# Centering-based Anaphora Resolution in Danish Dialogues

Costanza Navarretta

Center for Language Technology
Njalsgade 80
2300 Copenhagen S - DK
tel. 45 + 35 32 90 65
fax. 45 + 35 32 90 89
costanza@cst.ku.dk

**Abstract.** In this paper we present the results of applying two different centering algorithms ([1] and [9]) to Danish discourses. Then we describe how we have adapted the algorithm for resolving anaphora referring to both individual NPs and discourse deictics presented in [3] so that it covers Danish discourse deictics. The modified algorithm has been manually tested on Danish dialogues and the obtained results have been evaluated.

## 1 Introduction

Resolving anaphors is an important task in many NLP applications. Most of the current anaphora resolution algorithms only deal with coreference between anaphors and individual NPs in written texts. In particular algorithms based on centering theory [4] have been studied and tested extensively in many languages, but not for Danish. Recently centering has also been applied to dialogues, i.a. [2] and [3]. Eckert and Strube [3], in particular, present an algorithm, henceforth the ES99-algorithm, for the resolution of anaphors in English dialogues. The ES99-algorithm which applies to both anaphors referring to individual NPs and discourse deictics is based on rules for discriminating among individual NPs and discourse deictics, mainly determined by the syntactic constructions in which the anaphors occur. After having tested whether centering works on Danish texts, we have adapted the rules of the ES99-algorithm to Danish and applied the modified algorithm to Danish dialogues.

This paper is organized as follows. In section 2 we shortly describe two centering algorithms, the BFP-algorithm [1] and the S98-algorithm [9] and we present the results of their application on Danish texts. In section 3 we outline the ES99-algorithm. In section 4 we shortly present the Danish discourse deictics found in the two dialogue corpora Bysoc and SL[1] and we describe the modified ES99-

[1] Both corpora have been collected by researchers at the Department of General and Applied Linguistics at the University of Copenhagen.

algorithm accounting for Danish data. Finally (section 5) we evaluate the results obtained by manually testing the ES99-algorithm on the SL dialogues.

## 2 Centering and Danish data

The centering theory fits into Grosz and Sidner's model of discourse structure [5], 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. The basic assumption behind centering is that some entities in an utterance are more central than others and this fact influences the use of referring expressions. The entities which link an utterance $U_n$ to the others in the same discourse segment are called the *centers* of that utterance. Each utterance is assigned a set of *forward-looking centers*, $C_f$, and, with the exception of the initial utterance of the segment, a *backword-looking center*, $C_b$. The $C_b$ of an utterance $U_n$ connects with one of the *forward-looking centers* of the preceding utterance $U_{n-1}$, while the *forward-looking centers* exclusively depend on the expressions in $U_n$. The *forward-looking centers* are partially ordered to reflect relative prominence. In the BFP-algorithm *forward-looking centers* are ranked according to the obliqueness of the grammatical relations of the subcategorized functions of the main verb (subject > object > object2 > complements > adjuncts). The first element in the $C_f$ list is called the preferred center, $C_p(U_n)$. In BFP four types of transition relations across pairs of utterances, *continue, retain, shifting-1, shifting*, are identified. The discriminating elements between the transitions are given in table 1 [1][p. 157]. The following two rules constrain

| | $C_b(U_n) = C_b(U_{n-1})$ OR no $C_b(U_{n-1})$ | $C_b(U_n) \neq C_b(U_{n-1})$ |
|---|---|---|
| $C_b(U_n) = C_p(U_n)$ | *continue* | *shifting-1* |
| $c_b(U_n) \neq c_p(U_n)$ | *retain* | *shifting* |

**Table 1.** Transition States

center realization in BFP:

**Rule 1:** If any element of $C_f(U_{n-1})$ is realized by a pronoun in $U_n$, then $C_b(U_n)$ must also be realized by a pronoun

**Rule 2:** The center transitions have the following ranking: *continue* > *retain* > *shifting-1* > *shifting*

The BFP-algorithm consists of three steps:

1. **construct** the proposed anchors for an utterance and possible *Cb*-*Cf* combinations
2. **filter** by i.a. contra-indices, sortal predicates, centering rules and constraints
3. **rank** by transition orderings.

