Anaphoricity in Connectives: A Case Study on German

Manfred Stede and **Yulia Grishina**
Applied Computational Linguistics
FSP Cognitive Science
University of Potsdam, Germany
`stede|grishina@uni-potsdam.de`

Abstract

Anaphoric connectives are event anaphors (or abstract anaphors) that in addition convey a coherence relation holding between the antecedent and the host clause of the connective. Some of them carry an explicitly-anaphoric morpheme, others do not. We analysed the set of German connectives for this property and found that many have an additional non-connective reading, where they serve as nominal anaphors. Furthermore, many connectives can have multiple senses, so altogether the processing of these words can involve substantial disambiguation. We study the problem for one specific German word, *demzufolge*, which can be taken as representative for a large group of similar words.

1 Introduction

The vast majority of the research on anaphoricity in Computational Linguistics has been done on nominal anaphora; it is arguably the most important for many purposes, and also the most frequent phenomenon. Nonetheless, *event anaphors*[1] are also highly relevant for text understanding, but they have proven to be much more difficult to resolve than nominal anaphors; see, e.g., (Dipper and Zinsmeister, 2012). In this paper, we zoom in on a specific subclass of event anaphors, namely on anaphoric *c*onnectives: They pick up an abstract-object antecedent from the previous context, and at the same time signal a semantic or pragmatic coherence relation between that antecedent and their host clause.

A principal distinction between 'anaphoric' and 'structural' connectives has been made by Webber et al. (2003) in the context of Computational Linguistics; similar observations have been made by linguists working on the German 'Handbook of connectives' (Pasch et al., 2003). While structural connectives (conjunctions) take their arguments qua the syntactic configuration they appear in, anaphoric connectives (certain adverbials) pick up their 'external' argument (or the 'Arg1' in the terminology of the Penn Discourse Treebank, PDTB) (Prasad et al., 2008) by means of anaphora resolution. Often, this argument is present in the clause preceding the anaphoric adverbial, but it need not be; Prasad et al. report that in the PDTB, 9% of the 'Arg1' arguments of connectives in fact appear not in the same or in the previous sentence, but farther away. For illustration, here is a fictitious example:

(1) *[Tom didn't go to the café.]*$_{Arg1}$ *It would close soon anyway. [He chose to sit at the beach]*$_{Arg2}$ *[instead]*$_{conn}$.

In English, a few connective adverbials make their anaphoricity explicit, as they contain a morpheme that overtly refers backward: ***there****fore*, ***where****by* etc. In other languages, this phenomenon is more widespread. In this paper, we will especially look at German, where a large number of connectives exhibit such a morpheme; Section 2 will provide an overview. Afterwards, in Section 3, we present a case study on one specific German word, which can act both as a nominal anaphor and as an event anaphor (in which case it is a connective) and thus poses an additional ambiguity problem. Then,

---

[1] In this paper, we use *event* anaphora interchangeably with *abstract* anaphora.

Section 4 discusses the disambiguation task and sketches a path toward a solution.

2 Anaphoric connectives in German

A connective, according to Pasch et al. (2003), is a closed-class lexical item expressing a two-place relation whose arguments denote eventualities and can, in principle, be expressed as full sentences. Connectives do not form a syntactically homogeneous class but contain both conjunctions (coordinate or subordinate) and certain adverbials. Due to this, they are usually regarded as a discourse phenomenon, and there are not many comprehensive linguistic studies that survey the connectives of a language. A notable exception is the aforementioned handbook for German, which lists about 350 different connectives. In terms of machine-readable lexicons, one for German connectives (*DiMLex*) had been introduced by Stede (2002), which in its current version[2] contains 274 entries. For French, *LexConn* (Roze et al., 2012) is slightly bigger (328 entries). For English, a list has been derived from the PDTB corpus, consisting of 100 connectives.

