readers both that she engages in a particular mental practice (of interpreting, of concentrating) and that she communicates that practice in language or text.

<sup>3</sup> The category of 'spatial-physical practices' is subsumed here and other categories such as 'logical and practical action in research' and 'describing and presenting'; it is problematic as a label as it focuses on the everyday meaning of terms like *herausarbeiten* or *bearbeiten* rather than their actual use in CLA, which must be the primary focus of categorisation.

<sup>4</sup> The category of 'visual representation' as it is sometimes used is misleading. Actual figures, images etc. are not to be confused with verbs that allude to visual representation, such as *sketch* or *skizzieren*, but actually denote speech acts.

<sup>5</sup> This category includes various linguistic means of conveying controversy and debate, positioning and stance-taking.



As pointed out by Steinhoff (2008: 87) the focus on lexical items alone is not sufficient to grasp the domain-specific usage of many items. Collocations must be included in the analysis using sufficiently large corpora to show, for instance, whether a particular verb is used to comment on the same text or discuss literature. In addition, stylistic variation and preferences may mean that certain collocations are preferred over others with equivalent meanings and functions (Steinhoff 2007: 88). In particular where it refers to scientific or academic practice, CLA is determined by collocational patterns. This includes noun-verb collocations such as *einer Frage nachgehen* or the pattern of *im Folgenden*, which is used for text organisation at transitional places (Steinhoff 2009: 100).

It has also been observed that some of the challenges faced in acquiring CLA stem from the fact that many items of CLA are derived from everyday language, i.e. words that are used outside academia and without any academic meaning. It follows that CLA on the lexical level is not so much 'specific' as it is 'typical', i.e. we can expect the lexical items of CLA and their preferred collocations to be more frequent in but not exclusive to academic discourse (Adamzik et al. 1997, Feilke and Steinhoff 2003 and Feilke 2010). In terms of corpus linguistics, this means that relative frequency rather than exclusive occurrence needs to be used as defining criterion for compiling an inventory.

## Research Design

Our project consists of three main stages:

- Stage 1 Inventory: compile CLA in German and English
- Stage 2 Categorisation and Matching: establish functional equivalences and identify problems
- Stage 3 Dictionary: build a dictionary based on a functional comparison

In Stage 1, to inventory CLA in German and English we use a functional approach to corpus linguistics. In order to address the shortcomings of previous work outlined above, we use two comparable corpora of written academic texts totalling 39.8 million tokens. For German, we use the corpus of the Berlin-Brandenburg Academy of Sciences and Humanities, accessible online as DWDS (the Digital Dictionary of German Language), which is the only publicly available representative corpus of German academic texts. The academic corpus comprises 24.371.647 tokens from a variety of disciplines. For English, we use the British National Corpus maintained by Mark Davies at Brigham Young University, which comprises 15.431.668 tokens across the disciplines. Both corpora are suitable because of their size and accessibility, but also meet the criteria of Tschirner (2005) and Jone and Tschirner (2005) for 'corpus-based word frequency analysis'. Both corpora allow for the differentiation of disciplines (social sciences, humanities/arts, technical science, natural science and medicine), tracking collocations and word types.

The following items will be included in the inventory: verbs (reflexive verbs, verbal bracket), nouns, adjectives, conjunctions and phrases. Verbs, nouns, adjectives and conjunctions appear in different realisations in the context of CLA. Verbs can be realised, e.g., as indicative, reflective or participle. Nouns are either used as simple noun or compound, in singular or plural. Adjectives are used as adjectives or adverbs, comparative or superlative.

To provide an inventory of CLA as defined, several filtering routines are necessary to exclude

---

<sup>6</sup> The verbs of this category comprise various means of positioning, but represent more positive aspects of drawing on and incorporating the work of others. Elements like *According to Miller* and *Müller zufolge*, which represent perfunctory relations to the respective source, need to be included in the analysis as well. Indeed, nouns, adjectives and adverbials share many of the functions represented by the verb categories listed here.

items not specific to or inappropriate in academic contexts. We developed filtering routines using statistical operators for the following items:

1. Function words like pronouns and modal verbs have been excluded, because they can be reliably translated and have no specific academic meaning.
2. Discipline-specific words have been excluded, because they are not part of CLA and well defined in specialized dictionaries.
3. Non-academic words have been excluded, because they are part of other language domains and have no specific academic meaning or usage.