The S98-algorithm [9] treats both intrasentential and intersentential anaphors. In S98 the functions of the *backward-looking center* and the *transitions* in the centering theory are replaced by the order of elements in a list of salient discourse entities, the **S-list**. The ranking criteria for the elements in the S-list are based on [7] and [8], where discourse entities are classified into *hearer-old* (OLD), *mediated* (MED) and *hearer-new* (NEW). The two tuples ($x$, $utt_x$, $pos_x$) and ($y$, $utt_y$, $pos_y$) in the S-list indicate that the entity $x$ is evoked in utterance $utt_x$ at position $pos_x$ and that $y$ is evoked in the utterance $utt_y$ at position $pos_y$ respectively. Given that $utt_x$ and $utt_y$ refer to $U_n$ or $U_{n-1}$, the following ranking constraints on the S-list entities are valid [9][p.1253]:[2]

1. if $x \in$ OLD and $y \in$ MED, then $x \prec y$
   if $x \in$ OLD and $y \in$ NEW, then $x \prec y$
   if $x \in$ MED and $y \in$ NEW, then $x \prec y$
2. if $x$,$y \in$ OLD or $x$,$y \in$ MED or $x$,$y \in$ NEW,
   then if $utt_x > utt_y$ then $x \prec y$
   if $utt_x = utt_y$ and $pos_x < pos_y$ then $x \prec y$

Strube's algorithm for anaphora resolution consists in testing a referring expression against the elements in the S-list from left to right until the test succeeds. The S-list is then updated so that new elements are inserted according to the S-list ranking criteria. When the analysis of an utterance is finished all the entities which were not realized in the utterance are removed from the S-list.

We have applied the two algorithms to randomly chosen chapters of a pc-manual (10,715 words) and newspaper articles (9,914 words). In the test discourse segments were paragraphs and utterances were clauses. Following [9] we have extended the BFP-algorithm to cover complex clauses following the strategy described in [6].[3] We manually marked expletives and discourse deictics.

The success rate for the BFP-algorithm was 72,5 % while the S98-algorithm had a success rate of 91,67 %. The difference between the results obtained with the two algorithms is mainly due to the fact that the BFP-algorithm does not account for intrasentential anaphors. The cases where both algorithms failed

[2] We mark ranking precedence with $\prec$.

[3] Kameyama treats tensed clauses as independent utterances, while untensed clauses are treated as part of the main clause. Tensed clauses comprise reported speech, which is not accessible to the superordinate level, non-report complements and relative clauses which are accessible to the superordinate level, but less salient. The remaining types are processed at the same level as the main clause.

in resolving pronominal anaphora comprise complex plural antecedents (coordinated and split ones), generic use of the neuter pronoun *det* (it), plural pronouns without antecedents, ambiguity of antecedents. Although the results obtained in tests applied to different kinds of discourse in different languages cannot be compared, the results obtained in our test are similar to those obtained in other languages (i.a. [9], [10]). This indicates that centering also works for Danish.

## 3 The ES99-algorithm

In the ES99-algorithm the types of anaphor identified are individual anaphors, discourse deictics, inferrable-evoked anaphors and vague anaphors. Predicates that are preferably associated with abstract objects are marked as **i-incompatible** (*I) while predicates that are preferably associated with individual objects are marked as **a-incompatible** (*A). As an example we quote the ***I** predicates given in [3][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, xis false, x is correct, x is right, x isn't right.*
- Arguments of verbs describing propositional attitude which only 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.*

Individual anaphors are resolved with the S98-algorithm, while abstract objects are treated in a so called A-list. The A-list is filled when discourse deictics occur and elements remain in the list only for one dialogue act (I, Initiation).[4] A context ranking procedure describes the order in which the parts of the linguistic contexts are accessed.

## 4 The Modified Algorithm

We have adapted the ES99-algorithm so that it covers Danish data identified in our dialogue corpora. The focus in our description is on discourse deictics. In Danish the most used discourse deictic is *det* which corresponds to both *it* and *that.*[5] *Dette* (this) is another discourse deictic, but it is mainly used in written language and did not occur at all in our dialogues.

[4] In [3] grounded acts are used as domain for the resolution algorithm instead of clauses. We have followed the same discourse model.

[5] The pronoun *det* usually co-refers with (from now on we simply write *refers to*) nominals in neuter gender. It is also used as expletive. *Det* is also the neuter definite article (the) and the demonstrative adjective (that).

