In "Centering in Discourse", Oxford University Press, Walker, Joshi, Prince editors, 1998

Centering in Italian

Barbara Di Eugenio
Learning Research and Development Center
University of Pittsburgh
Pittsburgh, PA, 15260 USA
dieugeni@cs.pitt.edu

Abstract. This chapter explores the correlation between centering and different forms of pronominal reference in Italian, in particular zeros and overt pronouns in subject position. In previous work (Di Eugenio, 1990), I proposed that such alternation could be explained in terms of centering transitions. In this chapter, I verify those hypotheses by means of a small corpus of naturally occurring data. In the process, I extend my previous analysis in several ways, for example by taking possessives into account; I also provide a more detailed analysis of CONTINUE: more specifically, I show that pronouns are used in a markedly different way in CONTINUE's preceded by another CONTINUE or by a SHIFT, and in CONTINUE's preceded by a RETAIN.

1 Introduction

Italian is a *pro-drop* language, in that the subject of a clause need not be overt. Thus, an Italian speaker has a variety of choices in realizing a subject, including using a null subject or an overt pronoun. In previous work (Di Eugenio, 1990), I proposed that the alternation of null and overt pronominal subjects could be explained in terms of centering transitions. However, the hypotheses I put forward in my earlier work were supported only by a few constructed examples. In this chapter, I report on a corpus study that I conducted in order to find more solid evidence for those hypotheses. Analyzing real data had the added benefit of bringing me to address issues still problematic for centering, such as how possessives and subordinates affect the ordering on the Cf list, and to provide a more detailed analysis of CONTINUE's.

The version of centering that I use is basically the one described in the overview by Walker, Prince and Joshi (this volume). However, the ordering for the Cf list I use is modified with respect to the usual one postulated for Western languages:

(1) SUBJECT $>$ OBJECT2 $>$ OBJECT $>$ OTHERS

Various researchers (Kameyama, 1985; Walker, Iida, and Cote, 1994) had pointed out that in languages such as Japanese, *empathy* and *topic* marking affect Cf ordering. Turan in (1995) argues that the notion of *empathy* is relevant to Western languages as well, because of non-agentive psychological verbs, such as *interest, seem*; perception verbs, such as *feel, appear*; and in general, expressions that refer to a character's point of view, such as *The thought crossed her mind*.[1] Turan points out that, with such expressions, it is the experiencer, which is often in object position, rather than the grammatical subject, that should be ranked higher. Moreover, Turan points out that in her Turkish corpus quantified indefinite subjects (qis) and arbitrary plural pro's (pro_{arb}) rank very low. Therefore, the Cf ranking in (1) should be amended as follows

[1] For a thorough treatment of subjective expressions and tracking characters' points of view, see (Wiebe, 1994).

(Turan, 1995):[2]

(2) **empathy** $>$ SUBJECT $>$ OBJECT2 $>$ OBJECT $>$ OTHERS $>$ **qis, pro**$_{arb}$

In this chapter, I will adopt (2).

Another difference between the standard notion of centering and the way I applied it is that I don't explicitly take discourse segment boundaries into account. According to (Grosz, Joshi, and Weinstein, 1995), centering is a local mechanism that applies within a single specific discourse segment, but not across segment boundaries. However, segmenting discourse is an active research area in itself, and there are no texts with agreed upon discourse structures. It seems that, when analyzing naturally occurring text, two approaches are possible:

- The first is to postulate a segment structure for the text of interest, for example exploiting paragraph boundaries, cue words etc (Walker, 1989).
- The second is to disregard segment boundaries and apply centering between every two adjacent utterances. It is in fact possible that the absence of a centering transition between two utterances indicates a segment boundary — see Passonneau (this volume), Walker (this volume). This is the approach I adopted, as my interest is in using centering as an explanatory tool for the distribution of pronominal forms. Notice that the cues that (Walker, 1989) uses to provide a discourse structure prior to applying centering include whether the first sentence of a new paragraph

  > has a pronoun in subject position or a pronoun where none of the preceding sentence-internal noun phrases match its syntactic features.

  In these cases, Walker doesn't consider the paragraph as constituting a discourse segment separate from the preceding one.

Finally, note that in this paper, as in my previous work, I take the same position advocated in (Walker, Iida, and Cote, 1994),

> that the interpretation of zeros is an inferential process, but that syntactic information provides constraints on this inferential process.

I will suggest that it is the syntactic context up and including the verbal complex that affects the interpretation of null subjects.[3]

The chapter is organized as follows. In Sec. 2 I discuss the Italian pronominal system and the hypotheses from my previous work. Sec. 3 describes the corpus I used and details the distributions of various referring expressions in subject position. In Sec. 4 I first discuss assumptions and extensions I had to make in order to apply centering to naturally occurring text; and then report on the correlations between pronouns and centering transitions. Sec. 5 analyzes such correlations: in particular, I refine the notion of CONTINUE in order to account for a non negligible number of occurrences of strong pronouns. Finally, Sec. 6 presents conclusions and future work.

2 Italian pronouns and Centering

2.1 The Italian pronominal system

In Italian, there are two pronominal systems, characterized by a different syntactic distribution: weak pronouns, that must always be cliticized to the verb (la, lo, li, le, gli - respectively her, accusative; him, accusative; them, masculine accusative; them, feminine accusative or her, dative; him, dative), and strong

[2] The first researcher to propose that *empathy* should precede *subject* in the Cf ranking was (Kameyama, 1985).

[3] I'm using here the term *inferential processing*, and later terms such as *strategies*, in their intuitive sense. Susan Brennan (p.c.) brought to my attention the difference between *strategic* or *inferential* processing and *automatic* processing, and the fact that syntactic cues of the kind I discuss in this paper affect the latter, not the former.

pronouns (lui, lei, loro - respectively he or him; she or her; they or them).[4] The null subject is considered part of the system of weak pronouns.

In Italian there is no neuter gender: nouns referring to inanimate objects are masculine or feminine. The weak pronouns used in this case are those of the corresponding gender. However, strong pronouns can't refer to inanimate objects, so that paraphrases or deictics are used: a strong pronoun for inanimate objects does exist — esso / essi for masculine, singular and plural, essa / esse for feminine, singular and plural — but it is not much used in current Italian: there is only one instance of esse in my corpus, in (Pagetti, 1993).[5]

Weak and strong pronouns are often in complementary distribution, as strong pronouns have to be used in prepositional phrases, e.g. per lui, *for him*, as in Ex. (3b). However, this syntactic alternation doesn't apply in subject position: the choice of null versus strong pronoun depends on pragmatic factors, and can be accounted for in terms of centering transitions.

