

A contrastive analysis of abstract anaphora in Danish, English and Italian

Costanza Navarretta

Centre for Language Technology
University of Copenhagen
Njalsgade 80
2300 Copenhagen - DK
costanza@cst.dk

Abstract

This paper describes the results of an investigation of pronominal abstract anaphora in parallel and comparable data. Abstract anaphora are anaphors that refer to abstract entities. The languages in the investigation are Danish, English, and Italian. The results of the study indicate that there are many similarities, but also differences in the way different types of pronoun are used to point back to the same type of linguistic expression (antecedent) in the three languages. In all three languages the presence of a demonstrative pronoun in contexts which allow both abstract and non-abstract reference indicates that the referred entity is abstract. While in English the most frequently used abstract pronouns are demonstrative, especially with clausal antecedents, as noticed by i.a. Webber (1988), in Danish and Italian personal pronouns are also frequently used in abstract reference. Thus pronominal types cannot be used to individuate different types of anaphora in Italian and Danish in the same way as this is done in English (Eckert and Strube 2001; Byron 2002). Furthermore the use of personal pronouns in abstract reference in Danish and Italian indicates that abstract entities are often brought “in focus” using the terminology proposed by Gundel et al. (2004, 2005).

1. Introduction

This paper describes the results of an investigation of the use of pronominal abstract anaphors¹ (which we call abstract anaphors in the following) in parallel and non-parallel corpora in three languages: Danish, English, and Italian. In the paper we distinguish *antecedents* from *referents*. We call *antecedents* the most recent linguistic expressions in discourse which the anaphors point to (the demonstratum); while we call *referents* the entities that the anaphors refer to. We have focussed on anaphors whose antecedents in discourse are verb phrases, predicates in copula constructions, clauses, discourse segments or abstract pronouns. The referents of these anaphors are abstract entities such as events, states, situations, facts and propositions. Anaphors referring to non-abstract entities will be called individual anaphors in the following.

The focus in our study has been to individuate similarities and dissimilarities in the way the three languages use abstract reference in similar contexts and with the same type of antecedent. In literature it is generally recognized that personal pronouns refer to the most relevant entities in discourse, while demonstrative pronouns refer to entities that are less prominent (see i.a. Prince, 1981, 1992; Ariel 1988, 1994; Givón, 1983, Gundel et al. 1993). In particular Gundel et al. suggest (1993) that personal pronouns refer to the entities which are “in focus” in the addressee’s attentional state, while demonstrative pronouns can and usually refer to entities that are simply “activated” in the addressee’s attentional state.

Webber (1988) notices that in English abstract entities introduced by a clause are accessible to immediate subsequent reference with demonstrative pronouns, while

they are often inaccessible to reference with the personal pronoun *it*. Gundel et al. (2001, 2003) suggest that abstract entities are activated in the addressee’s attentional state, while individual entities are in focus, because in ambiguous contexts demonstrative pronouns are used to refer to abstract entities, while personal pronouns refer to individual entities. They illustrate this with example 1, where the demonstrative pronoun *that* in 1a refers to the act of destroying the leaf collection (an abstract entity), while the personal pronoun *it* in 1b refers to the leaf collection (an individual entity).

- (1)
a. Max destroyed his leaf collection last night. **That** was dumb.
b. Max destroyed his leaf collection last night. **It** was dumb.
(Gundel et al., 2001:38)

According to Gundel et al., however, also personal pronouns can refer to abstract entities if the abstract entities are brought “in focus” in the preceding discourse by syntactic, semantic and pragmatic factors.

Gundel et al. (2003) suggest that the “degree of word immanence” as proposed by Asher (1993) is one of the factors one must look at when analyzing the attentional state of entities. More specifically the authors propose that events, activities and states which according to Asher (1993) have the highest degree of world immanence among abstract entities can more easily be brought “in focus” than entities with a lower degree of world immanence such as facts and propositions.

