THE SYNTAX-DISCOURSE INTERFACE: EFFECTS OF THE MAIN-SUBORDINATE DISTINCTION ON ATTENTION STRUCTURE

Eleni Miltsakaki

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2003

Ellen Prince

Co-Supervisor of Dissertation

Aravind Joshi

Co-Supervisor of Dissertation

Don Ringe

Graduate Group Chairperson

Acknowledgments

My first year at graduate school, after a long break of a professional career, was challenging in many respects. My heartfelt thanks to Sabine Iatridou, Tony Kroch, Don Ringe and Ellen Prince for their invaluable help, emotional support and encouragement during that time. I would not have made it without you.

As a graduate student in the Linguistics Department at Penn, I have had the rare privilege to "grow up" in a truly interdisciplinary academic environment. I am deeply and sincerely grateful for that to everyone, faculty and students, from the departments of Linguistics, Computer Science and Psychology, and of course, the Insitute of Research in Cognitive Science.

I don't know where to begin thanking my advisors Ellen Prince and Aravind Joshi. I feel particularly fortunate for the privilege and honor of learning from them. My intellectual debt to both is too large to fit in the acknowledgments.

Ellen Prince has constantly supported, encouraged and guided me throughout my studies. She has been an unfailing source of inspiration and I have benefited enormously from her courses, seminars, long discussions inside and outside the classroom, and her brilliant remarks on linguistic analyses. I have learned from her what it means to be a linguist and I can't thank her enough for that. I am immensely proud to have been her student.

Aravind Joshi has been a great influence on my intellectual development in graduate school. He is solely responsible for my appreciation of computational models of language processing. I have learned from him the scientific importance and intellectual merit of combining insights from different disciplines. I am deeply grateful to him for his support

of my research interests, for his generous contributions to my work and thinking, and for letting me be his student.

I am grateful to the members of my committee, Robin Clark, Maribel Romero and Martha Palmer for their insightful comments and criticism on my dissertation work and throughout my years in graduate school. Robin Clark has taught me all I know about semantics and I have greatly benefited from his insights on complexity. Maribel Romero's patience, expertise, and clarity of thought has helped me a lot in identifying weaknesses in my arguments. Martha Palmer has always been encouraging and generous with fresh ideas about my current work and future extensions of it.

Special thanks to Lila Gleitman, John Trueswell, Jesse Snedeker, Felicia Hurewitz and the participants of the CHEESE seminars for sharing their expertise in experimental design and analysis. A big thank you also goes to the members of the Institute of Language and Speech Processing in Athens, Greece, who participated in the Greek experiment and Maria Gavriilidou for her invaluable help in organizing the experimental session.

I am immensely grateful to Karen Kukich, the director of the NLP group at ETS, for her collaboration on the *e-rater* experiment, her enthusiasm, the very useful and extended discussions, and for sharing with me her expertise on systems for automated essay scoring systems.

My work and thinking on topics, anaphora and discourse structure has been influenced a great deal by several visitors at Penn. I would like to extend my special thanks to Alistair Knott, Barbara Di Eugenio, Massimo Poesio, Julia Hirschberg, Jennifer Venditti, Jennifer Arnold, Bonnie Webber, Barbara Grosz and Candace Sidner. I am especially indebted to Anna Papafragou for insightful discussions on subordination and for her detailed comments on an earlier draft of my dissertation.

I have enjoyed immensely working with Cassie Creswell, Kate Forbes and Rashmi Prasad on the DLTAG project. Thank you for being such great team workers and good friends. Special thanks to Aravind Joshi and Bonnie Webber for being such inspiring leaders of this and other projects. I would also like to acknowledge all the participants of

the CLUNCH meetings for very stimulating discussions on computational issues as well as comments on my work. Special thanks to Tom Morton for discussions on pronoun resolution.

I gratefully acknowledge the School of Arts and Sciences, the Department of Linguistics, the Institute of Research in Cognitive Science, Aravind Joshi and Martha Palmer for their financial support during my graduate studies.

Finally, there are a few personal debts that I would like to express. First of all to Maria Angelou, who knew long before I did what my dreams would be and made sure they would all come true. Thank you for being an amazing friend, unparalleled mentor, and, above all, for believing in me and continuously raising my personal and professional standards.

My family back in Athens have also been an endless source of love and security for me. Heartfelt thanks to my parents Miltiadis and Maria Miltsakaki and to my brother Christos Miltsakakis for their continuous love and support.

My deepest thanks to my husband Kostas Daniilidis for being the extraordinary man that he is, for his rare integrity, intellectual depth, and deep respect for his principles, as well as for choosing to love me and support me in his unique unconditional way. I truly do not know what I have done to deserve you.

Last but not least, I would like to thank little Melina for her big smiles, tender hugs, and the countless restful nights. I couldn't have dreamed of a more supportive and loving baby.

ABSTRACT

THE SYNTAX-DISCOURSE INTERFACE: EFFECTS OF THE
MAIN-SUBORDINATE DISTINCTION ON ATTENTION STRUCTURE

Eleni Miltsakaki

Ellen Prince

Aravind Joshi

The central claim of this thesis is that, unlike main clauses, adjunct subordinate clauses do not form independent processing units in the computation of entity-based topic continuity (attention structure) in discourse. This claim has two primary consequences. First, discourse entities in adjunct subordinate clauses are assigned lower salience than main clauses entities, especially subjects. Second, the process that selects antecedents for pronouns in main clauses is qualitatively different from the process of anaphoric interpretation in subordinate clauses. The former is affected by the mechanism responsible for directing attention in discourse. The latter depends heavily on verbs semantics and the effect of connectives. The claims of this thesis are empirically tested for English and Greek. Primary evidence for the low salience of entities in adjunct subordinate clauses comes from corpus studies, which show that entities in adjunct subordinate clauses a) make poor competitors in the selection of antecedents for subject pronouns in main clauses, and b) are unlikely to be referred to in the subsequent discourse with a pronoun (unless they are already old). Primary evidence for the two level anaphora resolution mechanism comes from psycholinguistic experiments designed to test if there is a consistent difference in the way we interpret pronouns in main and adjunct subordinate clauses. These findings form the basis for the specification of two new NLP models, a system for the automated evaluation of coherence in student essays and a two level anaphora resolution algorithm. They also make two significant contributions to the Centering Model, a) Centering's "utterance" is formally defined on the basis of empirical evidence, and b) Centering's Rough-Shift transition is for the first time validated as a reliable estimator of poor coherence, empirically tested on an operable essay scoring system.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Thesis Statement

This dissertation is an investigation into attention management and related issues in discourse organization and the interpretation of anaphoric forms. Attention management is a discourse phenomenon too complex to attempt to address in a single project. Here, we have set ourselves more limited goals focusing on a set of related questions whose answers will hopefully shed some light on the poorly understood role of the syntactic main-subordinate distinction in attention management and anaphoric interpretation. For cross-linguistic comparison, the studies designed for the investigation of the topic of this dissertation are conducted in both English and Greek. We have chosen research methods across disciplines as we saw fit, i.e., psycholinguistic experimentation, corpus annotation, and computational evaluation.

In particular, this dissertation has set itself the following two main goals. First, we investigate the topicality status of entities in main and adjunct subordinate clauses and the implications of the findings for theories of pronominalization and anaphora resolution. For topic identification we use a specific algorithm based on Centering Theory. Second, based on our conclusions regarding topic management and its partial relationship to pronominalization, we build two computational models: one for the automatic evaluation of discourse

coherence and one for pronominal interpretation.

With respect to topic management, we show that topicality is computed across larger syntactic units than previously thought. Specifically, we argue that, for the purpose of topic continuity, adjunct subordinate clauses are processed as a single unit with the main clause to which they adjoin. This claim is justified by our studies on adverbial and relative clauses in English and Greek, which show that entities evoked in adjunct clauses a) do not obtain topical status by virtue of their *grammatical role* or *recency of mention*, and, relatedly, b) are less likely to be referenced in the subsequent discourse. To assess the role of adjoined subordinate clauses in topic continuity, we adopt a Centering-based measure of discourse coherence which enables quantification of the degree of connectivity in discourse. We have found that more coherent, and arguably easier to process, topic transitions are computed when the adjunct clause is processed as part of the unit containing the main clause. This result obtains even in discourses with no pronominal references, thus providing support for the applicability of Centering Theory as a model of local discourse coherence in pronoun-free text.

With respect to pronominal interpretation, we show that while it can trivially be maintained that pronouns refer to *accessible* entities, the notion of *accessibility* is not a homogeneous phenomenon. The homogeneous nature of pronominalization has been assumed by well known accessibility theories which associate pronouns with topical/salient entities and full NPs with less salient entities in a uniform model of anaphoric interpretation. We argue that anaphoric interpretation is not a uniform process. Based on our results of the interpretation of subject pronouns in main and adjunct clauses, we propose that pronouns in adjoined subordinate clauses are interpreted locally, primarily retrieving their antecedents from within the sentence in which they occur. Filtering out syntactic constraints, their interpretation is markedly sensitive to the semantics of the verb predicates and subordinate connectives. On the other hand, subject pronouns in main clauses are associated with the highest ranked entity in the previous sentence, the entity that is the topic of that sentence and often the most likely topic of the next sentence. For English and Greek this entity is

often the main clause subject, with the exception of certain subjects that are marked by the language for low or zero salience in the computation of topic continuity (e.g., impersonal pronouns). Entities evoked in the subordinate clauses of the sentence rank lower than their main clause counterparts, even if they have prominent grammatical roles (e.g., subjects) or are linearly closer to the pronoun.

The computational model for the evaluation of discourse coherence is based on Centering Theory. Drawing on our conclusions on the role of main-subordinate syntax in discourse organization, we define Centering's processing unit, i.e., the *utterance*, as the unit containing a single main clause and its associated adjunct subordinate clauses. A measure of topic continuity is then devised based on frequency of occurrence of Centering's Rough-Shift transitions. Rough-Shift transitions are shown to most reliably identify poor topic development in a corpus of student essays obtained from the Educational Testing Service (ETS) (Miltsakaki & Kukich, 2000b). The proposed model is tested on a currently operational system of electronic essay scoring developed at ETS.

Finally, a two level anaphora resolution algorithm is specified, based on our insights on the distinction between inter- and intrasentential anaphora as summarized above. We compare the merits of the proposed algorithm with other related algorithms and demonstrate how the new algorithm handles a number of problematic cases.

1.2 The Subject Matter of the Thesis

1.2.1 Attention Structure

Understanding attention structure in discourse is tantamount to understanding how the speaker communicates to the hearer which of the multiple entities evoked in discourse he selects to talk about and how he navigates the hearer's attention while retaining topic continuity, one of the factors contributing to the perceived coherence in discourse.

In previous work, Grosz and Sidner (1986) proposed that a component of discourse structure, responsible for "keeping a record of the objects, properties and relations that are

salient at any given point in discourse", is what they called the *attentional state*. Modeled as a stack of focus spaces, attentional state "restrains the availability of possible referents for definite descriptions and pronouns". This thesis is closely related to the study of the *attentional state* in discourse in that it contributes directly to the mapping of what Grosz and Sidner call *focus spaces* to linguistic units and the linguistic encoding utilized to determine the relevant salience of the entities which reside in it. In other words, the study of how speakers manipulate the main-subordinate syntax to organize discourse addresses the question of how often the hearer expects to look out for instructions to reassign saliences and how these instructions are realized linguistically.

Before proceeding any further a caveat to the reader is in order. An important assumption of this thesis is that a single entity is selected to be the topic at any given point in discourse, following e.g., (Reinhart, 1981), (Grosz, Joshi, & Weinstein, 1995). The assumption of a single topic receives support from empirical studies across languages. Most notably, Prince (1998) shows that certain instances of subject-drop in Yiddish refer to topics (as formalized in Centering Theory). Similar observations have been made for English (Walker, Joshi, & Prince, 1998), Japanese, e.g., (Kameyama, 1985) and (Walker, Iida, & Cote, 1994a), Turkish (Turan, 1995), Italian (Di Eugenio, 1998) and Greek (Miltsakaki, 2001), among other languages.

The notion of selecting a single entity to serve as the discourse topic has also received support in the computational literature from Joshi and Kuhn (1979). Joshi and Kuhn proposed an *almost* monadic calculus approach to discourse semantics. An *almost* monadic calculus approach to discourse can be achieved in "entity centered sentence representation" by singling out an entity among all the entities which are the arguments of the main predicate. In this way, an n-ary predicate can be made to look like monadic because its internal structure is temporarily hidden in the representation. As is well known, inferencing in monadic calculus is much easier than in full predicate calculus. Further, it can be shown that applying this representation facilitates natural language inferences.