Filtering of function words was easily accomplished as they are clearly defined in both languages. For the filtering of discipline-specific words, i.e. words entirely specific to or with a specific use in one discipline (or group of disciplines), filtering routines had to be developed using statistical operators. Firstly, we compared the word frequencies across the disciplines to identify words that are significantly more/less frequent in a single discipline (or group). To measure this, we used the statistical operators mean, median, mean deviation and the coefficient of mean deviation. Table 1 shows the usage of the words *law* and *show* in the disciplines Law (Law), Medicine (Med), Humanities and Arts (HumArts), Social Sciences (SoSc) and Engineering (Engi). Clearly, *law* is dominant (306,4) in the discipline Law and is therefore defined as discipline-specific and excluded from the inventory. For less obvious cases, e.g. *show*, the coefficient of mean deviation served as a second indicator. The cut-off point for this operator was chosen at the value 1. If the indicator for a word is below the limit, the word is included in the inventory. If the indicator exceeds the limit, the respective word is excluded. As the indicator for *show* (0,58) is below the limit, *show* is included in our inventory.

**Table 1: Filtering routine for discipline-specific words (law, show)**

Word	Total	per mio.	Indicator (coefficient)	Law	Med	HumArts	SoSc	Engi	Median	Mean dev.
law	11930	778,1	2,1	306,4	5,1	51,8	46,6	25,3	36,0	76,8
show	16024	1045,2	0,58	70,4	442,8	120,2	338,2	140,8	277,2	121,1

The filtering of non-academic words is more challenging. We identify non-academic words by comparing frequency and collocations to the non-academic reference corpora. Relative frequency is our first indicator. Table 2 shows that German *Tag* is twice as frequent in non-academic discourse, whereas *Frage* is nearly equally frequent in academic and non-academic discourse.

**Table 2: Filtering routine for non-academic words**

Word	Frequency/per mio. (academic)	Frequency/per mio. (non-academic)
<i>Tag</i>	7544 / 309,54	65053 / 660,81
<i>Frage</i>	13432 / 551,15	49157 / 499,34

In ambiguous cases, e.g. *Frage*, we also look at similarity/difference in collocations as a second indicator. If equal to or more than 1/3 of the most frequent collocations are different, we assume an academia-specific usage of the word. As shown in Figure 1, more than 1/3 of the most frequent collocations for *Frage* differ.

acad. coll. for <i>Frage</i>	frequency of bigramme	non-acad. coll. of <i>Frage</i>	frequency of bigramme
wichtige	118 (4,84)	deutsche	1412 (14,34)
andere	104 (4,27)	wichtige	975 (9,90)
entscheidende	62 (2,54)	solche	808 (8,21)
politisch	61 (2,50)	politische	765 (7,77)
Römische	54 (2,22)	ganze	750 (7,62)
schwierige	51 (2,09)	internationale	688 (6,99)
erste	50 (2,05)	unsere	673 (6,84)
praktische	46 (1,89)	große	653 (6,63)
letzte	45 (1,85)	beide	625 (6,35)
philosophische	44 (1,81)	kommende	513 (5,21)
soziale	44 (1,81)	offene	505 (5,13)
offene	41 (1,68)	soziale	499 (5,07)
gestellte	40 (1,64)	neue	483 (4,91)
große	40 (1,64)	folgende	482 (4,90)
grundsätzliche	40 (1,64)	wirtschaftliche	459 (4,66)
weitere	39 (1,60)	allgemeine	429 (4,36)
aufgeworfene	38 (1,56)	europäische	395 (4,01)
deutsche	38 (1,56)	letzte	380 (3,86)
allgemeine	35 (1,44)	verschiedene	324 (3,29)
zweite	34 (1,40)	grundsätzliche	317 (3,22)
solche	32 (1,31)	nationale	314 (3,19)
neue	31 (1,27)	entscheidende	288 (2,93)
folgende	30 (1,23)	gute	274 (2,78)
religiöse	30 (1,23)	besondere	269 (2,73)
umstrittene	30 (1,23)	mögliche	265 (2,69)
ganze	28 (1,15)	aktuelle	264 (2,68)