Gundel et al. (2004, 2005) analyze the occurrences of demonstrative pronouns in a corpus of spoken American English. In the corpus they annotate the number of demonstrative pronouns which have NP-antecedents and those that have non-nominal antecedents. Then, for each group, they mark the reference as direct if the anaphor and the antecedent have the same referent and indirect if the anaphor and the antecedent have different referents. The

¹ Abstract reference is also called i.a. *discourse deixis* (Levinson, 1983; Webber, 1991) and *situation reference* (Fraurud, 1992).

authors find that 72% of the occurrences of demonstrative pronouns in the analyzed corpus are used to refer to non-nominal constituents while only 6% of the occurrences of personal pronouns are used to refer to non-nominal constituents. The reference of pronouns having non-nominal antecedents is indirect in the data, confirming Webber's (1991) suggestion that abstract anaphors create their referent in the moment they are expressed by an act which she calls *ostention*.

In two studies of abstract reference in Danish Navarretta (2002, 2004) finds out that also in this language the use of demonstrative pronouns signals reference to a non-nominal constituent in ambiguous contexts that is when there are nominal and non-nominal antecedent candidates. However the pronoun, *det*, which in written language is ambiguous with respect to its pronominal type, is also the most frequently used abstract pronoun with clausal antecedents and in ambiguous contexts. Navarretta (2004) also analyzes abstract reference in few dialogues and notices that the unmarked personal pronoun *det* is also frequently used as abstract pronoun in these spoken data. This indicates that there are differences in the way Danish and English use various types of pronoun in abstract reference, especially when the antecedents are clauses. Navarretta also finds a number of language specific differences in the reference of various types of pronoun.

The question of which types of abstract pronoun usually point back to which type of antecedent is important to determine the connection between anaphor, antecedent and referent type and is also relevant for the automatic treatment of anaphora. In their anaphora resolution algorithms Eckert & Strube (2001) and Byron (2002) use the observation that in English demonstrative pronominal anaphors refer to abstract entities more often than to individual ones. This preference does not hold in Danish and thus cannot be used in the same general way to identify abstract anaphors (Navarretta, 2004a).

To discover possible connections between type of pronoun and type of antecedent, and to find similarities and differences in the way the three languages express abstract reference, we have analyzed the occurrences of abstract anaphors in fairy tales extracted from a parallel corpus of literary texts. The majority of the texts in the investigation belong to a multilingual corpus collected by researchers at the University of Copenhagen under the MULINCO-project (Maegaard et al., 2006).

The source language of the analyzed parallel texts is Danish, while English and Italian are the target languages. In translated texts the target language can be influenced by the source language, see i.a. Ehrensberger-Dow & Jekat (2005). Thus to compare linguistic phenomena in different languages it is also necessary to look at these phenomena in monolingual data. Abstract reference has been investigated in English i.a. by Gundel et al. (2004, 2005) and we compare our data with the results they report when discussing abstract reference in English. In discussing the Danish data we also take into consideration the results of the investigation reported in Navarretta (2002, 2004). Because we had no monolingual study of abstract reference in Italian, we have analyzed the occurrences of abstract anaphors in original Italian fairy tales from the above mentioned MULINCO corpus.

The rest of the paper is organized as follows: we first present the pronouns we have looked at in the three languages and describe the data used in the investigation; then we report the results of our investigation in the parallel data and in the Italian monolingual data. In the last section we make some general conclusions and discuss some aspects which still need be elucidated and which will be investigated in a recently started project.

2. The data

The English abstract pronominal anaphors are the personal pronoun *it*, and the demonstrative pronouns *this* and *that*. The Danish abstract anaphors include the pronouns *det* (it/this/that) and, in written language, *dette* (this)². The Italian anaphors comprise the personal pronouns *lo*, *ne* and *ci*, both as clitic particles and as independent forms, and the demonstrative pronouns *questo* (this) *quello* (that) and *ciò* (this/that). Being Italian a subject PRO-drop language abstract zero-anaphors occur in the data. An example of this is in 2.

- (2)
- si fece piccolo piccolo, si fece grosso, ma l' ombra non c' era, nessuno venne; disse uhm, uhm! Ma [Ø] non servì a niente.*
- He made himself little, he made himself long, but there was no shadow, and no shadow came. He said, "Here, here?" but **that** did no good.³
- [Andersen: *Skyggen* (The shadow)]

The fact that zero-anaphors are used in abstract reference is interesting because zero-anaphora are recognized to refer to entities which have even a higher accessibility status in the addressee's attentional state than personal pronouns (i.a. Givón, 1983; Ariel, 1988).

