Antecedent and Referent Types of Abstract Pronominal Anaphora

Costanza Navarretta
University of Copenhagen

Abstract

This paper is about the relation between pronominal types, syntactic types of the antecedent, semantic type of the referent and anaphoric distance in the Danish part of the DAD corpus comprising written and spoken data. These aspects are important to understand the use of abstract anaphora and to process them automatically and some of them have been investigated previously (see i.a. Webber (1988); Gundel et al. (2003); Navarretta (2010)). Differing from preceding studies, we extend the analysis of the syntactic types of the antecedent to include a fine-grained classification of clausal types and also investigate the anaphoric distance. The most common antecedent types in the data are *subordinate clause* and *simple main clause* and most abstract anaphora occurred in the clause which followed the antecedent or the clause in which the antecedent occurred. There is no clear dependence between the type of antecedent clause and the type of referent.

1 Introduction

In this paper we analyse the relation between abstract anaphora, syntactic types of antecedent, semantic types of referent and anaphoric distance in an annotated corpus of Danish texts, monologues and dialogues. Abstract anaphora indicate here third person singular pronouns which have as antecedents copula predicates, verbal phrases, clauses and discourse segments of varying size and refer to abstract types such as properties, events, situations and propositions. Abstract anaphors are also known as *impure textual deictics* (Lyons, 1977) and *discourse deictics* (Levinson, 1987; Webber, 1991) while reference to abstract entities independently of the type of antecedent has been called *situation reference* (Fraurud, 1992) and *abstract object reference* (Asher, 1993). An example of abstract anaphor in Danish is in (1) where the stressed pronoun *duet*[1] (this/that) has as antecedent the precedent utterance *She would certainly ensure that he came over there*:

(1) A: *Hun skulle nok sørge for han kom derover*
        'She would certainly ensure that he came over there'

     B: *d'et kunne jeg godt være sikker på*
        'of THAT I could be completely sure'
        (LANCHART)

Abstract anaphora are very frequent in languages such as English (Byron and Allen, 1998; Gundel et al., 2003) and Danish (Navarretta, 2000), but abstract pronominal reference varies from language to language (see i.a. Fraurud (1992); Borthen et al. (1997); Kaiser (2000); Navarretta (2002, 2010)).

Resolving abstract anaphora automatically is difficult because the antecedents belong to various syntactic types and have varying size. The anaphor can immediately

---

[1]We mark a stressed vowel with an apostrophe before its occurrence.

follow its antecedent, but it can also occur several clauses later. Furthermore there is not a one to one relation between the antecedent's syntactic type and the referent's semantic type (see Webber (1991); Gundel et al. (2003)).

Various algorithms for resolving abstract anaphora in English have been proposed (i.a. Eckert and Strube (2001); Byron (2002); Strube and Müller (2003); Müller (2007)). These algorithms rely on the pronominal type, on linguistic, semantic, domain-specific knowledge and/or on the annotations in domain-specific corpora. Their results are still not good enough to be used in practical applications and the evaluation of the algorithms indicates that the recognition of the anaphoric uses of the pronouns and the identification of the antecedents are some of the most problematic aspects. Although Danish anaphora have different characteristics than the English ones, the identification of the correct antecedent is also problematic in this language (Navarretta, 2002, 2004a).

The aim of the present work is to provide more knowledge about the use of abstract anaphora in Danish.

In previous investigations of abstract anaphora in Danish we have looked at the relation between type of pronouns, semantic referent types and syntactic types of antecedent. In the present work we extend the investigation to comprise a very fine-grained analysis of different types of clausal antecedents and to include anaphoric distance that is the number of clauses between the abstract anaphor and its antecedent. The anaphoric distance has been seen as one important salience indicator and it is the determining factor behind the accessibility hierarchy of nominal referring expressions proposed by Ariel (1988, 1994). Because the semantic type of the referent depends on the context in which the abstract anaphor occurs (see i.a. Webber (1991); Eckert and Strube (2001)), we believe that finding possible relations between these semantic types, the syntactic types of the antecedent, the anaphoric distance and the types of pronoun can contribute to the construction of resolution algorithms.

