Pronominal types and abstract reference in the Danish and Italian DAD corpora

Costanza Navarretta
Centre for Language Technology, University of Copenhagen, Karen Blixens vej 1, 2300 Copenhagen S, Denmark, costanza@hum.ku.dk

Abstract

In the paper we present Danish and Italian corpora of texts and dialogues which have been annotated with information relevant for the study and the resolution of abstract anaphora. Then we discuss differences and similarities between the use of abstract anaphora in these corpora and in English, i.a. (Webber, 1988; Gundel et al., 2003). Abstract anaphora, in this paper, refers to third person singular pronouns whose linguistic antecedents are copula predicates, verbal phrases, clauses and discourse segments and whose referents are abstract objects such as predicates, events, facts, and propositions.

The purpose of the described work is to study abstract reference in Danish and Italian systematically because previous research, i.a. (Fraurud, 1992; Navarretta, 2004; Navarretta, 2007), indicate that there are language specific characteristics of the phenomenon which do not fit into accounts of abstract reference based on English data. These characteristics must be explained and formalised in order to pave the way for the automatic treatment of abstract anaphora in these languages. In the paper we suggest that some differences in the use of abstract anaphora in Danish, English and Italian can be explained looking at the three languages’ pronominal system and syntactic structure.

1 Introduction

In this paper we describe the DAD corpora of Danish and Italian texts and dialogues to study and automatically treat abstract anaphora. Abstract anaphora refers in the paper to third person singular pronouns whose linguistic antecedents are predicates in copula constructions, verbal phrases, clauses and discourse segments. The referents of abstract anaphora are abstract objects such as properties, events, situations and propositions. English abstract anaphora are the personal pronoun it and the demonstrative pronouns this and that. An example of an abstract anaphor is in 1 where the antecedent of the pronoun that is the preceding clause the cake we produce is too small and not the nominal phrase the cake.

1. The cake we produce is too small and that is what we have to do something about.
   (English Financial Time - 1993)

Most theories and empirical studies on abstract anaphora are based on English data, i.a. (Webber, 1988; Asher, 1993; Hegarty, 2003; Gundel et al., 2003; Hedberg et al., 2007). Exceptions are studies by Fraurud (1992), Borthen et al. (1997), Navarretta (2002) and Navarretta (2007) indicating that there are many factors which can bring abstract entities in focus and that there are differences in the way various pronominal types are used to refer to abstract entities in different languages. Language specific uses of personal and demonstrative anaphora in general are also discussed in (Kaiser, 2000; Kaiser and Trueswell, 2004; Navarretta, 2002; Navarretta, 2004).

Because different pronominal types refer to objects having different degree of salience in the hearer’s cognitive status, see i.a. (Ariel, 1988; Givón, 1976; Gundel et al., 1993) the study of the relation between pronominal types and antecedent types is important to identify the antecedents and the referents of anaphora. Furthermore it is interesting from both a theoretical and a practical point of view to individuate the factors that influence the use of anaphors in various contexts in different languages. To discover some of these factors in Danish and Italian data and to provide annotated corpora for the automatic treatment of abstract anaphora in the two languages are the main aims of the DAD project.

The paper is organised as follows. We first present the background for our research and dis-
cuss related work (section 2) then we describe the information annotated in the DAD project (section 3) and the corpora which have been annotated until now (section 4). In section 5 and 6 we present and discuss some of the information extracted from the annotated corpora and finally in section 7 we make some conclusions and discuss work still to be done.

2 Background

All studies of referring nominal expressions indicate 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; Ariel, 1988; Givón, 1976; Gundel et al., 1993). Webber (1988) notices 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.

Gundel et al., i.a. (2003; 2007), confirm Webber’s observation in their studies of third person singular pronominal anaphors in English. Following Hegarty (2003) they explain this behaviour in terms of the cognitive statuses of nominal referring expressions as proposed in the Givenness Hierarchy by Gundel et al. (1993). According to this hierarchy demonstrative pronouns signal that the entities they refer to are activated in the cognitive status of the addressee while personal pronouns signal that the referred entities are both activated and in focus in the cognitive status of the addressee. Hegarty (2003) proposes that entities introduced in discourse by clauses are only activated in the cognitive status of the addressee while nominal phrases which occur in central syntactic positions in the current or in the preceding utterance are in focus, or are the most central according to the Centering theory, see i.a. (Brennan et al., 1987; Grosz et al., 1995). Because clauses often introduce entities such as facts, situations and propositions, these entities are seldom referred to by personal pronouns, according to Hegarty. Entities introduced in discourse by verbal phrases and which refer to states and events are often in focus in the addressee’s cognitive status. They have the same status as entities introduced in discourse by nominal phrases in prominent syntactic position and are often referred to by the personal pronoun it.