(3a) **Maria**$_i$ é andata in vacanza con **suo padre**$_j$:
Maria$_i$ is gone on holiday with **her father**$_j$:

(3b) ϕ é stato un vero piacere per ***lo**$_j$**/lui**$_j$.
(it) is been a real treat for **him**$_j$.

2.2 Previous results

In (Di Eugenio, 1990), I proposed that

(4a) Typically, a null subject signals a CONTINUE, and a strong pronoun a RETAIN or a SHIFT.

(4b) A null subject can be felicitously used in cases of RETAIN or SHIFT if U_i provides syntactic features that force the null subject to refer to a referent different from Cb(U_{i-1}). Moreover, it is the syntactic context up to and including the verbal form(s) carrying tense and / or agreement that makes the reference felicitous or not.

These claims stemmed from (constructed) examples such as the following, where I use referents of different gender — Maria, female proper name; Giovanni, Giorgio, male proper names — to show how gender and morphological markings come into play when resolving reference. These examples are not ambiguous in English, given that a null subject is not an option available to a speaker. Boldface is just meant to highlight referential expressions, not to indicate stress; pronouns in parentheses in English correspond to zeros[6] in Italian; also remember that lui, gli, lo are masculine, and lei, le, la feminine.

(5a) **Maria**$_i$ voleva andare al mare.
Maria$_i$ wanted to go to the seaside.

(5b) ϕ_i Telefono' a **Giovanni**$_j$.
(She$_i$**)** phoned to **Giovanni**$_j$.

(5c) i. ϕ_i Si arrabbio' perche' ϕ_i non **lo**$_j$ trovo' a casa.
(She$_i$**)** self got angry because **(she**$_i$**)** not **him**$_j$ found at home.

ii. $\phi_{i/?j}$ Si arrabbio' perche' ϕ_j stava dormendo.
(She$_i$**)/(?He**$_j$**)** self got angry because **(he**$_j$**)** was sleeping.

[4] Traditionally, lui, lei, loro were the oblique forms of the strong system, with the nominative forms being respectively egli, ella, essi/e: however, in current Italian the latter forms are only rarely used as the oblique forms have replaced them in subject position — among the 33 instances of strong pronouns in Table 3, there are only three occurrences of egli, and all three of them occur in (Pagetti, 1993).

[5] This is the same article in which also the three occurrences of egli appear, see fn. 4.

[6] I will occasionally use the term *zero*: the speaker should keep in mind that Italian allows to drop only subjects, at least as a general rule.

iii. **Lui**$_j$ si arrabbio' perche' ϕ_j stava dormendo.
He$_j$ self got angry because **(he**$_j$**)** was sleeping.

iv. ϕ_j Si e' arrabbiato$_{masc}$ perche' ϕ_j stava dormendo.
(He$_j$**)** self is become angry$_{masc}$ because **(he**$_j$**)** was sleeping.

Consider the four (5c) variations; notice that Maria is both Cb(5b) and Cp(5b):

(5c).i The null subject refers to Maria, which is then both Cb(5c.i) and Cp(5c.i). (5c).i thus realizes a CONTINUE.

(5c).ii The most natural interpretation is that the null subject in the main clause refers to Maria — the null subject in the subordinate clause is forced to refer to Giovanni on pragmatic grounds.[7] For this same pragmatic reason, the null subject in the main clause may be interpreted as referring to Giovanni, but the discourse sounds less coherent.

(5c).iii As Giovanni was neither Cb(5b), nor Cp(5b), S performs a felicitous SMOOTH-SHIFT by referring to Giovanni with a strong pronoun.

(5c).iv Contrast this utterance with (5c).ii. They should have the same effect on the hearer, namely, the null subject should be interpreted as referring to Maria: instead in (5c).iv it is felicitously interpreted as referring to Giovanni. This is due to the fact that in (5c).iv the verb is present perfect;[8] the past participle agrees with the subject, and its masculine morphology forces the null subject to refer to Giovanni, and not to Maria. This last alternation lends support to my claim that it is the context up to and including the verb that is taken into account when interpreting a zero: it is the fact that the main verb is marked for masculine that allows the null subject to refer to something different from Cb(5b).

Further evidence for the importance of clues up to and including the verb comes from clitics, more specifically, from clitics embedded in a modal or control verb construction, as Exs. (6c).i through iii show. The crucial features is that clitics may be cliticized to the infinitival complement of the higher verb, as in (6c).i and (6c).ii, or can climb in front of the higher verb, as in (6c).iii.

(6a) ϕ_k ho parlato con **Maria**$_i$ ieri.
(I$_k$**)** have talked with **Maria**$_i$ yesterday.

(6b) ϕ_i É arrabbiata$_{fem}$ con **Giorgio**$_j$:
(She$_i$**)** is angry$_{fem}$ with **Giorgio**$_j$:

(6c) i. ϕ_i non vuole piú parlar**gli**$_j$.
(she$_i$**)** not wants any more to talk **to him**$_j$.

ii. # ϕ_j non vuole piú parlar**le**$_i$.
(he$_j$**)** not wants any more to talk **to her**$_i$.

iii. ϕ_j non **le**$_i$ vuole piú parlare.
(he$_j$**)** not **to-her**$_i$ wants any more to talk.

(6c).i realizes a CONTINUE, with the null subject referring to Cp(6b) = Cb(6b), namely Maria.

[7] In (Di Eugenio, 1990) I was not addressing the issue of interpreting null subjects in subordinate clauses.

[8] There is a grammatical temporal incoherence between (5c).iv and the preceding discourse, as the former is in the present perfect, while the latter is in the simple past. However, this temporal incoherence does not affect resolution of pronoun reference, as we can change the tenses in (5a) and (5b) to make the whole discourse temporally coherent, and the same kind of pronominal reference occurs. The coherent discourse is: **Maria**$_i$ *vuole andare al mare.* ϕ_i *Ha telefonato a* **Giovanni**$_j$. ϕ_j *Si e' arrabbiato perche'* ϕ_j *stava dormendo.* which translates to **Maria**$_i$ *wants to go to the seaside.* **(She**$_i$**)** *has phoned to* **Giovanni**$_j$. **(He**$_j$**)** *self is become angry*$_{masc}$ *because* **(he**$_j$**)** *was sleeping.* The verb *volere, to want*, in the first clause cannot be used in the present perfect in this context.

(6c).ii is incoherent. The preferred interpretation for the null subject is Maria; however, when the clitic le is found at the end of the sentence, the hearer is forced to change the interpretation of the null subject to Giorgio. The effect is similar to a syntactic "garden path".

(6c).iii is acceptable, as the clitic le, that in (6c).ii is cliticized onto parlare, climbs in front of the modal verb vuole: so the hearer is forced to exclude Maria as referent of the null subject. This happens *early enough* so that no "garden path" effect is registered.

