# Metaphor Annotation for German

**Markus Egg, Valia Kordoni**

Humboldt-Universität zu Berlin

Unter den Linden 6, 10117 Berlin

{markus.egg, evangelia.kordoni}@hu-berlin.de

## Abstract

The paper presents current work on a German corpus annotated for metaphor. Metaphors denote entities or situations that are in some sense similar to the literal referent, e.g., when *Handschrift* 'signature' is used in the sense of 'distinguishing mark' or the suppression of hopes is introduced by the verb *verschütten* 'bury'. The corpus is part of a project on register, hence, includes material from different registers that represent register variation along a number of important dimensions, but we believe that it is of interest to research on metaphor in general. The corpus extends previous annotation initiatives in that it not only annotates the metaphoric expressions themselves but also their respective relevant contexts that trigger a metaphorical interpretation of the expressions. For the corpus, we developed extended annotation guidelines, which specifically focus not only on the identification of these metaphoric contexts but also analyse in detail specific linguistic challenges for metaphor annotation that emerge due to the grammar of German.

**Keywords:** metaphor, register, annotation

## 1. Introduction

In this paper, we present our work on a German corpus annotated for metaphor. The corpus is compiled and annotated in a project on the interdependence of metaphors and register. Its parts represent a number of different registers, but we expect the corpus to be of interest to research on metaphor in general.

The paper is structured as follows. After a short outline of the underlying theoretical concepts and a review of previous work, we describe the composition, compilation, and annotation of the corpus in detail. We also present first results of the annotation, which show clear qualitative and quantitative differences in the metaphor usage in the different registers in the corpus. These differences are then correlated with specific properties of registers.

## 2. Theoretical background

This section introduces the two phenomena of metaphor and register, and the way in which they are related.

### 2.1. Metaphor

Metaphors use an expression to refer to an entity that is similar to the referent of the literal interpretation of the expression. Metaphor theories reconstruct this similarity in different ways (for an overview see e.g. the chapters 2-4 in Ritchie (2013)). In the following, we will assume the framework of Conceptual Metaphor Theory (CMT; Lakoff (1993)) and describe concrete specimens of metaphors in this framework, but the work presented in this paper does not presuppose a specific metaphor theory.

Following CMT, we base the similarity between the literal and the metaphorical referent on an underlying mapping that maps the structure of a 'source domain', of which the literal referent is a part, onto a 'target domain', which comprises the entity denoted in the metaphorical interpretation of the respective expression. The effect of this mapping is a transfer from properties of the literal referent to the referent of the metaphorical interpretation, which introduces the similarity between the two referents.

E.g., the mapping in (1) goes from containers to mental structures. This mapping transfers to mental structures such as thoughts the property that they can have a content just like physical containers.

(1)  in meinen Gedanken
'in my thoughts'

Metaphors can appear on different morphosyntactic levels. While most of them appear on the level of words, sometimes whole phrases are metaphorical as a whole, like the verb phrase in (2). Also, especially in compounds, only parts of words can be metaphorical, e.g., in (3), where the second part of the compound is used in an abstract sense:

(2)  im Weg stehen
'obstruct' (literally, 'stand in the way')

(3)  Geschäftsfelder
'areas of business' (literally, 'fields of business')

Sometimes several metaphors are based on the same mapping from a source to a target domain, which is called extended metaphor (also known as metaphoric chain). For instance, as soon as the word *Licht* 'light' is introduced as a metaphor for hope, other words from the same source domain like *anzünden* 'enkindle' or *Kerze* 'candle' can be understood metaphorically, too (as 'introduce hope' and 'source of hope', respectively).

Metaphors can be assigned a degree of conventionalisation, ranging from innovative to fully conventionalised. For example, *vorbeirasen* 'rush by' would count as an innovative and non-conventionalised metaphor:

(4)  das letzte Jahr ist nur so vorbei gerast
'the last year has rushed by'

For the purpose of annotation, we will distinguish between conventionalised and non-conventionalised metaphors, see Section 3.