Since our focus here is on German, we worked with *DiMLex* and determined how many connectives have an explicitly-anaphoric morpheme (as explained above). We found 11 different relevant prefixes and suffixes, and their frequencies are: *da-*: 21, *-dessen*: 17, *wo-/wes-*: 11, *hier-*: 7, *-dem*: 7, *dem-*: 6, *des-*: 4, *-dann*: 3, *-dies*: 2, *dessen-*: 1. Thus, in total 79 connectives have one of the morphemes in question, which amounts to 29%.[3]

We went through these explicitly-anaphoric connectives and determined how many of them also have a non-connective reading. This problem of connective ambiguity had been quantified by Dipper and Stede (2006) as applying to 40% of the words, on the basis of an earlier (smaller) version of *DiMLex*. Many connectives have additional readings as discourse particles, verb particles, or nominal anaphors. Since our 79 connectives carry anaphoric

morphemes, ambiguity can hold between nominal anaphor and event anaphor (= connective). We found that this applies to 40 words; for most of them, their other function is that of a relative pronoun. For example:

(2) *Sie schenkte mir ein Buch, <u>womit</u> ich nichts anfangen konnte.*
'She gave me a book, <u>with which</u> I could not do anything.'

(3) *[Sie schenkte mir ein Buch,]$_{Arg1}$ [womit]$_{conn}$ [sie mir einen großen Gefallen tat.]$_{Arg2}$*
'She gave me a book, <u>whereby</u> she did me a big favor.'

3 Case study: *demzufolge*

The 40 words that we identified in the previous section are ambiguous between nominal anaphor and event anaphor. In order to approach the tasks of (a) determining the correct reading in a given context, and (b) finding the antecedent (which for the event anaphor reading corresponds to the Arg1 of the connective), we decided to first inspect one word in detail and chose *demzufolge*.

3.1 Different readings

A good way to map out the ambiguity of *demzufolge* is to collect the variety of its English translations in a parallel corpus. We used InterCorp[4], where the first 50 hits yield the following: *accordingly, as a result, consequently, as a consequence, therefore, that* (as complementizer or relative pronoun), *which* (as relative pronoun), and the null translation. Making this systematic, we see two broad classes of usages:

1. **Nominal anaphor,** a contracted form of *dem zufolge*, which in German can be paraphrased as *laut dem* ('according to which'). We find two syntactic forms:

(a) Introducing a relative clause:

(4) *Ich las ein Buch, <u>demzufolge</u> die Welt in diesem Jahr untergehen wird.*
'I read a book <u>according to which</u> the world will collapse this year.'

(b) Free adverbial:

---

[2]https://github.com/discourse-lab/dimlex

[3]Most morphemes do not straightforwardly translate to English; they correspond to local, temporal, and event anaphors in dative or genitive case. The phenomenon occurs in other languages as well; in Dutch, for instance, there are connectives like **daar**om, **daard**oor, **waar**dor; French examples are *après **ca**, à part **ca***.

[4]https://ucnk.ff.cuni.cz/intercorp/

(5) *Ich habe ein interessantes Buch gelesen. Demzufolge wird die Welt in diesem Jahr untergehen.*
'I read an interesting book. <u>According to it</u> the world will collapse this year.'

2. **Connective** with two arguments that denote eventualities. The online grammar *grammis*[5] in its 'grammatical lexicon' section states that it can appear in three different positions, as modeled by topological-field theory:[6]

- *Vorfeld* (pre-field):

(6) *Peter war der beste Torschütze. Demzufolge bekam er den Pokal.*
'Peter was the best goal scorer. Therefore he received the tophy.'

- *Mittelfeld* (middle-field):
  (...) *Er bekam demzufolge den Pokal.*

- *Nullstelle* (zero position):
  (...) *Demzufolge: Er bekam den Pokal.*

Irrespective of the position, the coherence relation being signalled is 'cause-result' (in the PDTB terminology), and intuitively, we expect this to be the only one; but see below for an exception. When considering various examples, it becomes clear that the readings cannot be easily distinguished at the linguistic surface. To explore this in depth, we thus conducted a (small) corpus study.

3.2 Corpus Study

To investigate the ambiguity and its potential resolution in authentic contexts, we randomly collected 140 instances of *demzufolge* (using a case-insensitive search) from the DWDS corpus[7]. 50 are from the print and online editions of the weekly paper *Die Zeit* (1946-2014), and 90 from the 'Kernkorpus 20', a genre-balanced corpus of 20th-century German that includes narratives, non-fiction books, scientific text, and some newspaper text. The extracted material for each instance was a window of

---

[5]http://hypermedia.ids-mannheim.de
[6]Very briefly, the finite verb and the other parts of the predicate constitute the *Satzklammer* ('sentence bracket'). The middle-field is inside the bracket; the pre-field precedes the left bracket; the zero position precedes the pre-field.
[7]www.dwds.de

three sentences, the second one of which contains the target word *demzufolge*. Henceforth, we call the two collections 'zeit50' and 'kernel90', respectively.