The fact that demonstrative pronouns often indicate that the antecedent is a clause has been used in algorithms for resolving abstract anaphora in English (Eckert and Strube, 2001; Byron, 2002). The behaviour of English demonstrative pronouns however seems not to be the same as that of demonstrative pronouns in other languages.

Fraurud (1992) studies abstract pronominal reference1 in Swedish texts and notices that the most frequently used abstract anaphor is det (it/this/that) whose pronominal status is ambiguous in texts. Furthermore she does not find any difference in the use of det and of the demonstrative pronoun detta (this) in abstract reference.

Navarretta (2002) analyses pronominal anaphora in Danish and reports that abstract anaphora are used in more contexts in Danish than in English and that the most frequent abstract pronoun in texts is det (it/this/that). The Danish det, as the corresponding Swedish pronoun, is ambiguous regarding its pronominal type. Thus the type of this pronoun cannot be relevant to determine the cognitive status of the referred entities in written Danish. Spoken Danish distinguishes between personal and demonstrative uses of the pronoun det via stress: the personal det (it) is unstressed while the demonstrative det (this/that) is stressed. However Navarretta did not include prosodic information in her study.

Navarretta (2004) reports that there are differences in the way the Danish demonstrative pronoun dette (this) and personal pronoun det (it/this/that) are used as abstract anaphors. Dette can indicate as the English demonstrative pronouns that the antecedent is abstract in ambiguous contexts (Gundel et al., 2004) where the individual reading is the most expected, but it can also signal that the antecedent is not the preceding complex clause, but only the immediately preceding subclause. Finally Navarretta (2007) describes differences in the use of abstract anaphora in a parallel corpus of fairy tales (English, Danish, Italian). English demonstrative pronouns in this corpus refer to entities introduced in discourse by clauses consistently with the analyses of these pronouns by i.a. Webber (1988) and Gundel et al. (2003; 2004). However personal pronouns (both clitic and non-clitic) and zero anaphora2 are used in Italian and the ambiguous pronoun det is used in Danish in similar contexts. Because zero

---

1She calls it situation reference.
2Italian is a subject pro-drop language.
anaphora, clitic pronouns and personal pronouns signal the most accessible entities in discourse, i.a. (Givón, 1979; Ariel, 1988) the parallel data seem to imply that the cognitive status of entities introduced by non nominal phrases is different in the three languages.

Borthen et al. (1997) describe differences in the way Norwegian abstract pronouns are used respect to their English correspondents and explain these differences in terms of extra-linguistic factors influencing the salience of the referred abstract entities.

In the following we describe the DAD corpora which have been annotated to investigate more systematically the characteristics of abstract anaphora in Danish and in Italian texts and dialogues.

3 The annotated information

The annotation of abstract anaphora in the DAD project is done in XML using an extension to the GNOME/MATE annotation scheme presented in (Poesio, 2004). A description of the extended DAD scheme for abstract reference is in (Navarretta and Olsen, 2008).

In DAD we annotate all the occurrences of singular third-person personal and demonstrative pronouns which potentially can be abstract anaphors in order to facilitate their automatic recognition. We then annotate the type and the function of each of these pronouns.

3.1 Pronominal types

Pronominal types are language dependent information. The relevant types for Danish are the following: the ambiguous det (it/this/that) and the demonstrative dette (this) which occur in texts and the unstressed det (it), stressed det(this/that), det her (this) and det der (that) occurring in spoken language.

The relevant pronominal types in Italian are the following: the personal pronouns esso (it subject), lo, ne and ci (it non-subject), both as clitic particles and as independent forms, and the demonstrative pronouns questo (this) quello (that) and ciò (this/that). The masculine pronouns egli (he), lui (he/him), lo (him) and questi (this) are also annotated. Being Italian is a subject PRO-drop language third-person singular verbal forms in which is implicit the subject pronoun are also annotated. We call these implicit pronouns zero anaphora henceforth. An example of “abstract zero anaphor” is in 2.