Figure 1: Academic and non-academic collocations

### **Showing equivalences through categorisation and matching**

In the second stage we categorize the inventoried items according to semantic-functional fields (see above) and match lexical items of each field across English and German. This allows us to show functional equivalences, overlaps, ambiguities and gaps. For instance, *herausarbeiten* can be used to express elaboration, conceptual development and explication; but in each field, the English equivalent would be a different lexical item, phrase or usage, e.g., develop, elaborate, explicate. The items *herausarbeiten* and *develop* thus match only on one level of their meaning and usage (see Fig. 2).



<p><b>Deutlich machen</b></p> <p>“Im folgenden Abschnitt werden die Gemeinsamkeiten herausgearbeitet.”</p>	<p><b>Clarify</b></p> <p>“... bring out clearly.”</p>	<p><b>Self-sustained Process</b></p> <p>“The economic situation is developing rapidly...”</p>
<p><b>Konzeptuelles Handeln</b></p> <p>“In ihrer jüngsten Publikation hat sie diesen Ansatz weiter herausgearbeitet.”</p>	<p><b>Conceptual Activity</b></p> <p>“... developed further.” “... elaborated further.”</p>	<p><b>Conceptual Activity</b></p> <p>“She develops her argument more fully...”</p>
<p><b>Erklären, Argumentieren</b></p> <p>“Es wurden die Gründe für diese politische Entscheidung herausgearbeitet.”</p>	<p><b>Explain, Argue</b></p> <p>“... explicated the reasons...”</p>	<p><b>Production Process</b></p> <p>“Database system are being developed for various applications...”</p>

Figure 2: Categorisation and matching

Categorisation and matching allow us to identify which words are equivalent or non-equivalent in both languages and show the diversity within a specific functional field and alternatives in each CLA. This sheds light on equivalence problems between German and English and helps to identify solutions.

### Designing the dictionary

The third stage will implement this research in the form of a dictionary designed to support both German-speaking students acquiring academic English and researchers writing academic papers in English. It may also help in the rare instance that English-speaking students and researchers write academic papers in German. Each dictionary entry shall consist of the word, word type, usage – indicating usage in metonymical constructions or not – the most common collocations, respective functional field(s) as well as possible translations including common collocations in the target CLA and false friends (see Fig. 3).

<p><i>herausarbeiten</i> Verb (metonym.)</p> <p>eine Grundlage, einen Zusammenhang, eine Bedeutung, einen Unterschied, einen Begriff, Gegensatz oder Punkt (aus einem Text, aus den Daten) herausarbeiten</p> <p>Bezugsfeld: wissenschaftliche Handlung – konzeptuelles bzw. logisches/ praktisches Forschungshandeln bzw. Erkenntnisprozess/ Beweisführung bzw. wiss. Sprechhandlung – sprachlich fassen bzw. argumentieren; siehe auch <i>herausfinden, bestimmen, eruieren</i></p> <p>Englisch: <i>to develop sth.</i>, betonend/kontrastiv: <i>to accentuate</i>; fokussierend/ausführend: <i>to detail or elaborate on sth.</i>; klärend: <i>bring out clearly</i>; erklärend: <i>explicate</i>. Von den hier angeführten englischen Wörtern hat <i>develop</i> den Vorteil, dass der Arbeitscharakter und das Prozesshafte der Bedeutung deutlich wird, jedoch verliert man dabei die Plastizität der Vorstellung des deutschen <i>herausarbeiten</i>. Diese ergibt sich aus der Verbindung des Handlungsverbs <i>arbeiten</i> mit der räumlichen Angabe <i>heraus</i>, wodurch eine Bewegung zur/zum LeserIn hin vermittelt wird.</p>
--

Figure 3: Dictionary entry

### Preliminary results

We have so far completed the inventory by word type for verbs, adjectives and nouns for German and English (comprising 4.100 entries). The inventory has been filtered for function words as well as for discipline-specific words. We still have to complete the inventory for

conjunctions and phrases, the filtering for non-academic words and usage, the categorisation and matching as well as translations and collocations.