The paper is organised as follows. In Section 2 we discuss related work and in Section 3 we shortly describe the Danish abstract anaphora and the corpus upon which we base our research. In Section 4 we present and discuss the results of our investigation and in Section 5 we conclude.

2 Background work

Webber (1991) notices that abstract anaphors with the same antecedent can refer to objects of different semantic type depending on the context in which the anaphor occurs. She suggests that abstract pronouns create their referents in the moment they are uttered by an act of *ostension*.

The relation between the type of abstract pronoun and the syntactic type of the antecedent has been addressed by i.a. Webber (1988); Hegarty (2003); Gundel et al. (2003, 2004); Navarretta (2004, 2007, 2010). In particular Webber (1988) reports that

personal pronouns in English often cannot refer to abstract entities when the antecedent is a clause, because the clause is not accessible to the pronoun. In corpus-based studies of the occurrences of abstract anaphora in English Byron and Allen (1998), Gundel et al. (2003) and Hedberg et al. (2007) confirm Webber's observation .

Hegarty (2003) explains the frequency of occurrence of demonstrative pronouns with clausal antecedents in terms of the *Givenness Hierarchy* (Gundel et al., 1993). According to him entities introduced in discourse by clauses are only activated in the cognitive status of the addressee, while entities introduced in discourse by verbal phrases are similar to entities introduced in discourse by nominal phrases: they are often in focus and can be referred to by the personal pronoun *it*. Because the most common referent types when the antecedents are clauses are facts, situations and propositions, demonstrative pronouns refer much more often to facts, situation and propositions than personal pronouns do. On the other hand the referents with verbal phrases antecedents are events or states and they are thus often referred to by personal pronouns.

Navarretta (2007, 2010) reports that, differing from the English *it*, Danish and Italian personal pronouns have often clausal antecedents and thus in numerous cases they refer to facts, situations and propositions. She explains one of the differences in abstract reference between Danish, English and Italian in terms of the three languages' pronominal systems and syntactic structure. With respect to syntax she notices that in Danish constructions such as left dislocation and clefts are much more frequent than in English. These constructions put the clauses in focus in information structure terms because and thus the clauses are in these cases very salient, or *in focus* in terms of the *Givenness Hierarchy* (Gundel et al., 1993).

In our preceding studies we accounted for the characteristics of anaphora with nominal phrase and verbal phrase antecedents opposed to anaphora with clausal antecedents and on the differences between Danish, English and Italian. In the present work we focus on the Danish data and add to our investigation the analysis of the types of clausal antecedent and the anaphoric distance and their relation to the referent types and the pronominal types.

3   The data

3.1   Abstract anaphora in Danish

In Danish texts two abstract pronouns are used: the pronoun *det* (it/this/that) which is ambiguous with respect to its pronominal type and the demonstrative pronoun *dette* (this). In spoken Danish abstract pronouns comprise the unstressed personal pronoun *det* (it) and the stressed demonstrative pronouns *d'et* (this/that), *d'et h'er* (this) *d'et d'er* (that). The demonstrative pronoun *dette* (this) occurs extremely seldom in spoken language and in our data it only occurred two times and in both cases it had a nominal phrase antecedent.

3.2 The corpus

Our study is mainly based on the Danish part of the DAD corpora (Navarretta and Olsen, 2009) which consist of the following data:

- Transcriptions of the DANPASS corpus (Grønnum, 2006) which is the Danish version of the MAPTASK corpus (Anderson et al., 1991) and comprises both dialogues and monologues. The DANPASS dialogues contain 52,145 running words while the monologues consist of 21,224 words.

- Transcriptions of multiparty spontaneous dialogue extracts from the LANCHART corpus (Gregersen, 2007) comprising 24,112 running words.

