

Annotating Abstract Anaphora

Stefanie Dipper · Heike Zinsmeister

Received: date / Accepted: date

Abstract In this paper, we present first results from annotating abstract (discourse-deictic) anaphora in German. Our annotation guidelines provide linguistic tests for identifying the antecedent, and for determining the semantic types of both the antecedent and the anaphor. The corpus consists of selected speaker turns from the Europarl corpus. To date, 100 texts have been annotated according to these guidelines. The annotations show that anaphoric personal and demonstrative pronouns differ with respect to the distance to their antecedents. A semantic analysis reveals that, contrary to suggestions put forward in the literature, referents of anaphors do not tend to be more abstract than the referents of their antecedents.

Keywords Abstract anaphora · Abstract entities · Coreference annotation · Semantic annotation

1 Introduction

An important component in understanding text is *anaphora resolution*, i.e., determining the reference of constituents whose interpretation depends on (the reference of) other textual elements. The majority of anaphora are instances of noun phrase anaphora (*NP anaphora*), which relate an anaphor to a nominal (NP) antecedent. Prevalent examples of anaphors are personal and demonstrative pronouns. The resolution process of such pronominal anaphors is guided by grammatical restrictions

We would like to thank the anonymous reviewers for their helpful comments, and our student annotators: Iris Bräuning, Christine Rieger, and Melanie Reiß. The work reported here was in part supported by Europäischer Sozialfonds in Baden-Württemberg, by the Rectorate's Fund of Ruhr-University Bochum, and by the Young Scholar Fund of the University of Konstanz's Excellence Initiative.

S. Dipper
Institute of Linguistics, Ruhr University Bochum, 44780 Bochum, Germany
Tel.: +234-32-25112, Fax: +234-32-14137
E-mail: dipper@linguistics.rub.de

H. Zinsmeister
Institute of Linguistics, Konstanz University, 78457 Konstanz, Germany
E-mail: Heike.Zinsmeister@uni-konstanz.de

(gender, number agreement) and saliency (grammatical function, recency). In addition to pronouns, definite noun phrases can be viewed as anaphoric in that they may corefer with some other NP in the given text. To resolve the latter type of anaphora, or coreferencing of proper nouns, simple pattern matching methods on the head nouns can be applied, or lexical semantic knowledge can be exploited, such as that provided by an ontology or a database like WordNet [14] (cf. [26] for a recent survey).

Another type of anaphora is discourse deixis [35,36], which relates an anaphor to a verbal or (multi-)clausal antecedent. The discourse entities that are referenced by antecedents of discourse deictic anaphors are called *abstract objects* since they refer to propositional entities, such as events or facts [3]. Grammatical restrictions cannot apply to them because the antecedent is non-nominal.

We assume that, in addition to saliency, the resolution process needs to take *semantic restrictions* into account, which are imposed by properties of the antecedent as well as the context of the anaphor. For instance, in Ex. 1a, the clause *penguins migrate to Fiji* denotes an event, and the anaphor *it* takes up the event while adding further specifications, namely that the migration event happens before another event, the hatching event. However, as shown, e.g., by [35,6], abstract anaphors can take up abstract objects *in different ways*. In Ex. 1b, the anaphor *that* does not refer to the migration event itself but to *the fact* that this event takes place each fall.

Example 1 (from [6])

- a. Each fall, penguins migrate to Fiji. *It* happens just before the eggs hatch.
- b. Each fall, penguins migrate to Fiji. *That's* why I'm going there next month.

In this paper, we use the terms *abstract anaphor/antecedent* for anaphors or antecedents that refer to abstract objects. The *semantic type of an antecedent* relates to semantic properties of the referent, as introduced in its original context (e.g., the semantic type of the first clause in Ex. 1a is “event”). The *semantic type of an anaphor* relates to the way the anaphor is used to characterize the propositional discourse referent denoted by its antecedent (e.g., the semantic type of *that* in Ex. 1b is “fact”). We report results from an annotation experiment comprising the following steps: (i) identifying the abstract antecedent; (ii) determining the semantic type of the antecedent and (iii) of the anaphor. We exclusively rely on *linguistic tests*, in the form of paraphrase tests, as the guiding principle of our annotation guidelines.

2 Related Work

Corpus-based studies have shown that abstract objects are less salient than other discourse referents, which effects the choice of the anaphoric element [17]. Whereas in English, anaphors referring to abstract entities are predominantly realized by demonstrative pronouns [35,15], the situation is different for other languages, e.g., Spanish and Catalan [30], or Danish and Italian [24].

The semantic type of the antecedent (e.g., *event*, *fact*) and that of the anaphor do not necessarily coincide (cf. Ex. 1b). The data suggests that reference to other types (referred to in the literature as *coercion*) is possible only in accordance with an abstractness hierarchy [17,8,9]. The hierarchy starts with events being defined as the most concrete type, which are anchored in spatial-temporal dimensions, and ends with propositions as the most abstract types, as illustrated in Fig. 1 [3]. The *Abstractness*

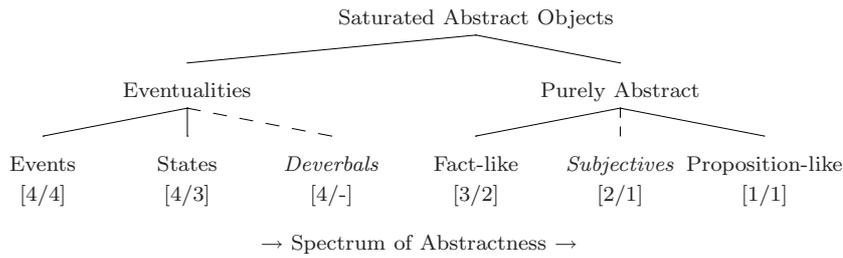


Fig. 1 Extract from the abstractness hierarchy from [3, p. 57]. Types in italics have been added by the authors. The numbers in square brackets are the scores assigned to our/Asher’s categories (cf. Table 2 and Sec. 5).