The assumption that a single topic can be identified at any point is not universal, however. Some researchers have argued that it is possible to have two types of topics, or otherwise defined salient entities, at each time (Sidner, 1979). And yet other researchers, most notably Givón (1983), also Arnold (1998), view topicality as a graded phenomenon, i.e., all entities are topical to a greater or lesser degree.

An important terminological clarification is in order, too. In the linguistic literature, the term *topic* has created a great deal of confusion because the same term has been used to explain a variety of linguistic phenomena. Most prominently the term *topic* has been used to describe a) what the discourse is about, e.g., (Reinhart, 1981) and b) the part of the sentence that is old information, e.g., (Hajicova & Sgall, 2001)). We use the term to refer exclusively to the entity that a sentence is *about*. We do not use the term to refer to the part of the sentence that represents old information. On the other hand, defining the topic as the entity the sentence is *about* has its own problems. As Prince (1999) and other researchers have noted, well-known attempts to define the topic via association with certain positions in the sentence, i.e., subject or initial position, quickly run into circularity, e.g., (Halliday, 1967). A non-circular definition of what we perceive as the topic of a sentence is offered in Centering Theory. Centering theory defines the notion of the backward-looking center as the most salient entity in the preceding discourse that is realized in the current processing unit. By definition, the backward-looking center is the entity that provides a link between the previous and the current discourse.

In addition to the notion of backward-looking center, Centering Theory provides an exceptionally suitable framework for the investigation of attention structure. Centering posits that each processing unit evokes a set of entities which are ranked according to their salience status in the unit. The ranking rule is not specified as it may vary cross-linguistically. The highest ranked member of this set is assumed to be the most likely topic of the subsequent unit. Centering also defines four topic transitions based on the relationship between the backward-looking center and the highest ranked entity of the

processing unit. The four Centering transitions reflect four degrees of discourse connectivity. The most coherent transition is the Continue transition which is computed when the topic of the current unit is the same as the topic of the previous unit and is also the highest ranked entity in the current unit. The least coherent transition is called Rough-Shift and it is identified when the topic of the current unit is not the same as the topic of the previous unit and is not realized as the highest ranked entity in the current unit. For each unit a single topic is identified. Centering also defines the Pronoun rule which states that, if any entity is realized with a pronoun, then the topic is also realized as a pronoun.

The set of ranked entities is useful in two ways. First, specifying the ranking rule for a language help us identify the strategy that the speakers of a language adopt to mark some entities as the most likely topics in the subsequent discourse. It is for this reason that the highest ranked entity is, not surprisingly, called the "preferred center". At the same time, the ranking of the remaining entities may prove useful for the interpretation of other pronominal references within the processing unit, or across units when the subsequent unit contains more than one pronouns that are not resolved locally.

Centering's Pronoun rule recognizes the special status of topics in discourse but at the same time does not associate every instance of a pronominal reference with a topical entity. If a unit contains a single pronoun then it is predicted that the pronoun refers to the topic of that unit but additional pronominal references are recognized whose interpretation is not expected to relate to topic structure. Discourses may include pronominal references to non-topical entities. On the other hand, centering transitions are not dependent on the Pronoun rule, which means that they can be used to compute topic structure even in discourses that do not contain pronouns. As will be shown in Chapter 6, this property of the Centering framework will prove essential for the computation of discourse connectivity in discourses which do not contain pronouns.

The notion of the processing unit was left unspecified in Centering. The Centering processing unit is referred to as the *utterance*. Defining the size of the processing unit is crucial because it affects directly the identification of topics and the computation of

topic transitions. Assuming that main clauses form at minimum a single processing unit, the question arises as to what the status of subordinate clauses is in the computation of topic structure. This thesis addresses this question directly and provides the first empirical investigation of the role of adjunct subordinate clauses in topic management. We hypothesize that adjunct subordinate clauses are processed together with the main clause to which they adjoin. This claim is supported by the studies we have conducted. Our studies include primarily written texts. Spoken discourses are different in that they may contain clauses that are not common in written discourse, e.g., elliptical clauses, and allow prosodic strategies to be used for topic management. While further investigation into the complexities of spoken discourse is necessary, an obvious first step to take in that direction is the study of the role of subordinate clauses in written text. Our better understanding of topic management in carefully planned discourse will shed light on important aspects of linguistic encoding of attention in discourse and will better guide future research on spoken data.

1.2.2 Subordinate Clauses

The term "subordinate clause" covers a wide variety of clause types, ranging from non-finite verb forms such as infinitives and participles, to complement clauses, free relatives, conditionals, relatives, and adverbial clauses. In this thesis, we focus on tensed adjunct subordinate clauses, specifically, adverbial and relative clauses. As a shortcut, we may refer to this group as simply subordinate clauses but the reader should be aware that all claims in this dissertation pertain solely to adverbial and relative clauses.

Tensed subordinate clauses are formally distinguished from main clauses. In English, they are introduced by a finite set of subordinate conjunctions and they adjoin to higher syntactic constituents. Adverbial clauses may be introduced with temporal, causal, and concessive conjunctions and they can be preposed relative to the main clause to which they attach, e.g., (1) and (2). Relative clauses modify primarily nominal phrases, e.g., (3), but they may also modify clausal constituents, as in (4) and they may or may not be a

constituent of the noun phrase that they modify, as in cases of extraposition, shown in (5).

(1) Mary is looking for a job because she needs money.

(2) Because she needs money, Mary is looking for a job.

(3) Mary is looking for a job which will not bore her.

(4) Mary found a new job, which made me very happy.

(5) The plumber arrived who we had called earlier.

An important class of tensed subordinate clauses that has been excluded from the current study is complement clauses. Unlike adjunct subordinate clauses, complement clauses serve as arguments of certain verb classes (e.g., verbs of *saying*, *knowing*, *belief*, *doubt*, etc.) and occupy the relevant argument position in the syntactic structure. The study of complement clauses is important for any theory of attention structure. This is because complement clauses seem to open up new discourses that may span across multiple clauses in the subsequent discourse. In (6), for example, the discourse (6b)-(6f) is associated with the complement clause in (6a), thus shifting the attention of the hearer to *Mary*. This is in contrast to (7). In (7), the subsequent discourse (7b)-(7f) is associated with the higher clause thus retaining *John* as the topical entity in the discourse.

(6) John said that

a. Mary could not come to the party.

b. She had prior engagements

c. which would hold her until late.

d. She tried to reschedule her appointments

e. but she was not successful

f. because her clients were very impatient.

(7) a. John said that Mary could not come to the party.

b. He decided to hold the party anyway.

c. He invited a lot of people

d. who accepted his invitation.

e. However, he was still unhappy

f. because his guest of honor would not be there.

Therefore, when processing discourse following a complement clause, one has to identify first whether the subsequent discourse is associated with the complement or the higher clause. This ambiguity is also present when the complement clause contains another subordinate clause. In (8) for example, the *because* clause is associated with the higher verb in the main clause while in (9) it is associated with the complement clause. [1]

(8) John said that Mary could not come to the party because he was embarrassed to admit that he hadn't invited her.

(9) John said that Mary could not come to the party because she was sick.

It would also be interesting to look at how topic continuity is established in discourses associated with the complement clause, what the attentional properties of such discourses are and how or how often a return to a topic introduced in the higher clause is realized. We suspect that once complement clauses open up a new discourse, they probably behave as main clauses. A local computation of topic continuity would then apply until the end of the subdiscourse is signalled, possibly through a mechanism similar to Grosz and

[1]Note that, in some cases, there is true ambiguity which is hard to be resolved even by the hearer, e.g., (1). Example (1) is ambiguous between analyses (2) and (3).

(1) Moritz said Monday his leg feels fine and , as a result , he hopes to start practicing field goals this week .

(2) Moritz said Monday [that his leg feels fine and, as a result, he hopes to start practicing field goals next week.]

(3) [Moritz said Monday his leg feels fine] [and, as a result, he hopes to start practicing field goals this week.]

Sidner's focus pops which signal return to a higher discourse segment. Preliminary indications that complement clauses are sensitive to discourse phenomena associated with main clauses rather than adverbial or relative clauses come from Prince (1988), who has found in her study of postposed subjects in Yiddish that the distribution of brand-new subjects in complement clauses is similar to the distribution of brand-new subjects in main clauses. Also, in their analysis of verb ellipsis, Romero and Hardt (2003) (also personal communication), have observed that the antecedent of verb ellipsis following a main and a subordinate clause is located in the main clause whereas the antecedent of verb ellipsis following a complement clause is located in the complement clause itself. The relevant examples are given in (10) and (11). This line of investigation, however, as well as the details of the salience status of entities evoked in complement clauses, lies beyond the scope of this dissertation.

(10) Agnes arrived after John ate. But Bill didn't *(eat)/(arrive after John ate).

(11) Agnes said that John ate. But Bill didn't (eat)/*(say that John ate).

1.3 Methodology and Contributions

The topic of this dissertation is the effect of the main-subordinate distinction on attention structure. The overall goal is to understand how speakers manage topic continuations and shifts while maintaining coherence. The problem of attention structure and associated issues (e.g., pronominal interpretation, discourse structure, information structure, etc.) is too broad and complicated to be studied at this level of generalization. For this reason, we have approached it in terms of more specific questions that we can investigate and define with reasonable accuracy. For the investigation of these subproblems, we have chosen research methods across disciplines as we saw fit, i.e., psycholinguistic experimentation, corpus annotation, and computational evaluation.

We **start** with the crucial assumption that a significant component of discourse organization is topic management and that each discourse unit has a single topic. We acknowledge that the relative salience of entities can be defined on a grading scale. We **conjecture** that main and subordinate clauses form a single attention (or topic) update unit. The entities which are evoked in each unit are ranked according to their grammatical role in English and Greek (or possibly other salience factors in other languages). A crucial consequence of our conjecture is that entities evoked in subordinate clauses rank lower than entities evoked in main clauses. It also follows that subjecthood is not a uniform measure of salience, as subjects of main clauses rank higher than subjects of subordinate clauses. Given that pronouns appear in both main and subordinate clauses and that each attention unit has a single topic, it follows that only a subset of pronouns refer to topical entities. Therefore in our framework, well-known accessibility hierarchies are relativized to the notion of attention update unit. Given the syntactic locality defined by subordination, pronouns in subordinate clauses almost by definition refer to *accessible* entities but not necessarily to *topical* entities. Given that discourse coherence is attained even in pronominal free text (e.g., newspaper and scientific articles, cf. Chapter 6 of this dissertation), it follows that pronominalization is not the sole factor for the evaluation of entity-based coherence and topic continuity.

We **hypothesize** that:

1. The interpretation of subject pronouns in subordinate clauses differs from the interpretation of subject pronouns in main clauses. Intersententially, pronouns are resolved according to the mechanism responsible for topic structure, which marks with high salience entities appearing in structurally prominent positions, e.g., main clause subjects. We also formulate the hypothesis that intrasententially pronouns are interpreted locally, most likely according to preferences projected by the verbs and subordinate conjunctions.

2. Entities evoked in subordinate clauses are less likely to be referenced in the subsequent discourse than entities evoked in main clauses.

3. Entities evoked in subordinate clauses are more likely to be referenced with a full NP in the subsequent discourse, unless they already appear pronominalized in the subordinate clause.

4. Centering Theory provides a suitable model for the evaluation of local discourse coherence even in text that does not utilize pronominal references to signal topic continuity. The Centering Model is capable of capturing sources of low coherence in discourse due to missing links between discourse units.

To test the above hypotheses we perform the following studies:

First, in controlled experimental conditions, we establish that pronouns in main clauses are interpreted differently from pronouns in adverbial clauses. This result obtains for both English and Greek. In particular, a main clause subject is consistently interpreted as the subject of the preceding main clause while the interpretation of a subordinate subject varies. This result is consistent with the hypothesis that pronoun interpretation intersententially, but not intrasententially, is determined by structural salience.

Second, we provide corpus-based support for the experimental findings. We extract sets of main-main clauses and main-adverbial clauses from a Greek corpus, which contain a pronoun (dropped subject or weak pronoun for Greek) in the second clause and two morphologically competing antecedents in the first. Pronominal forms most consistently refer to the structurally highest ranked entity in the preceding unit in the main-main condition but not in the main-subordinate condition.