It is from the contrast between (5c).ii and (5c).iv, and between the three (6c) variations that my claim about the importance of the context up to and including the verbal complex stems. Clearly, what exactly the *context up to and including the verbal complex* amounts to is not clear, as it includes agreement features, such as in (5c).iv, but not the clitic which hasn't climbed, as in (6c).i and (6c).ii: it apparently includes all the verbal forms carrying tense and agreement features, which explains why the past participle, marked for gender and number, is included, while the infinitival complement of a modal or control verb is not. The observation that the context up to and including the verbal complex affects the interpretation of the subject also makes sense from a lexical semantics point of view, given that the lexical semantics of the verb affects pronoun interpretation.

As we will see in Sec. 5.2, I found only few examples of such configurations in the corpus: while they support the hypothesis in (4b), more data is required to come to definitive conclusions. Moreover, it is clear that psycholinguistic experiments are needed to determine, among others: whether cases such as (5c).iv, a SMOOTH-SHIFT, require more time to process than cases such as (5c).i, a CONTINUE, even if both involve a null subject; whether a SMOOTH-SHIFT overtly marked by a strong pronoun such as in (5c).iii requires less processing time that one encoded by a zero and supported by syntactic features, such as in (5c).iv; whether indeed effects analogous to syntactic garden paths occur in cases such as (6c).ii.

3 Corpus

The corpus amounts to about 12,000 words (roughly 25 pages of text). It is composed of excerpts from two books (von Arnim, 1989; Fallaci, 1964), a letter (Mila, 1993), a posting on the Italian electronic bulletin board (SCI, 1994), a short story (Nichetti, 1993), and three articles from two newspapers (del Buono, 1993; Pagetti, 1993; La Nazione, 1994). The excerpts are of different lengths, with the excerpts from the two books being the longest, (von Arnim, 1989) with 3,641 words[9] and (Fallaci, 1964) with 1,918, and the posting on the Italian bulletin board, with only 603 words, the shortest. Texts were chosen to cover a variety of contemporary written Italian prose, from formal (newspaper articles about politics and literature), to informal (posting on the Italian electronic bulletin board).

The corpus I'm reporting about is a subset of the initial materials I assembled. In fact, I had to choose prose that describes situations involving several animate referents, as strong pronouns can refer only to animate referents.[10] Moreover, I eliminated texts that contain direct speech, another thorny issue for centering and in general for theories of discourse processing; excerpts from the two books don't contain dialogues. (von Arnim, 1989) is a book written in the form of a diary, which explains the presence of first person pronouns; the diary format has the advantage that there are almost no dialogues, which instead appear in the usual novel format. The excerpts from (von Arnim, 1989) involve at least two people, and possibly at least two are of the same sex; the chosen excerpts are discontinuous because, first, the diary format is in itself discontinuous; second, as the book revolves around the author's interest in gardening, many pages discuss plants and flowers or describe the landscape, and they obviously don't provide the required animate referents. The excerpt from (Fallaci, 1964) was chosen because the situation described involves four people, two men and two women.

[9] Most of my examples will come from (von Arnim, 1989), as it is the largest source of pronominal expressions: from now on assume that the source of an example is (von Arnim, 1989), unless otherwise noted.

[10] On the contrary, null subjects can refer to inanimate subjects as well and even be used for discourse deixis, i.e. to refer to a whole preceding discourse segment (Di Eugenio, 1989).

3.1 Quantitative data

Tables 1 through 4 provide the quantitative data for each text.

Table 1 gives the total number of grammatical subjects, and out of these, the total number of animate subjects: I counted subjects in both main and subordinate tensed clauses, but I excluded impersonal constructions and relative clauses where the relative pronoun is the subject.

Text	Total subjects	Animate subjects
(von Arnim, 1989)	241	229
(Fallaci, 1964)	27	24
(Mila, 1993)	63	50
(SCI, 1994)	45	32
(Nichetti, 1993)	77	60
(del Buono, 1993)	73	42
(Pagetti, 1993)	39	23
(La Nazione, 1994)	65	37
Total	630	497

Table 1: Total and animate subjects

Table 2 partitions animate subjects according to grammatical person. I didn't distinguish between singular and plural pronouns, as no phenomenon I will talk about seems to be affected by such distinction. About 90% of referential expressions are singular, as there are 60 plural subjects out of 630 total subjects.

Text	Total	1st	2nd	3rd
(von Arnim, 1989)	229	73	0	156
(Fallaci, 1964)	24	0	0	24
(Mila, 1993)	50	23	18	9
(SCI, 1994)	32	9	1	22
(Nichetti, 1993)	60	0	0	60
(del Buono, 1993)	42	0	0	42
(Pagetti, 1993)	23	0	0	23
(La Nazione, 1994)	37	0	0	37
Total	497	105	19	373

Table 2: First, second and third person animate subjects

Table 3 shows third person subjects partitioned into four classes: full NPs — this category also covers NPs that include a possessive adjective referring to an animate entity, which I will discuss below; strong pronouns; null subjects; and other anaphors, such as uno, one_{masc}, or tutte, all_{fem}.

Looking at Table 3, it is apparent that the percentage of full NPs versus pronouns is not constant through the eight texts. The percentages vary from 11% in (Mila, 1993),[11] to between 20% and 30% in (von Arnim, 1989), (Fallaci, 1964), and (SCI, 1994), 21%, 25% and 27% respectively. Then there is an increase for the last four texts, from 43% in (Nichetti, 1993), to 66% in (del Buono, 1993), 72% in (La Nazione, 1994) and

[11] (Mila, 1993) is probably a case in itself as it is a personal letter, and so it employs many more first and second person pronouns than third person ones — see Table 2.

Text	Total	Full NPs	Strong	Null	Other
(von Arnim, 1989)	156	45	23	81	7
(Fallaci, 1964)	24	6	2	16	0
(Mila, 1993)	9	1	2	5	1
(SCI, 1994)	22	7	0	11	4
(Nichetti, 1993)	60	26	1	33	0
(del Buono, 1993)	42	28	1	12	1
(Pagetti, 1993)	23	19	3	1	0
(La Nazione, 1994)	37	27	1	7	2
Total	373	159	33	166	15

Table 3: Distribution of 3rd person subjects

Text	strong	null
(von Arnim, 1989)	23	36
(Fallaci, 1964)	2	9
(Mila, 1993)	2	4
(SCI, 1994)	0	7
(Nichetti, 1993)	1	13
(del Buono, 1993)	1	6
(Pagetti, 1993)	3	0
(La Nazione, 1994)	1	5
Total	33	80

Table 4: Strong pronouns and null subjects

finally 82% in (Pagetti, 1993). Intuitively, it makes sense that more formal prose employs longer and more elaborate constructions.