Table 1 Survey of corpus-based research on abstract anaphora. The “Lang” column lists the codes of the language that the studies deal with. The “Anaphor-Form” column displays the type of NPs (including pronouns) considered as anaphor candidates. The “Antecedent-Form” column shows the restrictions on the extension of the antecedent. The “Sem(antics)” columns indicate whether semantic types are annotated.¹

Study	Lang	Anaphor		Antecedent	
		Form	Sem.	Form	Sem.
[13]	EN	Dem, Pers	–	Clauses	–
[34]	FR,PT	Dem-NPs	–	Arbitrary seq.	–
[7]	EN	Dem, Pers	–	Clauses	–
[19]	CZ	Dem, Pers, Poss, Zero	–	V-head (depend. str.)	–
[28]	EN	Dem (<i>this/these</i>), <i>this</i> -NPs	–	–	–
[4, 5]	EN	Dem, <i>this/that</i> -NPs	+	–	–
[16]	EN	<i>it, this, that</i>	+	Arbitrary seq.	(+)
[22, 23]	EN	<i>it, this, that</i>	–	V-head	–
[29]	EN	Dem, Pers, Poss, spec. NPs	–	V-head	–
[25, 24]	DA,IT	Dem, Pers, Zero	+	Clauses and larger seq.	–
[27, 1]	EN	all NPs/Prons	(+)	Arbitrary seq./ clauses	–
[10]	DE	Dem (<i>dies</i> ‘this’)	+	Arbitrary seq.	+
[32, 31]	CA,ES	all NPs/Prons (incl. Cl, Zero)	(+)	Verbs, clauses, and larger seq.	–

Hypothesis states that anaphoric reference is possible to antecedents that are of the same type as the anaphor, or are of a less abstract type than the anaphor [17, 8].

Compared to work on NP anaphora, considerably less corpus-based research has been devoted to abstract anaphora. Table 1 presents a survey of studies involving abstract anaphora. In the table, we classify annotation research efforts on discourse deixis on the basis of formal and semantic criteria which we consider crucial for the task of resolving abstract anaphora.

As can be seen from Table 1, most corpus-based works on anaphora resolution do not annotate semantic types. Notable exceptions are [4, 16, 25, 10], who distinguish be-

¹ Abbreviations used: *Cl*: clitics, *Dem*: demonstrative pronouns, *Expl*: expletives, *Pers*: personal pronouns, *Poss*: possessive pronouns, *Zero*: zero pronouns, *Dem-NP*: NP with a demonstrative determiner. *Clauses* means that antecedents are syntactically defined, e.g. as sentences, infinitives, gerunds; *V-head*: only the verbal head is marked. (+) means that semantic properties others than the types that we discuss are annotated.

tween semantic types such as *events*, *processes*, *states* (eventualities) and *facts*, *propositions* (factualities). [32] use a different classification; they define the types *token*, *type*, *proposition*. [27] annotate selected semantic properties such as *person*, *animate*, *concrete*, *space*, *time*, etc. [13,7] do not annotate semantic-type information but incorporate corresponding semantic constraints in their resolution algorithms. Only two projects also annotate semantic information of the antecedent: [16], who annotate cognitive status, and our own project, in which semantic types are annotated [10].

This paper focuses on criteria and *linguistic tests* for the annotation of anaphors and antecedents and their *semantic types*. We define different paraphrase tests that aid annotators in annotating formal and semantic properties.

3 The Corpus

Our corpus consists of texts from the Europarl corpus [18]. The Europarl corpus consists of verbatim reports of debates in the European Parliament, both in the original language, as delivered by the speaker, as well as in translations into ten other languages, as delivered by the translation services of the European Union. In the Europarl corpus, individual contributions (“turns”) are marked by SGML elements, along with the names, parties and languages of the respective speakers. As our basis, we selected all contributions whose original language is German (including Austrian German).

For the annotation task, we isolated medium-sized turns, consisting of 15–20 sentences. This choice was made to guarantee that the turns were not too lengthy but still provided enough information for the annotators to understand the broader context of the discussion, allowing them to resolve the anaphors without comprehension problems.

Note that these turns do not stem from interactive dialogues. Instead, they consist of complete individual contributions to heavily regulated parliamentary debates. Hence, they are usually self-contained and tend not to include references to previous utterances in the form of anaphoric pronouns.

4 The Guidelines

Our guidelines are based on theoretical research on discourse semantics as well as on work concerning the annotation of discourse phenomena. We follow [17,8] in assuming that the semantic type of the antecedent and anaphor are relevant for the resolution process.