By way of conclusion, we want to stress that the design parameters as well as the underlying research design and methodology can also be used as a model for developing dictionaries of CLA in other languages than German and English. Thus, a clear perspective for the future is to extend the dictionary to include other academic languages and build a multilingual dictionary for academic use.

## References

- Adamzik, K., Antos, G., and Jakobs, E.-M. (eds.) (1997) *Domänen- und kulturspezifisches Schreiben*. Frankfurt/Main: Lang
- Ammon, U., and McConnell, G. (2002) *English as an Academic Language in Europe. A Survey of its Use in Teaching*. Frankfurt: Lang.
- Bazerman, C. (1991) 'The Second Stage in Writing Across the Curriculum'. *College English* 53 (2), 209-212
- Belcher, D. (1995) 'Writing critically across the curriculum'. in *Academic Writing in a Second Language: Essays on Research and Pedagogy*. ed. by Belcher, D., Braine, G. Norwood, NJ: Ablex, 135-154
- Clyne, M. (1987) 'Cultural Differences in the Organisation of Academic Texts'. *Journal of Pragmatics* 11, 211-247
- Clyne, M., and Kreutz, H. (1987) 'The nature and function of digression and other discourse structure phenomena in academic German'. *Working Papers for Migrant and Intercultural Studies* 8, 1-22
- Coxhead, A. (2000) 'A New Academic Word List'. *TESOL Quarterly* 24 (2), 213-238
- Ehlich, K. (1993) 'Deutsch als fremde Wissenschaftssprache'. *Jahrbuch Deutsch als Fremdsprache* 19, 13-42
- Ehlich, K. (1994) 'Die Lehre der deutschen Wissenschaftssprache: sprachliche Strukturen, didaktische Desiderate'. in *Linguistik der Wissenschaftssprache* ed. by Kretzenbacher, H. and Weinrich, H. Berlin/New York: de Gruyter, 325-352
- Ehlich, K. (1998) 'Kritik der Wissenschaftssprachen'. in *Fachsprachen. Ein internationales Handbuch*. ed. by Hoffmann, L., Kalverkämper, H., Wiegand, H. E. Berlin/New York: de Gruyter, 856-866
- Ehlich, K. (1999) 'Alltägliche Wissenschaftssprache'. *InfoDaF* 26, 3-24
- Ehlich, K. (2000) 'Deutsch als Wissenschaftssprache für das 21. Jahrhundert'. *GFL* 1/2000 [online] available from <<http://www.gfl-journal.de/1-2000/ehlich.html>> [13 March 2014]
- Ehlich, K. (2002) 'Wissenschaftskomparatistik.' in *Mehrsprachige Wissenschaft – europäische Perspektiven*. ed. by Ehlich, K. München [online] available from <<http://www.euro-sprachenjahr.de>> [13 March 2014]
- Erk, H. (1972) *Zur Lexik wissenschaftlicher Fachtexte I: Verben – Frequenz und Verwendungsweise*. München: Hueber
- Fandrych, C. (2002) 'Herausarbeiten vs. Illustrate: Kontraste bei der Versprachlichung von Sprechhandlungen in der englischen und deutschen Wissenschaftssprache'. in *Europäische Wissenschaft: Europäische Perspektiven*. ed. by Ehlich, K. München: [online] available from <<http://www.euro-sprachenjahr.de>> [13 March 2014]
- Fandrych, C. (2004) 'Bilder vom wissenschaftlichen Schreiben. Sprechhandlungsausdrücke im Wissenschaftsdeutschen. Linguistische und didaktische Überlegungen'. *Materialien Deutsch als Fremdsprache* 73, 269-291
- Feilke H., Steinhoff, T. (2003) 'Zur Modellierung der Entwicklung wissenschaftlicher Schreibfähigkeiten'. in *Wissenschaftlich schreiben - lehren und lernen*. ed. by Ehlich,