It is clear that a full analysis should include full NPs as well, as about 60% of the full NPs in Table 3 are used referentially: for example, (Turan, 1995, ch. 6) discusses some intriguing results regarding the referential usage of full NPs in subject position in Turkish. Turan notices that, in her Turkish corpus, ROUGH-SHIFT's are realized 99% of the times by means of full NPs, and never by means of a null subject — this is consistent with the absence of ROUGH-SHIFT's in my pronominal data, see below. For SMOOTH-SHIFT's, the picture is more complicated: Turan notices that the shift to the object of the previous utterance is performed by means of a full NP if the object is inanimate, of an overt pronoun if the object is animate. I have started analyzing full NPs and their relation to centering: while I won't discuss full NPs in this chapter, some preliminary results can be found in (Di Eugenio, 1996).

Finally, Table 4 shows just the data of interest, namely strong pronouns and null subjects. Notice that the null subjects in Table 4 amount to about half of those appearing in Table 3: to analyze centering transitions, I only considered those null subjects whose antecedents are not determined by contraindexing constraints (Lasnik, 1976; Chomsky, 1981). I also excluded those that appear in a conjoined main clause which is not the first conjunct, and such that the null subject corefers with the subject of the preceding conjunct:

(7) **Lui**$_i$ non sembra mai demoralizzato, e ϕ_i va avanti ...
He$_i$ not appears ever frustrated, and ϕ_i carries on ...

In this case I consider the null subject to be constrained as if by contraindexing. Conjunctions do impose syntactic constraints that are different from those derived from simply juxtaposing clauses,[12] as shown for example by the fact that this is one of the rare contexts in which subject pronouns are sometimes dropped even in English.

4 Subject pronouns and centering transitions

4.1 Applying centering to real text

When analyzing real text, one realizes that many issues are still open. I will comment here on how deictics, possessives, and subordinate clauses affect centering.

Deictics. In texts such as (von Arnim, 1989), (Mila, 1993) or (SCI, 1994) there is an abundance of first and second person pronouns, most of them singular (see Table 2). The problem is whether *situational deictics* such as *I* and *you* are part of the Cf list or not; moreover, *I* in (von Arnim, 1989) often appears with verbs of thought, so that the problem of how to deal with situational deictics compounds with the problem of how to take subordinates into account. Consider the following example, where, in the utterance preceding (8a), Cb = Cp = pastore (pastor$_{masc}$), and lui in (8b) refers to pastore:

(8a) **Mi**$_i$ e' venuto spesso di pensare che cosa terribile sarebbe
To me$_i$ is come often to think what thing terrible would be

(8b) se **lui**$_j$ si sentisse male nella sua bussola
if **he**$_j$ self$_j$ felt bad in the his small room.

The issue is whether *I* belongs to the Cf list of (8a), or of (8a) and (8b) taken together, if a complement clause such as (8b) is not an independent centering unit. I follow (Walker, 1993) in assuming that deictics are always available as part of global focus, and therefore are outside the purview of centering.

[12] Contra (Kameyama, 1993).

Possessives. As noted above, the full NP category in Table 5 includes NPs that include a possessive adjective referring to an animate entity, such as *i suoi sforzi — his efforts*. Possessives frequently occur — they constitute about one fifth of the full NPs that perform centering transitions — and provide another means of keeping the center of attention.

The problem is deciding how possessives affect Cb computation and the order on the Cf list. An NP of type *possessive* in fact refers to two entities, the possessor P_{or} and the possessed P_{ed}: in the following example, P_{or} = Irais_i and P_{ed} = $\mathsf{husband}_k$ — in the previous sentence, Cb=Irais_i, Cf=[Irais_i > $\mathsf{English\ gentleman}_j$]:

(9a) **Suo**$_i$ **marito**$_k$ non ha piú avuto pace,
Her$_i$ **husband**$_k$ not has any longer had peace,

(9b) e ogni volta che ϕ_i deve uscire da una stanza ...
and every time that **(she**$_i$**)** has to leave from a room ...

While Cb computation does not appear to be affected by a possessive, that behaves like a pronoun, the Cf ranking needs to be modified. P_{ed} corresponds to the full NP, and thus its position in Cf is determined by the NP's grammatical function; as regards P_{or}, my working heuristics is to rank it as immediately preceding P_{ed} if P_{ed} is inanimate, as immediately following P_{ed} if P_{ed} is animate. Consider the following (contrived) discourse:

(10a) I met **Mary**$_i$ yesterday.

(10b) **She**$_i$ was worried.

(10c) i. **Her**$_i$ **husband**$_j$ was in the hospital.
ii. **Her**$_i$ **car**$_k$ wasn't working.

In both (10c).i and (10c).ii the Cb is $\mathtt{Mary}_i$; as regards the Cf list, in (10c).i it is $[\mathtt{husband}_j\ (\mathtt{P}_{ed}) > \mathtt{Mary}_i\ (\mathtt{P}_{or})]$, while in (10c).ii it is $[\mathtt{Mary}_i\ (\mathtt{P}_{or}) > \mathtt{car}_k\ (\mathtt{P}_{ed})]$. Clearly this heuristics needs to be rigorously tested.

Subordinates. Another important issue, that has not been extensively addressed yet — but see (Kameyama, 1997), (Suri and McCoy, 1993) — is how to deal with complex sentences that include coordinates and subordinates. The questions that arise concern whether there are independent Cb's and Cf lists for every clause; if not, how the Cb of the complex sentence is computed, and how semantic entities appearing in different clauses are ordered on the global Cf list.

A simple example is the following discourse, for which I provide a literal, but not word by word, translation; for the utterance preceding (11a), we have Cb(U_{i-1}) = vicina_j (neighbor$_{fem}$), Cf(U_{i-1}) = [vicina_j].

(11a) Prima che **i pigroni**$_i$ siano seduti a tavola a far colazione,
Before **the lazy ones**$_i$ sit down to have breakfast,

(11b) **lei**$_j$ e' via col **suo**$_j$ calessino alle altre cascine della tenuta.
she$_j$ has left with **her**$_j$ buggy for the other farmhouses on the property.

The issue is whether the Cb and Cf list are updated after the whole sentence, or whether a new Cb and Cf list are computed after (11a): these new items would then be the input to a new computation of Cb and Cf list after (11b). It is my impression that preposed adjuncts, such as (11a), do affect centering transitions: the fact that an overt subject is used after a preposed adjunct seems to support the fact that this is a SHIFT or a CENT-EST — see below — and not a simple CONTINUE from the previous utterance that could be encoded with a null subject, which on the contrary is not particularly felicitous here. As a working hypothesis, I've loosely adopted Kameyama's proposal (1993; 1997), that sentences containing conjuncts and tensed adjuncts are broken down into a linear sequence of centering "units", while tenseless adjuncts don't generate independent centering units[13]. Thus (11a) and (11b) have each distinct Cb's and Cf lists.