Our investigations are currently restricted to abstract anaphors in the form of *das*, *dies* ‘that, this’ (demonstrative pronouns) and *es* ‘it’ (personal pronoun). The guidelines define (i) how to identify abstract anaphora, (ii) how to identify the antecedent, (iii) how to determine the semantic type of the antecedent, and (iv) how to determine the semantic type of the anaphor. For each of these tasks, the guidelines provide linguistic tests [11].

4.1 *Pronoun Check*: Identifying abstract anaphora

The first task is to determine whether the pronoun refers to an abstract entity. That is, annotators are asked to exclude expletives, or pronouns referring to nominal an-

tecedents, or pronouns whose referent is salient in the context without being mentioned explicitly in the text (e.g., an action plan under discussion in a debate).

4.2 *Namely Test*: Identifying the antecedent

To determine the scope of the anaphoric relation, i.e., its antecedent, the *Namely Test* is applied: The annotator supplements the anaphor by a paraphrase in the form of *nämlich X* ‘namely X’. The placeholder X corresponds to the sought antecedent in the text.

The following example illustrates the application of the *Namely Test*. The antecedent as determined by the test is underlined; both the anaphor and its antecedent are labelled by their semantic types (cf. Secs. 4.3 and 4.4). The last item shows the results from the *Namely Test*: the anaphor is followed by the inserted test paraphrase (in italics).

Example 2

German: Die Zusammenführung der nationalen und europäischen Ebene ist sehr wohl notwendig, obwohl natürlich die Haupttätigkeit in den Mitgliedstaaten stattfinden sollte und nur dann auf europäischer Ebene eingegriffen werden_{neg} sollte, wenn [dies]_{prop/ev/proc} unbedingt notwendig ist.

English: It is indeed necessary to bring the national and European levels together, even though, of course, the main work should be done in the Member States, with the European level intervening only when **this** is absolutely necessary.

Namely Test:

... wenn dies — *nämlich dass auf europäischer Ebene eingegriffen wird* — ...
 ‘... when this — *namely that the European level intervenes* — ...’

Antecedents consist of VPs, (discontinuous fragments of) main or subordinate clauses, or multiple sentences. For instance, the anaphor *dies alles* ‘all this’ often refers to an antecedent consisting of multiple sentences. In some cases, the antecedent is an abstract NP, usually involving a deverbal noun, such as *diese Einschränkung* ‘this limitation’, as in the following example.

Example 3

German: Zweitens geht es in diesem Kontext um [die Einschränkung gewerkschaftlicher Rechte von Polizisten]_{deverb}. Auch [dies]_{deverb} ist nicht hinnehmbar.

English: Secondly, there is also, in this context, the issue of the limitation of police officers trade union rights. **This** is not acceptable either.

Namely Test:

... dies — *nämlich diese Einschränkung* —
 ‘... This — *namely this limitation* —’

When formulating the paraphrase, annotators are asked to adhere as closely as possible to the original text. Nevertheless, the actual antecedent can diverge from the one constructed by the *Namely Test* in minor aspects. Minor aspects are differences that only involve inflectional endings or functional words, i.e., non-lexical material. Examples of such variations are active-passive-alternations or bare infinitive vs. *zu*-infinitive vs. participle.

In some cases, the divergences are more fundamental and could, for instance, involve inserting or modifying the main verb. In such cases, annotators are asked to record the differences. For instance, in the following example, they would record that their choice of a main verb, *herstellen* ‘establish’ or *pflügen* ‘maintain’, differs from the main verb in the given text (*brauchen* ‘need’).

Table 2 Noun classes of the *NP-Replacement Test* for anaphors, and the corresponding semantic types. The “OS” and “AS” columns display the semantic scores according to our scale and a scale based on Asher’s categories (cf. Sec. 5). Deverbal nouns have not been included in Asher’s scale.

Sem. type	OS	Label(s)	Exemplary head nouns	AS
Eventualities	4	<i>ev</i>	<i>Ereignis, Aktion, Maßnahme</i> (‘event, action, measure’)	4
		<i>proc</i>	<i>Vorgang, Aktivität, Maßnahme</i> (‘process, activity, measure’)	
		<i>state</i>	<i>Zustand</i> (‘state’)	
		<i>circ</i>	<i>Sachverhalt/Umstand, Tatsache</i> (‘circumstance, fact’)	
		<i>deverb</i>	Deverbal nouns, e.g. <i>Einschränkung</i> (‘limitation’)	
Facts	3	<i>fact</i>	<i>Tatsache</i> (‘fact’)	2
Subjectives	2	<i>subj</i>	<i>Einschätzung, Meinung</i> (‘assessment, opinion’)	1
		<i>neg</i>	<i>Problem</i> (‘problem’)	
Propositions	1	<i>mod</i>	<i>Möglichkeit, Notwendigkeit</i> (‘possibility, necessity/need’);	
		<i>wh</i>	<i>Frage, Fragestellung</i> (‘question, issue’);	
		<i>prop</i> (= art+X)	Combinations of specific determiners with the above nouns: <i>diese Art von X, eine solche X</i> (‘this kind of, such a’)	

Example 4

German: Es geht nicht darum, in Einflussphären zu denken, dies ist nicht gegen Russland gerichtet, denn auch [eine demokratische Ukraine wird gute Beziehungen zu Russland brauchen]_{circ}, und [das]_{prop} sollten wir als Europäer unterstützen.