## 2.2. Register

Register refers to the influence of situational and functional context on intra-individual variation (Biber and Conrad, 2009). Systemic-Functional Linguistics (SFL) decomposes register into field, tenor, and mode (Halliday and Hasan, 1985). 'Field' refers to the nature of an interaction that involves language, including subject matter and purpose of the interaction. 'Tenor' targets the participants, in particular, their statuses and social relationships.[1] Finally, 'mode' is about the role of text or discourse in the interaction, e.g., its organisation and properties (such as orality vs. literacy or a monologic vs. a dialogic nature).

Metaphors are alternatives to reference via literal expressions, hence, they are optional ways of referring to an entity. This allows intra-individual variation in establishing reference to be influenced by – and to influence – the situational and functional embedding of a discourse, viz., register. Thus, the choice between (co-referring) literal expressions or metaphors can contribute to establishing a specific register or indicate compliance with it.

This interdependence of metaphors and registers is due to the fact that the function of a metaphor depends on the situational embedding of the discourse it is a part of (Goatly, 1994; Goatly, 2011). Consider for instance the role of tenor: The function of metaphors is tied to the relations of the interlocutors, e.g., building and maintaining rapport between peers, and explanation when experts communicate with non-experts. Such differences can result in different realisations of the metaphors. For instance, Skorczynska and Deignan (2006) and Deignan et al. (2013) report that explicating a metaphorical mapping in the form of a simile ('A is like B') is more likely in expert-non-expert communication than in the exchange between peers, and that source domains are more concrete in the first kind of communication.[2]

## 3. Previous work

The interdependence between metaphor and register was approached in previous work from different perspectives. The first perspective focuses on the function of metaphors. Different functions were found, including 'framing', which exploits the fact that metaphors highlight specific aspects of (target) domains to influence an audience's view on these domains (Musolff, 2016).

Other work investigates metaphor in specific registers, e.g., academic discourse (Littlemore, 2001; Herrmann, 2015; Beger, 2015), fiction (Dorst, 2015), newspapers (Krennmayr, 2011) or educational discourse (Cameron, 2003).

Finally, functions of metaphors were correlated with SFL features of metaphors (Goatly, 1994; Goatly, 2011; Steen et al., 2010). E.g., the latter claim that informational registers (news, fiction, or academic discourse) use metaphor to express content to a much larger extent than conversation.

Berber Sardinha (2015) investigates the influence of metaphor-related features in multidimensional approaches to register variation. He uses such features (e.g., metaphor density or metaphors with specific source domains like perception) to calculate dimensions of register variation.

Gerald Steen and his associates created the VU Amsterdam Metaphor Corpus (VUAMC), which consists of 187,000 words taken from the British National Corpus (BNC). The corpus, which was annotated for metaphors, comprised four registers (academic discourse, newspaper texts, fiction, and conversations) (Steen et al., 2010, ch. 10). The annotation also extends to so-called 'metaphor flags', linguistic material that indicates metaphors but is not in itself metaphorical, e.g., the preposition *like* or 'scare quotes'.

There are also a 100,000 word corpus of Dutch (half conversations and half newspaper texts) (Pasma, 2012) and a corpus of scripted sitcom conversations of 16,500 words (Skorczynska and Giménez-Moreno, 2017) annotated along the same lines.

Shutova and Teufel (2010) and Shutova et al. (2013) annotated first verbs and then tokens from further parts of speech in a corpus of 13,700 words (taken from the BNC) according to whether they were used in a metaphorical or a non-metaphorical sense. In addition, for the metaphorical senses, they annotated source and target domain, using a subset of the metaphorical mapping listed in the 'Master Metaphor List' (Lakoff et al., 1991). The corpus contains material from different text types, for which they report different frequencies of metaphors, in particular, a very low frequency of metaphor in spoken language.

Bizzoni and Lappin (2018) compiled a corpus of 200 sets of five sentences, in which one sentence contains a metaphor (noun, verb, adjective, or multi-word expression), and the other four sentences are evaluated with respect to the degree to which they paraphrase the first sentence. Zayed et al. (2020) created a corpus of around 1,500 verb-direct object combinations in which the verb is used metaphorically, and include interpretations of these combinations. Dipper et al. (2021) are preparing a corpus of German sermons annotated for metaphor.

