Chapter 1

Anaphora Resolution in Danish Dialogues

COSTANZA NAVARRETTA
Center for Sprogteknologi
Copenhagen - Denmark
costanza@cst.ku.dk

ABSTRACT. In this paper we describe how we have modified Eckert and Strube's algorithm [Eckert and Strube,1999a, Eckert and Strube,1999b] for resolving anaphora in English dialogues, so that it applies to Danish data. The first results of the application of the modified algorithm to Danish dialogues are discussed and evaluated.[1]

1 Introduction

The resolution of anaphors is an important task in NLP and, in particular, in dialogue systems. One of the theories for anaphora resolution that has been studied and tested on many languages is centering [Grosz et al.,1995]. Centering is interesting because it reflects the assumption that some entities in an utterance are more central than others and this fact influences the use of referring expressions. Centering fits into Grosz and Sidner's model of discourse structure [Grosz and Sidner,1986] according to which a discourse is composed of discourse segments which exhibit *global coherence*. A discourse segment is composed of a sequence of utterances which exhibit *local coherence*. The latter phenomenon is accounted for by the centering theory.

Centering-based algorithms, as most of the existing anaphora resolution algorithms, have been designed for dealing with anaphors having individual NPs as antecedents and have been tested on written texts. Recently centering has also been applied to dialogues, i.a. [Byron and Stent,1998, Eckert and Strube,1999a, Eckert and Strube,1999b]. In particular Eckert and Strube describe an algorithm for the resolution of anaphors in English dialogues which applies to both anaphors referring to individual NPs and

[1]The research described in this paper has been done under the Staging project funded by the Danish Research Councils.

discourse deictics. The algorithm is based on rules for discriminating among anaphors referring to individual NPs and discourse deictics, mainly determined by the syntactic constructions in which the anaphors occur. The individual anaphors are then resolved with a modified centering algorithm [Strube,1998] while a new algorithm is developed for resolving discourse deictics. The precision of the algorithm for discourse deictics was 63,6%, the recall 70%. The low precision was mainly due to the fact that the algorithm cannot distinguish between deictics which refer to the preceding clause and vague deictics, i.e. deictics which vaguely refer to concepts or more clauses in the preceding discourse. Eckert and Strube point out that looking at different classifications of verbs and nominals could improve the discriminating rules their algorithm is based on. We believe that the algorithm is a good starting point for investigating how far one can go in resolving anaphors in dialogues exclusively on the basis of the local contexts in which the anaphors occur. We have adapted the algorithm to Danish data and applied this modified algorithm to Danish dialogues.[2]

This paper is organized as follows. In section 2 we outline Eckert and Strube's algorithm, henceforth the ES99-algorithm. In section 3 we shortly present Danish personal and demonstrative pronouns focusing on discourse deictics found in the two dialogue corpora Bysoc and SL, both collected by researchers at the University of Copenhagen. In section 4 we describe how we have adapted the ES99-algorithm so that it can be applied to Danish data. In section 5 we evaluate the results obtained by manually testing the modified ES99-algorithm on SL dialogues and in section 6 we outline future work for improving the algorithm and make some concluding remarks.

2 Eckert and Strube's Algorithm

The ES99-algorithm is based on a classification of anaphor types in dialogues, based on the analysis of a collection of telephone conversations. The types which Eckert and Strube identify are individual anaphors, discourse deictics, inferrable-evoked anaphors and vague anaphors. Inferrable-evoked pronouns refer to the use of the plural pronoun *they* indirectly co-specifying with a singular NP, indicating a country or an institution. Vague anaphors are anaphors which do not refer to a specific linguistic object in the discourse. Other types of pronoun, such as first and second person pronouns, expletives and other types of pronoun are not included in the study. Predicates that are preferentially associated with abstract objects are marked as **I-incompatible** (*I) while predicates that are preferentially associated with individual objects are marked as **A-incompatible** (*A). Eckert and

[2]Tests of two centering-based algorithms for resolving anaphors with individual nominal antecedents on Danish texts are reported in [Navarretta,forthcoming].