4.2 Centering Transitions

Table 5 shows the distribution of null and strong pronouns with respect to centering transitions, while Table 6 gives the distribution of transitions per text.

Type	Total	CONTINUE	RETAIN	SHIFT	CENT-EST	OTHER
zero	80	56	4	6	12	2
strong	33	13	4	5	10	1
Total	113	69	8	11	22	3

Table 5: Distribution of centering transitions

Text	Type	Total	CONTINUE	RETAIN	SHIFT	CENT-EST	OTHER
(von Arnim, 1989)	null	36	20	4	5	6	1
	strong	23	7	3	3	9	1
(Fallaci, 1964)	null	9	7	0	0	2	0
	strong	2	0	0	1	1	0
(Mila, 1993)	null	4	4	0	0	0	0
	strong	2	1	0	1	0	0
(SCI, 1994)	null	7	6	0	0	1	0
	strong	0	0	0	0	0	0
(Nichetti, 1993)	null	13	11	0	0	2	0
	strong	1	1	0	0	0	0
(del Buono, 1993)	null	6	4	0	1	0	1
	strong	1	1	0	0	0	0
(Pagetti, 1993)	null	0	0	0	0	0	0
	strong	3	2	1	0	0	0
(La Nazione, 1994)	null	5	4	0	0	1	0
	strong	1	1	0	0	0	0
Total		113	69	8	11	22	3

Table 6: Distribution of centering transitions per text

Tables 5 and 6 require some explanation, as they don't distinguish between SMOOTH- and ROUGH-SHIFT, and include new transitions such as CENT-EST.

First of all, I don't distinguish between SMOOTH- and ROUGH-SHIFT, as ROUGH-SHIFT's involving pronouns can appear only in very specific conditions, that do not occur in my data. In fact, the conditions for a ROUGH-SHIFT are:

1. $Cb(U_i) \neq Cb(U_{i-1})$ and
2. $Cb(U_i) \neq Cp(U_i)$

[13] The situation for complements is more complicated, and space prevents me from discussing it.

Notice that given the Cf ranking in (2), the null or strong pronoun p_i in subject position will always be Cp(U$_i$).[14] Thus, for a ROUGH-SHIFT to arise, p_i must not be Cb(U$_i$), otherwise a SMOOTH-SHIFT would occur. For p_i not to be Cb(U$_i$), U$_i$ must have at least another pronoun (otherwise if p_i is the only pronoun, it is Cb(U$_i$), and being also Cp(U$_i$), condition 2 does not obtain). A configuration in which a ROUGH-SHIFT obtains in U$_i$ is, schematically — both e$_2$ and e$_3$ are pronouns, e$_3$ corresponds to p_i:

$$\begin{array}{lll} U_{i-1}: & Cb = e_1, & Cf = [e_1 > e_2 > e_3] \\ U_i: & Cb = e_2, & Cf = [e_3 > e_2] \end{array}$$

A constructed example where (12b) and (12c) instantiate this configuration is:

(12a) **Giorgio**$_i$ e' amico di **Maria**$_j$.
Giorgio$_i$ is friend of **Maria**$_j$.

`Cb =?; Cf = [Giorgio`$_i$ `> Maria`$_j$`]`

(12b) ϕ_i **l**$_j$**'** ha presentata$_{fem}$ a **Giovanni**$_k$.
(He$_i$**) her**$_j$ has introduced$_{fem}$ to **Giovanni**$_k$.

`Cb = Giorgio`$_i$`; Cf = [Giorgio`$_i$ `> Giovanni`$_k$ `> Maria`$_j$`]`

(12c) **Lei**$_j$ **lo**$_k$ trova antipatico.
She$_j$ **him**$_k$ finds unpleasant.

`Cb = Giovanni`$_k$`; Cf = [Maria`$_j$ `> Giovanni`$_k$`]`

However, there are no examples of this sort in my data.[15]

Moving now to CENT-EST and OTHER, also included in Tables 5 and 6, CENT-EST — for CENTER ESTABLISHMENT — captures the fact that sometimes pronouns (even the null subject!) can be used to refer to an entity in the global focus, and not on the Cf list of the previous utterance. Also (Grosz, Joshi, and Weinstein, 1995, p.216) notices this phenomenon:

> The second case [of quasi violations of Rule 1] concerns the use of a pronoun to realize an entity not in the $C_f(U_n)$; such uses are strongly constrained. The particular use that have been identified involve instances where attention is shifted globally back to a previously centered entity (e.g. (Grosz, 1977), (Reichman, 1985)).

However, not all occurrences of CENT-EST represent a global focus shift. For example, if one postulates that adjuncts constitute centering units in themselves, it is possible that the shift in U$_i$ is to an entity that belonged to Cf(U$_{i-2}$), where U$_{i-1}$ is an adjunct preposed to U$_i$: clearly such a shift seems to have a less dramatic effect (in terms of inference load) than a shift to an entity that had been mentioned much earlier in the discourse.

I suspect that there may be some correlation between how global the shift is and the usage of a specific form, in particular the usage of a full NP. Moreover, when the null subject is used for CENT-EST, the resulting discourse may be slightly incoherent. For example, in (von Arnim, 1989, p. 70-71), the author describes the pastor and his wife. After discussing the virtues of both, the author devotes the next 10 (complex) sentences only to the pastor; in fact, the 10th sentence doesn't talk about either of them. When in the 11th sentence

[14] Empathy effects don't occur in my data when the subject is a pronoun: rather, they arise when the subject is a full NP's pertaining to a character's point of view, as in *Le sue convinzioni lo trascinano fuori dalla casetta a tutte le ore — His beliefs drag him out of his house at all hours*.

[15] Note that (12c) also has another interpretation, a RETAIN with **lo** referring to Giorgio$_i$ rather than to Giovanni$_k$: in this case though, the ROUGH-SHIFT is preferred to RETAIN. It is clear that the semantics of the situation comes into play.

the author uses a null subject to refer to both, the effect is slightly incoherent. Sentences 9 through 11 are reported here with a literal, but not word by word, translation.

(13a) **Lui**$_i$ non parla mai di **queste cose**$_j$, ma come ϕ_j potrebbero rimanere nascoste?
He$_i$ never talks about **these things**$_j$, but how could **(they**$_j$**)** remain hidden?