Third, we perform a series of corpus studies for relative clauses in English (introduced by *who, which, that*) and Greek (introduced by *opoios* 'who/which' and *pou* 'that'). In a series of corpus annotation studies we annotate formal features of relative clauses (definiteness/indefiniteness, restrictive/non-restrictive). Crucially, we code for whether the referent of the head noun and other referents evoked in the relative clauses are referenced in the subsequent discourse, as well as the type of referring expression used. The results of these studies show that the referent of the head noun is subsequently referenced with a pronoun when it is the highest ranked entity in the main clause that contains it, i.e., before

processing the relative clause. With regard to other entities evoked in relative clauses, they are less likely to be referenced. In almost all the cases where they *are* referenced, a full noun phrase is used, indicating that their position in the relative clause is not sufficient to establish them as topical. The results from these studies confirm our hypothesis that entities in adjunct subordinate clauses are less salient than entities in main clauses and that topics are established primarily in main clauses. Entities evoked in relative clauses may become topical in the subsequent discourse but they first have to be promoted to a prominent structural position, such as the main clause subject position. This frequently happens via repetition of the relevant entity in the subject position of the subsequent sentence with a full noun phrase.

Fourth, to investigate whether subordinate clauses form an independent attention update unit, we conduct two studies in English and Greek based on a Centering-based metric of topic continuity. We focus on non-restrictive relative clauses, which in the syntactic literature have been claimed to behave as main clauses. In the two aforementioned studies we compute Centering transitions in two conditions. In the first condition, relative clauses are processed together with the main clause they are associated with. In the second condition, relative clauses are processed as independent attention update units. Comparing the transition type computed after the relative clause is processed in the two conditions yields significantly less coherent transitions in the second condition than in the first condition. Assuming that the authors of the texts in the corpus opt for high coherence, these results indicate that relative clauses belong to the same unit as the main clause. In many cases, a subsequent pronominal reference resolved to an entity in the main clause skipping competing entities evoked in the relative clause, indicating that entities evoked in relative clauses do not affect topicality. These studies offer empirical evidence for the definition of Centering's "utterance", a crucial part of Centering Theory that has been left unspecified. Our conclusions so far are the following: a) topic management involves the use of pronouns for reference to topical entities but this process is distinct from anaphora resolution at large and b) the unit for computing attention updates is best defined at the sentence and

not at the clause level.

Fifth, we present an ideal corpus for the computational evaluation of the resulting model of attention management. The corpus consists of a set of student essays written for the writing section of the GMAT examination. These essays were scored by both humans and *e-rater*, an electronic essay scoring system developed at ETS. Using a Centering-based metric of discoure coherence, we tested its contribution on the performance of *e-rater*. Our positive results show that our proposed model for topic management articulated in the Centering framework is successful in improving the evaluation of the essays. This study makes two important contributions to the Centering model. It validates the Rough-Shift transition which was unattested in previous Centering-based research. Given that Centering's empirical testability relied only on the Pronoun rule, this study provides empirical support for the theory for cases that do not involve pronominal references. Such were the cases of the attested Rough-Shift transitions.

Finally, based on the results from this thesis, we are able to propose a two level anaphora resolution algorithm. Intrasententially, the algorithm opts to resolve pronouns locally according to the semantic focusing properties of verbs and connectives. Intersententially, the algorithm opts to resolve pronouns according to structural salience which reflects the topic structure of the text. The merits of the proposed algorithm are compared and contrasted to related algorithms.

1.4 Outline of the Thesis

The thesis is organized as follows. Chapter 2 provides a brief overview of Grosz and Sidner's (1986) approach to modeling discourse structure with special emphasis on the component of attentional state. Next, the basic concepts of the Centering model are reviewed in some detail setting the theoretical background for the methodology adopted for the study of subordinate clauses. The remainder of Chapter 2 contains an overview of related approaches to salience in discourse, topichood, and anaphoric interpretation.

Chapter 3 demonstrates the puzzle posed for theories of salience and anaphoric interpretation. A number of anaphora resolution algorithms based on current theories are discussed in detail. Crucial insights from their shortcomings are highlighted. In this chapter we offer a possible explanation and propose a new discourse processing model. The proposed model crucially relies on the hypothesis that subordinate clauses do not form independent center update units in topic management.

Chapters 4 and 5 are devoted to the investigation of adjunct subordinate clauses, adverbial clauses, and relative clauses, respectively. Chapter 4 reports the results of three experimental and one corpus study for English and Greek. The studies are designed to test the following two hypotheses: a) subject pronouns in main clauses pick their referents from the previous discourse via structure-driven assignment of salience (e.g., subjecthood), whereas subject pronouns in adverbial clauses pick their referents locally in accordance with the focusing properties of the matrix predicate and the semantics of the subordinate conjunction; b) entities introduced in sentence final adverbial clauses are dispreferred as antecedents for subject pronouns in subsequent main clauses. Chapter 5 reports corpus based results of eight studies in English and Greek designed to evaluate the salience status of entities evoked in relative clauses. Special emphasis is given to non-restrictive subordinate clauses which in the syntactic literature are often treated on a par with main clauses. We also take a close look at the distinction between restricting relative clause with a definite versus indefinite head.

Chapter 6 reports a Centering study for the evaluation of coherence in students' essays. Assuming that the center update unit contains a single main clause and all its dependent subordinate clauses, we propose a Centering-based metric for the evaluation of textual coherence. The proposed metric is evaluated on *e-rater*, an automated essay scoring system developed by the NLP group at ETS. The results of this study show that Centering's Rough-Shift transition is a good indicator of coherence in student's essays. We discuss the role of Centering's Rough-Shift transition in capturing a source of incoherence, poor topic development, and abrupt topic shifts, that is captured even when the text contains

no pronouns. The crucial insight is that Centering captures low textual coherence due to the lack of topic connectivity which makes it harder for the reader to establish links between the current and previous discourse. This finding offers additional support for our claim that topic management is a process distinct from (albeit overlapping with) anaphora resolution.

In Chapter 7, we specify a new anaphora resolution algorithm based on the findings of the thesis. The proposed anaphora resolution algorithm interleaves two mechanisms: the mechanism responsible for resolving pronouns that refer to topical entities (intersententially) and the mechanism responsible for resolving other pronouns (intrasententially). The proposed algorithm is compared with other related algorithms, specifically those proposed by Hobbs (1978) and Lappin and Leass (1994). We conclude in Chapter 8.

Chapter 2

Related Work

2.1 Discourse Structure, Coherence, and Reference

2.1.1 Grosz and Sidner

According to Grosz and Sidner (1986), discourse structure is composed of three components: the linguistic structure, the intentional structure and the attentional state. The attentional state is a record of the objects, properties and relations that are salient at any given point in the discourse. It is modeled by a set of focus spaces, the available collection of which at each time is the focusing structure. The focusing structure is a stack. The stacking of focus spaces reflects the relative salience of the entities in each space during the corresponding segment's portion of discourse. Each focus space can also be thought of as a repository of the contextual information needed to process utterances at each point in discourse. On one hand, its primary role is to constrain the discourse segment purposes relating to the current processing segment. On the other hand, the focusing structure constrains the possible referents of definite descriptions and pronouns.

Grosz and Sidner laid out the components of a theory of discourse structure providing thus a solid basis for further investigation of its components. However, several basic questions have remained open. With regard to the attentional state, for example, a model

of defining, identifying and ranking entities remain to be fully specified. The notion of focus space is, also, elusive. What is a focus space and how is it identified? Is the focus space equivalent to an abstract segment associated with a discourse purpose or is it an attentional update unit?

A first attempt to model aspects of attentional structure yielded a reformulation of Centering as a model of local discourse coherence, to which we now turn.

2.1.2 Centering Theory

Centering was developed as a model of the center of attention between speakers in natural language discourse. The model aimed at modeling the interaction between "attentional state", inferential complexity and the form of referring expressions. The formulation of Centering Theory resulted from the synthesis of two main lines of work. Originally, Joshi, Kuhn, and Weinstein (Joshi & Kuhn, 1979; Joshi & Weinstein, 1981) proposed Centering as model of the complexity of inferencing involved in discourse when speakers process the meaning of an utterance and integrate it into the meaning of the previous discourse. Grosz and Sidner (Sidner, 1979; Grosz, 1977; Grosz & Sidner, 1986) recognized what they called the "attentional state" as a basic component of discourse structure and proposed that it consisted of two levels of focusing: global and local. A synthesis of these two approaches yielded the Centering model for monitoring local focus of attention, which was designed to account for those aspects of processing that are responsible for the difference in the perceived coherence of discourses as those demonstrated in (12) and (13) below (examples from (Grosz et al., 1995)).

(12) a. John went to his favorite music store to buy a piano.

b. He had frequented the store for many years.

c. He was excited that he could finally buy a piano.

d. He arrived just as the store was closing for the day.

(13) a. John went to his favorite music store to buy a piano.

b. It was a store John had frequented for many years.

c. He was excited that he could finally buy a piano.

d. It was closing just as John arrived.

Discourse (12) is intuitively more coherent than discourse (13). This difference may be seen to arise from the different degrees of continuity in what the discourse is about. Discourse (12) centers a single individual, *John*, whereas discourse (13) seems to center in and out on different entities, *John, store, John, store*. Centering is designed to capture these fluctuations in continuity.

Contra earlier assumptions based on purely semantic or inferential theories of discourse understanding (Hobbs, 1985), Centering also predicts that discourses (14) and (15) below differ in coherence despite the fact that there is no semantic ambiguity at the time the discourses are fully processed.

(14) a. Jeff helped Dick wash the car.

b. He washed the windows and Dick waxed the car.

c. He soaped a pane.

(15) a. Jeff helped Dick wash the car.

b. He washed the windows and Dick waxed the car.

c. He buffed the hood.

The pronominal subject in (15c) can be interpreted only as *Dick* because the semantics of *buffing* is associated with the *waxing* event. Still, by using a pronoun in (15c), the speaker is only confusing the reader because up to utterance (15a) *Jeff* has been the center of attention and therefore the most likely referent of the pronoun in (15c). It is only when the hearer gets to the word *buff* that s/he realizes that the referent must be *Dick*.

In what follows we present the basic concepts and data structures of the model to demonstrate how Centering evaluates discourse coherence and its interaction with choice of referring expression.

According to Centering, discourse consists of a sequence of textual segments and each segment consists of a sequence of utterances. Utterances are designated by $U_i - U_n$. Each utterance U_i evokes a *set* of discourse entities, the Forward-looking Centers, designated

	Cb(U_i)=Cb(U_{i-1})	Cb(U_i)≠Cb(U_{i-1})
Cb(U_i)=Cp	Continue	Smooth-shift
Cb(U_i)≠Cp	Retain	Rough-shift

Table 2.1: Table of Centering transitions

by $Cf(U_i)$. The members of the Cf set are ranked according to discourse salience. The highest-ranked member of the Cf set is the Preferred Center, Cp. A Backward-looking Center, Cb, is also identified for utterance U_i. The highest ranked entity in the previous utterance, $Cf(U_{i-1})$, that is *realized* in the current utterance, U_i, is its designated Backward-looking Center, Cb. The Backward-looking Center is a special member of the Cf set because it presumably represents the discourse entity that U_i is about, what in the literature is often called the "topic" (Reinhart, 1981; Horn, 1986).

The Cp for a given utterance may be identical with its Cb, but not necessarily so. This distinction between looking back in the discourse with the Cb and projecting preferences for interpretations in the subsequent discourse with the Cp is the key element in computing local coherence in discourse within the Centering framework.

Centering rules and transitions. Since Centering is designed to model attentional state, it follows that it also defines changes or shifts in attention. Four transitions from one attentional state to another are defined which reflect four degrees of center continuity: Continue, Retain, Smooth-Shift, and Rough-Shift. The rules for computing the transitions between two adjacent utterances are shown in Table 2.1. They correspond to the four combinations of two variables: whether the "topic" of the current utterance, i.e., Cb(U_i), is the same as the "topic" of the previous utterance, i.e., Cb(U_{i-1}), and whether the "topic" of the current utterance, Cb(U_i), is realized in a position saved for salient entities, Cp(U_i), the highest ranked entity in the Cf set. In English, for example, that position has been argued to be the subject position. Finally, Centering transitions are ordered according to degree of coherence as defined in the Transition Ordering rule, shown below.

Transition Ordering Rule:
Continue is preferred to Retain, which is preferred to Smooth-shift, which is preferred to Rough-shift.