Strube assume the following ***I** predicates for English
[Eckert and Strube,1999a][p. 40]:

- Equating constructions where a pronominal referent is equated with an abstract object, e.g., *x is making it easy, x is a suggestion.*
- Copula constructions whose adjectives can only be applied to abstract entities, e.g., *x is true, x is false, x is correct, x is right, x isn't right.*
- Arguments of verbs describing propositional attitude which take S'-complements, e.g., *assume.*
- Object of *do.*
- Predicate or anaphoric referent is a "reason", e.g., *x is because I like her, x is why he's late.*

In a parallel way Eckert and Strube define the following ***A** predicates [Eckert and Strube,1999a][p. 40]:

- Equating constructions where a pronominal referent is equated with a concrete individual referent, e.g., *x is a car.*
- Copula constructions with adjectives which can only be applied to concrete entities, e.g., *x is expensive, x is tasty, x is loud.*
- Arguments of verbs describing physical contact/stimulation, which cannot be used anaphorically, e.g., *break x, smash x, eat x, drink x, smell x* but NOT **see x*

Eckert and Strube use grounded acts as domain for the anaphor resolution algorithms in dialogues. In particular they recognize two simple dialogue acts, **Initiations** (**I**'s) and **Acknowledgments** (**A**'s). **I**'s have semantic content, while **A**'s are only used to ground the preceding **I**. **Acknowledgments/Initiations** (**A/I**'s) are dialogue acts that have both the function of grounding the preceding **I** and of establishing a new **I**. An **I** and the corresponding **A**, together with longer **I**'s in the same turn-taking which do not need to be acknowledged, constitute a **Synchronizing Unit** (**SU**). Short **I**'s which are not acknowledged are ignored by the resolution algorithms.

Individual anaphors are resolved with the modified centering algorithm described in [Strube,1998]. The algorithm consists in testing the referring expression against the elements in the so-called S-list[3] from left to right until the test succeeds. The S-list contains relevant entities from the previous **SU**

[3]The S-list substitutes both the backward-looking center and the forward-looking centers in the original centering theory.

(in the original algorithm the preceding utterance) and is continuously updated so that new elements are inserted according to the S-list ranking criteria. These criteria are based on the distinction between *hearer-old*, *mediated* and *hearer-new* discourse entities proposed in [Prince,1981, Prince,1992]. When the analysis of an **SU** is finished all the entities which were not realized in the **SU** are removed from the S-list.

In their treatment of reference to abstract objects Eckert and Strube follow i.a. [Webber,1991] in assuming that anaphoric discourse deictic reference involves reference coercion and that only discourse sections adjacent to the anaphor or, using Webber's terminology, sections on the right frontier of the discourse structure tree, are available for discourse-deictic reference. Furthermore they assume, as Asher [Asher,1993], that the type of abstract object is determined by the context in which the anaphor occurs. To resolve anaphors referring to abstract objects Eckert and Strube use a second list, called the A-list. The A-list is only filled when discourse deictics occur and its elements remain only for one **I**. The following context ranking describes the order in which the parts of the linguistic contexts are accessed [Eckert and Strube,1999a][p. 40]:

1. the A-list
2. within the same **I**: the clause to the left of the clause which contains the anaphor.
3. within the previous **I**: rightmost main clause and subordinated clauses to its right
4. within previous **I**'s: rightmost complete sentence, if previous **I** is an incomplete sentence.

The anaphora resolution algorithm for third person singular neuter personal pronouns is the following [Eckert and Strube,1999b]:

case PRO is **I-incompatible**
    if *resolveDiscourseDeictic*(PRO)
      then classify as discourse deictic
      else classify as vague pronoun;
case PRO is **A-incompatible**
    if *resolveIndividual*(PRO)
      then classify as individual pronoun
      else classify as vague pronoun;
case PRO is ambiguous
    if *resolveIndividual*(PRO)
      then classify as individual pronoun

else if *resolveDiscourseDeictic*(PRO)
then classify as discourse deictic
else classify as vague pronoun;

The same algorithm is used for demonstratives, with the exception that the last two **if**s constructions in the algorithm for pronouns are reversed reflecting the preference for demonstratives to be discourse deictics [Webber,1991].