English: This is not about thinking in terms of spheres of influence, nor is it anti-Russian, for a democratic Ukraine too will need good relations with Russia, and we Europeans need to support it in **this**.

Namely Test:

... das — nämlich dass eine demokratische Ukraine gute Beziehungen zu Russland **herstellt/pflegt** —

... this — namely that a democratic Ukraine too will need **to establish/maintain** good relations with Russia —’

Anaphors that precede their antecedents (= cataphors) or are intraposed — usually embedded in parantheses — also occur, as illustrated by the next example.

Example 5

German: ... sondern dass klar ist, dass [wir] — und [das]_{circ} hat die Berichtstatterin auch sehr deutlich gemacht — [im Falle, dass unsere Gesetzgebungskompetenz beschnitten wird, klar und deutlich dagegen vorgehen]_{circ}.

English: ... but it must be clear that we — the rapporteur has made **this** very clear — must be firm in taking action to deal with any threat to curtail our legislative competence.

4.3 *NP-Replacement Test*: The semantic type of the anaphor

Once the antecedent is localized, the next step is to determine the semantic type of the anaphor. For this task, we defined the *NP-Replacement Test*. With this test, the anaphors *dies, das, es* are replaced by suitable NPs with prototypical head nouns, such as *dieses Ereignis, diese Tatsache, diese Ansicht* ‘this event, this fact, this opinion’.

We use the head noun as an indicator of the semantic type of the propositional entity (e.g., event, fact). See Table 2 for the classes of replacement nouns that we distinguish between.

Annotators are presented with a list of nouns from which they may choose. They may also add new nouns if necessary; however, our current list turned out to be adequate for most cases. The NP-Replacement Test is divided into two parts: First, annotators are asked to pick the first noun that comes to their mind which fits the context well. Second, they can select (or add) further well-fitting nouns. Suitable determiners have to be picked along with the nouns, e.g. *diese*, *eine solche*, *diese Art von* ‘this, such a, this kind of’.

In Ex. 5, a suitable NP would be *dieser Umstand* ‘this circumstance’. Accordingly, the anaphor is labelled as ‘circ’ (cf. Table 2).

Example 6

Namely Test: (applied to Ex. 5)

... und [**das**]_{circ} (= *diesen Umstand*) hat die Berichterstatteerin auch sehr deutlich gemacht
 ‘... the rapporteur has made **this** (= *this circumstance*) very clear’

We consider the determiners ‘this kind of’ (Ex. 2) and ‘such a’ (Ex. 4) as indicative of the semantic type ‘prop(osition)’, since they are used to abstract away from concrete instances of events, situations, etc. Replacement NPs and their semantic types for the remaining examples are shown below. Ex. 2 is annotated twice, and, furthermore, the noun *Maßnahme* ‘measure’ is assigned two semantic types; it can denote an event or a process.

- Ex. 2: *diese Art von Aktion*, ‘this kind of action’, and *diese Maßnahme* ‘this measure’; type ‘prop’ and types ‘ev/proc’, respectively.
- Ex. 3: *diese Einschränkungen* ‘these limitations’; type ‘deverb’.
- Ex. 4: *eine solche Entwicklung* ‘such a development’; type ‘prop’.

4.4 Colon Test: The semantic type of the antecedent

The last annotation step is to determine the semantic type of the antecedent. The annotators are instructed to read the immediate context before the sentence containing the antecedent. Next, they are asked to express the speaker’s view towards the antecedent’s content. For instance: Is it a fact that the speaker reports? Is it a goal that s/he wants to be pursued? Is it a problem that s/he wants to point out? To express the speaker’s view, the annotators formulate an introductory phrase that expresses this view and ends with a colon, followed by the antecedent (hence, the Colon Test).

The annotators are provided with a list of phrase patterns they can select from, such as *I state the following needs*. The object NPs in these phrases are selected from the same list already used in the *NP-Replacement Test*. Example verb patterns are shown in Table 3.

Applied to Ex. 5, a potential introductory phrase could be *Ich weise auf den folgenden Umstand hin* ‘I point to the following circumstance’.

Example 7

Colon Test: (applied to Ex. 5)

German: *Ich weise auf den folgenden Umstand hin: Wir werden im Falle, dass unsere Gesetzgebungskompetenz beschnitten wird, klar und deutlich dagegen vorgehen.*

English: *I point to the following circumstance: We must be firm in taking action to deal with any threat to curtail our legislative competence.*

The semantic type of the antecedent is determined as follows: If the main verb is a “neutral” verb, such as ‘point to, state’, we use the object NP as an indicator of

Table 3 Introductory phrase patterns for the *Colon Test* for antecedents. The slot <NP> is to be filled with the same NPs as those used in the *NP-Replacement Test*.

Sem. type	OS	Label(s)	Exemplary verb patterns	AS
“Neutral” verbs	–	–	<i>Ich weise auf <NP> hin: X</i> (‘I point to’) <i>Ich stelle <NP> fest: (‘I state’)</i>	–
“Subjective” verbs	2	<i>subj</i> <i>neg</i>	<i>Ich stelle mir <NP> vor: (‘I imagine’)</i> <i>Ich plädiere für <NP>: (‘I advocate’)</i> <i>Ich lehne <NP> ab: (‘I reject’)</i>	1

the semantic type (as in the NP Replacement Test). Hence, the antecedent in Ex. 7 is labeled as ‘circ’. That is, in this example, the semantic types of the anaphor and its antecedent coincide.