Centering, also, defines a rule, known as the **Pronoun Rule**, which constrains the choice of referring expression under certain conditions and at the same time makes a testabe prediction for the theory:

> **Pronoun Rule:**
> If some element of the Cf of the previous utterance is realized as a pronoun in the current utterance,
> then so is the Cb of the current utterance.

The Pronoun Rule captures the intuition that pronominalization is one way to indicate discourse salience and that Backward-looking centers are often deleted or pronominalized. Later studies in pro-drop languages like Japanese (Kameyama, 1985) or Turkish (Turan, 1995) showed that the Pronoun Rule for such languages must be reformulated to accommodate zero pronouns: If some element of the Cf of the previous utterance is realized as a zero pronoun in the current utterance, then so is the Cb of the current utterance.

The Pronoun Rule and the Centering Transitions predict that the interpretations that hearers will prefer when processing discourse are those requiring minimal processing effort. For example, an instance of a Continue transition followed by another Continue transition requires minimal effort for interpretation, as the hearer only needs to keep track of one main entity which is both the Cb and the Cp of the current utterance. Below, we demonstrate how the Centering Rules apply to discourses (14) and (15), shown in Table 2.2.

Utterance (b) is a Continue transition because the Cb is the same as in (a) and the Cp in (b) is the same as the Cb in (b), namely *Jeff*. In contrast, (c') is a Smooth-shift transition, because the Cb has changed from (b), but the Cp is the same as the Cb. According to the Centering Model, the discourse with the (c') continuation is less coherent that the one with (c). The Continue transition identified in the (b) utterances is interpreted as an indication by the speaker that s/he intends to Continue talking about *Jeff*. Instead, the speaker, shifts attention (with a Smooth-shift transition) to *Dick*. This is misleading for the hearer who first interprets the pronoun *he* in (c') as the Cp of the previous utterance (cf the Pronoun Rule) and then has to revise this interpetation. Hudson-D'Zmura and Tanenhaus (1998)

a. Jeff helped Dick wash the car.	a. Jeff helped Dick wash the car.
Cb=none Cf=Jeff,Dick, car Transition=none	Cb=none Cf=Jeff,Dick, car Transition=none
b. He washed the windows and Dick waxed the car.	b. He washed the windows and Dick waxed the car.
Cb=**Jeff** Cf=Jeff,windows, Dick, car Transition= Continue	Cb=**Jeff** Cf=Jeff,windows, Dick, car Transition= Continue
c. He soaped the pane.	c'. He buffed the hood.
Cb=**Jeff** Cf=Jeff Transition= Continue	Cb=**Dick** Cf=Dick Transition= Smooth-shift

Table 2.2: Sample discourses

and Walker et al. (1998) show that this corresponds to both an increase in processing time and an increase in subjects' judgment that the discourse with the (c') continuation does not make sense.

2.2 Topics, Subjects, and "Topicalization"

2.2.1 Prince

The view that the current thesis takes on the relationship between topics and what in the syntactic literature has been characterized as "topicalization" follows on work of Prince (1999). "Topicalization" is the term used to describe constructions with a non-canonical object-initial form, containing a gap in the canonical position of the object in English (an SVO language). An example of "topicalization", shown in boldface, is given below, taken from (Prince, 1999):

(16) Thanks to all who answered my note asking about gloves. I didn't look at this bb for serveral days and was astounded that there were 11 answers. **Some I missed**, darn. Don't know if I'm brave enough to work gloveless.

Assuming the topic to be the entity in the center of attention, Prince takes Centering's definition of Cb as the definition of topic to investigate if indeed "topicalized" entities are also topics. Prince examines a set of naturally-occurring data and shows that, in English OVS sentences, the leftmost constituent is *not* typically the topic, or Centering's Cb, of the relevant unit. The same conclusion is reached by applying other tests for topichood that have been proposed in the literature, e.g., (Gundel, 1974), (Gundel, 1985), (Reinhart, 1981). Further, Prince observes that in the examples of the corpus of "topicalizations" collected by Gregory Ward, which contained a third person pronoun, there were no instances of pronominal reference to the "topicalized", leftmost constituent but many to the subject. In Centering terms, the leftmost constituent also did not represent the Preferred Center, Cp. Prince then goes on to a detailed investigation of the discourse function of fronted constituents in English (and Yiddish) and shows that English "topicalizations" a) trigger an inference on the part of the hearer that the entity represented by the fronted constituent stands in a salient partially ordered set relation to some entity or entities already evoked in the discourse model, and b) triggers the inference that the proposition is to be structured into a focus and focus-frame. Similar but more general discourse functions have also been identified for Yiddish.

This thesis adopts the notion of topic as explicated in the work of Prince and the Centering framework. Further, Prince's work on the distinction between the notion of topic and the functions of "topicalization", which is essentially an information structure phenomenon, is supportive of our claim that topic structure in discourse is, possibly, orthogonal to information structure (the details of our proposal are given in Chapter 3).

Inevitably, topics are related to both information status and, at least for a number of languages, subjecthood. In her work on the taxonomy of given-new information, Prince (1981a) shows that "evoked" entities tend to appear in subject positions much more frequently than in non-subject positions, whereas "new" entities appear almost categorically in non-subject positions. Subjects, also, tend to be represented by definite descriptions, but as Prince (1992) shows this is because subjects tend to be discourse-old entities. So,

while topics naturally correlate with subjects (for languages which mark subjects as the most salient entity) and discourse-old entities (to establish an entity as the topic, you first have to evoke that entity), they are not *defined* by either of the two.

2.2.2 Chafe

Chafe's (1976) view on topics and subjects is quite different from what we have discussed so far. Chafe defines the "subject" as "what we are talking about", its main function being "adding-knowledge-about" its referent. He then investigates the realization of "subject" in various languages. In English, Chafe's "subject" coincides with the grammatical subject, whereas Dakota marks the "subject" on the agent role. Further, he claims that, in Dakota as well as in other languages, subjecthood might be expressed in alternative ways, e.g., word order. Clearly, in Chafe's view, the term "subject" is defined on functional terms and it seems to correspond to what elsewhere in the literature has been understood as the "topic". Conversely, Chafe seems to define "topic" on surface syntactic grounds, as the fronted constituent in "topicalization". He claims that such positions in English serve to represent the focus of contrast. On the other hand, in Chinese, or other "topic prominent" languages, the function of "topics" is to "limit the applicability of the main predication to a certain domain", or in yet other languages, such as Caddo, they behave as a premature "subject". So, according to Chafe, "subject" is a functional category which may have different realizations in different languages. "Topics", on the other hand, are defined syntactically for each language and they may have different functions in different languages. The terminological choices that Chafe has made and the definitions he provides are confusing. Defining linguistic forms on functional grounds renders research on the interpretation of linguistic form impossible because linguistic form is not defined independently. Further, as Prince has extensively argued in a series of corpus investigations, e.g., (Prince, 1998), a one-to-one mapping of linguistic form to discource function cannot be maintained, as the same discourse function can be expressed with more than a single syntactic option and a single linguistic form may serve more than one function. Even if

we treated Chafe's "subject" as a mere terminological issue and equated it to our notion of "topic", the "aboutness" definition would still leave us short of an objective characterization of discourse entities as "subjects" in Chafe's terms, or topics in ours.

2.3 Theories of Salience and Referent Accessibility

2.3.1 Topicality as a Continuum

2.3.1.1 Givón

Givón does not distinguish between topical and non-topical entities. He proposes that all entities are topical to a higher or lower degree. In his framework topicality is seen as a continuum rather than a discreet notion. The following scale of topicality has been proposed, in which zero anaphors refer to the most topical entities and, at the other end of the scale, referential indefinite NPs refer to the least topical entities.

Givón's *Scale of Topicality*

Most continuous/accessible topic

1. zero anaphor
2. unstressed/bound pronouns or grammatical agreement
3. stressed/independent pronouns
4. right dislocated definite NPs
5. neutrally ordered definite NPs
6. left dislocated definite NPS
7. Y-moved NPs ("contrastive topicalization")
8. cleft/focus constructions
9. referential indefinite NPs
10. discontinuous/inaccessible topic
11. Most discontinuous/inaccessible topic

Givón recognizes that the above scale of topicality is too language-specific, as it overlooks other means of coding topicality attested in languages, such as word-order, morphology, intonation and phonological size. Still, to support his claim that the syntactic coding of topic identification obeys basic principles of *iconicity* in language he, then, proposes the following generalized scale of topic continuity that underlies the *grammar* of topic identification across languages. Again, at one end of the scale, zero anaphora marks the most topical entities and full NPs the least topical entities.

Givón's grammar of topic identification

1. zero anaphora
2. unstressed/bound pronouns ("agreement")
3. stressed/independent pronouns
4. full NPs

Forms of reference are correlated with three discourse measures of topicality: a) referential distance (how recently the entity has been mentioned), b) potential interference (how many other potential antecedents of the referring forms there are), and c) persistence (how long the entity will remain in the discourse). These measures have been applied in various studies to identify statistical correlations with form of reference.

Referential distance is counted with respect to the number of clauses intervening between the current reference to an entity and the most recent reference to the same entity in the previous discourse. This measure of topicality is not supported by the studies in this dissertation. In Chapters 4 and 5, we see that entities introduced in adjunct subordinate clauses do not override the salience of the entities introduced in the main clause. Crucially, in many cases we have seen a pronoun in a main clause finding its antecedent in the preceding main clause skipping competing antecedents evoked in an intervening relative clause. Relatedly, in Chapter 4, we see that not all pronouns refer to topical entities. Pronouns in adjunct subordinate clauses tend to find their antecedents locally, i.e., within the boundaries of the sentence. The anaphora resolution model that we propose involves two

mechanisms: one applying intrasententially and one applying intersententially. According to our model, the interpretation of pronouns that appear intrasententially is affected by the semantics of the verbs and the connectives. It is therefore possible that a pronoun in such positions will not resolve to a topical entity. Pronouns in main clauses however, especially those in subject position, tend to resolve to topical entities. If our claims are correct, then the statistical correlations that Givón reports between linguistic form and referential distance are not relevant to the form-topicality mapping. As our model suggests, if a pronoun appears in a subordinate clause, it will be interpreted locally. So, the referential distance will be short but the pronoun may not resolve to the topic but to some other entity introduced in the sentence. If a pronoun appears in a sentence-initial main clause and the immediately preceding discourse is a main clause, then again the antecedent will be found at a short distance. The crucial test for the importance of referential distance and its association with topichood and pronominalization is found in cases where the preceding discourse includes subordinate clauses that evoke entities other than what we identify as the topic. In such cases, as we will see, the intervening competing antecedents and the distance created by the presence of subordinate clauses do not obstruct resolution to the main clause subject, for example, which often represents the topic. Such cases are crucial but are relevant to a small set of pronouns and are not likely to be captured by correlation measures of distance and form across all pronominal expressions. Our data, therefore, which include pronouns that resolve to topical entities and others that resolve to non-topical entities do not support the *iconicity* hypothesis for topic identification, according to which topic identification is syntactically coded in phonologically reduced forms in grammar.

"Interference" as one of the measures of topicality indicates that Givón is moving away from the notion of "topicality" as traditionally conceived and going closer to the notion of referent accessibility as depicted by accessibility hierarchies. The topicality of an entity is intuitively understood as a characteristic of the role of an entity in the discourse. It is not clear how "interference" from other entities should affect the topicality of an entity. As

Arnold (1998) has pointed out, "interference from other entities is only relevant insofar as it may hinder the interpretation of referring forms". We will argue that a more useful way of exploiting interference is to treat it as a test for the topical status of an entity. Successful pronominal reference to an entity that was introduced before "interfering" material provides additional evidence for the topical status of such an entity. This is in contrast with Givón's conception of the role of interference. In his account, the more interference is attested in the discourse, the less topical an entity preceding the interfering entities is and therefore Givón predicts that in such cases the form of referring expression for such an entity will be chosen from the lower end of the topicality scale (i.e., with a full NP or stressed/independent pronoun).

Givón's measure of "persistence" is in line with the view of topicality as a scalar notion. It depicts the lifetime of an entity in discourse and yields a measure of more or less topicality by virtue of the number of times that an entity is referenced in the discourse. While we also hold that subsequent reference to an entity is relevant in identifying topic continuity in discourse, the number of times that an entity is referenced is only relevant for the identification of topic structure in discourse. At any point in processing discourse an entity is either topical or not. Topic switches are recognized as shifts of attention to a new topic. Such shifts, however, do not make previous entities *less* topical. A short-lived topic is as good a topic as a long-lived one. Abrupt shifts to a new topic, with no links to the previous discourse may render a discourse more or less *coherent*.