The data we have used in our investigation are the following:

- Seven Danish fairy tales by H.C. Andersen: *Den flyvende Kuffert* (The flying trunk), *Keiserens nye Klæder* (The emperor's new clothes), *Lykkens Kalosker* (The galoshes of fortune), *Den grimme Ælling* (Den ugly duckling), *Sneedronningen* (The snow queen), *Skyggen* (The shadow) and *Sneemanden* (The snow man). These fairy tales consist of 40,596 running words (tokens).
- Duncan's English translations of the same fairy tales, consisting of 43,190 running words.
- Italian translations (by two different translators) of the same fairy tales, consisting of 40,943 running words.
- 34 Italian fairy tales by Luigi Capuana belonging to the two collections *C'era una volta... : fiabe* (Once upon a time...fairy tales) and *Raccontafiabe - seguito a C'era una volta* (Fairy-teller – the following of Once upon a time), consisting of 90,048 running words.

² In spoken language abstract reference is expressed by the stress unmarked *det* (it), the stress marked *det* (it/this/that), *det her* (this) and *det der* (that). The latter pronoun is mainly used as cataphor (Navarretta, 2002).

³ The English translations in the examples from Andersen's fairy tales are by Duncan.

All texts are POS-tagged and lemmatized automatically⁴. The relevant pronouns have been extracted from the data automatically, but they have been checked and annotated manually.

3. A contrastive analysis of the abstract anaphors

In this section we present some of the results of our investigation. In tables 1-3 are the occurrences of the relevant pronominal forms in Andersen's fairy tales (table 1) and in their English and Italian translations (table 2 and 3 respectively).

In the first column of each table a pronoun is given; the second column contains the number of occurrences of the pronoun independently from its function which can be deictic, pleonastic, anaphoric (individual or abstract) and cataphoric. In the third column the occurrences of the pronoun when it is used as abstract anaphor are reported while the fourth column gives the percentage of the abstract occurrences.

| <i>pronoun</i> | <i>all</i> | <i>abstract</i> | <i>%</i> |
|------------------------------|------------|-----------------|----------|
| <i>det</i> (it/this/that) | 787 | 244 | 31% |
| <i>dette</i> (this) | 17 | 4 | 24% |
| total | 804 | 248 | 30% |

Table 1: Occurrences of the pronouns *det* and *dette* in the Danish fairy tales

| <i>pronoun</i> | <i>all</i> | <i>abstract</i> | <i>%</i> |
|----------------|------------|-----------------|----------|
| <i>it</i> | 633 | 80 | 13% |
| <i>this</i> | 59 | 27 | 46% |
| <i>that</i> | 189 | 116 | 61% |
| total | 881 | 223 | 25% |

Table 2: Occurrences of *it*, *this* and *that* in the English translation of the fairy tales

In table 3 we have excluded the occurrences of so called double pronouns. An example of double pronouns is in 3. In the example the personal pronoun *lo* (it) doubles the demonstrative pronoun *questo* (this). Only the first occurring pronouns are accounted for in table 3. Thus in our example only the demonstrative pronoun is accounted for in table 3.

(3)

"Ah, se solo fossi libero di nuovo, allora starei certo meglio!" Ecco **questo** lo avrebbe dovuto dire un po' prima;

"If I were free, most likely it would pass over." That's what he ought to have said at first.

[Andersen: *Lykkens Kalosker* (The galoshes of fortune)]

| <i>pronoun</i> | <i>all</i> | <i>abstract</i> | <i>%</i> |
|------------------|-----------------|-----------------|----------|
| <i>questo</i> | 48 | 20 | 42% |
| <i>quello</i> | 22 | 2 | 9% |
| <i>ciò</i> | 3 | 2 | 67% |
| <i>ci</i> | 30 | 4 | 13% |
| clitic <i>ci</i> | 15 | 0 | 0 |
| <i>ne</i> | 34 | 15 | 44% |
| clitic <i>ne</i> | 7 | 1 | 14% |
| <i>lo</i> | 104 | 28 | 27% |
| clitic <i>lo</i> | 44 | 19 | 43% |
| 0-anaphor | 59 ⁵ | 25 | 42% |
| total | 366 | 116 | 32% |

Table 3: Occurrences of the relevant pronouns in the Italian translation of the fairy tales

In table 4 the percentage of abstract occurrences of each type of pronoun (demonstrative or personal) in the three languages is given. The Danish pronoun *det* is classified as a personal pronoun in the table, because, as noticed in section 2, it is not possible to distinguish stressed occurrences of it in written language.