When used as discourse deictic *det* can refer to an infinitive or a finite clause, as it is the case in the following examples:

(1) a. *At ryge er farligt og det er også dyrt*
(Smoking is dangerous and it is also expensive)
b. *Jeg skal måle dit blodtryk.*
(I have to measure your blood pressure.)
*Hvorfor det?*
(Why (that)?)

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

(2) a. *Jeg faldt, men det gjorde hun ikke*
(I fell, but she did not)

*Det* refers to a clause in constructions with verbs such as *tro* (think), *sige* (say), *vide* (know):

(3) *Han lyver.*
(He is lying)
*Det tror jeg ikke*
(I do not think so)

*Det* can also refer to more clauses, or to something that can be vaguely inferred from the previous discourse. On the basis of the deictics in the Danish dialogues we have defined the following types of ***I** predicate for Danish:

- constructions where a pronoun is equated. with an abstract object, e.g., *x gør det svært* (x is making it difficult)
- copula constructions with adjectives which can only be applied to abstract entities, such as *x er sandt* (x is true)
- arguments of verbs indicating propositional attitudes which take S'-complements, such as *tro* (believe), *antage* (assume)
- arguments of verbs such as *sige* (say) and *vide* (know)
- object of *gøre* (do)
- 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...* (x is because)

Our ***A** predicates are the following:

- constructions where a pronominal referent is equated with a concrete individual referent, such as *x er en legemsdel* (x is a body part)

– copula constructions with adjectives which can only be applied to concrete entities, such as *x er dyr* (x is expensive), *x er rød* (x is red)
– arguments of verbs describing physical contact/stimulation, e.g., *slå x* (hit x), *spise x* (eat x)

## 5 Evaluation and Conclusion

We have applied the modified ES99-algorithm to randomly chosen SL dialogues (9,728 words). It must be noted that we only used one annotator in our test, while in the test reported in [3] there were two annotators. The precision and recall of the modified algorithm on our dialogues were of 64.7% and 70,4 %, respectively. These results are similar to those reported in [3]. Most of the wrongly resolved anaphors are due to the fact that the algorithm cannot distinguish between discourse deictics and vague anaphors. Some errors are due to missing information on nominals referring to abstract objects, some depended on the chosen discourse model.

In conclusion, both centering algorithms and the ES99-algorithm seem to perform as well for Danish as for English. Future work consists in testing the algorithms on more types of dialogue, identifying more discriminating predicates and adding more lexical and domain knowledge to the modified ES99-algorithm.

## References


1. Brennan, S. F., Friedman, M. W. and Pollard, C. J.: A Centering Approach to Pronouns. In *Proceedings of the ACL-87.* Stanford University,California, USA (1987) 155–162.
2. Byron, D., Stent, A.: A Preliminary Model of Centering in Dialog. In *Proceedings of the ACL-98.* Université de Montréal, Quebec, Canada (1998) vol II, 1475–1477.
3. Eckert, M., Strube, M.: Resolving Discourse Deictic Anaphora in Dialogues. In *Proceedings of the EACL-99.* (1999) 37–44.
4. Grosz, B., Joshi, A. K. and Weinstein, S.: Centering:A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics.* **21(2)** (1995) 203–225
5. Grosz, B. J., Sidner, C. L.: Attention, Intentions, and the Structure of Discourse. *Computational Linguistics.* **12(3)** (1986) 175–284.
6. Kameyama, M.: Intrasentential centering: A case study. In Walker, M., Joshi, A., and Prince, E. editors. *Centering Theory in Discourse.* Oxford University Press, Oxford, U.K. (1998) 89–112.
7. Prince, E. F.: Toward a taxonomy of given-new information. In Cole, P., editor, *Radical Pragmatics.* Academic Press (1981) 223–255.
8. Prince, E. F.: The ZPG Letter: Subjects, definiteness, and Information-status. In Mann, W., Thompson, S. A., editors, *Discourse Description. Diverse Linguistic Analyses of a Fund-Raising Text.*. John Benjamins (1992) 295–325.
9. Strube, M.: Never look back: An alternative to centering. In *Proceedings of the ACL-98*, Université de Montréal, Quebec, Canada (1998) vol II, 1251–1257.
10. Tetreault, J.R.: Analysis of Syntax-Based Pronoun Resolution Methods. In *Proceedings of the 37th ACL.* Maryland, USA (1999) 602–605. 1999.