2.3.2 Accessibility Hierarchies

2.3.2.1 Gundel et al.

Gundel, Hedberg, and Zacharski (1993) have proposed that there are six cognitive statuses that are related to the form of referring expression and these statuses are implicationally related in the *Givenness Hierarchy*, shown below.

in focus > *activated* > *familiar* > *uniquely identifiable* > *referential* > *type identifiable*

it > *that, this, this N* > *that N* > *the N* > *indefinite* this *N* > *a N*

Each status on the hierarchy is a necessary and sufficient condition for the appropriate use of a linguistic form. The corresponding English forms are given below the givenness scale, with a pronoun at the *in focus* end of the scale and an indefinite noun phrase at the *uniquely identifiable* end of the scale.

Of special concern to us is the category *in focus*, which provides the necessary and sufficient conditions for the use of a pronoun. According to Gundel et al., the referent of an entity *in focus* is in the short term memory (as are *activated* entities), and in addition it is also in the current center of attention. "The entities in focus at a given point in the discourse will be that partially-ordered subset of activated entities which are likely to be continued as topics of subsequent utterances. Thus, entities in focus generally include at least the topic of the preceding utterance, as well as any still relevant higher-order topics." By the above definition, it seems that entities in focus partially overlap with Centering's list of forward-looking centers. It is only a partial overlap because, according to the definition, only a set of entities evoked in a clause may claim *in focus* status. According to Gundel et al., subject and object positions bring entities *in focus* but entities evoked in prepositional phrases, for example, do not obtain *in focus* status and therefore cannot be referenced with a pronoun. Since both topical and non-topical entities can be *in focus*, Gundel et al.'s hierarchy correctly predicts that both topical and non-topical entities can be referenced with a pronoun. However, the givenness hierarchy cannot account for the fact that the pronoun in the following two discourses receives a different interpretation.

(17) John$_i$ criticized George$_j$ because he$_j$...

(18) John$_i$ criticized George$_j$. Then, he$_i$...

With regard to subordinate clauses, Gundel and her collaborators recognize that, like PPs, subordinate clauses do not bring their entities in focus. Our results are consistent with this observation. However, as we will see, the notion of the center update unit missing in Gundel et al.'s account is an important factor in the use and interpretation of pronouns. Consistent with Gundel et al.'s observation about subordinate clauses, we, too, claim that

entities in subordinate clauses are less salient than entities in main clauses. Instead of creating a list of *in focus* entities which includes only as subset of the evoked entities, we adopt Centering's notion of the Cf set, which includes all the entities evoked in the relevant unit. The advantage of this approach is that the use of a pronoun for reference to an entity appearing in a subordinate clause (presumambly not *in focus*) is expected if, for example, an entity already appears pronominalized in the subordinate clause. We have found several such cases in our corpus, e.g., (19). Also, for languages that rank animate referents higher that inanimate referents, e.g., Turkish (Turan, 1995), the use of a pronoun for reference to an entity evoked in a subordinate clause can be accounted for if the only animate referent of the unit appears in the subordinate clause.

(19) a. Indeed, apart from the nature of the investigation which **my friend** had on hand, there was something in **his** masterly grasp of a situation, and **his** keen, incisive reasoning, which made it a pleasure to me to study **his** system of work, and to follow the quick, subtle methods by which **he** disentangled the most inextricable mysteries.

b. So accustomed was I to **his** invariable success that the very possibility of **his** failing had ceased to enter into my head.

Finally, for the givenness hierarchy to correctly predict the distribution of forms in a language, other factors of discourse organization, in addition to the notion of center update unit, must be taken into account. One such factor, for example, is the use of a full NP to mark a segment or tense boundary, e.g., (20), even when Gundel et al.'s necessary and sufficient conditions for the use of a pronoun are met. Further, an NP is also used when a pronoun would be otherwise appropriate when the writer intends to modify the referent with a relative clause or an appositive. These less frequently cited uses of NPs seem to be at work in (21) and (22), respectively (all examples below were extracted from the Wall Street Journal available on line at http://www.ldc.upenn.edu).

(20) a. **Mr. Nixon** is traveling in China as a private citizen, but **he** has made clear that

he is an unofficial envoy for the Bush administration.

b. *Mr. Nixon* met Mr. Bush and his national security adviser, Brent Scowcroft, before coming to China on Saturday.

(21) a. Mr. Trudeau's attorney, Norman K. Samnick, said the harassment consists mainly of the guild's year-long threats of disciplinary action.

b. **Mr. Samnick** said a guild disciplinary hearing is scheduled next Monday in New York.

c. **Mr. Samnick**, **who** will go before the disciplinary panel, said the proceedings are unfair and that any punishment from the guild would be unjustified.

(22) a. **Michael R. Bromwich**, a member since January 1987 of the three-lawyer trial team in the prosecution of Oliver North, became a partner in the Washington, D.C., office of the 520-lawyer firm.

b. **He** will specialize in white-collar criminal defense work.

c. **Mr. Bromwich**, 35, also has served as deputy chief and chief of the narcotics unit for the U.S. attorney's office for the Southern District of New York, based in Manhattan.

2.3.2.2 Ariel

Ariel (1988, 1990) suggests that the accessibility of a referent is determined by multiple factors. She proposes that the four most important ones are:

1. Distance: The distance between the antecedent and the anaphor (between the antecedent and the anaphor).
2. Competition: The number of competitors on the role of antecedent.
3. Saliency: The antecedent being a salient referent, mainly whether it is a topic or a non-topic.
4. Unity: The antecedent being within vs without the same frame/world/point

of view/segment of paragraph as the anaphor.

The first two factors, i.e., *distance* and *competition*, are similar to Givón's measurements of *referential distance* and *potential antecedents*. Her *saliency* factor is closer to our notion of *topic* in that it is treated as a binary property of entities. Entities either are or are not topical in the segment that they are processed. While topichood is a binary property for Ariel, her accessibility hierarchy reflects a graded accessibility scale. Topicality is treated as one of the many factors determining that scale. Based on her own and other scholars' empirical measurements of the distribution of forms in texts, Ariel proposes the following graded *Accessibility Marking Scale* with forms preferred for entities of low accessibility appearing at the top of the scale and forms preferred for highly accessible entities appearing at the bottom.

Low accessibility

1. Full name + modifier
2. Full name
3. Long definite description
4. Short definite description
5. Short definite description
6. Last name
7. First name
8. Distal demonstrative + modifier
9. Proximal demonstrative + modifier
10. Stressed pronoun + gesture
11. Stressed pronoun
12. Unstressed pronoun
13. Cliticized pronoun
14. Extremely high accessibility markers (gaps including pro, PRO and *wh*-traces, and Agreement)

15. High accessibility

Ariel's accessibility scale is similar to Givón's scale of topicality only much more detailed. As with Givón's scale, our main and most relevant criticism for Ariel's accessibility scale is that it, too, is based on the assumption that a one-to-one mapping of linguistic form and usage can be achieved. As we saw in the review of Prince's work in Section 2.3.1, such one-to-one mappings are hard to establish as the same linguistic form may, and often does have more than one function. It is beyond the scope of this dissertation to explore the variety of functions that such forms may have. Previous work by Prince (Section 2.3.1) and Fox (1987), for example, have investigated such one-to-many and many-to-one mappings of form to function and vice-versa. With respect to NPs, specifically, Fox (1987) has found in her corpus studies that an NP may be used for what would be for Ariel a highly accessible entity as a stylistic option for adding information about the referent (e.g., *the smart editor*). Passonneau (1998), among others, has also observed that NPs can also be used for reference to highly accessible entities to mark a discourse boundary. The current work also shows that in some, but not all, cases pronouns are used to signal continuation on the same topic, thus drawing a clear distinction between the role of *accessibility* and *topicality* in the use of pronouns and relativizing both to the notion of center update unit in discourse organization.

2.3.3 Semantic Focusing Approaches

2.3.3.1 Stevenson et al.

Stevenson, Knott, Oberlander, and McDonald (2000) investigate the interaction between structural, thematic, and relational preferences in interpreting pronouns and connectives in discourse. Stevenson, Crawley, and Kleinman (1994) have argued that the crucial factors underlying focusing mechanisms in discourse are semantic/pragmatic factors. Semantic/pragmatic focusing assumes that verbs and connectives project their own focusing preferences. Verbs project focus preferences to the entities associated with the endpoint or

consequence of the described event. The focusing preferences of the connective depend on its meaning. For example, connectives like *because* direct attention to the cause of the previously described event, connectives like *so* direct attention to the consequences of the event. Thus, in a sentence like (23), the verb projects a focus preference for *Bill*, because *Bill* is the person associated with the endpoint of the event of criticizing. The connective, *so*, directs attention to the consequences reinforcing the focus on *Bill* which is then picked as the most preferred antecedent for the interpretation of the subsequent pronominal.

(23) John criticized Bill so he tried to correct the fault.

As we discuss in more detail in Chapter 3, the semantic/pragmatic focusing account runs into the type of problem demonstrated in (24), where the preferred interpretation for *he* is *John*, i.e., the structural subject, independent of semantic/pragmatic factors which would otherwise be responsible for making *Bill* the most salient referent in the subsequent discourse.[1] In such discourses it seems that a structural account is at play, in the sense of (Grosz & Sidner, 1986) (further elaborated in Chapter 3).

(24) a. John criticized Bill.

b. Next, he insulted Susan.

In the current work, we recognize the effects of verb semantics and connectives but we show that they take precedence over other factors of salience *within* the center update unit where pronominalization is not used as a strategy to manage topic structure.

Other work on the effects of the semantics of verb causality on the interpretation of subsequent pronouns includes, among many others, (Caramazza & Gupta, 1979) and (McDonald & MacWhinney, 1995).

[1] Experimental results regarding these cases are reported in (Stevenson et al., 2000).

2.4 Other Work on Subordinate Clauses

The study of subordinate clauses has received much attention in the narrative literature. Reinhart (1984) discusses the relationship between subordinate clauses and event structure in the context of the relationship between the temporal organization of narratives and the principals of gestalt perception. She defines "foreground" as the sequence of narrative clauses as defined by Labov (1972), and suggests that a "powerful means" for marking background is "the use of syntactic embedding". As Labov puts it, "once a clause is subordinate to another, it is not possible to disturb semantic interpretation by reversing it." Thus it is only independent clauses which can function as narrative clauses. Reinhart further claims that "material presented in subordinate clauses cannot normaly be foregrounded", but that writers can sometimes "play" with the foreground-background relations such that a narrative clause can "function as background" if it is marked syntactically as subordinate, just as long as the events are still ordered "on the same time axis" as in the represented world. Dispensing with the "reversability" criterion, other linguists have been inclined to consider certain subordinate clauses as part of the temporal sequence. Polanyi-Bowditch (1976) for example discusses the following example.

(25) When she began to arrange the flowers in a ball, a small fly flew out.

Because the two events are presented in iconic order it is reasonable to consider them to be part of the "temporal structure". However, as McCleary (1982) points out, there is some difference between presenting them this way and, say, as a pair of coordinate clauses.

Thompson (1987) discusses the correlation between "subordinate clauses" in English written narrative and the notion of temporal sequencing. She poses the question of whether, in written English narrative discourse, "subordination" is inversely correlated with "foregrounding" in the sense of "sequentially ordered". If not, why not? To answer this question she looks at two narrative passages from a narrative by Herbert Terrace, *Nim*. She identifies the predicates which are in temporal sequence, or "on the time line", and those

which occur in “subordinate” clauses. Indeed, 89% of the subordinate predicates occur in non-temporally sequenced clauses. She analyzes the remaining 11% to test her hypothesis that these are performing some other discourse function in addition to indicating a successive event in temporal sequence. Further, she attempts to address the question of why a writer should decide to present a temporally sequenced event in a subordinate clause. In one case a dependency of one event on the other is created which cuts across their temporal relationship.

(26) Only after he *STOPPED SMILING* and *SHRIEKING* did he GO to Stephanie and HUG her.

In the example above (predicates on the event line are capitalized), the “hugging” is reported to have been possible only after Nim had calmed down enough to stop smiling and shrieking. Many of the tokens fall under this category. In other cases, there is intervening descriptive material, unrelated to the temporal line, between the predicate in question and the preceding temporally sequenced event predicate. In these cases, Thompson conjectures that what the adverbial clause does that could not be done by an independent clause is to relate the clause following it back to the ongoing temporal line.