(13b) Qui **tutti**$_k$ sanno tutto prima che la giornata volga al termine, e quel che ϕ_l mettiamo in tavola é assai piú d'interesse generale del piú sbalorditivo capovolgimento politico.
Here **everybody**$_k$ knows everything before the day comes to an end, and what **(we**$_l$**)** have for dinner is of much more general interest than the most surprising political change.

(13c) ϕ_m Hanno un cottage spazioso, carino, con un bel pezzetto di terreno attiguo al cimitero.
(They$_m$**)** have a roomy, nice cottage with a sizable piece of land next to the cemetery.

Strong pronouns are sometimes used to establish a new center by selecting a member of a set available on the Cf list of the previous utterance, as in (14) and (15). In the utterance preceding (14a), Cb = Cp = {pastore & sua moglie}, {pastor & his wife}:

(14a) ϕ_i Sono entrambi di un'austera devozione.
(They$_i$**)** are both of an austere devotion.

(14b) **Lui**$_j$ lavora nella sua parrocchia con nobile dedizione, and ...
He$_j$ works in the his parish with noble dedication, and ...

Another such example:

(15a) ϕ Avevamo ormai finito il té
(We$_j$**)** had already finished the tea

(15b) e **lei**$_i$ era salita di sopra a cambiar**si**$_i$ quando ...
and **she**$_i$ was gone upstairs to change **herself**$_i$ when ...

It is debatable whether, once a set is available on the Cf list, also its members are: however, a null subject would be infelicitous both in (14b) and in (15b), thus providing weak evidence that the members of sets on the Cf list are not themselves available on the Cf list. I consider such usages of strong pronouns as CENT-EST's.

Finally, OTHER refers to configurations that I have left unanalyzed for the time being: they are characterized by parallelism, or by expressions that build a set out of Cb(U_{i-1}) and some other entity, such as *sia lui che sua moglie — both him and his wife*. It is not clear how to deal with these constructions within the centering framework yet.

5 Discussion

The reader will recall that the reason I conducted my corpus analysis was to verify the strategies in (4), repeated here for convenience:

(16a) Typically, a null subject signals a CONTINUE, and a strong pronoun a RETAIN or a SHIFT.

(16b) A null subject can be felicitously used in cases of RETAIN or SHIFT if U_i provides syntactic features that force the null subject to refer to a particular referent and not to Cb(U_{i-1}). Moreover, it is the syntactic context up to and including the verbal form(s) carrying tense and / or agreement that makes the reference felicitous or not.

I will now detail the results.

5.1 CONTINUE **after** RETAIN

The first part of (16a) — null subjects used for CONTINUE — is strongly supported. Zeros are used 80% of the times, and there is a significant difference (χ^2 = 9.204, $p < 0.01$) between zeros and strong pronouns used in CONTINUE and zeros and strong pronouns used in all other transitions taken together — see the following contingency table.[16] Thus, in its use of null subjects for CONTINUE, Italian behaves in the same

	CONTINUE	ALL OTHERS
zero	56	24
strong	13	20

Table 7: CONTINUE vs. all other transitions

way as languages as diverse as Japanese (Kameyama, 1985; Walker, Iida, and Cote, 1994; Shima, 1995) and Turkish (Turan, 1995), (Turan, this volume). In fact, the usage of zeros for CONTINUE seems to be a robust cross-linguistic phenomenon.

However, as the 20% percentage of strong pronouns used for CONTINUE is not negligible, I set out to investigate which factors may affect such a choice. I analyzed the CONTINUE's in my corpus, and I did find that one relevant factor is the transition preceding the CONTINUE in question. Consider Table 8, that shows the different possible transitions in the utterance U_{i-1} preceding the utterance U_i in which a CONTINUE occurs. The configuration in which a CONTINUE is preceded by a RETAIN, which I call RET-CONT, differs from the other two because of the constraint $Cp(U_{i-1}) \neq Cb(U_{i-1})$ in the RETAIN: this in a sense predicts that the center will shift. But if a RETAIN is followed by a CONTINUE, as in a RET-CONT, such prediction is not fulfilled.

	CONTINUE	RETAIN	SHIFT
U_{i-1}	$Cb_{i-1} = Cb_{i-2}$ $Cp_{i-1} = Cb_{i-1}$	$Cb_{i-1} = Cb_{i-2}$ $\mathbf{Cp_{i-1} \neq Cb_{i-1}}$	$Cb_{i-1} \neq Cb_{i-2}$ $Cp_{i-1} = Cb_{i-1}$
U_i	$Cb_i = Cb_{i-1}$ $Cp_i = Cb_i$		

Table 8: Centering transitions preceding a CONTINUE

Before providing quantitative support for the distinct behavior of RET-CONT, I will illustrate this configuration with Ex. (17), which provides two examples of RET-CONT: the first, in (17c), is realized with a strong pronoun; the second, in (17e), is realized with a null subject. In the utterance preceding (17a), Cb = Irais

[16] χ^2 test results are reported here more as a source of suggestive evidence than as strong indicators, as the observations in the corpus, which come from only 8 authors, are not totally independent.

and Cf = [Irais] — the translation is literal but not word by word.

(17a) ϕ_i Incomincerό a ricondurre il **suo**$_j$ **pensiero** sui **suoi**$_j$ **doveri** chiedendo**le**$_j$ ogni giorno
(I$_i$) will start to bring **her**$_j$ **thoughts** back to **her**$_j$ **duties** by asking **her**$_j$ every day
`Cf: [Irais`$_j$ `> Irais`$_j$`'s thoughts, Irais`$_j$`'s duties], Cb:Irais, continue`

(17b) *come sta* **suo**$_j$ **marito**$_k$.
how **her**$_j$ **husband**$_k$ is.
`Cf: [husband`$_k$ `> Irais`$_j$`], Cb:Irais`$_j$`, retain`

(17c) Non è che **lei**$_j$ **gli**$_k$ voglia granché bene,
It's not the case that **she**$_j$ cares much about **him**$_k$
`Cf: [Irais`$_j$ `> husband`$_k$`], Cb:Irais`$_j$`, continue`

(17d) perché **lui**$_k$ non corre ad aprir**le**$_j$ la porta
because **he**$_k$ doesn't run to open the door **for her**$_j$
`Cf: [husband`$_k$ `> Irais`$_j$`], Cb:Irais`$_j$`, retain`

(17e) ogni volta che Φ_j si alza per lasciare la stanza;
whenever **(she**$_j$**)** gets up to leave the room.
`Cf: [Irais`$_j$`], Cb:Irais`$_j$`, continue`

Moving now to the quantitative analysis of RET-CONT, Table 9 shows how RET-CONT's affect the usage of null and strong pronouns — CONT-CONT and SHIFT-CONT respectively refer to a CONTINUE preceded by another CONTINUE or by a SHIFT.