- K., and Steets, A. Berlin: de Gruyter, 112-128
- Feilke, H. (1998) 'Idiomatische Prägung'. in *Zwischen Grammatik und Lexikon. Forschungen zu einem Grenzbereich*. ed. by Barz, I. and Öhlschläger, G. Tübingen: Narr, 69-81
- Feilke, H. (2010): 'Aller guten Dinge sind drei – Überlegungen zu Textroutinen & literalen Prozeduren' [online] available from <[www.festschrift-gerd-fritz.de/files/feilke\\_2010\\_literale-prozeduren-und-textroutinen.pdf](http://www.festschrift-gerd-fritz.de/files/feilke_2010_literale-prozeduren-und-textroutinen.pdf)> [13 March 2014]
- Graefen, G. (1997) 'Wissenschaftssprache – eine Thema für den Deutsch-als-Fremdsprache-Unterricht?'. *Materialien Deutsch als Fremdsprache* 47, 31-44
- Graefen, G. (1999) 'Wie formuliert man wissenschaftlich?' in *Alternative Vermittlungsmethoden und Lernformen auf dem Prüfstand*. ed. by Barkowski, H., and Wolff, A. Regensburg: FaDaF, 222-239
- Graefen, G. (2000) 'Einführung in den Gebrauch der Wissenschaftssprache'. in *Wie schwer ist die deutsche Sprache wirklich?* ed. by Wolff, A., Winters-Ohle, E. Regensburg: FaDaF, 191-210
- Graefen, G. (2009) 'Versteckte Metaphorik – Ein Problem im Umgang mit der fremden deutschen Wissenschaftssprache'. in *Wissenschaftliche Textsorten im Germanistikstudium deutsch-italienisch-französisch kontrastiv*. ed. by Dalmas, M., Foschi, M. and Neuland, E. Deutsch-Italienisches Zentrum Centro Italo-Tedesco Via Giulio Vigoni, 149-166
- Hund, E. (1999) 'naheliegend oder weithergeholt – unterrichtspraktische Überlegungen zu Raum und Körper in der Wissenschaftssprache'. in *Alternative Vermittlungsmethoden und Lernformen auf dem Prüfstand*. ed. by Barkowski, H. and Wolff, A. Regensburg: FaDaF, 290-320
- Hyland, K. (2003) *Second Language Writing*. Cambridge: Cambridge University Press.
- Hyland, K. (2004) *Genre and Second Language Writing*. Ann Arbor: University of Michigan Press.
- Johnen, T. (2008) 'Alltägliche Wissenschaftssprache im studienbegleitenden Fachsprachenunterricht'. in *Tagungsakten des VI. Brasilianischen Deutschlehrerkongresses*. ed. by Dill, C., Dietrich, G., Kaufmann, G., Alt, L., Silva, R., and Theis, W. São Paulo: Abrapa/Goethe-Institut, 1-13
- Jones, R., and Tschirner, E. (2005) *Frequency dictionary of German*. London: Routledge
- Kruse, O., Battaglia, S. (1998) 'Trockenübung Hausarbeit: Wissenschaftliches Schreiben im Studium'. in *Handbuch Studieren: Von der Einschreibung bis zum Examen*. ed. by Kruse, O. Frankfurt: Campus, 193-205
- Leki, I. (2007) *Undergraduates in a Second Language: Challenges and Complexities of Academic Literacy Development*. Mahwah, NJ: Lawrence Erlbaum
- Meißner, C. (2009) 'Figurative Verben in der alltäglichen Wissenschaftssprache des Deutschen. Eine korpuslinguistische Pilotstudie'. *Apples – Journal of Applied Language Studies* 3/1, 93-116
- Nesi, H., Sharpling, G., and Ganobcsik-Williams, L. (2004) 'Student papers across the curriculum: Designing and developing a corpus of British student writing'. *Computers and Composition* 21, 439-450