(27) When he *FINISHED GROOMING* Josh, Nim TURNED to Stephanie and her family and repeated SIGNED ”PLAY”.

She finds that this “orienting” function is the most prominent function of initial adverbial clauses. The reason why they are able to function this way, she suggests, is precisely because a temporally sequenced event is being coded in a *marked* form, that is, in a form which makes it grammatically dependent on another clause. She concludes that temporally sequenced events in a written narrative do not form a homogeneous class in terms of their grammatical representation. While the majority are coded by independent clauses, a significant subset are “subordinate”. In each of these cases the use of a subordinate clause allows the writer to accomplish a “text-creation goal” in addition to the obvious one of maintaining the temporal line. As Reinhart says, this is part of “what makes a text a work

of art". Thompson suggests that it is also part of what makes a text readable. A strictly linearly organized written narrative text would be not only boring but hard to attend to, for "the well-known reasons discussed in the gestalt perception literature".

Another approach to the study of main-subordinate clauses which appeals to gestalt principles of perception is Talmy's (2000). Talmy argues that in language there exist "two fundamental cognitive functions, that of the Figure, performed by the concept that needs anchoring, and that of the Ground, performed by the concept that does the anchoring. This pair of concepts can be of two objects relating to each other in space in an event of motion or location—and represented by nominals in a single clause. Or the pair of concepts can be of two events relating to each other in temporal, causal, or other type of situation—and represented by the main and subordinate clauses of a complex sentence." With regard to temporal sequences, Talmy proposes the following universal sequence principle which states that "The unmarked (or only possible) linguistic expression for any particular relation between two events in temporal sequence treats the earlier event as a reference point, or Ground, and the later event as requiring referencing—that is, as the Figure. Where the complete syntactic form is a full complex sentence, the two events are in the subordinate and the main clause, respectively." Assuming that linguistic universals reflect innate organizational and functional characteristics of the language-related portions of the brain, he then proposes that some of these characteristics are continuous with those of more general cognition-related areas. From this perspective, according to Talmy, the proposed universal about sequential events can be analyzed as follows.

> At times, a newly cognized item will illuminate or necessitate the rearrangement of items already in memory. But generally, cognitive effects seem to operate in the other direction: items already in memory constitute the basis, afford the analytic categories, and function as the reference points by which a newly cognized item is assessed, characterized, and analyzed. In particular, of two concurrent events, both cognized, the earlier one will, of course, already be in memory when the later one is newly occurrent, and so is generally

> to be used as the basis for the latter's assessment. The parallelism between this cognitive characteristic —the earlier being used as basis for assessing the later— and the linguistics characteristic —earlier and later treated semantically/syntactically as Ground/subordinate clause and Figure/main clause, respectively — suggests the following possibility. This feature of cognitive functioning may well have become incorporated in the innate structuring for conceptual/grammatical organization of the brain's language system, as the latter evolved. (Talmy, 2000)

An important weakness of the above argument, as Talmy himself recognizes, is that on hearing a complex sentence of the "temporal sequence" type, for example, the reader does not "cognize", in Talmy's terms, two actually occurring events but two adjacent descriptions of them. So unless we assume than linguistic descriptions are iconic in the sense that they also inherit some of the same effects as the original phenomena that are being "iconized", the argument loses its force as it would have to appeal to the experience of the descriptions, not the experience of the referents. Having to assume that linguistic organization is iconic further weakens the argument as there is ample evidence showing that it is not possible to maintain a close pairing-up of form-function correlations in a single language, let alone across different languages with different syntactic properties. To that we would like to add that it would be impossible to maintain a mapping from how we actually experience events temporally and their linguistic description as, even in a narrative text, more events precede others than those that are expressed in subordinate clause. While it is possible to maintain that in a complex sentence the event in the subordinate clause may be presented as a background event with respect to the matrix predicate, which events are chosen to be expressed as background events seems to be a choice made by the author. For example, both in (28) and (29), the earlier event serves as a "reference point" by which the "newly cognized item is assessed." However, only in the second discourse is this earlier event expressed in a subordinate clause.

(28) Mary finished the report. Then, she went to the movies.

(29) After Mary finished her report, she went to the movies.

A more plausible explanation that would also be consistent with the results of this thesis would be that given the choice that speakers have in presenting full propositions in both main and subordinate clauses, they may use this linguistic distinction to help them organize the discourse according to their intentions. With respect to entity salience, for example, we find that entities evoked in subordinate clauses are perceived of lower salience than entities in main clauses. It is, therefore, likely that, given a choice, speakers opt for a subordinate clause to accommodate entities that are not currently intended to play a topical role in the discourse.

Chapter 3

Attention, Anaphora, and the Main-Subordinate Distinction

3.1 Introduction

The problem of proposing referents for anaphoric expressions has been extensively researched in the literature and significant insights have been gained by the various approaches. However, no single model is capable of handling all the cases. We argue that this is due to a failure of the models to identify two distinct processes. Drawing on current insights and empirical data from English and Greek (Chapters 4 and 5), an aposynthetic[1] model of discourse is specified where topic continuity, computed across units, and resolution preferences internal to these units are subject to different mechanisms. The observed resolution preferences across the units (i.e., intersententially) are modeled structurally, along the lines suggested in Centering Theory. The resolution mechanism within the unit is subject to preferences projected by the semantics of the verbs and the connectives in the unit as suggested in semantic/pragmatic focusing accounts. This distinction not

[1] "Aposynthesis" is a Greek word which means "decomposition", i.e., pulling apart the components that constitute what appears to be a uniform entity.

only overcomes important problems in anaphora resolution but it also reconciles seemingly contradictory experimental results reported in the literature. A model of anaphora resolution can then be specified (Chapter 7) which interleaves the two mechanisms. A Centering-based model of the contribution of attention structure to discourse coherence is defined and tested on student essays (Chapter 6). In what follows, we briefly sketch the data that motivate the main-subordinate distinction and the proposed solution. Section 3.3 discusses in some detail the strengths and weaknesses of previous approaches to anaphoric interpretation which attempt to handle anaphora resolution in a uniform model. A detailed outline of the proposed model is given in Section 3.5.

3.2 Puzzles in Anaphora Resolution

Extensive research reported in the anaphora resolution literature has focused on the problem of proposing referents for pronominals.[2] First, Centering, formulated as a model of the relationship between attentional state and form of referring expressions, was utilized as the basis of an algorithm for binding pronominals on the intersentential level (Brennan, Walker-Friedman, & Pollard, 1987). The proposed algorithm, henceforth the BFP algorithm, gives the correct interpretation for the pronominal *he* in example (30) below, stating a preference to resolve the pronominal to *Max* rather than *Fred*.

(30) a. Max is waiting for Fred.

b. He invited him for dinner.

However, it was soon observed that the BFP algorithm was not capable of handling cases of intrasentential anaphora such as in (31), adapted from (Suri, McCoy, & DeCristofaro, 1999).

(31) a. Dodge was robbed by an ex-convict.

[2]While a significant amount of research in anaphora resolution has been carried out in statistical approaches, reviewing such approaches is beyond the scope of this thesis.

b. The ex-convict tied him up

c. because he wasn't cooperating.

d. Then he took all the money and ran.

The Centering based BFP algorithm would have a preference to resolve *he* in (31d) to *Dodge* and not to the *ex-convict*, based on a preference for a Continue transition.

Alternative approaches to anaphora resolution have sought to account for the resolution facts by proposing a semantic/pragmatic rather than structural mechanism. Stevenson et al. (2000) argue that both verbs and connectives have focusing properties affecting the preferred interpretation of pronominals. So, in (32a), the verb focusing highlights *Bill*, since *Bill* is the person associated with the endpoint of the event of criticizing. The connective, *so*, directs attention to the consequences and hence reinforces the focus on *Bill*.

(32) a. John criticized Bill,

b. so he tried to correct the fault.

The semantic/pragmatic focusing account runs into the type of problem demonstrated in (33), where the preferred interpretation for *he* is *John*, i.e., the structural subject, independently of semantic/pragmatic factors.[3] In such discourses it seems that a structural account is at play, in the sense of (Grosz & Sidner, 1986).

(33) a. John criticized Bill.

b. Next, he insulted Susan.

This chapter sets out to explicate the behavior of pronominals demonstrated in the above examples. Gaining significant insights from current research in anaphora resolution, seemingly contradictory findings are reconciled in a model according to which inter- and intrasentential anaphora are not subject to the same mechanism. We argue that the shortcomings of the proposed algorithms are due to confounding two distinct processes, namely

[3]Experimental results regarding these cases are reported in (Stevenson et al., 2000).

topic continuity and the internal structure of the sentence.[4] Intersentential anaphora is subject to structural constraints whereas intrasentential anaphora is subject to grammatical as well as semantic/pragmatic constraints. A discourse model informed of this distinction is specified in Section 3.5.

3.3 Related Work

3.3.1 The BFP Algorithm

Brennan et al. (1987) were the first to use the Centering Model as the basis for an anaphora resolution algorithm. The Centering Model (Grosz & Sidner, 1986), (Grosz, Joshi, & Weinstein, 1983) makes the following assumptions:

a) a discourse segment consists of a sequence of utterances $U_1, ..., U_n$,
b) for each utterance a ranked list of evoked discourse entities is constructed, designated as the Cf list,
c) the highest element of the Cf list is called the Preferred Center, Cp, and
d) the highest ranked entity in the Cf list of U_{i-1} realized in U_i is the Backward-Looking Center, Cb.

There are several types of topic transitions from one utterance to the next depending on whether the Cb is retained over two consecutive utterances U_{n-1} and U_n and whether this Cb is also the Cp of U_n (see Table 3.1. The distinction between a Smooth-Shift and a Rough-Shift is due to Brennan et al. (1987) who observed that the Centering Model

[4] As discussed in Chapter 1, we use the term "topic" to describe a centered entity, i.e., the entity that the discourse is "about". The notion of a centered entity is a discourse construct distinct from "topic" or "theme" as defined in information structure. Topic continuity is derivative of attention structure in discourse. We have opted for the more transparent term "topic continuity" as it describes the phenomenon we are mostly concerned with in pre-theoretical terms.

generated ambiguity in cases such as in (34):

(34) a. Brennan drives an Alfa Romeo.

b. She drives too fast.

c. Friedman races her on weekends.

d. She often beats her.

Table 3.1: Table of Centering transitions

	Cb(U_i) = Cb(U_{i-1})	Cb(U_i) $\neq$ Cb(U_{i-1})
Cb(U_i) = Cp	Continue	Smooth-Shift
Cb(U_i) $\neq$ Cp	Retain	Rough-Shift

Adding weight to the status of the Cp in (34c) makes it possible to successfully resolve the pronominal *she* in (34d) to *Friedman*. We return to the issue of ambiguity shortly. Their algorithm consists of three basic steps:

- Generate possible Cb-Cf combinations.
- Filter by constraints, e.g. contra-indexing, sortal predicates, Centering rules, and constraints.
- Rank by transition orderings (Continue>Retain>Smooth-Shift>Rough-Shift).

Some of the shortcomings of the BFP algorithm are discussed by Prasad and Strube (2000) who observe that it makes two strategic errors. Their observations are made with respect to Hindi but hold in English and Modern Greek, as shown in (35) and (36), respectively.

The first error occurs in cases when Cb(U_{i-1}) is different from Cp(U_{i-1}). In such cases, the preference for a Continue transition is responsible for resolving the pronominal in U_i to the Cb(U_{i-1}), and not to the Cp(U_{i-1}).

(35) a. $Ellen_i$ saw $Mary_j$ at school.

b. $Mary_j$ didn't talk to her_i.

c. She_j took her_j friends and walked away.

(36) a. I $Eleni_i$ ide ti $Maria_j$ sto sholio.
the Eleni saw the Maria at-the school.
'$Eleni_i$ saw $Maria_j$ at school.'

b. I $Maria_j$ den tis_i milise.
the Maria not to-her talked.
'$Maria_j$ didn't talk to her_i.'

c. $NULL_j$ pire tis files tis_j ki $NULL_j$ efige.
NULL took the friends her and NULL left.
'She_j took her_j friends and left.'

There is an important observation to be made here which is presented as the first indication for the distinction between topic continuity and anaphora resolution. On the one hand, the BFP Centering-based algorithm makes a resolution error opting for a Continue transition in (35c) and (36c). On the other hand, anaphora aside, the topic transition identified by Centering is, intuitively, correct. In (35) and (36), the discourse is initiated with *Ellen/Eleni* as the current topic, *Mary/Maria* is introduced as an entity related to the current topic, and then the discourse shifts to *Mary/Maria* to elaborate on her doings. The shift is in fact anticipated by the promotion of *Mary/Maria* from the object position in (35a) and (36a) to the subject position in (35b) and (36b).