Type	Total	CONT-CONT+ SHIFT-CONT	RET-CONT
zero	56	51	5
strong	13	7	6
Total	69	58	11

Table 9: Pronoun occurrences for RET-CONT

Compared to strong pronouns, null subjects are used 87% of the times for CONT-CONT and SHIFT-CONT taken together and only 45% of the times for RET-CONT. Moreover, if a zero is used in a CONTINUE, that CONTINUE is ten times more likely to be a CONT-CONT or SHIFT-CONT than a RET-CONT; in contrast, if a strong pronoun is used in a CONTINUE, that CONTINUE is as likely to be a CONT-CONT or a SHIFT-CONT as a RET-CONT. These trends in usage are confirmed by a strongly significant difference between zeros and strong pronouns used in CONT-CONT plus SHIFT-CONT, and zeros and strong pronouns used in RET-CONT (χ^2 = 10.910, $p < 0.001$). Moreover, there is a very strongly significant difference between zeros and strong pronouns used in CONT-CONT plus SHIFT-CONT, and zeros and strong pronouns used in all other transitions, including RET-CONT — χ^2 = 16.922, $p < 0.001$, see Table 10. Consistently, there is no significant

	CONT-CONT + SHIFT-CONT	RET-CONT + ALL OTHERS
zero	51	29
strong	7	26

Table 10: CONT-CONT + SHIFT-CONT vs. all other transitions

difference between zeros and strong pronouns used in RET-CONT and zeros and strong pronouns used in transitions different from CONTINUE — χ^2 = 0.292, $p < 0.7$, see Table 11. This suggests that RET-CONT's pattern more like transitions different from CONTINUE than like other CONTINUE's; in fact, all transitions other than CONTINUE in Table 5 present a rough half-half split between zeros and strong pronouns, as do RET-CONT's in Table 9.

	RET-CONT	ALL OTHERS (excluding CONTINUE)
zero	5	24
strong	6	20

Table 11: RET-CONT vs. transitions different from CONTINUE

My results on different pronominal distributions for CONT-CONT and SHIFT-CONT on the one hand, and RET-CONT on the other, seem to be yet another source of evidence for the hypothesis that a RETAIN signals an upcoming shift: namely, not fulfilling the prediction given by the Cp seems to require the explicit signal provided by a strong pronoun. [17]

Also (Turan, 1995) independently noticed the existence of RET-CONT's, and her results are compatible with mine: she found that in RET-CONT's, zeros decrease from 97% to 68% while strong pronouns increase from 1% to 11% with respect to their percentages of use for CONT-CONT and SHIFT-CONT.

RETAIN, SHIFT and CENT-EST. As far as RETAIN's and SHIFT's are concerned, the numbers are too small to draw any definitive conclusion. The tentative one is as follows: the examples I found do seem to support (16b), as I will discuss below; namely, the null subject can be used in cases of RETAIN's and SHIFT's if there are enough "early" clues that force the null subject to refer to a particular referent. However, the numbers in themselves do not identify any preferred usage for strong pronouns for these transitions, contrary to what claimed by (16a).

CENT-EST's pattern like RETAIN and SHIFT (and RET-CONT!), in that zeros and strong pronouns appear to be evenly distributed; moreover, there is a significant difference between zeros and strong pronouns used for CENT-EST's and zeros and strong pronouns used for CONT-CONT plus SHIFT-CONT (χ^2 = 10.624, $p < 0.01$).

A topic for future work is to investigate which factors, if any, affect the choice between null and strong pronouns in these configurations, especially because null subjects used for SMOOTH-SHIFT or for CENT-EST sometimes result in a slightly less coherent discourse — see (13c).

5.2 Verb agreement, clitics, and null / strong pronouns

The second part of my claim, (16b), namely, that a null subject can be used if U_i provides syntactic clues that force the null subject not to refer to $Cb(U_{i-1})$, is indeed borne out — however, given the small number of occurrences of null subjects encoding these transitions (four RETAIN's and six SHIFT's) this conclusion can just be tentative. The most frequent clue is agreement in gender and / or number; in some examples, clitics are useful for disambiguation as well, but I found no example of clitic climbing as discussed with respect to Ex. (6).

[17] The only researcher I know of who argues against the prediction associated with a RETAIN is Linson: in (1993), he presents evidence based on a corpus study, in which a RETAIN is followed by a CONTINUE 50% of the times, and by a SHIFT only 15% of the times. However, Turan notices (p.c.) that Linson used the Cf ranking in (1), and that if (1) is amended as in (2), i.e. taking "empathy" into account, Linson's results may be different, in that certain RETAIN's may in fact be CONTINUE's.

However, I hoped to be able to verify a stronger claim, that whenever such clues are available a null subject is used. But the data only partly support this stronger claim. In fact, of the 9 instances of strong pronouns realizing a RETAIN or SHIFT, 4 do have clues that should make a null subject possible. Two of the four examples, both RETAIN's, are :

(18a) **Io**$_i$ faccio visita **a lei**$_j$ una volta all'anno,
I$_i$ pay visit **to her**$_j$ one time per year,

(18b) e **lei**$_j$ **mi**$_i$ ricambia la visita quindici giorni dopo.
and **she**$_j$ **to me**$_i$ returns the visit fifteen days later.

(19a) ϕ_i É pronta a difender**lo**$_j$ in ogni occasione contro di **noi**$_k$.
(She$_i$**)** is ready to defend **him**$_j$ in every occasion against of **us**$_k$.

(19b) **Lui**$_j$ non **le**$_i$ parla mai.
He$_j$ not **to her**$_i$ talks ever.

Both (18b) and (19b) have a clitic available before the main verb, analogously to (6c).iii: thus, substituting the strong pronoun with a zero should result in a coherent discourse, but this is not the case. (Turan, 1995) notices that in Turkish the rule that prescribes using zeros in a CONTINUE is overridden if the pronominal expression has to carry additional pragmatic information, such as phonetic prominence or a listing reading. In the case of (18b), clearly parallelism comes into play. In (19b), there is indeed a contrast between **lei** (a female guest) defending **lui** (the author's husband), and **lui** trying to ignore **lei** as much as possible.

In contrast to (18b) and (19b), I would like to mention two examples, again both RETAIN's, that pattern like Ex. (6c).ii, namely, where the clitic is found "too late" to allow the usage of a null subject.

The first example can be found in (17d) above. The clitic **le** (for her) is cliticized to the infinitive *aprire* (to open), which is an adjunct to the main verb *corre* (runs). The second example is (20d): the clitic **le** (to her) doesn't climb in front of the modal *vuole* (wants), but is cliticized to the lower verb *correr(e)* (to run).