- Nesi, H., Gardner, S. (2006) 'Variation in Disciplinary Culture: University Tutors' Views on Assessed Writing Tasks'. in *Language, Culture and Identity in Applied Linguistics*. ed. by Kiely, R., Rea-Dickens, P., and Woodfield, H. London: Equinox Publishing, 99-117
- Pearson-Casanave, C. (2004) *Controversies in Second Language Writing. Dilemmas and Decisions in Research and Instruction*. Ann Arbor: University of Michigan Press
- Phillipson, R. (1992) *Linguistic imperialism*. Oxford: Oxford University Press
- Schepping, H. (1976) 'Bemerkungen zur Didaktik der Fachsprache des Deutschen als Fremdsprache'. in *Beiträge einer Arbeitstagung an der RWTH Aachen vom 30. September bis 4. Oktober 1974*. ed. by Rall, D., Schepping, H., and Schleyer, W. Bonn: DAAD, 13-34
- Schlieben-Lange, B., Kreuzer, H. (1983) 'Probleme und Perspektiven der Fachsprachen- und Fachliteraturforschung'. *Fachsprache und Fachliteratur* 51/52, 7-26
- Steinhoff, T. (2007) *Wissenschaftliche Textkompetenz. Sprachgebrauch und Schreibentwicklung in wissenschaftlichen Texten von Studenten und Experten*. Tübingen: Niemayer
- Steinhoff, T. (2009) 'Alltägliche Wissenschaftssprache und wissenschaftliche Textprozeduren. Ein Vorschlag zur kulturvergleichenden Untersuchung wissenschaftlicher Texte'. in *Wissenschaftliche Textsorten im Germanistikstudium deutsch-italienisch-französisch kontrastiv*. ed. by Dalmas, M., Foschi, M., and Neuland, E. Deutsch-Italienisches Zentrum Centro Italo-Tedesco Via Giulio Vigoni, 97-10
- Strauß, G., Zifonun, G. (1985) *Die Semantik schwerer Wörter im Deutschen. 1 Lexikologie schwerer Wörter*. Tübingen: Gunter Narr Verlag
- Swales, J. (1990) *Genre Analysis: English in Academic and Research Settings*. New York: Cambridge University Press
- Thielmann, W. (1999) 'Begründungen versus advance organizers - Zur Problematik des Englischen als lingua franca der Wissenschaft'. *Deutsche Sprache* 4/99, 370-378
- Thielmann, W. (2009) 'Wissenschaftliches Sprechen und Schreiben an deutschen Universitäten'. in *Wissenschaftliche Textsorten im Germanistikstudium deutsch-italienisch-französisch kontrastiv*. ed. by Dalmas, M., Foschi, M., and Neuland, E. Deutsch-Italienisches Zentrum Centro Italo-Tedesco Via Giulio Vigoni, 47-54
- Tschirner, E. (2005) 'Korpora, Häufigkeitslisten, Wortschatzerwerb'. in *Deutsch als Fremdsprache – Konturen und Perspektiven eines Faches*. ed. by Heine, A., Hennig, M., and Tschirner, E. München: Iudicium, 133-149
- Ventola, E., and Mauranen, A. (eds.) (1996) *Academic Writing: Intercultural and Textual Issues*. Amsterdam, Philadelphia: Benjamins
- Vossen, H. (1985) 'Die Bedeutung umgangssprachlicher Lexik in Fachvorlesungen am Beispiel Chemie'. in *Lexik im Fremdspracherwerb. Sprachlehrveranstaltungen für Postgraduierte*. ed. by Wolff, A., and Hornung, W. Regensburg: FaDaF, 180-201