The second error observed by Prasad and Strube (2000) is that the BFP algorithm generates ambiguity when U_{i-1} is discourse initial. Example (37) is given as illustration.

(37) a. John gave a lot of his property to George.

b. His current salary exceeded the average salary by a lot.

Given that the Cb in the discourse initial (37a) is unspecified, Continue transitions are generated when resolving *his* to either *John* or *George*. At this point, the BFP algorithm is not capable of reaching a decision.

The solution we propose for the two problems is simple: the preferred antecedent for the pronominal in U_i is the highest ranked entity in U_{i-1} that is compatible with the anaphoric expression. Compatibility is defined in terms of agreement features, number

and gender in the case of English. The proposed solution is consistent with the Centering model. The most relevant Centering notion for anaphora resolution is the Pronoun Rule which stipulates that if an entity is realized as a pronoun then so is the Cb. Opting for resolution to the highest ranked entity in the previous entity is precisely supported by the Pronoun Rule because the highest ranked entity realized in the following utterance *is* the Cb. On the other hand, using Centering transitions for anaphora resolution does not necessarily follow from the original formulation of Centering. Centering transitions, as originally formulated and as confirmed by the data discussed above, are best at identifying degrees of topic continuity. There is no a priori reason to expect that they will perform equally well in identifying pronominal referents. This is because assuming maximal coherence (preference for Continue transitions) overlooks properties of attention structure in discourse: strategies that hearers use to signal attention shifts to new centers while maintaining coherence. A Smooth-Shift may be intended and signaled appropriately by, for example, promoting a proper name from object to subject position. Interpreting pronominals in accordance with the Pronoun Rule as suggested here exploits precisely such strategies.

The conclusion from this section is that while Centering transitions identify successfully topic continuity in the discourse, in the domain of anaphora resolution the most useful Centering notion is not the transitions themselves but the Cf list ranking in combination with the Pronoun rule.

3.3.2 Functional Centering

Strube and Hahn (1996, 1999) elaborate on the nature of the Cf list and propose a Centering-based model of anaphora resolution where the Cf ranking is not based on grammatical function but on functional information status. They recast Centering notions in terms of Daneŝ's (1974) trichotomy between given information, theme, and new information. The Cb(U_i), the most highly ranked element of Cf(U_{i-1}) realized in U_i, corresponds to the element which represents given information. The Cp(U_i) corresponds to the theme of U_i. The rhematic elements of U_i are the ones not contained in U_{i-1}. While the original

motivation for the functional recast of Centering was due to German, a free word order language, Strube and Hahn (1996) claim that the functional framework is superior because fixed and free word order languages can be accounted for by the same principles. They argue against Walker, Iida, and Cote (1994b) who view the Cf ranking as a language-specific parameter that needs to be set.

In what follows we remain agnostic as to the suitability of the functional Centering framework for German. We will argue, however, that functional Centering is not the appropriate framework for all free word order languages, much less for languages universally. Preliminary evidence comes from Modern Greek, a free word order language.

To identify the factors determining the Cf ranking in Greek, we employ Rambow's (1993) diagnostic.[5] Rambow's diagnostic is used to test whether surface word order or grammatical function is the most reliable indicator of salience. The relevant examples for the Greek version of Rambow's diagnostic are shown in (38) and (39). The null pronominal in (38b) and (39b) resolves to the subject irrespective of its surface position. Gender and lexical considerations are controlled. Both *economical policy* and *arrangement* are feminine and they can both be *inadequate*. This judgment has been confirmed with a sizable group of native speakers of Greek attending the 15th International Symposium of Theoretical and Applied Linguistics (Miltsakaki, 2001). It seems, then, that the relevant

[5]Rambow suggests that the order of entities in the position between finite and non-finite verbs in German (Mittelfeld) affects their salience. Gender in German is grammaticized so he constructs an example with two same-gender entities in Mittelfeld and uses an ambiguous pronoun in subsequent discourse to determine which of the two entities is more salient. The constructed example is given below.

(1) a. Glauben Sie, dass [eine solche Massnahme]$_i$ [der russischen Wirtshaft]$_j$ helfen kann?
think you that a such measure-Fem the Russian economy-Fem help can?
'Do you think that such a measure can help the Russian economy?'

b. Nein, sie_i ist viel zu primitiv.
no, she is much too primitive.
'No, it's much too primitive.'

indicator of salience in the Cf list is grammatical function, or, at least subjecthood.[6]

(38) a. I prosfati diefthetisi$_i$ veltioni tin ikonomiki politiki$_j$?
the recent arrangement improve the economic policy?
'Does the recent arrangement improve the economic policy?'

b. Ohi, (null$_i$) ine aneparkis.
No, (it) is inadequate.
'No, it is inadequate.'

(39) a. Tin ikonomiki politiki$_j$ ti$_j$ veltioni i prosfati diefthetisi$_i$?
the economic policy it-(clitic) improve the recent arrangement?
'Does the recent arrangement improve the economic policy?'

b. Ohi, (null$_i$) ine aneparkis.
No, (it) is inadequate.
'No, it is inadequate.'

Further evidence for the role of grammatical function in Greek comes from syntactic objects.[7] In Greek, as in Turkish (Turan, 1995), a strong pronominal or a full NP must be

[6]It is interesting that in Turkish, another free word order language, it has also been shown (Turan, 1998) that the strongest indicator is subjecthood.

[7]Greek has two pronominal systems: weak pronouns that must cliticize to the verb and strong pronouns that are syntactically similar to full NPs. Dropped-subjects are considered part of the system of weak pronouns. In (Miltsakaki, 2000), it is argued that speakers of various languages use available nominal and pronominal forms and prosodic features in spoken language to signal attention structure in discourse. Greek speakers with a 3-way distinction in their nominal system (i.e. full noun phrases, weak and strong pronominals) use strong pronominals to signal reference to an entity previously evoked in discourse, which, however, is not the most salient entity. This use of strong pronominals is in some cases equivalent to certain prosodic effects in English. For example, prominent stress on the pronominals in (1) yields co-specification of *he* with *Bill* and *him* with *John*.

(1) John criticized Bill. Then, HE criticized HIM.

The need to recruit special prosody to achieve resolution to *Bill* indicates that structural focusing is indeed at work projecting strong "default" focusing preferences. In (1), there is sufficient semantic information to help the hearer arrive at the intended interpretation. If there was no default interpretation available at hand there would be no need to evoke prosodic effects. Once the linguistic encoding of speakers' strategies for building attention structure in discourse are identified, incorporating them in the Centering framework

used to promote the object of U_{i-1} to the subject position of U_i.[8] As the infelicitous interpretations (indicated by the pound sign) show in (40b), reference to the object *Yorgo* becomes felicitous only with the use of name repetition or a strong pronominal, shown in (40c) and (40d).[9] We take this as further evidence that objects rank lower than subjects in Greek.

(40) O Yannis$_i$ proskalese ton Yorgo$_j$.
the John invited the Yorgo.
'John invited George.'

a. null$_i$ tu$_j$ prosfere ena poto.
he him offered a drink.
'He$_i$ offered him$_j$ a drink.'

b. #null$_j$ #tu$_i$ prosfere ena poto.
he him offered a drink.
'He$_j$ offered him$_i$ a drink.'

c. O Yorgos tu$_i$ prosfere ena poto.
the George him offered a drink.
'George offered him$_i$ a drink.'

d. Ekinos$_j$ tu$_i$ prosfere ena poto.
he-strong him offered a drink.
'HE$_j$ offered him$_i$ a drink.'

Finally, to test the current results against the Functional Centering alternative, the definite subject in (38) has been replaced with an indefinite noun phrase. Shown in (41), the subject is an indefinite noun phrase representing new (hearer-new) information and the object is a definite phrase, encoding old (hearer-old) information. The null pronominal in (41b) resolves to the subject of (41a) disregarding the information status of the potential antecedents.

---

should be trivial.

[8] A "full NP" is any noun phrase that contains a head noun, either common or proper.

[9] Empirical evidence for the use of strong pronominals to signal reference to non-salient entities in Greek is provided in (Dimitriadis, 1996). Further functions of strong pronominals in Greek are identified in (Miltsakaki, 1999) and (Miltsakaki, 2001).

(41) a. Mia kenurgia diefthetisi$_i$ tha veltiosi tin ikonomiki politiki$_j$?
a new arrangement will improve the economic policy?
'Will a new arrangement improve the economic policy?'

b. Ohi, (null$_i$) tha ine aneparkis.
No, (it) will inadequate.
'No, it will be inadequate.'

That the information status is not the relevant factor in discourse salience, at least not cross-linguistically, is also confirmed in (Turan, 1998) for Turkish and in (Prasad & Strube, 2000) for Hindi. In both languages, the relevant factor for the ranking of elements in the Cf list is grammatical function.

In conclusion, information status (or hearer-status) is not universally the most important factor determining discourse salience (in Cf ranking). Given the pronominalization facts, at least for English, Greek, Hindi and Turkish, grammatical function can most reliably determine the relative salience of entities.

3.3.3 The S-list Algorithm

A further modification of the Centering model is proposed by Strube (1998), who replaces the functions of the Backward-Looking Center and the Centering transitions by the ordering among elements of what he calls the S-list, i.e., the list of salient discourse entities. The S-list ranking criteria define a preference of hearer-old over hearer-new discourse entities and is intended to reflect the attentional state of the hearer at any given point in discourse processing. The S-list is generated incrementally and is updated every time an anaphoric element is resolved. Anaphoric elements are resolved with a look-up in the S-list. The elements of the S-list are tested in the given order until one test succeeds. When the analysis of the utterance is finished (processed left to right), the discourse entities which are not realized in the utterance are removed. Strube (1998) claims that the incremental generation and processing of the S-list enables his system to handle inter- and intrasentential anaphora without any further specifications.

While the S-list has the merit of avoiding ambiguities caused by the way the Cb and the Centering transitions interact, it is not capable of handling intrasentential anaphora without any further specifications as claimed in (Strube, 1998). Stevenson et al. (2000) report experimental results pointing out cases where focus preferences are projected by verbs and connectives. Neither a grammatical function ordering nor an information based ordering is adequate to handle such cases. To illustrate the point, we quote an example, shown in (42), from (Stevenson et al., 2000). We construct the S-list ranking the elements according to grammatical function (information status would not distinguish between the two proper names).[10]

(42) a. Ken$_i$ admired Geoff$_j$ because he$_j$ won the prize

b. Geoff$_j$ impressed Ken$_i$ because he$_j$ won the prize

In both (42a) and (42b) the pronominal resolves to *Geoff*, the verb argument with the *stimulus* role. However, the ordering in S-list in (42a) is *Ken*>*Geoff* so the S-list algorithm will resolve the subsequent pronominal to the higher ranked element at the time of processing, in this case *Ken*. In fairness to the S-list algorithm, this is a problem for any Centering based algorithm which attempts to handle intrasentential anaphora according to a fixed ranking of entities in a salience list.[11]

Apparently, for certain discourses, algorithms relying on a fixed ordering of potential antecedents are not capable of resolving anaphora successful. We propose that such cases are most commonly identified intrasententially.

[10] This strategy was also adopted by Prasad and Strube (2000) in the implementation of the S-list algorithm for Hindi.

[11] It is conceivable that a discourse can be constructed where the semantics will force a similar pattern of resolution intersententially. However, Hudson-D'Zmura and Tanenhaus (1998) report experimental results which show that in such cases sentence processing is slowed down.

3.3.4 The RAFT/RAPR Algorithm

Based on previous work, Suri and McCoy (1994) and Suri et al. (1999) propose a methodology for extending their RAFT/RAPR[12] algorithm to handle focusing properties of complex sentences.

To determine how their framework should be extended to handle complex sentences, they develop a methodology specifically designed to determine how people process complex sentences. The central question they pose is whether a complex sentence should be processed as a multiple sentence or as a single sentence. They specifically investigated the "SX because SY" type of complex sentence as well as its interaction with the sentences occurring in the immediately previous and subsequent discourse.

(43) (S1) Dodge was robbed by an ex-convict the other night.

(44) (S2) The ex-convict tied him up because he wasn't cooperating.

(45) (S3) Then he took all the money and ran.