Some verbs contribute relevant meaning, such as ‘advocate’. In these cases, it is the verb that determines the semantic type of the antecedent.

Introductory phrases and their semantic types for the remaining examples are shown below. The type combinations of antecedent–anaphor that occur in these examples are then: neg–prop or neg–ev/proc (Ex. 2), deverb–deverb (Ex. 3), circ–prop (Ex. 4), and circ–circ (Ex. 5). That is, in all examples (except for the ambiguous case neg–ev/proc), the Abstractness Hypothesis (cf. Sec. 2) is indeed valid.

- Ex. 2: *Ich lehne folgende Möglichkeit ab: Es wird auf europäischer Ebene eingegriffen*, ‘I reject the following possibility’; type ‘neg’. Alternative: *Ich weise auf folgendes Problem hin*: ‘I point to the following problem’; type ‘neg’.
- Ex. 3: *die Einschränkung . . .* ‘the limitation’; type ‘deverbal’ (no introductory phrase possible).
- Ex. 4: *Ich weise auf folgenden Umstand hin*: ‘I point to the following circumstance’; type ‘circ’.

5 Results

5.1 Rationale of the annotation procedure

The idea underlying our annotation method is reliance entirely on linguistic tests. For antecedents, this test consists of a “container test” in the sense of Vendler: we create a suitable context (the container) which allows us to embed the antecedent. The context serves as a diagnostic for the semantic type. That is, rather than examining properties of the antecedent directly, we design an annotation procedure that makes explicit the way the speakers conceptualize the entity denoted by the antecedent.²

In a similar way, the NP-Replacement test also serves to make the speaker’s view on the antecedent’s content explicit, by replacing “neutral” pronouns (*dies* ‘this’) with more specific NPs, such as *diese Tatsache*, *diese Ansicht* ‘this fact, this opinion’.

² In an earlier version of the guidelines (cf. [10]), annotators had to determine the semantic types of the antecedents via reference to a table that listed 10 types of propositional entities, such as *event*, *process*, *state*, *fact* (see e.g., [33,3,21]). Annotators had to select an entity on the basis of features like *world-dependent*, *time-dependent*, *dynamic*, *telic*, and by applying linguistic tests from theoretical semantic work, such as “Is it possible to add frequency adverbials like ‘three times’, or time frames like ‘within one hour’, or time spans like ‘for one hour’?” [12]. These tests, however, proved very difficult to apply to naturally occurring sentences. Therefore, the annotation criteria were redesigned as described in the text.

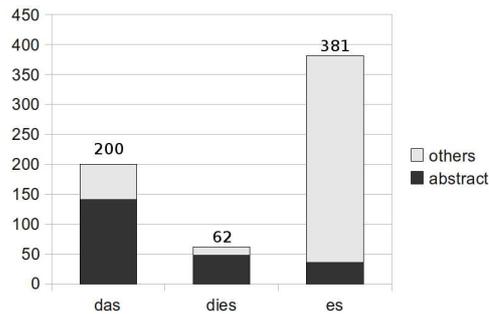


Fig. 2 Usages of *das* ‘that’, *dies* ‘this’, and *es* ‘it’ as abstract anaphors vs. other uses, e.g., expletives or NP anaphors.

However, it is important to distinguish between both acts of reference. The speakers can, of course, refer to and conceptualize one and the same event or situation etc. in different ways. As a consequence, anaphors and their antecedents can differ with regard to their semantic type [17,8]. Hence, annotators of antecedents must not be influenced by prior anaphor-related annotations. Accordingly, the annotators carrying out the Colon Test are presented the plain text, with the antecedent underlined, but without any information related to the anaphor.

To operationalize the Abstractness Hypothesis, we map the semantic types to scores (similar to the approach by [20]). We use two different scales: (i) Our scale is directly derived from, and is a rather natural consequence of, the list of prototypical head nouns (combined with article and verb information), which are mapped to the semantic classes defined by us. (ii) A scale based on Asher’s abstract entity classes which results from mapping our head nouns to Asher’s classes. Some of the nouns cannot be easily mapped to Asher’s scale (e.g. deverbals), and have to be left out from this scale. Further research is needed, however, to confirm our mapping from nouns to semantic classes.

5.2 Corpus statistics

Our current corpus consists of 100 texts (turns) from the Europarl corpus. Two annotators annotated the texts independently from each other using the tool MMAX2.³ Annotating a complete turn with all four tests takes between 5 and 10 minutes per annotator, excluding adjudicating the double-annotated texts into a gold standard.

Pronoun usages: In total, the texts contain 643 instances of anaphor candidates. Among them, 225 (35%) are abstract anaphors (plus 4 (<1%) “deictic anaphors” without an explicit antecedent in the text). Fig. 2 displays frequency information about the usages of pronouns.⁴ In the remaining section, we focus on analyses of abstract anaphors.

Antecedents: Mean length of the antecedents is 13.9 tokens; standard variation is 10.6. This varies considerably among the different types of pronouns: The shortest

³ MMAX2: <http://mmax2.sourceforge.net/>.

⁴ For the task of classifying abstract anaphor vs. the rest, we achieve an inter-annotator agreement of $\kappa=.79$ (computed according to [2]).