As our first step, to get an initial overview, one author of this paper annotated kernel90: For each instance of *demzufolge* we marked its antecedent and identified the syntactic type. These are the frequencies of the various antecedent types (we also indicate the English translation equivalent of *demzufolge*):

- NP antecedent: 42 (47%)
  Roles of *demzufolge*:
  - relative pronoun ("according to which"): 33 (37%)
  - other function ("therefore"): 9 (10%)
- VP antecedent ("therefore"): 19 (21%)
- S antecedent ("therefore"): 29 (32%)
  Subtypes:
  - one or more full sentences: 22 (24%)
  - sentential complement: 4 (4%)
  - sentences in coordinate structures: 2 (2%)
  - subordinate sentence: 1 (1%)

The relatively balanced distributions between syntactic antecedent types and also between readings/translations (33 non-connectives; 57 connectives) shows that disambiguation cannot be avoided by means of a simple majority baseline.

Next, we were interested in inter-annotator agreement regarding class (non-/connective), connective sense (PDTB taxonomy) and extension of the two arguments. One author of this paper and two trained annotators, who are familiar with German connectives but previously had not studied *demzufolge* in particular, labelled the 50 instances in zeit50. We can subsume the non-/connective decision under the sense labeling, where a non-connective receives the sense 'none'. Another special label annotators could use was 'missing context', indicating that a judgement is not possible because of the restricted context information available.

Results: With three annotators, there are 150 pairs of annotations to be compared. 103 (69%) of the decision pairs were completely identical (i.e., two annotators agreed on the connective sense and on the extensions of both arguments). For the senses, there were 25 cases of pairwise disagreement, and the vast majority (21) concerned the non-/connective

distinction. 'Missing context' was used on only one instance (by two annotators). Among the connective senses, 'cause-result' was used 39 times, and 'specialization' four times. Given these two relations plus 'none' and 'missing context', we can see sense labeling as a four-way classification task, and we computed the chance-corrected Fleiss-κ for the 3 raters, which is 0.55.

The presence of the 'specialization' sense seems to contradict our initial expectation of non-ambiguity. But, upon reflection, 'specialization' indeed can be quite compatible with a causal or justifying relation, so this is not an extraordinary finding. To illustrate, here is one (abbreviated) instance that received the 'specialization' sense:

(7) *[Im ARD-Deutschlandtrend liegt Merkel in der Wählergunst deutlich hinter ihren möglichen Herausforderern Steinbrück und Steinmeier.]*$_{Arg1}$ *[Bei einer Direktwahl des Regierungschefs würde sie [demzufolge]*$_{conn}$ *im Duell gegen Steinbrück zurzeit mit 37 zu 48 Prozent klar unterliegen.]*$_{Arg2}$
'In the ARD poll, Merkel clearly lags behind her challengers Steinbrück and Steinmeier. In a direct election of the chancellor, she would <u>thus</u> currently lose to Steinbrück with 37 against 48 percent.'

When the disagreement on senses pertains to the non-/connective reading, it – unsurprisingly – correlates with disagreement on Arg1 extension. Overall, among the 150 pairs of instance annotations, there are 32 disagreements on Arg1 extension, and 18 on Arg2 extension. Both of these disagreements are largely restricted to the connective usage, which illustrates the finding (also well-known from the PDTB) that the extension of the spans of causal relations can be quite vague: Is the Arg1 just the preceding clause or sentence, or more than that? For Arg2, as indicated, disagreement is relatively rare. However, our results on argument extension are preliminary, as the annotators had only a three-sentence extract from the host texts to make their judgements.[8] In a larger study, these annotations need to be done on full texts.