Zero-anaphors have been counted together with personal pronouns in the table.

| <i>language</i> | <i>pronominal type</i> | <i>abstract %</i> |
|-----------------|------------------------|-------------------|
| Danish | personal/demonstrative | 31% |
| Danish | proximal demonstrative | 24% |
| English | personal | 13% |
| English | demonstrative | 58% |
| Italian | personal | 31% |
| Italian | demonstrative | 34% |

Table 4: Abstract use of different pronominal types in the parallel fairy tales

In the data we have annotated the following information: the type of pronoun, the syntactic type of the antecedent and the distance between the anaphor and its antecedent.

The most frequently used abstract anaphor in the Danish fairy tales is the pronoun *det*. This was also the case in the study of newspaper articles and dialogues reported by Navarretta (2002, 2004). There are only four occurrences of the proximal demonstrative pronoun *dette* in the analyzed fairy tales. This is not surprising being the pronoun *dette* mostly used in formal written language. Two of the occurrences of *dette* in the fairy tales refer to the preceding clause, one refers to the last occurring (less distant) clause among coordinated clauses and one refers to the subclause in the preceding complex clause⁶. The latter two occurrences of *dette* are in 4 and 5 respectively.

4)

Justitsraaden antog, at hun ikke kunde den danske Tunge og fremførte derfor sit Ønske i Tydsk; dette

⁴ Only the annotation of part of the data has been manually verified by the participants of the Mulinco project.

⁵ Only zero-anaphors which occur in contexts that may allow an abstract reading have been counted.

⁶ By complex clause we both mean a main clause and its subclauses and more coordinated main clauses.

tilligemed hans Dragt bestyrkede Konen i, at han var en Udlænding;

The councillor now supposed that she did not understand Danish, so he repeated his wish in the German language. *This*, and his costume, convinced the woman that he was a foreigner.

[Andersen, Lykkens Kalosker (The Galoshes of Fortune)]

In 4 the antecedent of *dette* (this) is the preceding clause *og fremførte derfor sit Ønske i Tydsk* (so he repeated his wish in German), and not the whole complex clause *Justitsraaden antog, at hun ikke kunde den danske Tunge og fremførte derfor sit Ønske i Tydsk* (The councillor now supposed that she did not understand Danish, so he repeated his wish in the German language).

5)

Vi mærke allerede at han er blevet Digter; at antyde dette, vil i de fleste Tilfælde være, hvad Tydskeren kalder "abgeschmackt", det er en taabelig Forestilling, at tænke sig en Digter anderledes end andre Mennesker, der kan mellem disse være langt mere poetiske Naturer, end mangel stor erkjendt Digter er det;

We already notice that he has become a poet. To point **this** out would, in most cases, be what the Germans call "mawkish". It is a foolish fancy to imagine a poet different from other people, for among the latter there may be natures more poetical than those of many an acknowledged poet.

[Andersen, Lykkens Kalosker (The Galoshes of Fortune)]

In 5 the antecedent is the immediately preceding subclause, *at han er blevet Digter* (that he has become a poet) and not the whole complex clause, *Vi mærke allerede at han er blevet Digter* (We already notice that he has become a poet).

The use of the abstract *dette* to refer to parts of complex clauses is also discussed in Navarretta (2004). The author suggests that this "abstract" use of the proximal pronoun *dette* is parallel to one of its uses as individual anaphor, which is to point back to the latter nominal in coordinated nominal phrases (Navarretta, 2004). Also in the English translations of the Danish 4 and 5, and in the Italian translation of 4 a proximal demonstrative pronoun is used. It should be investigated whether proximal demonstrative pronouns in English and Italian, as in Danish, consequently have the function of pointing back to the constituent of complex constituents which is nearest to the anaphor.

In the English version of the fairy tales the most frequently used pronouns in abstract reference are the demonstrative pronouns and especially the distant *that*. This is in line with the results reported i.a. by Gundel et al. (2004, 2005). As noticed by Webber (1988) and Gundel et al. (2001), reference to clauses is often done using demonstrative pronouns in English. Our data confirms this observation.

The antecedents of the abstract pronoun *it* in our texts are abstract pronouns (37 cases), verb phrases, and, in few cases, clauses.