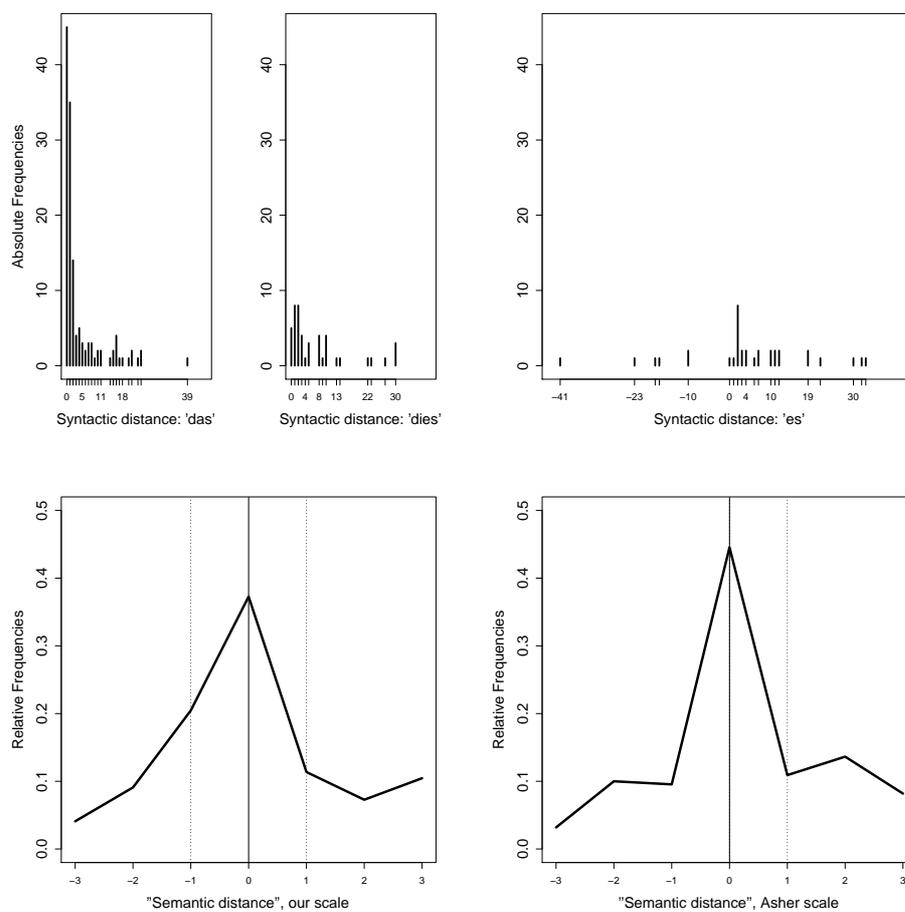


Fig. 3 Distances between different types of anaphors: In the upper figures: “Syntactic distance”, measured as the number of intervening tokens (with a few outliers having been removed). Bottom figures: “Semantic distance”, measured according to two different abstractness scales (on the left, our scale; on the right, the scale according to the abstractness hierarchy from [3]); solid line: median, dotted lines: quartiles.

antecedents occur with *es* ‘it’: $\bar{O}7.9 \pm 5.2$; the value for *das* ‘that’ is $\bar{O}15.4 \pm 11.6$, for *dies* ‘this’ $\bar{O}14.1 \pm 8.9$.^{5,6}

Syntactic Distance: Figure 3 displays interesting differences between the three types of anaphors with regard to the distance to their antecedent. The diagram on

⁵ According to the Kolmogorov-Smirnov test, antecedent lengths for *das* vs. *dies* do not differ significantly, but they do if compared to *es*: $D=.36$, $p<.01$ (*das* vs. *es*); $D=.36$, $p<.05$ (*dies* vs. *es*).

⁶ Inter-annotator agreement for the task of identifying the antecedent is .40 (observed agreement on exact matches). If only the head verb of the antecedent is considered, agreement improves to .55, and measuring simple overlap yields .84.

the left illustrates “syntactic distance”, measured in the number of intervening tokens: demonstrative *das*, which is the most frequent abstract anaphor (with more than 100 instances), occurs most frequently in close adjacency to its antecedent. At the other extreme, pronominal *es* often occurs in front of its antecedent (cataphoric use). This is indicated by negative distance values. A distance of 0 indicates that the anaphor is intraposed within the antecedent, usually embedded in a parenthesis (see Ex. 5). Finally, demonstrative *dies* tend to occur with larger distance to its antecedent than *das*.⁷

Semantic Distance:⁸ The bottom plots display “semantic distance”, which compares the semantic types of antecedents and anaphors by computing the “distance” on a given abstractness scale. The plot on the left is based on the scale that we established. This scale starts with the highest score for the most concrete entities: eventualities are scored with 4, facts with 3, down to propositions as the most abstract entities, which are scored with 1 (see Table 2). The plot on the right is based on Asher’s abstractness hierarchy. The scale again assigns the highest score to the most concrete entities.

The anaphor and its antecedent are first scored separately. The final score is the result of subtracting the anaphor’s score from the antecedent’s score.⁹ A score of 0 means that the anaphor and its antecedent are of the same semantic type. Positive scores mean that the anaphor is of a more abstract type than its antecedent. That is, a score of 0 or more means that the Abstractness Hypothesis from [17, 8] holds true.