- Transcriptions of two TV-interviews from the Danish public television *DR* (2,192 words).

- Translations from Italian of three Pirandello (1922-1937) stories (11,280 words).

- EU texts (24,389 words).

- Danish juridical texts (11,600 words).

- Extracts of newspaper and journal articles, novels and reports (12,570 words) from the Danish general language PAROLE corpus (Keson and Norling-Christensen, 1998).

3.3 The annotations

The DAD data contain many types of annotation such as structural information for the texts and speaker and turn information for the dialogues, PoS and lemma information, information about the functions and uses of third person singular neuter pronouns and especially their anaphoric uses (Navarretta, 2010). A description of the annotation schemes used and a report of inter-coder agreement measures for the various annotation types are in Navarretta and Olsen (2008); Navarretta (2009); Navarretta and Olsen (2009). The annotations which are relevant to the present work are the following:

- the type of pronoun, e.g. *det*, *dette*, unstressed *det*, stressed *det*;

- the antecedent;

- the syntactic type of the antecedent;

- the semantic type of the referent;

- the anaphoric distance in term of clauses.

The syntactic types of antecedent relevant to this work are *verbal phrase* (VP), *adjectival phrase*, *prepositional phrase* and *nominal phrase* in copula constructions (these three types are called *CPR* henceforth), *discourse segment* (DS) and *clause*. The

clausal type is furthermore distinguished in the following subtypes: *simple main clause* (CL) which covers main clauses which do not have subordinate clauses, *matrix clause* (MCL), *subordinate clause* (SCL) and *complex clause* (CCL) which comprise co-ordinated clauses and main clauses with their subordinated clauses. The choice of these clausal types was inspired by a classification of clauses which Kameyama (1998) adopts in an extended version of *Centering*.

The semantic types we consider in the following are: *property*, *eventuality*, *fact-like object*, *proposition-like object*. The latter three types are taken from the middle layer of the hierarchy of saturated abstract objects proposed by Asher (1993). The type *eventuality* comprises the types *state* and *event* which includes *activity*, *process*, *accomplishment* and *achievement*. *Fact-like object* includes *possibility*, *situation*, *fact* and *state of affairs*, while *proposition-like object* comprises *pure proposition*, *question*, *command* and *desire*. For simplicity in this paper we have included referents coded as *speech act* in the *proposition-like object* type.

Referent type	Det	Dette	Total
CL			
eventuality	15	2	17
fact-like	13	17	30
proposition-like	5	1	6
total	33	20	53
CCL			
fact-like	4	7	11
proposition-like	1	1	2
total	5	8	13
SCL			
eventuality	16	11	27
fact-like	14	17	31
proposition-like	6	4	10
total	36	32	68
MCL			
eventuality	2	0	2
fact-like	1	0	1
total	3	0	3
DS			
eventuality	2	1	3
fact-like	3	5	8
proposition-like	1	1	2
total	6	7	13
VP			
eventuality	17	3	20
CPR			
property	13	1	14

Table 1: Pronouns, antecedent and referent types in texts

4 Investigating the data

We have extracted from the annotated corpora the pronominal types, their antecedent and referent types and the anaphoric distance. Differing from Navarretta (2010) we do not consider here abstract anaphora chains, that is abstract anaphors having abstract pronominal anaphors as antecedents.

4.1 Syntactic antecedent and semantic referent types

Table 1 contains the occurrences of personal and demonstrative pronouns, the syntactic type of their antecedents and the semantic type of their referents in the Danish texts.

These data indicate that the majority of the abstract anaphors in the texts had a subordinate clause or a simple main clause antecedent. The ambiguous pronoun *det* is the most common abstract anaphor with all antecedents, but especially with copula predicate or verbal phrase antecedents. Both personal and demonstrative pronouns occur with all types of referent, but the demonstrative pronoun *dette* only seldom refers to a property. Demonstrative pronouns refer more frequently to *fact-like objects* than to other types of objects.