2.

Occorre tempo per approntare queste misure? Non è vero, ha detto Abete.

(lit. Does it take time to take these measures? Ø is not true, Abete said)

(Does it take time to take these measures? This is not true, Abete said)

(II Sole 24 Ore - 1992)

3.2 Pronominal functions

The following pronominal functions are recognised:

- pleonastic as in det regner (it rains), jeg har det fint (lit. I have it fine) (I am fine), det er forbudt at ryge (lit. it is prohibited to smoke) (smoking is not allowed);
- cataphoric, i.e. the pronoun precedes the linguistic expression necessary to its interpretation in discourse as in 3.

3.

Det at han kom for sent til mødet, skabte alvorlige problemer for hans kollegaer.

(lit. It that he came too late to the meeting gave serious problems to his colleagues);
- deictic. The pronoun refers to something in the physical word as in the utterance Hvad er det her? (What is this?) accompanied by a pointing gesture to an object;
- individual anaphoric: anaphors with nominal phrase antecedents;
- individual vague anaphoric: anaphors whose antecedents are implicit in discourse;
- abstract anaphoric;
- textual deictic (Lyons, 1977), as in 4.

4.

“I am happy” - Det råbte han, mens han gik.

(lit. “I am happy”- It/This/That he shouted while he walked)

(“I am happy”- He shouted this while he walked);

5. The zero anaphor is marked with a “Ø” in the English translation.

Johansson, C. (Ed.)
3.3 Other information

If a pronoun is an individual anaphor its antecedent and the relation between anaphor and antecedent are marked. We have only distinguished between two relation types: "identity" and "non-identity". If the anaphors are abstract, their antecedents and the syntactic type of the antecedents are marked. The anaphoric distance (distance between abstract anaphor and antecedent) in terms of clauses, the semantic type of the referent and the referents are also annotated. The antecedents and the anaphoric distance are annotated for textual deictic pronouns, while the semantic types of the referent and the referents are individuated for vague anaphors.

The semantic types of referent which we distinguish are mainly taken from the middle layer of the hierarchy of abstract objects proposed by (Asher, 1993) and comprise eventuality, fact-like and proposition. Similar types have been used by Hedberg et al. (2007) in their annotation work of abstract anaphora in English. To these types we have added property, which is assigned to entities introduced in discourse by copula predicates. We have also tentatively used the two types question and speech act in some of the dialogues.

Following the MATE/GNOME scheme nominal phrases are annotated in de XML elements, while other syntactic constructions, such as the antecedents of abstract anaphora, are annotated in seg elements. We have added to the MATE/GNOME scheme an explet element to mark up pleonastic pronouns. We have also tentatively used the two types question and speech act in some of the dialogues.

In the DAD scheme an XML-link is established between the anaphors and their antecedents. Ambiguous antecedents and/or ambiguous interpretations of the referents are marked in special comment elements. Finally a seg1 element is introduced to annotate clitics and zero anaphora in Italian, see for more details (Navaretta and Olsen, 2008). The annotation is made using the PALinkA tool (Orasan, 2003). An example of the DAD annotation is in 5:

```xml
<seg ID="s5" SYN-TYPE="scl">
  <W id="w19.277">at</W>
  <W id="w19.278">tr\{ae\}et</W>
  <W id="w19.279">er</W>
  <W id="w19.280">delamineret</W>
</seg>
```

5. The annotated corpora

In the project texts and transcriptions of spoken language in Danish and Italian have been annotated.

The transcriptions of spoken Danish contain information about stress so that it is possible to distinguish between the unstressed det and the stressed one d'et. The corpora transcribed until now in the two languages are the following:

- Danish dialogues and monologues from the DanPASS corpus (Gronnum, 2006) consisting of 52,145 and 21,224 running words respectively;
- three of Pirandello’s stories (1912 1937) (11,139 words) and their translations to Danish (11,280 words);
- Danish and Italian parallel EU texts (24,389 and 25,303 running words respectively);
- Danish texts from the juridical domain consisting of 11,600 words;
- extracts of newspaper and journal articles, novels and reports from the Danish Parole corpus (Keson and Norling-Christensen, 1998) (12,570 words);
- dialogues from the Italian AVIP corpus consisting of 70,054 running words.