In the Italian translations personal and demonstrative pronouns are equally frequent in abstract reference. The most used demonstrative pronoun in Italian is the proximal *questo* (this). The distal demonstrative pronoun *quello* (that) is only used twice as an abstract anaphor and in one case the reference in Italian is not felicitous⁷ (example 6).

(6)

Han sagde ikke Noget, han klagede ikke, og det er det rigtige Tegn.

He said nothing and made no complaint - and **that's** an infallible sign.

Non diceva niente, non si lamentava, e quello è il segno giusto.

[Andersen, *Sneemanden* (The Snowman)].

There are more reasons, in our opinion, for the infelicity of the reference in 6. First of all the use of a singular pronoun to point back to more clauses, i.e. to refer to a set of situations, is not common in Italian. In the English translation the comma between the two antecedent clauses has been substituted with the coordinating conjunction "and" which probably makes it easier to interpret the two clauses as a whole. The use of a distal demonstrative pronoun to point back to clauses which are in the same complex clause as the anaphor makes also the reference hard to resolve in Italian. Finally the translation of the adjective *rigtig* with *giusto* (right) does not help resolution. More acceptable translations of the Danish text in 6 could be "Non diceva niente, non si lamentava, e questi sono segni infallibili." (He said nothing, made no complaint and these are infallible signs) or "Non diceva niente e non si lamentava. (Quelli/Questi/Ø) sono segni infallibili." (He said nothing and made no complaint. Those/These are infallible signs.).

Comparing the uses of abstract pronouns in the three languages with respect to their antecedent type we discovered the following regularities.

The majority of abstract pronouns in the texts have immediately preceding antecedents.

When the Danish *det* is used in contexts where the referent must be abstract and the antecedent is a verb phrase or a predicate of a copula construction (an adjectival or a nominal phrase), the English translator uses elliptic constructions, adverbial anaphors (*so, too*) or translates the text without pronominal anaphors, see also Navarretta (2002). In the same contexts the Italian translators use personal pronouns (both independent and clitic forms), nouns or constructs which do not contain anaphoric expressions.

When the antecedents of the Danish *det* are clauses, the English translator uses a demonstrative pronoun (mainly *that*), or a construction which does not contain pronominal anaphors. In Italian anaphors with clausal antecedents are translated with both personal and demonstrative pronouns, and in four cases, with a noun and a demonstrative determiner, such as *queste cose* (these things). If the anaphoric context is not ambiguous with respect to whether the anaphor is abstract or individual, the antecedent is the preceding clause and the anaphor is

⁷ More Italian informants have confirmed this.

the subject in the current clause, zero-anaphors can “occur” as in 2.

In cases of textual deixis, that is when the pronoun points literally back to a piece of text, the pronoun *det* is used in Danish, while the English and the Italian translators always use a demonstrative pronoun. An example of textual deixis in the three languages is in 7.

(7)

...jeg skal fortælle saadant noget, Enhver har oplevet; det kan man saa rart sætte sig ind i, og det er saa fornøieligt: »Ved Østersøen ved de danske Bøge!« »Det er en deilig Begyndelse!« sagde alle Talerknerne, »det bliver bestemt en Historie, jeg kan lide!«

...I will tell a story from real life, something that every one has experienced, so that we can easily imagine the situation, and take pleasure in it. “On the Baltic, by the Danish shore –” “That 's a pretty beginning!” cried all the Plates. “That will be a story we shall like.”

...vi racconterò una storia che ciascuno di noi ha vissuto: è così utile approfondire le proprie esperienze! Ed è anche molto divertente! Dunque: «sulle sponde del mar Baltico, all' ombra dei faggi di Danimarca ...» «Che bell' inizio», dissero i piattini in coro, «questa storia ci piacerà senz' altro!»

[Andersen, *Den flyvende kuffert* (The flying trunk)]

In the Danish fairy tales *det* is also used in contexts which are ambiguous with respect to the abstract state of the antecedent. In English both proximal and distal demonstrative pronouns are used in these cases, while in Italian only the proximal demonstrative pronoun *questo* (this) is used. These uses again confirm Gundel et al.’s (2001) observation that in ambiguous contexts demonstrative pronouns indicate reference to abstract entities.

In Danish and Italian personal pronouns are used in pronominal chains, that is when the pronoun has another abstract pronoun as antecedent. In English both personal and demonstrative pronouns are used in this case⁸. In the Danish fairy tales there are only two cases of abstract pronouns with a pronominal antecedent, while in English these cases are 39, and in Italian 9.