Table 2 shows the abstract anaphors, the antecedent and referent types which occurred in the monologues.

Referent type	Unstressed pronoun	Stressed pronoun	Total
CL			
eventuality	1	1	2
fact-like	0	1	1
proposition-like	3	9	12
total	4	11	15
CCL			
fact-like	2	0	2
proposition-like	8	1	9
total	10	1	11
SCL			
proposition-like	1	0	1
DS			
proposition-like	1	0	1
VP			
eventuality	1	2	3
CPR			
property	3	3	6

Table 2: Pronoun, antecedent and referent types in monologues

The data in the table indicate that the most common antecedent types of abstract anaphors in the monologues are *simple main clause* and *complex clause* and that the most common referent type of the unstressed pronoun *det* is *proposition-like object*. The occurrences of pronominal, antecedent and referent types in the dialogues are in Table 3.

Referent type	Unstressed pronoun	Stressed pronoun	Total
CL			
eventuality	26	26	52
fact-like	51	44	95
proposition-like	35	7	42
total	112	77	189
CCL			
eventuality	6	6	12
fact-like	11	14	25
proposition-like	12	5	17
total	29	25	54
SCL			
eventuality	8	4	12
fact-like	7	10	17
proposition-like	6	7	13
total	21	21	42
MCL			
fact-like	1	0	1
proposition-like	0	1	1
total	1	1	2
DS			
eventuality	4	2	6
fact-like	9	4	13
proposition-like	4	4	8
total	17	10	27
VP			
eventuality	52	35	87
CPR			
property	21	11	33

Table 3: Pronouns, antecedent and referent types in dialogues

Also in the dialogues the most frequently occurring antecedent syntactic type of the abstract anaphors is *simple main clause*. *Complex clause* and *verbal phrase* are also quite frequent antecedent types. The latter type is especially frequent in the maptask dialogues. This is not surprising given that maptask dialogues are interactions between a speaker (the giver) who gives instructions to a second speaker (the follower) on how to reach a place on a map. The two speakers cannot see each other and they have two slightly different maps.

As in the other types of the DAD data also in the monologues the personal pronoun, the unstressed *det*, is the most frequently occurring abstract pronoun. The demonstrative pronouns refer more frequently to *fact-like* objects than to other types of abstract objects, while proposition-like objects are more often referred to by the personal pronouns than by the demonstrative ones, as also described by Navarretta (2010).

Concluding most abstract anaphors in our texts have a single subordinate clause or simple main clause antecedent, while in dialogues they have a simple main clause

or a complex clause antecedent. Clauses are the antecedents of both demonstrative pronouns and personal pronouns as also reported in Navarretta (2010). The analysis of the referents' semantic types also confirms the results in Navarretta (2010) and shows that proposition-like objects are more often referred to by personal pronouns than by demonstrative pronouns and that the preferred referent type of demonstrative pronouns is *fact-like object*. Matrix clauses are only seldom the antecedents of abstract anaphora.

4.2 Anaphoric distance

In Table 4 we show the distance between abstract anaphors and their antecedents in terms of clauses.

Distance	Texts	Monologues	Dialogues	Total
zero	163	34	266	463
one	13	2	119	134
two	3	1	28	32
three	2	0	13	15
four	1	0	3	4
five	1	0	0	1
six	1	0	0	1
nine	0	0	1	1
ten	0	0	1	1
eleven	0	0	2	2

Table 4: Anaphoric distance in clauses

Table 5 contains the most frequent combinations (more than 8 occurrences) of referent type, antecedent syntactic type, pronominal type and a certain anaphoric distance in the texts. Table 6 shows the same data for the dialogues and the monologues.