The plots show that our data does not confirm the hypothesis. We do *not* find the large majority of the data to lie on the positive side (including 0). Instead, the plots resemble a normal distribution. Most often, the anaphor and its antecedent are of the same semantic type; the remaining cases are distributed more or less equally on both the negative and positive sides. The shape of the Asher-based plot is more irregular, and the first quartile equals the median. Still, however, a considerable amount of cases lie on the negative side of the plot, thus contradicting the Abstractness Hypothesis. An example is given in Ex. 8. The antecedent is of type ‘neg’ (score 1; test phrase: *I imagine the following problem*), and the anaphor is of type ‘ev/proc’ (score 4; test phrase: *this event*).

Example 8

German: Wir, die Vertreter der Bürger, der Menschen Europas, erreichen innerhalb der Institutionen etwas, und wir debattieren darüber kurz vor Mitternacht und werden ermahnt, wir sollen uns kurz fassen_{neg(1)}. [Das]_{ev/proc(4)} ist ein Wiederholungsfall.

English: Within the institutions, we — the representatives of the citizens, of the people of Europe — are actually achieving something, and, as we debate it shortly before midnight, we are enjoined to keep things brief. This is not the first time **this** has happened.

⁷ According to the Kolmogorov-Smirnov test, distances associated with *dies* vs. *es* do not differ significantly. Significant differences exist between *das* vs. *dies*: $D=.29$, $p<.01$; and between *das* vs. *es*: $D=.36$, $p<.01$.

⁸ For the annotation of semantic types in the NP Replacement Test, we achieved inter-annotator agreements of .75 for head noun classes and .77 for article classes. In the Colon Test, we achieved observed agreements of .70 for noun classes, .92 for article classes, and .86 for verb classes. For evaluating agreement, we consider both the first-to-the-mind choice as well as the alternative choice as equal candidates. If the annotators agree on any candidate, we consider it to be a match. As a consequence, we can only compute observed agreement but not expected agreement. The same holds for antecedent identification (Fn. 6).

⁹ If multiple or ambiguous paraphrases have been annotated, the most concrete type is chosen for antecedents, and the most abstract type for anaphors. This is done in order to satisfy the Abstractness Hypothesis.

6 Conclusion

In this paper, we presented a corpus-driven approach to abstract anaphora in German. We introduced annotation guidelines that provide linguistic tests for identifying the antecedent, and for determining the semantic types of both the antecedent and the anaphor.

Our analysis of “syntactic distance” reveals clear differences between the demonstrative pronouns *dies*, *das* and the personal pronoun *es*: *es* is often used cataphorically, whereas *dies*, *das* usually follow their antecedents. In analyzing “semantic distance”, our data does not confirm the hypothesis that anaphors are of the same type as their antecedents or more abstract. However, semantic types can be helpful for resolving abstract anaphora since in most cases antecedents and anaphors agree in their types.

As next steps, we will exploit automatic ways of determining semantic types of anaphors. To generate training data automatically, we plan to take into account anaphors of the form *dieses Ereignis*, *dieser Umstand*, etc. (‘this event, this circumstance’), which explicitly name the semantic type of the anaphor. In a later step other, more ambiguous, types of anaphors will be included in the investigation. Further work will also include exploitation of contextual information in combination with the semantic types to confine the set of potential antecedents.

The annotations will eventually be applied in an automatic procedure that can be envisaged as follows: Given some text, we first identify abstract anaphors. Next, the semantic type of each anaphor is determined, based on contextual features derived from annotated corpus data. The anaphor’s semantic type restricts the semantic type of the antecedent, and thus narrows the search space. Finally, the antecedent is identified with the help of these semantic restrictions and, again, with contextual features derived from the corpus.

Even if abstract anaphora are not able to be reliably resolved in the near future, their identification would be an important preprocessing step in a general coreference resolution system. Performance would be improved if abstract anaphors were not assigned incorrectly to an NP antecedent.

References

1. Artstein, R., Poesio, M.: Identifying reference to abstract objects in dialogue. In: Proceedings of Brandial 2006, pp. 56–63 (2006)
2. Artstein, R., Poesio, M.: Inter-coder agreement for computational linguistics (survey article). *Computational Linguistics* **34**(4), 555–596 (2008)
3. Asher, N.: *Reference to Abstract Objects in Discourse*. Boston: Kluwer Academic Publishers (1993)
4. Botley, S.: Indirect anaphora: Testing the limits of corpus-based linguistics. *International Journal of Corpus Linguistics* **11**(1), 73–112 (2006)
5. Botley, S., McEnery, T.: Demonstratives in English: A corpus-based study. *Journal of English Linguistics* **29**, 7–33 (2001)
6. Byron, D.K.: Resolving pronominal reference to abstract entities. In: Proceedings of ACL-02, pp. 80–87 (2002)
7. Byron, D.K.: Annotation of pronouns and their antecedents: A comparison of two domains (2003). Technical Report, University of Rochester
8. Consten, M., Knees, M.: Complex anaphors — ontology and resolution. In: Proceedings of the 15th Amsterdam Colloquium, pp. 65–70 (2005)
9. Consten, M., Knees, M., Schwarz-Friesel, M.: The function of complex anaphors in texts: Evidence from corpus studies and ontological considerations. In: *Anaphors in Text*, pp. 81–102. Amsterdam: John Benjamins (2007)