(20a) Ma **lui**$_k$ doveva sposare **la cuoca**$_{j,fem}$,
But **he**$_k$ wanted to marry **the cook**$_{j,fem}$,

(20b) e **la cuoca**$_j$ ha visto un fantasma
and **the cook**$_j$ has seen a ghost

(20c) e ϕ_j **se**$_j$ n'é andata$_{fem}$ su due piedi,
and ϕ_j **self**$_j$ is gone on two feet,

(20d) e **lui**$_k$ vuole correr**le**$_j$ dietro ...
and **he**$_k$ wants to run **to her**$_j$ after ...

6 Conclusions

The work presented in this chapter aims at explaining the different usages of Italian pronominal subjects in terms of centering transitions. The current research follows up on and extends (Di Eugenio, 1990). My goal was to test the claims made in my earlier work and based on constructed examples against naturally occurring data. Not surprisingly, conducting a corpus analysis was useful not just to verify those hypotheses but also to extend the analysis in a variety of ways. The hypothesized strong preference for null subjects in the case of CONTINUE is verified. Furthermore, taking the transition preceding a CONTINUE into account provides an elegant explanation for about half of the strong pronouns used in CONTINUE's: a CONTINUE preceded by a RETAIN behaves differently from one preceded by a CONTINUE or by a SHIFT.

The results regarding the usage of strong pronouns for RETAIN and SHIFT are mixed: in fact, the numbers don't indicate any preference for one pronominal form over the other. Somewhat to my surprise, I found that what is supported, at least tentatively given the small numbers, is the second part of my claim, (16b): a null subject can be used for RETAIN or SHIFT if the context up to and including the verbal forms marked for tense and agreement provides "early enough" clues that prevent pronominal interpretation garden paths.

It is clear that it is necessary to refine the analysis first of all by collecting more instances of RETAIN's and SHIFT's, and of CONTINUE's occurring after RETAIN. Moreover, other pragmatic factors, such as parallelism and contrast, should be examined, in order to understand how they affect the choice of referring expressions.

I think a fruitful direction in which to move is to study the functions of referential full NP's in terms of centering transitions. Some preliminary results, available in (Di Eugenio, 1996), show that the percentage of CONTINUE's realized by means of full NP's is not negligible at all, as it amounts to 16%; and that full NP's account for the majority of CENT-EST: the preference for full NP's over other referring expressions for CENT-EST is statistically significant. If CENT-EST's do correspond, at least in part, to shifts in global focus, as mentioned in Sec. 4, an issue to tease apart concerns the conditions under which full NP's, strong pronouns and zeros are used. In general, by including full NP's in the analysis, a more complete account of the choices of referring expressions will be possible.

Acknowledgements

This research was carried out while the author was with the Computational Linguistics program, Carnegie Mellon University, Pittsburgh, PA, USA. I wish to thank Susan Brennan, Ümit Turan and Marilyn Walker for their insightful comments on earlier versions of this chapter.

References

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Foris, Dordrecht.

del Buono, Oreste. 1993. Avanti a tutto Metz contro la retorica fascista. *La Stampa, TuttoLibri*, December.

Di Eugenio, Barbara. 1989. Clausal reference in Italian. In *Proceedings Penn Linguistics Colloquium*.

Di Eugenio, Barbara. 1990. Centering Theory and the Italian Pronominal System. In *Proceedings 13th International Conference on Computational Linguistics (COLING 90)*, pages 270–275.

Di Eugenio, Barbara. 1996. The discourse functions of Italian subjects: a centering approach. In *Proceedings 16th International Conference on Computational Linguistics (COLING 96)*, pages 352–357.

Fallaci, Oriana. 1964. *Penelope alla guerra*. Rizzoli, Milano.

Grosz, Barbara. 1977. The representation and use of focus in dialogue understanding. Technical Report 151, SRI International, 333 Ravenswood Avenue, Menlo Park, CA 94025.

Grosz, Barbara, Aravind Joshi, and Scott Weinstein. 1995. Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics*, 21(2):203–225.

Kameyama, Megumi. 1985. *Zero anaphora: the case of Japanese*. Ph.D. thesis, Stanford University.

Kameyama, Megumi. 1993. Intrasentential Centering. In *Proceedings of Workshop on Centering,* University of Pennsylvania.

Kameyama, Megumi. 1997. Intrasentential Centering. *Centering in Discourse*, Ellen Prince, Aravind Joshi, Lyn Walker editors, Oxford University Press.

La Nazione. 1994. Il declino di Ross Perot: la nazione piu' capitalista del mondo sa di non essere un'azienda, March.

Lasnik, Howard. 1976. Remarks on coreference. *Linguistic Analysis*, 2(1).

Linson, Brian. 1993. A Distributional Analysis of Discourse Transitions. In *Proceedings of Workshop on Centering,* University of Pennsylvania.

Mila, Massimo. 1993. Letter to his mother. *(Reprinted in) La Stampa, TuttoLibri*, December.

Nichetti, Maurizio. 1993. La tv delle formiche. *Comix*, 84.

Pagetti, Carlo. 1993. Dick l'illuminato. *La Stampa, TuttoLibri*, December.

Reichman, Rachel. 1985. *Getting Computers to Talk like You and Me*. MIT Press.

SCI. 1994. Posting on the soc.culture.italian electronic newsgroup, March.

Shima, Kaori. 1995. Anaphora Resolution in Japanese: a Centering Approach. Master's project, Carnegie Mellon University, May.

Suri, Linda Z. and Kathleen F. McCoy. 1993. Focusing and Pronoun Resolution in Particular Kinds of Complex Sentences. In *Proceedings of Workshop on Centering,* University of Pennsylvania.

Turan, Ümit Deniz. 1995. *Subject and Object in Turkish Discourse: a Centering Analysis*. Ph.D. thesis, University of Pennsylvania. Institute for Research in Cognitive Science report IRCS-96-13.

von Arnim, Elizabeth. 1989. *Il giardino di Elizabeth*. Bollati Boringhieri, Torino.

Walker, Marilyn. 1989. Evaluating discourse processing algorithms. In *Proc. 27th Meeting, Association for Computational Linguistics*, pages 251–261.

Walker, Marilyn. 1993. Initial Contexts and Shifting Centers. In *Proceedings of Workshop on Centering,* University of Pennsylvania.

Walker, Marilyn, Masayo Iida, and Sharon Cote. 1994. Japanese Discourse and the Process of Centering. *Computational Linguistics*, 20(2):193–232.

Wiebe, Janyce M. 1994. Tracking Point of View in Narrative. *Computational Linguistics*, 20(2):233–287.