Total	Distance	Referent type	Antecedent type	Pronominal type
16	ZERO	FACT-LIKE	CL	dette
14	ZERO	EVENTUALITY	SCL	det
13	ZERO	EVENTUALITY	CL	det
13	ZERO	FACT-LIKE	SCL	dette
12	ZERO	EVENTUALITY	VP	det
12	ZERO	FACT-LIKE	CL	det
11	ZERO	FACT-LIKE	SCL	det
10	ZERO	EVENTUALITY	SCL	dette

Table 5: Anaphoric distance, referent, antecedent and pronominal types in texts

The data in these tables indicate that the majority of the antecedents of abstract anaphors (70.8% of the cases) in our corpora occur in the clause or are the clause which immediately precedes the clause where the anaphor occurs. In 20.5% of the cases the anaphoric distance is one, in 4.9% of the cases there are two clauses in between the anaphor and the antecedents, in 2.3% of the cases the anaphoric distance is of three clauses and in 0.6% of the cases there are four clauses in-between the anaphor and its antecedent. Larger anaphoric distance (up to 11 clauses/utterances) occur very

Total	Distance	Referent type	Antecedent type	Pronominal type
Dialogues				
30	ZERO	FACT-LIKE	CL	unstressed
28	ZERO	EVENTUALITY	VP	unstressed
27	ZERO	EVENTUALITY	VP	stressed
27	ZERO	FACT-LIKE	CL	stressed
21	ZERO	PROP-L	CL	unstressed
17	ONE	EVENTUALITY	VP	unstressed
16	ZERO	EVENTUALITY	CL	stressed
16	ONE	FACT-LIKE	CL	unstressed
13	ZERO	EVENTUALITY	CL	unstressed
10	ZERO	PROP	CPR	unstressed
10	ONE	EVENTUALITY	CL	unstressed
9	ZERO	FACT-LIKE	CCL	stressed
9	ZERO	PROP-L	CCL	unstressed
Monologues				
9	ZERO	EVENTUALITY	CL	unstressed

Table 6: Anaphoric distance, referent, antecedent and pronominal types in dialogues and monologues

seldom and only in the multiparty dialogues. It must be noticed that in dialogues short utterances like *yes* and *okay* are counted as clauses and there might occur more of them uttered by different speakers. This partly explains the occurrences of long distance abstract anaphors in our data. The syntactic types which occurred more frequently when the anaphoric distance is more than one clause are subordinate clause and discourse segments in the texts, simple main clauses and discourse segments in the dialogues. Only referents of the types *eventuality* and *fact-like objects* occur when the anaphoric distance is of more than one clause and in nearly all cases the anaphors with distant antecedents are occurrences of the unstressed pronoun in dialogues and the ambiguous pronoun *det* in texts.

5 Conclusion

In this paper we have presented an investigation of the relation between syntactic antecedent types, semantic referent types, pronominal types and anaphoric distance in Danish written and spoken data. The results of our study show that the most frequently occurring antecedent types of both abstract personal and demonstrative pronouns in texts are *subordinate clause* and *simple main clause*, while in spoken data they are *simple main clause* and *complex clause*. In the maptask dialogues *verbal phrase* is also a common antecedent type and this is not surprising given the interaction type. Matrix clauses occur only seldom as the antecedents of abstract anaphors.

The investigation of the semantic types of the referents of abstract anaphors in the data confirms the studies reported in Navarretta (2010) showing that proposition-like

objects are more often referred to by personal pronouns than by demonstrative pronouns and that the preferred referents of demonstrative pronouns are fact-like objects.

The study of the anaphoric distance indicates that 70.8% of the abstract anaphors occur in the clause that immediately follows the clause in which the antecedent is (or the clause that is the antecedent) . In 20.5% of the cases there is a clause in between the anaphor and the antecedent and in 4.9% of the cases there are two clauses in between the anaphor and its antecedent. Larger anaphoric distance occur seldom and mostly in the dialogues. The syntactic types which occurred most frequently at long distance are subordinate clause and discourse segments in the texts, simple main clauses and discourse segments in the dialogues. Only eventuality and fact-like referents are referred to when the anaphoric distance is of more than one clause and in nearly all cases the anaphor with distant antecedents is the unstressed pronoun in dialogues and the ambiguous pronoun *det* in texts.