3 Danish Data

In the following we shortly present Danish third person personal and possessive pronouns and demonstrative pronouns. The focus is on their discourse deictic use based on occurrences in the Danish dialogue collections. The Danish third person singular personal and possessive pronouns can be found in table 1, while the third person plural personal and possessive pronouns can be found in table 2.

gender	**subject**	**object**	**reflexive**	**possessive**	**pos.refl.**
feminine	*hun* she	*hende* her	*sig* herself	*hendes* hers	*si-n/t/ne* hers
masculine	*han* he	*ham* him	*sig* himself	*hans* his	*si-n/t/ne* his
common	*den* it	*den* it	*sig* itself	*dens* its	*si-n/t/ne* its
neuter	*det* it	*det* it	*sig* itself	*dets* its	*si-n/t/ne* its

Table 1: Third person singular pronouns

subjective	**objective**	**reflexive**	**possessive**
de they	*dem* them	*sig* themselves	*deres* their/ theirs

Table2: Third person plural pronouns

It must be noted that *den*, *det* and *de* are also used as definite articles (the) and demonstrative adjectives (this/that and these/those). In spoken language the demonstratives are always stressed. *Den*, *det*, *de* are demonstratives if followed by the adverbials *her* and *der*, in which case they correspond to the English *this/these* and *that/those* respectively. Furthermore the demonstratives *denne, dette* (this) and *disse* (these) exist.

Feminine and masculine pronouns generally refer to persons. Common gender pronouns refer to common gender nouns which do not denote humans. There are many common gender nouns denoting humans, but they are, in most cases, neutral as to the sex of the person they refer to. The neuter gender pronouns are used to refer to neuter nouns. They refer also to person nouns in neuter gender, where the sex is unknown or irrelevant, and in this case the agreement is a syntactic one. In the case were the sex is known or relevant, the appropriate feminine or masculine pronouns are used, i.e. the agreement is semantic.

In Danish the most frequently used discourse deictic is *det* which corresponds to both *it* and *that*. Other discourse deictics are *det her* (this) and *det der* (that).[4]. *Det* can also be used as expletive. The neuter demonstrative *dette* (this) has also a discourse deictic use, but is mostly used in written language and there were no occurrences of it in our dialogues.

As discourse deictic *det* co-refers[5] with an infinitive or a clause, as it is the case in the following examples:

(1) a. *At rejse er spændende men det er også dyrt*
(Travelling is exciting but it is also expensive)

b. *Du skal tage en blodprøve.*
(You have to take a blood test)
Hvorfor det?
(Why is that?)

Det is also used as the subject complement of *være* (be) and *blive* (become) in answers.

(2) a. *Er du træt?*
(are you tired?)
Det er jeg
(I am)

b. *Blev du færdig med opgaven?*
(Did you finish the task?)
Ja, det blev jeg
(Yes, I did)

Det refers to a verb phrase when it is used as the object complement for the verb *have* (have), *gøre* (do) and modal verbs as in

(3) a. *Peter dumpede og det gjorde jeg også*
(Peter failed, and so did I)

b. *Jeg gider ikke spise*
(I do not feel like eating)
Det skal du

[4]These two deictics can be used in most of the same contexts as *det*, but it seems to be a preference for using them to refer to more clauses.

[5]From now on we will simply write "refers to".

(You have to)

Det refers to a clause in constructions with attitude verbs and other verbs which take clausal complements, such as *synes* (think), *tro* (believe) and *vide* (know), *sige* (say), *håbe* (hope):

(4) *Jeg tror at det snart begynder at regne.*
(I think it will soon begin to rain)
Det håber jeg ikke
(I hope not)

In the latter two cases the pronoun *det* is nearly always topicalized, i.e. it appears before the main verb, in the place that usually is occupied by the subject[6]. *Det* can also refer to more clauses, or to something that can vaguely be inferred from the discourse.