-
10. Dipper, S., Zinsmeister, H.: Annotating discourse anaphora. In: Proceedings of the ACL-IJCNLP Linguistic Annotation Workshop (LAW III), pp. 166–169 (2009)
 11. Dipper, S., Zinsmeister, H.: Annotation guidelines “Discourse-Deictic Anaphora”. Draft (2009). Universities of Bochum and Konstanz
 12. Dowty, D.: Word meaning and Montague Grammar. Dordrecht: Reidel (1979)
 13. Eckert, M., Strube, M.: Dialogue acts, synchronising units and anaphora resolution. *Journal of Semantics* **17**(1), 51–89 (2000)
 14. Fellbaum, C.: WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press (1998)
 15. Gundel, J.K., Hedberg, N., Zacharski, R.: Demonstrative pronouns in natural discourse. In: Proceedings of the 5th Discourse Anaphora and Anaphora Resolution Colloquium (DAARC-2004), pp. 81–86 (2004)
 16. Hedberg, N., Gundel, J.K., Zacharski, R.: Directly and indirectly anaphoric demonstrative and personal pronouns in newspaper articles. In: Proceedings of the 6th Discourse Anaphora and Anaphora Resolution Colloquium (DAARC-2007), pp. 31–36 (2007)
 17. Hegarty, M., Gundel, J.K., Borthen, K.: Information structure and the accessibility of clausally introduced referents. *Theoretical Linguistics* **27**(2-3), 163–186 (2001)
 18. Koehn, P.: Europarl: A parallel corpus for statistical machine translation. In: Proceedings of the 10th Machine Translation Summit (MT Summit), pp. 79–86 (2005)
 19. Kučová, L., Hajičová, E.: Coreferential relations in the Prague Dependency Treebank. In: Proceedings of the 5th Discourse Anaphora and Anaphora Resolution Colloquium (DAARC-2004), pp. 97–102 (2004)
 20. Lappin, S., Leass, H.J.: An algorithm for pronominal anaphora resolution. *Computational Linguistics* **20**(4), 535–561 (2004)
 21. Maienborn, C.: Die logische Form von Kopula-Sätzen. Berlin: Akademie Verlag (2003)
 22. Müller, C.: Resolving *it*, *this*, and *that* in unrestricted multi-party dialog. In: Proceedings of ACL-07, pp. 816–823 (2007)
 23. Müller, C.: Fully automatic resolution of ‘it’, ‘this’, and ‘that’ in unrestricted multi-party dialog. Ph.D. thesis, University of Tübingen (2008)
 24. Navarretta, C.: Pronominal types and abstract reference in the Danish and Italian DAD corpora. In: Proceedings of the 2nd Workshop on Anaphora Resolution (WAR II), pp. 63–71 (2008)
 25. Navarretta, C., Olsen, S.: Annotating abstract pronominal anaphora in the DAD project. In: Proceedings of LREC-08, pp. 2046–2052 (2008)
 26. Ng, V.: Supervised noun phrase coreference research: The first fifteen years. In: Proceedings of ACL-10, pp. 1396–1411 (2010)
 27. Poesio, M., Artstein, R.: Anaphoric annotation in the ARRAU corpus. In: Proceedings of LREC-08, pp. 1170–1174 (2008)
 28. Poesio, M., Modjeska, N.N.: Focus, activation, and *this*-noun phrases: An empirical study. In: A. Branco, T. McEnery, R. Mitkov (eds.) *Anaphora Processing*, pp. 429–442. Amsterdam/Philadelphia: John Benjamins (2005)
 29. Pradhan, S.S., Ramshaw, L., Weischedel, R., MacBride, J., Micciulla, L.: Unrestricted coreference: Identifying entities and events in OntoNotes. In: Proceedings of the International Conference on Semantic Computing (ICSC 2007), pp. 446–453 (2007)
 30. Recasens, M.: Discourse deixis and coreference: Evidence from AnCora. In: Proceedings of the 2nd Workshop on Anaphora Resolution (WAR II), pp. 73–82 (2008)
 31. Recasens, M.: Towards coreference resolution for Catalan and Spanish. Master’s thesis, University of Barcelona (2008)
 32. Recasens, M., Martí, M.A.: AnCora-CO: Coreferentially annotated corpora for Spanish and Catalan. *Language Resources and Evaluation* pp. 1–31 (2009). DOI 10.1007/s10579-009-9108-x
 33. Vendler, Z.: *Linguistics in Philosophy*, chap. Verbs and Time, pp. 97–121. Ithaca: Cornell University Press (1967)
 34. Vieira, R., Salmon-Alt, S., Gasperin, C.: Coreference and anaphoric relations of demonstrative noun phrases in a multilingual corpus. In: A. Branco, T. McEnery, R. Mitkov (eds.) *Anaphora Processing: Linguistic, Cognitive and Computational Modelling*, pp. 385–403. Amsterdam: John Benjamins (2005)
 35. Webber, B.L.: Discourse deixis: Reference to discourse segments. In: Proceedings of ACL-88, pp. 113–122 (1988)
 36. Webber, B.L.: Structure and ostention in the interpretation of discourse deixis. *Language and Cognitive Processes* **6**, 107–135 (1991)