Steen et al. (2010) developed detailed guidelines for the annotation of their corpus. They address the question of how to identify metaphors. The basic idea is that the context-based sense of an expression must be different from another, more 'basic' sense of the expression (e.g., one which is more concrete or related to bodily action), which is currently in use. These senses must be related in terms of similarity but not subsumable under a common, more general hypernym. E.g., the contextual temporal sense of *vorbeirasen* 'pass quickly' in (4) shares with its spatial, basic sense the element of speed.

Senses are defined as synchronically relevant when they appear in a suitable dictionary. Concretely, Steen et al. (2010) used the *Macmillan English Dictionary for Advanced Learners* and the *Longman Dictionary of Contemporary English*. If both senses appear in the dictionaries, the metaphor is conventionalised, if only the basic sense does, the metaphor is non-conventional. The inverse phenomenon (only the non-basic sense is listed in the dictionary) counts as a case of a 'dead' metaphor, which is no

---

[1]Note that in metaphor research this term is sometimes used to refer to non-literal, metaphorical meaning.

[2]For instance, both kinds of communication used 'flow' as a source domain, but whereas this domain was introduced exclusively by the word *flow* in peer communication, expert-non-expert communication used a variety of words for this task, including *bottle up* and *trickle*.

longer metaphorical.

These guidelines were applied to data in a wide range of languages in order to adapt the guidelines to these languages (Nacey et al., 2019), including German (Herrmann et al., 2019). There are as yet no publicly available corpora annotated for metaphors resulting from these activities, however.

# 4. Our approach

## 4.1. The corpus

For the purpose of investigating the interdependence between metaphor and register, we are compiling a corpus that intends to integrate a wide range of different registers variation, where 'range' is defined as variation in register dimensions defined in previous work.

The corpus consists of five parts, parliament speeches from the German Parlamentsreden-Korpus (Blaette, 2017), newspaper commentaries, which include the Potsdam Commentary Corpus (Stede, 2004), sermons (compiled by contacting the authors individually), light fiction (available online under CC licence)[3] , and debates from debating competitions of the organisation 'Jugend debattiert' (Kemmann, 2013). Each subcorpus will eventually comprise 100,000 words, which we expect will suffice for quantitative evaluations, since we assume approx. 10-15% metaphorically relevant words in the corpora, in parallel to the results of Steen et al. (2010) and Dorst (2015), which is confirmed by our initial results (see Section 5.).

| subcorpus | hierarchical/ equal | distant/ close | oral/ literal | dialogue/ monologue |
|---|---|---|---|---|
| speeches | E | D | L | M |
| sermons | H | C | L | M |
| commentaries | H | D | L | M |
| light fiction | E | C | L | M |
| debates | E | D | O | D |

Table 1: SFL register properties of the subcorpora

Table 1 shows the distribution of SFL register properties in the corpus. We first vary two dimensions of *tenor*, hierarchy vs. equality and distance vs. closeness. The second factor that we took into account belongs to *mode*; we include spoken and written registers in the corpus. Relevant for the spoken/written dimension is the domain of 'conceptual literality/orality' in the model of Koch and Oesterreicher (1994). Hence, both speeches and sermons are classified as literal (they are prepared and fixed in advance), despite their oral presentation. Debates are conceptually oral and dialogic whereas the other registers are monologic.

The subcorpora also represent variation along two important Biber (2009) dimensions (Table 2). For 'situation-dependent vs. elaborated reference' (how dependent is reference on the situational context), we expect that commentaries and speeches relate to concrete extralinguistic situations and individuals, whereas debates and sermons are more abstract deliberations, and fiction is highly detached

from reality. Thus, the anticipated level of situation dependence for reference is low for fiction, medium for debates and sermons, and high for commentaries and parliamentary speeches. As regards 'overt expression of persuasion', the expected level is high for debates, sermons, and commentaries, moderate for speeches (whose influence on actual decision making in politics is low), and low for light fiction.

| subcorpus | reference | persuasion |
|---|---|---|
| speeches | + | o |
| sermons | o | + |
| commentaries | + | + |
| light fiction | - | - |
| debates | o | + |

Table 2: Biber dimension properties of the subcorpora

After completion, we will make the annotated corpus and the annotation guidelines available to the research community at the end of the research project in which it is being prepared. To this end, we took great care in the compilation of the corpus to include only material that can be republished in this form.