4 The Modified ES99-algorithm

On the basis of the deictics in the two Danish dialogue corpora we have established the following ***I** predicates for Danish:

- constructions where a pronoun is equated with an abstract object, e.g., *x er et forslag* (x is a suggestion)
- copula constructions with adjectives which can only be applied to abstract entities, such as *x is sandt* (x is true), *x er usandt* (x is untrue), *x er rigtigt* (x is correct)
- arguments of verbs which take S'-complements, e.g., *tro* (believe), *antage* (assume), *mene* (think), *sige* (say)
- object of *gøre* (do)
- subject complement with *være* (be) and *blive* (become) in answers
- object of *have* (have) if the verb was not used as a main verb in the previous clause
- object of modal verbs
- predicate or anaphoric referent is a "reason", such as *x er fordi du er holdt op med at ryge* (x is because you have stopped smoking) *x er på grund af at du er gravid* (x is because you are pregnant)

[6]This place is called *fundamentfelt* (actualization field) by [Diderichsen,1946].

We have assumed the following ***A**:

- constructions where a pronominal referent is equated with a concrete individual referent, such as *x er en legemsdel* (x is a body part), *x er et barn* (x is a baby)
- copula constructions with adjectives which can only be applied to concrete entities, such as *x er sygt* (x is sick)
- arguments of verbs describing physical contact/stimulation, which cannot be used anaphorically, e.g., *slå x* (hit x), *spise x* (eat x), *drikke x* (drink x)

5 Evaluation of the Algorithm

We have applied the modified ES99-algorithm to three randomly selected SL dialogues for a total of 6,305 words.[7] It must be noted that only one annotator identified dialogue acts and classified the anaphors in the dialogues, while in [Eckert and Strube,1999b] the classification was accomplished by two annotators.

Because it is not always possible to distinguish between *den, det, de* used as personal or demonstrative pronouns without having access to stress information, we have classified them as personal pronouns unless they are topicalized, or occurred in syntactic constructions where demonstratives are normally used. The classification of pronouns and demonstratives in the three SL dialogues can be found in table 3:

	A5	AA10	AA11	$\sum$
Individual Pro	39	43	34	116
Discourse Deictics Pro	25	16	17	58
Vague Pro	4	6	0	10
Inferrable Evoked Pro	1	0	0	1
Individual Dem	1	5	0	6
Discourse Deictics Dem	27	20	19	66
Vague Dem	2	5	3	10

Table 3: Classification of Pronouns and Demonstratives

[7]The SL dialogues are conversations between doctors and patients collected by Jesper Hermann, associate professor at the Department of General and Applied Linguistics at the University of Copenhagen, from 1993 to 1995 under a project funded by the Danish Research Council for the Humanities and the Danish Social Science Research Council.

The results of the individual anaphora resolution algorithm can be found in table 4, while the results of the discourse deictics resolution algorithm is given in table 5.

	A5	AA10	AA11	$\sum$
No. resolved correctly	26	37	16	79
No. of Individual Pro	41	48	34	123
Precision	0.634	0.77	0.471	0.642

Table 4: Results of the Individual Anaphora Resolution Algorithm

	A5	AA10	AA11	$\sum$
No. resolved correctly	41	23	33	97
No. of Discourse Deictics	58	47	39	144
Precision	0.707	0.489	0.846	0.674

Table 5: Results of the Discourse Deictics Resolution Algorithm

The obtained results for discourse deictics are slightly better than those reported in [Eckert and Strube,1999b], while those for individual anaphors are slightly worse than those obtained by Eckert and Strube. Most of the instances of wrongly resolved anaphors are due to the fact that the algorithm cannot distinguish between discourse deictics and vague anaphors. Few errors are due to the fact, already noticed by Eckert and Strube, that the defined I* and A* predicates do not contain information about i.a. nominals referring to abstract objects. In the case of individual anaphors many errors occur because many anaphors, in the considered dialogues, referred back to entities which were not evoked in the immediately preceding **SU** and thus had been removed from the S-list. To avoid this problem we have extended the domain for the resolution of individual anaphors with the exception of neuter gender singular pronouns in order not to interfere with the resolution algorithm for discourse deictics. If the individual anaphora resolution algorithm fails to resolve an anaphor using the contents of the S-list it looks back in the saved contents of the S-list for the last N[8] **SU**'s.[9] This change allows the correct resolution of further 15 individual anaphors in our test giving a precision of 0.764 for the algorithm resolving anaphors with individual NP antecedents.

[8] Five **SU**'s in the current dialogues.