The Danish DanPASS and the Italian AVIP dialogues have the same type as the MapTask dialogues.

---

6 The transcription conventions used in the DanPASS corpus are in http://www.cphling.dk/ng/danpass_webpage/danpass.htm.
8 http://www.hcrc.ed.ac.uk/maptask/.

---
The source language of the EU texts is not registered, but it is probably English.

5 The results
The Danish corpora have been annotated independently by two annotators, following the project’s coding manual (Navarretta, 2007a). The results have been compared and an agreed upon version of the annotated data has been made. Only part of the Italian corpora has been annotated by more than one annotator. Intercoder agreement measured in terms of $\kappa$ score (Carletta, 1996) was over 85% for the majority of the mark-ups (Navarretta and Olsen, 2008) and would be slightly higher using Krippendorff’s $\alpha$ (1995) because partially overlapping antecedents in the two annotations are counted as disagreement using the $\kappa$ score, see also (Passonneau, 2004).

In the following we present some of the results extracted from the annotated data. The number and type of pronoun encoded are given in table 1.

5.1 Results for Danish

Texts
The most frequently used abstract pronoun (85% of cases) in the Danish texts is det (it/this/that). The demonstrative pronoun dette (this) is used in the remaining cases and it is most frequently used in the juridical domain.

The annotated data confirm Navarretta’s (2004) suggestion that dette can signal that the antecedent is a part of the preceding utterance. More precisely dette is often used in the data when the antecedent is the last subordinate clause or the last clause in a group of coordinated clauses instead of the preceding complex clause (a complex clause being the preceding main clause and its subordinate clauses and/or a group of coordinated clauses). Differing from English demonstrative pronouns, the Danish dette is also used to signal that the antecedent is an individual object and not an abstract one, as usually expected in Danish. An example of this use is in 6 where the pronoun dette can both refer to the infinitive clause at etablere omfangsdren (to establish a circumferential drain) and to the nominal phrase omfangsdren (circumferential drain). Six out of seven native speakers have chosen the individual reading in this example.


(In order to decide the damage cause the property was again inspected by the damage adviser on the 27/12 2005. It was considered to establish a circumferential drain. Still there was no guarantee of this (the circumferential drain) to have the necessary effect.)

(Order of court about an insurance claim, 2006)

All the described uses of dette are compatible with Ariel’s (1994) proposal that demonstrative pronouns in general mark that the antecedent is not the most expected one.

Our data indicate that both det and dette are used with all types of antecedents and they refer to all types of referents. Reference to eventualities was done in 90% of the cases with det, reference to facts by det occurred in 63% of the cases and reference to propositions by det occurred in 82% of the cases. The demonstrative pronoun dette refers more often to facts than to propositions and events in the data. It never refers to properties.

Dialogues
The frequency of the abstract stressed (demonstrative) and unstressed (personal) det in the DANPASS dialogues is nearly the same (51% and 49% respectively). Reference to individual objects is done with a demonstrative (stressed det) in 44% of the cases and with a personal pronoun (unstressed det) in the remaining cases.

These results show that although Danish demonstrative pronouns are more frequent in abstract reference than in individual reference, they are not at all as frequent as demonstrative pronouns are in English, see i.a. (Hedberg et al., 2007; Navarretta, 2007).

Both stressed and unstressed det occur equally often when the antecedent is a clause in these dialogues. The pronoun det der (that) does not occur as abstract anaphor in the data and the pronoun det her (this) is nearly always used as cataphor. These results indicate that clauses are more often brought in focus in spoken Danish than in English. In the analysed dialogues the stressed and unstressed det refer to abstract objects belonging to all semantic types. However the stressed det is the preferred pronoun to refer to entities classified as eventualities (64% of the cases) and as fact-like (58% of the cases), while the
Table 1: Annotated pronouns