## 4.2. The annotation infrastructure

The first four subcorpora were available in written form. For the debates, the videos, which had been published on the Youtube channel of 'Jugend debattiert', had to be transcribed first. In a pre-test, Microsoft Azure emerged as the tool that produced the best raw results for automatic transcriptions, which were then corrected twice in order to obtain accurate transcripts. No attempts were made to correct anacolutha, repetitions, self-corrections, and other properties of spontaneous spoken language.

For the annotation, we use the INCEpTION tool (Klie et al., 2018). The tool allows annotating not only metaphorical words, but also complex constituents or parts of words as illustrated in the examples (2) and (3). It is also very easy to integrate external sources; we have integrated GermaNet (Hamp and Feldweg, 1997) for an extension of the annotation which will specify source and target domains of the metaphors. INCEpTION also offers – in the form of 'recommenders' – algorithms that learn from previous annotations and use this knowledge to provide recommendations for further annotations.

Metaphors are annotated independently by four annotators following our guidelines; differences between annotators are subsequently discussed and adjudicated. The annotation takes place on top of a layer of syntactic dependency structure, derived by parsing the texts with the Stanza package (Qi et al., 2020). This information will allow identifying syntactic constellations for analyses of their metaphorical potential.

## 4.3. The annotation guidelines

As our starting point, we chose the guidelines developed on the basis of the 'Metaphor Identification Procedure-VU' (MIPVU) (Steen et al., 2010; Herrmann et al., 2019). To distinguish different degrees of conventionalisation of the metaphors, we applied the MIPVU strategy of rely-

---

[3]The light fiction we investigate is written by amateurs for a general audience of their peers.

ing on suitable lexical resources (Steen et al., 2010). We used the *Duden* dictionary and the *Digitales Wörterbuch der deutschen Sprache* (Digital dictionary of the German language, DWDS), both of which are available online.[4]

For the purpose of the annotation, we introduced a binary distinction in conventionalised and a non-conventionalised metaphors based on the following criterion: Assuming that the context-based sense of the expression qualifies as metaphorical according to the MIPVU rules (see Section 3.), we check whether it is listed in at least one of the lexical resources along with the basic sense of the expression. If it is, the metaphor is classified as conventionalised, otherwise, as non-conventionalised. For example, *anpeilen* 'take a bearing' has the contextual sense 'envisage' in (5). Neither of our resources lists this context-based sense, hence, it qualifies as non-conventionalised. In contrast, *Facette* (literally, 'facet') has the context-based sense 'aspect' in (6), which is listed in both resources, thus, we annotate it as a conventionalised metaphor.

(5)     der angepeilte ausgeglichene Bundeshaushalt für 2015
        'the envisaged balanced federal budget for 2015'

(6)     alle Kriminalitätsfacetten und Probleme
        'all aspects of delinquency and problems'

We also annotate open signals of metaphoricity, like *wie* 'like' in similes or expressions such as *gewissermaßen* 'as it were' or *praktisch* 'in effect'.

(7)     Wir sind wie Blumen praktisch, geerdet.
        'In effect, we are like flowers, earthed.'

Our resource breaks new ground in that it not only identifies the metaphorical expressions themselves but also the two domains they are based on. To this end, we annotate, apart from metaphorical expressions, also elements in their immediate context that trigger the metaphorical interpretation. This context, which we call 'background', represents the target domain of the metaphor. The source domain can be reconstructed from the basic interpretation of the metaphorical element, which is indicated in the lexical resources.

Backgrounds and metaphorical expressions are located in the same clause. The relation between the two is defined in syntactic terms, as a head-complement or head-modifier structure, or as the arguments of copulative verbs like *sein* 'be', *werden* 'become', or *bleiben* 'remain'. In this way, we intend to make our corpus useful for approaches to the automatic detection of metaphor, which often rely on metaphor-background pairs defined in syntactic terms.