It is interesting to note that the non-/connective distribution differs between zeit50 and kernel90. In the former, the annotators labeled 34±2 instances as non-connectives, i.e., 68%. In kernel90, the corresponding figure is 37%. We attribute this difference to the genres: zeit50, as stated earlier, is taken from a newspaper, including its online edition, which to a large extent presents "instant news" that often involve citing other sources, so that the "according to which" reading is much more prominent than the "therefore" reading of *demzufolge*.

4 Toward disambiguation and resolution

Interpreting *demzufolge* and the 39 similar German words involves two subproblems: Disambiguate the reading (connective or non-connective), and resolve the argument(s) – either the antecedent of the NP-anaphor, or the two arguments of the connective.

For disambiguation, before embarking on full-fledged feature-based classification, it is advisable to check whether standard POS tagging can (partially) solve the problem. To this end, we experimented with two German taggers on the kernel90 set: clevertagger[9], which is integrated in the ParZu parser (Sennrich et al., 2009), and the tagger of the MATE tools (Bohnet, 2010). Both were used with their standard models, which for ParZu was trained on the TüBa-D/Z treebank[10] and for MATE on a dependency-converted version of the TIGER treebank[11]. They both make use of the STTS tagset[12] but in different versions. For our purposes, it is relevant that they use PROAV and PROP, respectively, for the German pronominal adverbs (contractions of a pronominal form and a preposition). Table 1 shows the tag distribution for the four groups of antecedent types; in each group, the top line gives the MATE results and the bottom line those of ParZu. The "other" column conflates a few obvious mistaggings as finite verb, adjective, etc. For the 29 instances with 'S' antecedents, both parsers failed to produce output in some cases (MATE: 5, ParZu: 4).

While we cannot really expect a tagger to differentiate between the types of antecedents (thus providing information for anaphora resolution), it is worth testing whether it can predict the non-/connective readings, which here means to split the relative pronoun uses from all others (as shown at the beginning of Sct. 3.2). It turns out that MATE correctly identifies only 6 of 33 relative pronouns (18%) as PRELS. ParZu tags 19 of them (58%) as subordinating conjunctions (KOUS), which is the wrong tag, yet it serves to distinguish them from the connective usages. Closer inspection reveals that 5 of the 6 MATE-PRELS instances are also ParZu-KOUS instances, so that for this task, on the whole ParZu is the better tool. If we assume that the ratios hold for *demzufolge* instances in general, then the upshot of the experiment is: ParZu can partially identify the non-/connective readings of *demzufolge*, when we interpret the KOUS tag as non-connective (with perfect precision, and recall of 19/33 = 58%), and the PROP tag as connective, with a precision of 50/61 (82%) and a recall of 50/57 (94%; counting also the four failed parses). For many purposes, this situation will not be good enough, so that classifiers using "deeper" features, in the spirit of Pitler and Nenkova (2009) have to be built.

Likewise, for the second problem of finding the arguments – of the nominal anaphor or of the connective – deeper features have to be used. Some work on Arg1 identification for English reports results around 80% accuracy based on surface and syntactic features (Elwell and Baldridge, 2008), but it seems not likely that this can be reached for the fairly complicated distinction between NPs, VPs, and sentences for the German connectives we are studying here. The most promising route might be to aim for identifying just the heads of the antecedents, as done for English, e.g., by Wellner and Pustejovsky (2007); also, it can help to consider semantic features, as proposed by Miltsakaki et al. (2003) for the anaphoric connective *instead*.

5 Summary and outlook

The distinction between structural and anaphoric connectives is well-established, but for the anaphoric ones it is an open question whether those with an explicit anaphoric morpheme be-

antecedent	PROAV PROP	KOUS	PRELS	other
NP (relpro)	22		6	5
	11	19		3
NP (other)	9			
	8			1
VP	19			
	19			
S	24			
	23			2

Table 1: kernel90 dataset: POS tags assigned to *demzufolge* by the parsers MATE (first row in a cell) and ParZu (second row).

have differently from those that do not have one, i.e., whether the group of anaphoric connectives should be split in two for purposes of argument identification. Entangled with this is the problem of non-connective ambiguity: many explicitly-anaphoric connectives also have a second reading as nominal anaphors. As a step toward resolving these issues, we started from a comprehensive lexicon of German connectives and determined that 79 of them have one of 11 different anaphoric morphemes. Of the 79 words, 40 are ambiguous between a connective and non-connective reading. We selected *demzufolge* for a pilot study and built a small corpus of 140 instances annotated with connective senses and argument spans. Experiments with POS taggers revealed that – at least for this word – they can help only to a limited extent for distinguishing the non-/connective readings.