[9] We have used the cache model described in [Walker,1998] to implement this extension.

6 Concluding Remarks

Although the modified ES99-algorithm has only been tested on one type of dialogue and classified by one annotator and the type of dialogue in the Danish test is quite different from that used by Eckert and Strube, the results reported in the previous section (5) indicate that the algorithm performs as well for Danish as for English. We have also tested the algorithm for resolution of anaphors with individual antecedents allowing a wider resolution domain that a single **SU** improving its performance. Eckert and Strube [Eckert and Strube,1999b] note that the given rules for discriminating between anaphors referring to individual NPs and discourse deictics could be improved by using information about metaphoric use of verbs and recording nominals which denote abstract objects. We believe that this kind of information would improve the performance of the algorithm also for Danish data. Furthermore we think that more lexical semantic knowledge and more precise descriptions of the contexts in which discourse deictics occur can be included in the ES99-algorithm.

Further work consists in incorporating this information in the modified ES99-algorithm using a Danish semantic NLP lexicon and analyzing in more detail our data. We will then test the improved algorithm on more types of dialogue. Including more domain-specific information in the algorithm could also allow other types of inferrable pronoun than the evoked ones treated in [Eckert and Strube,1999b] to be covered by the algorithm.

Bibliography

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

[Brennan et al.,1987] S. F. Brennan, M. W. Friedman, and C. J. Pollard. *A Centering Approach to Pronouns*. In *Proceedings of the ACL-87*, pages 155–162, California, USA, 1987. Stanford University.

[Byron and Stent,1998] D. Byron and A. Stent. A Preliminary Model of Centering in Dialog. In *Proceedings of the ACL-98*, volume II, pages 1475–1477, Montreal, Quebec, Canada, 1998. Université de Montréal.

[Diderichsen,1946] P. Diderichsen. *Elementær Dansk Grammatik*. Gyldendal, Copenhagen, 1946.

[Eckert and Strube,1999a] M. Eckert and M. Strube. Resolving Discourse Deictic Anaphora in Dialogues. In *Proceedings of the EACL-99*, 37–44, 1999.

[Eckert and Strube,1999b] M. Eckert and M. Strube. Dialogue acts, synchronising units and anaphora resolution. In J. van Kuppevelt, N. van Leusen, R. van Rooy, and H. Zeevat, editors, *Amstelogue'99 Proceedings - Workshop on the Semantics and Pragmatics of Dialogue*, 1999.

[Grosz et al.,1995] B. Grosz, A. K. Joshi, and S. Weinstein. Centering:A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics*, 21(2):203–225, 1995.

[Grosz and Sidner,1986] B. J. Grosz and C. L. Sidner. Attention, Intentions, and the Structure of Discourse. *Computational Linguistics*, 12(3):175–284, 1986.

[Navarretta,forthcoming] C. Navarretta. Centering-based Anaphora Resolution in Danish. In P. Sojka, I. Kopecek, and K. Pala, editors, *to appear in: Proceedings of TSD 2000*, Brno, Czech Republic, 2000.

[Prince,1981] E. F. Prince. Toward a taxonomy of given-new information. In P. Cole, editor, *Radical Pragmatics*, 223–255. Academic Press, 1981.

[Prince,1992] E. F. Prince. The ZPG Letter: Subjects, definiteness, and Information-status. In W. Mann and S. A. Thompson, editors, *Discourse Description. Diverse Linguistic Analyses of a Fund-Raising Text.*, 295–325. John Benjamins, 1992.

[Strube,1998] M. Strube. Never look back: An alternative to centering. In *Proceedings of the ACL-98*, volume II, pages 1251–1257, Montreal, Quebec, Canada, 1998. Université de Montréal.

[Walker,1998] M. A. Walker, A. K. Joshi, and E. F. Prince. Centering in Naturally-Occurring Discourse:An Overview. In M. Walker, A. Joshi, and E. Prince, editors, *Centering Theory in Discourse*. Oxford University Press, Oxford, U.K., 1998.

[Webber,1991] B. L. Webber. Structure and Ostension in the Interpretation of Discourse Deixis. *Natural Language and Cognitive Processes*, 6(2):107–135, January 1991.