Currently we are testing the relation between the various annotation types using machine learning algorithms.

References

Anne H. Anderson, Miles Bader, Ellen Gurman Bard, Elizabeth Boyle, Gwyneth Doherty, Simon Garrod, Stephen Isard, Jacqueline Kowtko, Jan McAllister, Jim Miller, Catherine Sotillo, Henry S. Thompson, and Regina Weinert. The HCRC Map Task Corpus. *Language and Speech*, 34:351–366, 1991.

Mira Ariel. Referring and accessibility. *Journal of Linguistics*, 24(1):65–87, 1988.

Mira Ariel. Interpreting anaphoric expressions: a cognitive versus a pragmatic approach. *Journal of Linguistics*, 30(1):3–40, 1994.

Nicholas Asher. *Reference to Abstract Objects in Discourse*, volume 50 of *Studies in Linguistics and Philosophy*. Kluwer Academic Publishers, Dordrecht, the Netherlands, 1993.

Kaja Borthen, Thorstein Fretheim, and Jeanette K. Gundel. What brings a higher-order entity into focus of attention? Sentential pronouns in English and Norwegian. In Ruslan Mitkov and Branimir Boguraev, editors, *Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts*, pages 88–93, 1997.

Donna K. Byron. Resolving pronominal reference to abstract entities. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL 2002)*, pages 80–87, 2002.

Donna K. Byron and James Allen. Resolving demonstrative pronouns in the TRAINS93 corpus. In *Proceedings of the Second Colloquium on Discourse Anaphora and Anaphor Resolution (DAARC 1998)*, pages 68–81, 1998.

Miriam Eckert and Michael Strube. Dialogue acts, synchronising units and anaphora resolution. *Journal of Semantics*, 17(1):51–89, 2001.

Kari Fraurud. *Processing Noun Phrases in Natural Discourse*. Department of Linguistics - Stockholm University, 1992.

Frans Gregersen. The LANCHART Corpus of Spoken Danish, Report from a corpus in progress. In *Current Trends in Research on Spoken Language in the Nordic Countries*, pages 130–143. Oulu University Press, 2007.

Nina Grønnum. DanPASS - A Danish Phonetically Annotated Spontaneous Speech Corpus. In N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odijk, and D. Tapias, editors, *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC)*, Genova, Italy, May 2006.

Jeanette K. Gundel, Nancy Hedberg, and Ron Zacharski. Cognitive status and the form of referring expressions in discourse. *Language*, 69(2):274–307, 1993.

Jeanette K. Gundel, Nancy Hedberg, and Ron Zacharski. Cognitive status, information structure, and pronominal reference to clausally introduced entities. *Journal of Logic, Language and Information*, 12:281–299, 2003.

Jeanette K. Gundel, Nancy Hedberg, and Ron Zacharski. Demonstrative pronouns in natural discourse. In A. Branco, T. McEnery, and R. Mitkov, editors, *Proceedings of DAARC-2004 - The 5th Discourse Anaphora and Anaphora Resolution Colloquium*, pages 81–86, Furnal, S.Miguel, Portugal, 2004. Ediçoes Colibri.

Nancy Hedberg, Jeanette K. Gundel, and Ron Zacharski. Directly and indirectly anaphoric demonstrative and personal pronouns in newspaper articles. In António Branco, T. McEnery, Ruslan Mitkov, and F. Silva, editors, *Proceedings of DAARC-2007 - the 6th Discourse Anaphora and Anaphora Resolution Colloquium*, pages 31–36, Lagos, Portugal, March 2007.

Michael Hegarty. Semantic types of abstract entities. *Lingua*, 113:891–927, 2003.

Elsi Kaiser. Pronouns and demonstratives in Finnish: Indicators of Referent Salience. In P. Baker, A. Hardie, T. McEnery, and A. Siewierska, editors, *Proceedings of the Discourse Anaphora and Anaphor Resolution Conference*, volume 12 of *University Center for Computer Corpus Research on Language - Technical Series*, pages 20–27, Lancaster, UK, 2000.