For instance, (8) [= (1)] comprises a metaphorical preposition with an NP complement, here, *meinen Gedanken* 'my thoughts', that functions as background. The literal meaning of the preposition *in* is spatial while the background refers to abstract ideas, thus, following the 'Master Metaphor List' (Lakoff et al., 1991), (8) emerges as an instance of the metaphor IDEAS ARE LOCATIONS.[5]

(8)     in meinen Gedanken
        'in my thoughts'

The more complex (9) shows that several arguments (complements or subject) of a head can simultaneously serve as background. The abstract subject *Angelegenheit* 'affair' as well as the abstract complement *Wendung* 'turn' together trigger the likewise abstract context-dependent metaphorical sense of *nehmen* 'take' in the sense of 'perform':

(9)     die Angelegenheit nahm eine unerwartete Wendung
        'the affair took an unexpected turn'

We also annotate cases in which metaphoricity is mediated through anaphoric relations like in (10), where the masculine 3rd person singular pronoun *er* forms the background for the metaphoric use of *weich* 'soft' because it corefers with *der EU-Jahresbericht* 'the annual report of the EU'.

(10)    der EU-Jahresbericht ... An vielen Stellen bleibt er zu weich
        'the annual report of the EU ... In many places, it is too soft'

Furthermore, we found that, in our corpus, metaphors frequently form the background for other metaphors, which we call metaphor chains. The phenomenon is illustrated in (11): Once the context-dependent sense 'base' of *Fundament* (literally, 'foundation') has been established (through the general context, which is about financial precautions for future generations), it serves as the background for the metaphoric senses of *fest* 'strong' and *bauen* 'build'.

(11)    ein festes Fundament bauen
        'to build a strong base'

To account for extended metaphor, metaphors can optionally be indexed. Coindexation indicates that metaphors share the specific mapping from source to target domain. For instance, one of the parliamentary speeches characterises social spending metaphorically as an investment. The underlying mapping from assets to families provides the basis for metaphorical interpretations of both *auszahlen* 'pay off' and *investieren* 'invest':

(12)    ... dass sich die gute Familienpolitik der letzten Jahre auszahlt, und dass sich das Geld, das wir investieren, lohnt
        'that the good family policy of the last years is paying off, and that the money which we invest is worthwhile'

Another phenomenon, to our knowledge not described as yet in the linguistic literature, is potential metaphor, the deliberate combination of tokens of an expression with basic and metaphorical senses in the same text. We annotate the tokens with the basic sense as potential metaphor and coin-

---

dex them with the tokens in the metaphorical sense. E.g., in one of the sermons the term *dunkel* 'dark' is used repeatedly in its basic sense 'without physical light' before it is used metaphorically in the sense 'bad':

(13)     in dem Dunkel, in dem Wurzelbereich bei dem Weizen
         'in the dark zone, in the rhizosphere of the wheat'

(14)     die dunkle Erde elterlicher Übermüdung
         'the dark soil of parental fatigue'

The existence of potential metaphor in a register has a very powerful effect, because even clearly non-metaphorical tokens of an expression evoke the possibility of a subsequent metaphorical used of the expression.

Finally, the guidelines also need to address some issues that are specific to German. For example, the issue of separable lexemes includes not only separable prefix verbs like *aufmachen* 'open' (cp. the 1st person singular past tense *ich machte auf* 'I opened') but also combinations of a demonstrative with a preposition like *darin*, literally 'that-in'. E.g., in (15), the demonstrative *dar* is an anaphoric complement of the preposition *in*, and its antecedent is the preceding sentence. Thus, *dar* refers to an abstract state of affairs, which forms the background for the metaphorically used preposition.

(15)     dass darin der Schlüssel liegt
         'that the key is to be found (lit. 'lies') in there'

I.e., in (15), only the prepositional part *in* of *darin* has a metaphorical sense. This shows that the anaphoric and the prepositional part can function independently in metaphor structures, which must be reflected in the annotation.

The guidelines will be published together with the corpus as part of the deliverables of the project on the interdependence of metaphor and register, but we will also make them available online.

## 5.    First results

In this section, we present first results of our annotation initiative. Table 3 summarises the counts for conventionalised, non-conventionalised, and extended metaphors (percentages are calculated with respect to word tokens).

| subcorpus | convent. metaphor | non-conv. metaphor | extended metaphor |
|---|---|---|---|
| speeches | 15.13% | .14% | .01% |
| sermons | 10.14% | .24% | .29% |
| commentaries | 11.34% | .24% | .10% |
| light fiction | 4.06% | .16% | .002% |
| debates | 10.38% | .13% | .02% |

Table 3: Metaphor counts for the subcorpora

The table reveals that there are clear differences between the different registers when it comes to the use of metaphors. The overall level of metaphors is high for speeches, moderate for sermons, commentaries, and debates, and low for light fiction, while non-conventionalised and extended metaphors pattern similarly, occurring mostly in sermons and commentaries. As for potential metaphors, they only occur in sermons (.41%), so we did not include them in the table.