In table 5 the results of the classification of pronouns in the monolingual Italian fairy tales by Capuana are shown. As in tables 1-3, the described pronoun is in the first column, the second column shows the occurrences of the pronoun, the third column indicates the occurrences of the pronoun when used in abstract reference and the fourth column gives the percentages of abstract use for the pronoun. In table 6 the percentage of abstract occurrences for different pronominal types in Capuana’s fairy tales is given.

The results of our investigation in the monolingual Italian data show that abstract anaphors are clearly less frequent in these texts than in the translated ones. This difference is probably due to the influence on the translated texts of the source language.

| <i>pronoun</i> | <i>all</i> | <i>abstract</i> | <i>%</i> |
|------------------|------------|-----------------|----------|
| <i>questo</i> | 79 | 19 | 24% |
| <i>quello</i> | 55 | 1 | 2% |
| <i>ciò</i> | 0 | 0 | - |
| <i>ci</i> | 141 | 15 | 11% |
| clitic <i>ci</i> | 57 | 2 | 4% |
| <i>ne</i> | 168 | 48 | 29% |
| clitic <i>ne</i> | 51 | 8 | 47% |
| <i>lo</i> | 281 | 31 | 11% |
| clitic <i>lo</i> | 132 | 24 | 18% |
| 0-anaphor | 63 | 19 | 30% |
| all | 1027 | 167 | 16% |

Table 5: Occurrences of the relevant pronouns in Capuana’s fairy tales

| Pronominal type | abstract % |
|-----------------|------------|
| personal | 17% |
| demonstrative | 15% |

Table 6: Percentage of abstract uses for each pronominal type in Capuana’s fairy tales

The different types of abstract anaphor in the Italian monolingual texts are used in manly the same way as in the translated data. Both in the Italian translations and in the monolingual texts there is no clear preference for a type of pronoun in abstract reference and the mostly frequently used demonstrative pronoun in abstract reference is the proximal *questo*. In conclusion also in Italian, as in Danish, demonstrative pronouns are not the preferred pronominal type in abstract reference and in this aspect they differ from the English demonstrative pronouns.

4. Conclusion and future work

The results of our investigation indicate that there are both differences and similarities in the way Danish, English and Italian use different types of pronoun in similar contexts and with the same kind of antecedent. In the Danish and English texts abstract pronominal reference is more frequent than in the monolingual Italian texts, while abstract reference in the translated Italian texts is as frequent as in the original Danish fairy tales. More Italian data should be analyzed to confirm whether the discrepancy in the frequency of abstract anaphors in original and translated texts is only due to the influence of the source language on the translated fairy tales.

In two third of their occurrences demonstrative pronouns in the English translations of the Danish fairy tales are abstract anaphors. These results are similar to those reported for a monolingual English spoken corpus by Gundel et al. (2004). The preference for demonstrative pronouns in abstract reference has not been observed in the Italian and Danish data, where abstract reference with personal pronouns is also common. Furthermore personal pronouns have also often clausal antecedents in Italian and Danish differing from their English counterparts.

In Danish the most frequently used abstract pronoun is *det*, which is ambiguous with respect to its type in written language. Thus the distinction in pronominal types

⁸ Pronouns in pronominal chains are classified as pronouns with NP-antecedent in Gundel et al (2005).

seems not to be indispensable for the identification of abstract reference in Danish. In Italian there is no significant difference in the occurrences of personal and demonstrative abstract pronouns. Thus the pronominal type cannot be used to individuate different types of anaphor in this two languages in the same way as in English (Eckert and Strube 2001; Byron 2002).

Although a first investigation of Danish abstract anaphors in spoken data indicates that the most frequently used abstract anaphor is the unmarked personal pronoun *det* more phonetically transcribed spoken data should be analyzed to investigate the role of the stressed *det* in abstract reference. The fact that in Italian and Danish personal pronouns (and in the case of Italian also zero-anaphors) are often used in abstract reference seems to indicate that abstract objects are more often brought “in focus” in these two languages than in English. The reasons of this must be investigated further.