Megumi Kameyama. Intrasentential centering: A case study. In M. Walker, A. Joshi, and E. Prince, editors, *Centering Theory in Discourse*, pages 89–112. Oxford University Press, Oxford, U.K., 1998.

Britt Keson and Ole Norling-Christensen. PAROLE-DK. Technical report, Det Danske Sprog- og Litteraturselskab, http://korpus.dsl.dk/e-resurser/parole-korpus.php, 1998.

Stephen C. Levinson. Pragmatics and the grammar of anaphora: A partial pragmatic reduction of Binding and Control Phenomena. *Journal of Linguistics*, 23(2):379–434, 1987.

John Lyons. *Semantics*, volume I-II. Cambridge University Press, 1977.

Christoph Müller. Resolving It, This and That in unrestricted multi-party dialog. In *Proceedings of ACL-2007*, pages 816–823, Prague, 2007.

Costanza Navarretta. Centering-based anaphora resolution in Danish dialogues. In P. Sojka, I. Kopeček, and K. Pala, editors, *Text, Speech and Dialogue - Proceedings of the 3rd International Workshop (TSD 2000)*, pages 345–350, Brno, Czech Republic, 2000.

Costanza Navarretta. *The Use and Resolution of Intersentential Pronominal Anaphora in Danish Discourse*. PhD thesis, Centre of Language Technology and Department of General and Applied Linguistics Copenhagen University, 2002.

Costanza Navarretta. The main reference mechanisms of Danish demonstrative pronominal anaphors. In A. Branco, T. McEnery, and R. Mitkov, editors, *Proceedings of DAARC-2004 - the 5th Discourse Anaphora and Anaphora Resolution Colloquium*, pages 115–120, Furnal, S.Miguel, Portugal, 2004. Ediçoes Colibri.

Costanza Navarretta. Resolving individual and abstract anaphora in texts and dialogues. In *Proceedings of the 20th International Conference of Computational Linguistics, COLING-2004*, pages 233–239, Geneva, Switzerland, 2004a.

Costanza Navarretta. A contrastive analysis of the use of abstract anaphora. In A. Branco, T. McEnery, R. Mitkov, and F. Silva, editors, *In Proceedings of DAARC-2007 - 6th Discourse Anaphora and Anaphora Resolution Colloquium*, pages 103–109, Lagos, Portugal, March 2007. Centro de Linguistica da Universidade do Porto.

Costanza Navarretta. Co-referential chains and discourse topic shifts in parallel and comparable corpora. *Revista de Procesamiento de Lenguaje Natural - La Sociedad Espaõla para el Procesamiento del Lenguaje Natural (SEPLN)*, 42:105–102, 2009.

Costanza Navarretta. The DAD parallel corpora and their uses. In *Proceedings of LREC 2010*, pages 705–712, Malta, May 2010. ELRA.

Costanza Navarretta and Sussi Olsen. Annotating abstract pronominal anaphora in the DAD project.

In *Proceedings of LREC-2008*, Marrakesh, Morocco, May 2008. ELRA.

Costanza Navarretta and Sussi Olsen. The annotation of pronominal abstract anaphora in Danish texts and dialogues. DAD report 1, Centre for Language Technology, University of Copenhagen, January 2009.

Luigi Pirandello. *Novelle per un anno*. Giunti, 1922-1937.

Michael Strube and Christoph Müller. A machine learning approach to pronoun resolution in spoken dialogue. In *Proceedings of the ACL'03*, pages 168–175, 2003.

Bonnie L. Webber. Discourse deixis and discourse processing. Technical report, University of Pennsylvania, 1988.

Bonnie L. Webber. Structure and ostension in the interpretation of discourse deixis. *Natural Language and Cognitive Processes*, 6(2):107–135, January 1991.