The fact that non-conventionalised and extended metaphors show up predominantly in commentaries and sermons (.24% and .10%, and .24% and .29%, respectively) allows the formulation of the hypothesis that these metaphors often occur in highly persuasive registers. The observation that the hypothesis seems not to extend to debates does not necessarily contradict this hypothesis, as this might be due to a conflicting factor, viz., the time pressure of oral discourse, which impedes the creation of more involved forms of metaphor.

The expectation that oral discourse has a lower degree of metaphoricity than literal discourse throughout could not be confirmed. The counts for conventional metaphors are comparable for the debates and the average of the literal registers (10.37%); for non-conventional and extended metaphors, the overall score is higher for literal registers (.2% and .11%, respectively), but Table 3 reveals that the numbers for debates are close to or even higher than those for some of the literal registers. This suggests that previous very low counts of metaphoricity for oral discourse as in Steen et al. (2010) might be related to the conversational nature of the data, which calls for further investigations of the differences within oral registers.

As for individual registers, our data first suggest an overall low degree of metaphoricity for light fiction, which converges with the results of Steen et al. (2010). The register in the corpus that conveys the highest degree of register marking is the one of sermons. Not only do they exhibit a high degree of non-conventional metaphors, extended metaphor emerges as a clear register marker for sermons, with 2.9 occurrences per 1,000 words of running text, and few occurrences at all in the other registers. This effect is enhanced in that many of these extended metaphors recur throughout the texts with high frequency and variety. What is more, potential metaphors occurred only in sermons.

## 6.    Conclusion

In this paper, we presented the ongoing construction of a German corpus annotated for metaphors in different registers. In an extension of our work, we intend to broaden the annotation to include metonymy, which has also been identified as a register-sensitive phenomenon (Littlemore, 2015).

## 7.    Acknowledgements

## 8.    Bibliographical References

Beger, A. (2015). Metaphors in psychology genres. Counseling vs. academic lectures. In Berenike Herrmann et al. editors, *Metaphor in specialist discourse*, pages 53–75. Benjamins, Amsterdam.

Berber Sardinha, T. (2015). Metapher and register variation. In Berenike Herrmann et al. editors, *Metaphor in*

*specialist discourse*, pages 17–51. Benjamins, Amsterdam.

Biber, D. and Conrad, S. (2009). *Register, genre, and style.* Cambridge University Press, Cambridge.

Biber, D. (2009). Multi-dimensional approaches. In Anke Lüdeling et al. editors, *Corpus linguistics. An international handbook*, pages 822–855. Mouton de Gruyter, Berlin.

Bizzoni, Y. and Lappin, S. (2018). Predicting human metaphor paraphrase judgments with deep neural networks. In *Proceedings of the Workshop on Figurative Language Processing*, page 45–55.

Blaette, A. (2017). GermaParl. Corpus of plenary protocols of the German Bundestag. TEI files. available at: https://github.com/PolMine/GermaParlTEI.

Cameron, L. (2003). *Metaphor in educational discourse.* Continuum, London.

Deignan, A., Littlemore, J., and Semino, E. (2013). *Figurative language, genre and register.* Cambridge University Press, Cambridge.

Dipper, S., Ehlert, A., Scholz, D., Theodor, F., and Weber, L. (2021). Recognizing deliberate metaphors. Poster presented at the Annual Meeting of the Deutsche Gesellschaft für Sprachwissenschaft.

Dorst, A. (2015). More or different metaphors in fiction? A quantitative cross-register comparison. *Language and Literature*, 24:3–22.

Goatly, A. (1994). Register and the redemption of relevance theory. The case of metaphor. *Pragmatics*, 4:181.

Goatly, A. (2011). *The language of metaphors.* Routledge, 2 edition.

Halliday, M. and Hasan, R. (1985). *Language, context and text: A social semiotic perspective.* Deakin University Press, Victoria.