<table>
<thead>
<tr>
<th>corpus</th>
<th>all</th>
<th>abstract / textual deictic</th>
<th>indiv</th>
<th>pleonastic</th>
<th>cataphor</th>
<th>deictic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danish dialogues</td>
<td>713</td>
<td>241 (34%)</td>
<td>358</td>
<td>62 (9%)</td>
<td>45 (6%)</td>
<td>7 (0.9%)</td>
</tr>
<tr>
<td>Danish monologues</td>
<td>282</td>
<td>51 (18%)</td>
<td>181</td>
<td>23 (8%)</td>
<td>26 (9%)</td>
<td>1 (0.3%)</td>
</tr>
<tr>
<td>Danish texts</td>
<td>686</td>
<td>221 (32%)</td>
<td>232</td>
<td>194 (28%)</td>
<td>39 (5%)</td>
<td>-</td>
</tr>
<tr>
<td>Italian dialogues</td>
<td>212</td>
<td>15 (7%)</td>
<td>148</td>
<td>1 (0.05%)</td>
<td>46 (22%)</td>
<td>2 (0.09%)</td>
</tr>
<tr>
<td>Italian texts</td>
<td>571</td>
<td>59 (10%)</td>
<td>487</td>
<td>2 (0.04%)</td>
<td>23 (0.4%)</td>
<td>-</td>
</tr>
</tbody>
</table>

unstressed *det* is the preferred pronoun when the referents are propositions (69% of the cases).

Monologues
Reference to individual objects in the monologues is done in 57% of the cases with the unstressed *det*. Stressed and unstressed pronouns occur equally often in reference to abstract objects and have equally often clausal antecedents. Reference to propositions is in most cases done by personal pronouns (90%), while reference to eventualities is in most cases done by demonstrative pronouns (75%).

5.2 Results for Italian
The Italian data confirm that abstract pronominal reference is not as frequent in this language as it is in English and Danish. In fact nominal phrases such as *tali situazioni, questi avvenimenti, l’incidente* (such situations, these events, the accident), are often used in Italian in constructions where pronouns usually occur in the other two languages.

Texts
The Italian texts contain 59 abstract anaphors. Of these only four are demonstrative pronouns while 21 are zero anaphors. All pronouns refer to all types of referents, but zero anaphors are the most frequently used pronouns when the referred entity has been classified as a proposition. All four demonstrative abstract anaphors in the texts have a clausal antecedent and all the referents of these anaphors are classified as fact-like.

Dialogues
There are 55 abstract anaphors in the AVIP dialogues. Of these anaphors only three are demonstrative pronouns while 42 are zero anaphors. Zero anaphors refer to all types of abstract object and usually have clausal antecedents. Two of the three demonstrative abstract anaphors have a clausal antecedent (one referent classified as fact-like, the other as proposition) and one has a verbal phrase as antecedent (referent classified as eventuality).

6 Discussion
The data extracted from the DAD corpora confirm that the occurrences of language specific uses of abstract anaphora in Danish and Italian are so frequent that they must be inherent to these languages and connected to language specific aspects such as the languages’ syntax and pronominal system.

Zero anaphors and personal pronouns (both clitics and independent forms) are often used in Italian in contexts where demonstrative pronouns occur in English. Although abstract pronominal reference in Italian is seldom, these data confirm that zero anaphora, clitics and personal pronouns are often used in contexts where English requires the use of demonstrative pronouns.

The most used abstract anaphor in Danish texts is the pronoun *det* which is ambiguous regarding its pronominal type; the demonstrative pronoun *dette* (this) is not frequently used as abstract anaphor and often signals that the antecedent is the last clause in the preceding complex clause. Demonstrative abstract anaphors are slightly more frequent than personal abstract anaphors in the DanPASS dialogues, but they are not at all as frequent as in English. Furthermore personal pronouns are often used with clausal antecedents in Danish. The same seems to be the case for the Norwegian unstressed pronoun *det*, but Borthen et al. (1997) explain these occurrences by extralinguistic factors which according to them influence the salience of abstract entities. Although we agree with the observation that many factors determine salience and that aspects such as information structure must be taken into account, see i.a. (Hajicová et al., 1990; Kaiser, 2000; Gundel et al., 2003; Navarretta, 2005),

ISSL 1736-6305 Vol. 2
http://dspace.utlib.ee/dspace
handle/10062/7129
we also believe that some of language specific uses of abstract anaphora are so frequent in our data that they cannot be explained in terms of extralinguistic factors, but can be accounted for looking at the languages’ pronominal systems and their syntactic structure.