In a just started Danish project founded by the Danish Research Councils, differences in abstract reference in the three languages will be investigated further in more types of text and, in the case of Danish and Italian, in more spoken data. The project’s aims are also to individuate referential differences between the different types of pronoun in especially Italian and Danish in order to facilitate the automatic identification and resolution of abstract anaphors in these two languages. In the project the semantic types of the referent will also be included in the analysis, although it should be considered that it is difficult to annotate these semantic types in a reliable way as noticed by Gundel et al. (2005)

5. Acknowledgements

Part of the research in the article is done under the MULINCO project financed by The Danish Research Councils. The authors want to thank all project participants.

6. References

- Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics*, 24(1):65–87.
- Ariel, M. (1994). Interpreting anaphoric expressions: a cognitive versus a pragmatic approach. *Journal of Linguistics*, 30(1):3–40.
- Asher, N. (1993). *Reference to Abstract Objects in Discourse*, volume 50 of *Studies in Linguistics and Philosophy*. Kluwer Academic Publishers, Dordrecht, the Netherlands.
- Byron, D.K. (2002). Resolving pronominal reference to abstract entities. In *Proceedings of ACL '02*, pages 80–87, 2002.
- Eckert, M., Strube, M. (2001). Dialogue Acts, Synchronising Units and Anaphora Resolution. In: *Journal of Semantics* 17(1):51–89.
- Ehrensberger-Dow, M., Jekat, S. J. (2005). Languages in contact: The influence of language activation and competing language patterns on translation performance. *Linguistik online*, 23, 2/05. http://www.linguistik-online.com/23_05/ehrJekat.html.
- Fraurud, K. (1992). *Processing Noun Phrases in Natural Discourse*. Department of Linguistics - Stockholm University.
- Givón, T. (ed.) (1983). *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. John Benjamin, Amsterdam.
- Gundel, J.K., Hedberg, N., Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69(2):274–307.
- Gundel, J.K., Hegarty, M., Borthen, K. (2001). Information structure and pronominal reference to clausal introduced entities. In *Proceeding of the ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*. Helsinki, 37–51.
- Gundel, J. K. , Hegarty, M., Borthen, K. (2003). Cognitive Status, Information Structure, and Pronominal Reference to Clausally Introduced Entities. *Journal of Logic, Language and Information*, 12:281–299, Kluwer Academic Publishers.
- Gundel J. K., Hedberg, N., and Zacharski, R. (2004). Demonstrative Pronouns in Natural Discourse. In: Branco, A., McEnery, T., Mitkov, R. (eds.): *Proceedings of DAARC 2004 - 2004*, Ediçoes Colibri, 81–86.
- Gundel, J. K., Hedberg, N., Zacharski, R. (2005). Pronouns without NP antecedents. In: Branco, A., McEnery, T. & Mitkov, R. (eds.): *Anaphora Processing – Linguistic, cognitive and computational modelling*. Current Issues in Linguistic Theory. John Benjamins Publishing Company. 351–364.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge, England: Cambridge University.
- Maegaard, B., Offersgaard, L., Henriksen, L., Jansen, H., Lepetit, X., Navarretta, C., Povlsen, C. (2006). The MULINCO corpus and corpus platform. In *Proceedings of the 5th International Conference on Language Resources and Evaluation*, Genova. 2148–2153.
- Navarretta, C. (2002). *The Use and Resolution of Intersentential Pronominal Anaphora in Danish Discourse*. Ph.D. thesis - Copenhagen University.
- Navarretta, C. (2004). The Main Reference Mechanisms of Danish Demonstrative Pronominal Anaphors. In: Branco, A., McEnery, T., Mitkov, R. (eds.): *Proceedings of DAARC 2004*, 2004, Ediçoes Colibri, 115–120.
- Navarretta, C. (2004a). Resolving Individual and Abstract Anaphora in Texts and Dialogues. In *Proceedings of the 20th International Conference of Computational Linguistics, COLING-2004*, Geneva, Switzerland, 233–239.
- Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole, editor, *Radical Pragmatics*, Academic Press, pages 223–255.
- Prince, E. F. (1992). The ZPG Letter: Subjects, Definiteness, and Information-status. In *Discourse Description. Diverse Linguistic Analyses of a Fund-Raising Text*, John Benjamins, pages 295–325.
- Webber, B.L. (1988). *Discourse deixis and discourse processing*. Technical report, University of Pennsylvania.
- Webber, B.L. (1991). Structure and Ostension in the Interpretation of Discourse Deixis. *Natural Language and Cognitive Processes*, 6(2):107–135, January.