Hamp, B. and Feldweg, H. (1997). GermaNet - a lexical-semantic net for German. In *Proceedings of the ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*, Madrid.

Herrmann, B., Woll, K., and Dorst, A. (2019). Linguistic metaphor identification in German. In Susan Nacey, et al. editors, *Metaphor identification in multiple languages. MIPVU around the world*, pages 113–135. Benjamins.

Herrmann, B. (2015). High on metaphor, low on simile. An examination of metaphor type in sub-registers of academic prose. In Berenike Herrmann et al. editors, *Metaphor in specialist discourse*, pages 163–190. Benjamins, Amsterdam.

Kemmann, A. (2013). Debatte. In Björn Rothstein et al. editors, *Kernbegriffe der Sprachdidaktik Deutsch. Ein Handbuch*, pages 41–43. Schneider, Hohengehren.

Klie, J.-C., Bugert, M., Boullosa, B., de Castilho, R. E., and Gurevych, I. (2018). The inception platform: Machine-assisted and knowledge-oriented interactive annotation. In *Proceedings of COLING 2018: system demonstrations*, pages 5–9.

Koch, P. and Oesterreicher, W. (1994). Schriftlichkeit und Sprache. In Hartmut Günther et al. editors, *Schrift und Schriftlichkeit. Writing and Its Use. Ein interdisziplinäres Handbuch internationaler Forschung. An Interdisciplinary Handbook of International Research*, volume 1, pages 587–604. de Gruyter, Berlin.

Krennmayr, T. (2011). *Metaphor in newspapers.* Ph.D. thesis, Vrije Universiteit Amsterdam.

Lakoff, G., Espenson, J., and Schwartz, A. (1991). Master metaphor list. Available from http://araw.mede.uic.edu/~alansz/metaphor/METAPHORLIST.pdf.

Lakoff, G. (1993). The contemporary theory of metaphor. In Andrew Ortony, editor, *Metaphor and thought*, pages 202–251. Cambridge University Press, New York, 2nd edition.

Littlemore, J. (2001). The use of metaphors in university lectures and the problems that it causes for overseas students. *Teaching in Higher Education*, 6:333–349.

Littlemore, J. (2015). *Metonymy: hidden shortcuts in language, thought and communication.* Cambridge University Press, Cambridge.

Musolff, A. (2016). *Political metaphor analysis. Discourse and scenarios.* Bloomsbury Academic, London.

Susan Nacey, et al. editors. (2019). *Metaphor identification in multiple languages. MIPVU around the world.* Benjamins, Amsterdam.

Pasma, T. (2012). Metaphor identification in Dutch discourse. In Fiona MacArthur, et al. editors, *Metaphor in use: Context, culture, and communication*, pages 69–83. Benjamins, Amsterdam.

Qi, P., Zhang, Y., Zhang, Y., Bolton, J., and Manning, C. (2020). Stanza: A Python natural language processing toolkit for many human languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*.

Ritchie, D. (2013). *Metaphor.* Cambridge University Press, Cambridge.

Shutova, E. and Teufel, S. (2010). Metaphor corpus annotated for source-target domain mappings. In *Proceedings of LREC 2010*, pages 3255–3261.

Shutova, E., Teufel, S., and Korhonen, A. (2013). Statistical metaphor processing. *Computational Linguistics*, 39:301–353.

Skorczynska, H. and Deignan, A. (2006). Readership and purpose in the choice of economics metaphors. *Metaphor and Symbol*, 21:87–104.

Skorczynska, H. and Giménez-Moreno, R. (2017). Analysing metaphor in the family register through scripted sitcom conversations. *Metaphor and the Social World*, 7:252–269.

Stede, M. (2004). The Potsdam Commentary Corpus. In Bonnie Webber et al. editors, *ACL 2004 Workshop on Discourse Annotation*, pages 96–102, Barcelona, Spain. Association for Computational Linguistics.

Steen, G., Dorst, A., Herrmann, B., Kaal, A., Krennmayr, T., and Pasma, T. (2010). *A method for linguistic metaphor identification: from MIP to MIPVU.* Benjamins, Amsterdam.

Zayed, O., Philip, J. M., and Buitelaar, P. (2020). Figure me out: a gold standard dataset for metaphor in-

terpretation. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 5810–5819, Marseille. European Language Resources Association.