6.1 The pronominal system

English pronouns referring to inanimate entities belong to only one gender, while in Danish and Italian the pronouns referring to inanimate entities belong to two genders. Only pronouns in one of the two genders, the neuter gender in Danish and the masculine gender in Italian, can be abstract anaphora. Intuitively it is natural that abstract pronominal reference is more restricted in English than in the other two languages and this can partly explain the more frequent use of demonstrative pronouns in English to signal that the antecedent is abstract.

6.2 The syntactic structure

One of the syntactic characteristics of Danish is that clefts are very frequently used, e.g. *det er farligt at ryge* (it is dangerous to smoke) opposed to *at ryge er farligt* (smoking/to smoke is dangerous). Thus Danish clauses are very often in focus.11 This is why objects introduced by clauses are often more in focus than objects introduced by nominal phrases in Danish and are referred to by a personal pronoun. This is completely in line with Gundel et al.’s (1993) Givenness Hierarchy.

Differing from English and Danish, Italian is a free-order language and this might partially explain why abstract reference by nominal phrases is preferred, in that the use of nominal phrases restricts the antecedent search space. Although it is not possible to make any conclusion about abstract reference in Italian without extending the study to abstract nominal phrases, a first analysis of the Italian data indicates that abstract anaphora in this language are used when the abstract reading is the expected one and mainly occur in unambiguous contexts. This again can be explained in terms of the Givenness Hierarchy.

7 Concluding remarks and future work

In the paper we have described the information chosen in the DAD project to study and automatically treat abstract pronominal anaphora in Danish and Italian. These information has been included in the MATE/GNOME annotation scheme (Poesio, 2004) and the resulting extended scheme has been applied to annotate a corpus of Danish and Italian texts and dialogues. The intercoder agreement obtained on the data suggests that the chosen annotation types can be identified by different annotators in a consistent way.

The data indicates that there are language specific characteristics in the way abstract pronominal reference is done in Danish and Italian. An explanation of some of these characteristics in terms of the languages’ pronominal system and of their syntactic structure has been proposed. One of the consequences of our account of the differences in the use of pronominal types in the three languages is that the “default” cognitive status of individual and abstract entities introduced by various antecedent types can be different from language to language, and that resolution systems must account for this.

The fact that language specific aspects such as word order and syntactic structure must be taken into account in anaphora resolution in general is not controversial as indicated by the numerous language specific Centering algorithms, see i.a. (Brennan et al., 1987; Grosz et al., 1995; Strube and Hahn, 1996).

Although the data we have analysed so far show clear tendencies in the way abstract pronominal reference occurs in Danish and Italian and confirm some of the observations done by the author in previous studies, much work still must be done to annotate abstract anaphora in more types of data and to analyse all the information in our corpora, such as the relation between the syntactic type of clausal antecedents and the type of referent and between type of antecedent and/or pronoun and anaphoric distance.

We are currently annotating different types of dialogue in Danish because the DANPASS dialogues contained a higher number of demonstrative pronouns than we expected, probably because they regard the accomplishment of specific tasks, such as finding a path on a map and building a house out of some geometric figures. We are now annotating dialogues from the LANGCHART corpus (Gregersen, 2007) which are free conversations about everyday subjects12. The data we have anno-

11This is of course also related to information structure.

12Prosodic information has been added to the original
tated until now confirm that these dialogues contain fewer occurrences of the stressed det than the DanPASS dialogues. Furthermore abstract anaphors are much more frequent in the LANGCHART dialogues than in the DanPASS ones.

In the rest of the project we plan to complete the annotation of different types of corpora and to use the annotated data to train machine learning algorithms to automatically recognise and treat abstract anaphora in Danish.

Future work, which is out of the scope of the DAD project, is to investigate abstract reference by nominal phrases in Italian which is the most frequent way to refer to abstract objects introduced in discourse by verbal phrases, clauses and discourse segments.

8 Acknowledgements

The work we describe has been funded by the Danish Research Councils in the on-going project DAD, Det abstrakte det (The abstract det) (2007-2009).

Our thanks go to Sussi Olsen who, together with the author, has been annotating the Danish data. We thank Maria Carmela Malatesta for having annotated part of the Italian data.

References


E. Hajičová, P. Kuboň, and V. Kuboň. 1990. Hierarchy of Salience and Discourse Analysis and Production. In H. Karlgren, editor, Pro-