Our next steps are to determine the parallelism between *demzufolge* and the other connectives and then to build sense/argument classifiers for groups of similar connectives. Since there are no large annotated resources for German, we will also look into the possibility of annotation projection, as suggested by Versley (2010) for English-German or Laali and Kosseim (2014) for English-French. For the connectives we study, this might be difficult, since English appears to have much fewer (explicitly-)anaphoric connectives; but if projection can also be done for AltLex instances (multi-word expressions in the PDTB), this might be helpful.

References

Bernd Bohnet. 2010. Top accuracy and fast dependency parsing is not a contradiction. In *Proc. of the 23rd International Conference on Computational Linguistics (COLING)*, pages 89–97, Beijing, China, August.

Stefanie Dipper and Manfred Stede. 2006. Disambiguating potential connectives. In Miriam Butt, editor, *Proc. of KONVENS '06*, pages 167–173, Konstanz.

Stefanie Dipper and Heike Zinsmeister. 2012. Annotating abstract anaphora. *Language Resources and Evaluation*, 46(1):37–52.

Robert Elwell and Jason Baldridge. 2008. Discourse connective argument identification with connective specific rankers. In *Proc. of the IEEE Conference on Semantic Computing (ICSC)*, Santa Clara/CA.

Klaus Krippendorf. 2004. Measuring the reliability of qualitative text analysis data. *Quality and Quantity*, 38(6):787–800.

Majid Laali and Leila Kosseim. 2014. Inducing discourse connectives from parallel texts. In *Proc. of the 25th International Conference on Computational Linguistics (COLING)*, Dublin/Ireland.

Eleni Miltsakaki, Cassandre Creswell, Katherine Forbes, Aravind Joshi, and Bonnie Webber. 2003. Anaphoric arguments of discourse connectives: Semantic properties of antecedents versus non-antecedents. In *Proc. of the 10 Conference of the European Chapter of the ACL*, Budapest.

Renate Pasch, Ursula Brauße, Eva Breindl, and Ulrich Herrmann Waßner. 2003. *Handbuch der deutschen Konnektoren*. Walter de Gruyter, Berlin/New York.

Emily Pitler and Ani Nenkova. 2009. Using syntax to disambiguate explicit discourse connectives in text. In *Proc. of the ACL-IJCNLP 2009 Conference Short Papers*, pages 13–16, Suntec, Singapore.

R. Prasad, N. Dinesh, A. Lee, E. Miltsakaki, L. Robaldo, A. Joshi, and B. Webber. 2008. The Penn Discourse Treebank 2.0. In *Proc. of the 6th International Conference on Language Resources and Evaluation (LREC)*, Marrakech, Morocco.

Charlotte Roze, Laurence Danlos, and Philippe Muller. 2012. Lexconn: A French lexicon of discourse connectives. *Discours*, 10.

Rico Sennrich, Gerold Schneider, Martin Volk, and Martin Warin. 2009. A new hybrid dependency parser for German. In C. Chiarcos, R. Eckart de Castilho, and M. Stede, editors, *From Form to Meaning: Processing Texts Automatically. Proc. of the Biennial GSCL Conference 2009*, Tübingen. Narr.

Manfred Stede. 2002. DiMLex: A lexical approach to discourse markers. In A. Lenci and V. Di Tomaso, editors, *Exploring the Lexicon - Theory and Computation*. Edizioni dell'Orso, Alessandria.

Yannick Versley. 2010. Discovery of ambiguous and unambiguous discourse connectives via annotation projection. In *Proceedings of Workshop on Annotation and Exploitation of Parallel Corpora (AEPC)*, Tartu, Estonia.

Bonnie Webber, Matthew Stone, Aravind Joshi, and Alistair Knott. 2003. Anaphora and discourse structure. *Computational Linguistics*, 29(4):545–587.

Ben Wellner and James Pustejovsky. 2007. Automatically identifying the arguments of discourse connectives. In *Proc. of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pages 92–101, Prague, Czech Republic.