Their findings indicate that the pronoun resolution facts within S2, given in (44) above, are consistent with the expectations of both Centering and RAFT/RAPR. However, on completing the processing of the SY clause, the most salient entity for the following discourse is not picked from SY. Based on these findings, they propose the "Prefer SX hypothesis" to extend RAFT/RAPR.

While the "Prefer SX hypothesis" fixes the algorithm with respect to the construction in question, it seems to be missing a generalization regarding inconsistencies observed within versus across sentences.

3.3.5 Stevenson et al.'s Semantic/Pragmatic Focusing

Stevenson et al. (2000) investigate the interaction between structural, thematic, and relational preferences in interpreting pronouns and connectives in discourse. Stevenson et al.

[12]RAFT/RAPR stands for Revised algorithms for Focus Tracking and and Revised Algorithms for Pronoun Resolution.

(1994) have argued that the crucial factors underlying focusing mechanisms in discourse are semantic/pragmatic factors. Semantic/pragmatic focusing assumes that verbs and connectives project their own focusing preferences. Verbs project focus preferences to the entities associated with the endpoint or consequence of the described event. The focusing preferences of the connective depend on its meaning. For example, connectives like *because* direct attention to the cause of the previously described event, connectives like *so* direct attention to the consequences of the event. Thus in a sentence like (46), the verb projects a focus preference for *Bill*, because *Bill* is the person associated with the endpoint of the event of criticizing. The connective *so* directs attention to the consequences reinforcing the focus on *Bill* which is then picked as the most preferred antecedent for the interpretation of the subsequent pronominal.

(46) John criticized Bill so he tried to correct the fault.

By way of demonstration, let us turn our attention to action and state verbs. The semantic/pragmatic focusing account predicts that sentences with action verbs focus the entity associated with the end point of the event, namely the patient, independently of its structural position. This focus is maintained when the connective is *so*. In one of Stevenson et al.'s (2000) experiments, it is shown that in cases such as (47a) the pronominal *he* picks the patient as its referent both when it is introduced in the previous clause as a subject and when it is introduced as an object, as in (47b).

(47) a. Patrick$_i$ was hit by Joseph$_j$ so he$_i$ cried.

b. Joseph$_j$ hit Patrick$_i$ so he$_i$ cried.

A similar pattern was observed with state verbs, shown in (48), where *he* in the continuation was interpreted as the experiencer of the event independent of its structural position.

(48) a. Ken$_i$ admired Geoff$_j$ so he$_i$ gave him the prize.

b. Ken$_i$ impressed Geoff$_j$ so he$_j$ gave him the prize.

So the experimental evidence supports Stevenson et al.'s view that the focusing properties of verbs affect the interpretation of pronominals.

However, Hudson-D'Zmura and Tanenhaus (1998) report experimental results which, at first blush, contradict this view. They conducted a similar experiment to test if subject-object or stimulus-experiencer is the crucial distinction for pronominal interpretation. The participants of the experiment were given sentence (49) followed by the continuations (49a)-(49b) and were asked to judge the continuations for naturalness.

(49) Max despises Ross

a. He always gives Ross a hard time.

b. He always gives Max a hard time.

Their results show that there is a strong preference for the subject interpretation independently of the thematic role.

What are we to conclude from these inconsistent results? The results show that the same type of verb, i.e., state verb, projects its own focus preference, e.g., the experiencer, but in other cases it does not. One option would be to continue stretching structural focusing to account for the facts. Another option would be to continue stretching semantic focusing. In the following section, we propose an aposynthetic model for anaphora resolution where we divide the labor between the two mechanisms and define the domains of their applicability. The proposed model assumes that discourse is structured hierarchically. Before presenting the basic outline of the proposed model, we will briefly discuss the hierarchical and linear view of discourse structure and argue that the hierarchical view gains support from empirical data.

3.4 Hierarchical vs. Linear Discourse

In Grosz and Sidner's (1986) model of discourse structure, the global level component of attentional state is modeled as a stack. Discourse consists of segments and each segment is associated with discourse segment goals (intentions). The fulfillment of discourse segments goals achieves an overall discourse goal. Processing a discourse segment creates a focus state containing the objects, properties and relations relevant to that segment.

The focusing structure is modeled as a stack, thus allowing segments to be ordered either hierarchically or linearly with respect to other segments. The intentional relationships between segments determine the pushes and pops of focus spaces on the stack.

The stack model of discourse has received empirical support in cases of long distance anaphora. Referents for anaphoric expressions cannot always be identified within the boundaries of a segment (Hitzeman & Poesio, 1998). The stack model predicts that once an embedded segment has been *popped out*, the entities evoked in the dominating segment become available again as antecedents of subsequent anaphoric expressions. By way of demonstration consider the following example from (Walker, 1998).

(50) a. Caller: OK Harry, I have a problem that uh my —with today's economy my daughter is working,

b. Harry: I missed your name.

c. Caller: Hank.

d. Harry: Go ahead Hank.

e. Caller: as well as her uh husband

According to the stack model, once the embedded interruption spanning over (50b)-(50d) is popped out the caller can felicitously refer to his daughter introduced in (50a) with the anaphoric expression *her*. The hierarchical structure of discourse plays a crucial role in our understanding of attention management (topic structure) and its interaction with linguistic form. The next section reviews Walker's (1998)'s counterproposal to the hierarchical model. Empirical data are presented in support of the hierarchical model.

3.4.1 The Cache Model

Walker (1998, 1996), argues that it is possible to integrate Centering with a model of global discourse structure and abandon the restriction that Centering applies within segments. While she recognizes the focus-pop phenomena supporting the stack model, she

observes that the hierarchical adjacency achieved with the stack model is not always sufficient for licensing the use of anaphoric expressions. In the following adaptation of (50), the anaphoric expression *her* is much harder to interpret despite the fact that its antecedent is located in a hierarchically adjacent utterance.

(51) a. Caller: OK Harry, I have a problem that uh my —with today's economy my daughter is working,

b. Harry: I missed your name.

c. Caller: Hank.

d. Harry: I'm sorry, I can't hear you.

e. Caller: Hank.

f. Harry: Is that H A N K?

g. Caller: Yes.

h. Harry: Go ahead Hank.

i. Caller: as well as her uh husband

Walker proposes the cache model of attention state which integrates Centering and at the same time replaces the stack model of global structure. In the cache model, there are two types of memory: the main memory from where entities can be cue-retrieved, as in cases of focus-pops, and the cache or working memory which is immediately available for referent search. Segment boundaries are abandoned and local coherence phenomena are handled by the "Cache Size Assumption" which limits the cache contents to two or three sentences. Referents in the cache survive over segment boundaries for as long as the size of the cache memory permits it. The least recently accessed items in the cache are displaced to main memory. So, the cache model can handle cases of anaphora across boundaries using Centering while accommodating long distance anaphora through the cue-based retrieval of old entities from the main memory.

Despite the many appealing properties of the cache model, empirical evidence suggests that it cannot replace the stack model. The discourse in (52), for example, reveals two main

weaknesses of the cache model (or any other model assuming linear structure). The first weakness is that it is prone to error in cases where a competing antecedent appears within the cache space. In (52e), the cache algorithm will erroneously resolve the null subject to *Elsa Piu*. The morphology of the verb *anelave* indicates that the dropped subject is a third person singular noun but it is not marked for gender. The selection properties of the verb *anelave kathikonta* require a human subject, which restricts the search considerably. However, the most recent antecedent fulfilling the selectional requirements is *Elsa Piu*. Utterance (10), shown in (52e), does not include any linguistic cues so the cue-retrieval mechanism proposed to identify focus-pops will not be helpful. The previously focused entity, *teacher*, was the center of the first seven utterances in this segment. U7 is the beginning of an embedded segment giving background background information, possibly identified by the hearer through the change of tense from past to past perfect. U9 closes off the embedded segment and the null subject of the subsequent utterance resolves to the only entity available in U7, namely the *teacher*. The hierarchical structure represented in the stack model would, in this case, enable picking the correct referent by projecting a preference to search for an antecedent in the super-ordinate segment by-passing the embedded segment.

(52) a. U1-U6: Cb=teacher

b. U7 ki NULL-i erhotan trehontas na sinehisume to mathima
U7 and he-i was-coming running to continue-we the lesson
'And he-i was coming running so we would continue the lesson.'

c. U8 Tu-i ihan pi oti o skopos tu seminariu itan na mporo sto telos na gazoso enan anthropo en kinisi apo apostasi saranta metron, apo ti mia akri, diladi, tis skinis stin alli

'They had told him that the purpose of the seminar was for me to be able to shoot a moving person from a distance of forty meters, that is to say, from one edge of the stage to the other.'

d. U9 Fisika i Elsa Piu tha apoteluse poli pio efkolo kinigi ma kalio gaidoroderne para gaidurogireve.

‘Of course Elsa Piu would be a much easier target but better be safe than sorry.’

e. U10 NULL-i Anelave kathikonta me kefi ke ipsilo esthima
U10 NULL-i undertook duties with eagerness and high sense
efthinis.
of-responsibility.
‘He-i took on his duty eagerly and a high sense of responsibility.

The second weakness of the cache model is due to the design of the cache memory. The Cache Size Assumption in collaboration with the Cache Replacement Policy of the model will either err in the cases involving competing antecedents as demonstrated above or will involve the system in an “expensive” processing of a considerable number of possible antecedents stored in the main memory in the event of lengthy interruptions. The discourse in (53)-(55) is demonstrative. To avoid lenghty glossing only segment translation is given, keeping it as close to the Greek text as possible. The form and grammatical role of the crucial entities (appearing in boldface) have not been altered in the translation. Lettered examples belong to the embedded segment.

(53) **Tus-i** ixera apo tin kali kai apo tin anapodi.

(54) I prosfati hiria mu, pu den apoteluse logo na min **NULL-i** epithimun diakaos na me pidixun, ofile ostoso na me empodisi na emfanisto sto theatro.

a. Sto kato-kato, o antras mu den itan kanenas tiheos gia na perifrono etsi ti mnimi tu.

b. Ihe iparxi apo tus **stilovates** tis parataxis tu ke iroas tis dimokratias.

c. Kamia dekaria **dromi** se oli tin Ellada ihan pari to onoma tu

d. ke tha kikloforuse sintoma ke gramatosimo me ti fatsa tu.

(55) I simperifora mu **tus-i** ihe prokalesi foveri amihania...

(56) I knew **them-i** like the back of my hand.

(57) My recent widowhood, which did not constitute a reason for **them** to not want to fuck me badly, ought, however, to stop me from appearing on stage.

a. After all, my husband was not an unimportant figure so that I could disrespect his memory in this way.

b. He had been one of the **pillars** of his party and a hero of democracy.

c. A dozen or so **streets** all over Greece had been named after him

d. and a stamp displaying his face would be released soon.

(58) My behavior had caused **them-i** a good deal of embarrassment...

The crucial point in this example is that, a four-utterance long intervention separates the pronominal *tus* in (55) from its most recent antecedent, NULL-i, in the relative clause in (54). Entities marked with the same number and gender specifications as the pronominal are shown in boldface. Clearly, neither the size of the interruption, nor the presence of a competing antecedent, *stilovates*, is blocking the intended interpretation of the anaphoric expression.[13] Cache's sensitivity to the size of interruption is not supported by the data. A word of caution is in order here: we do not want to make the claim that any size of embedding is likely to occur. Most probably memory limitations at some point will diminish the hierarchical structure effect. What we are claiming is that memory size is not the crucial factor in tracking global focusing and attempting to decide whether working memory has the capacity to hold three, four or five utterances is probably a misleading direction in understanding discourse structure (albeit practical and probably convenient for applications).

[13]The competing antecedent *dromi* could in principle be ruled out by the semantics of (54c): streets cannot be *embarrassed.*

3.5 The Proposal: Aposynthesis

3.5.1 Outline of the Discourse Model

We assume that the discourse is organized hierarchically in linear and embedded segments as specified in (Grosz & Sidner, 1986). We also adopt the Centering view of local discourse coherence to model topic continuity in discourse. According to the Centering model each segment consists of a sequence of *utterances*. The size of an *utterance*, however, was left unspecified. We define an *utterance* as the unit consisting of a matrix clause and all its associated subordinate clauses. We call this unit the *center update unit*. For each update unit a list of forward-looking centers is constructed and ranked according to their salience. Consistent with the proposed definition of unit, entities evoked in subordinate clauses are less salient than entities evoked in the matrix clause and are ranked accordingly. The proposed Centering specifications have the following corollaries:
a) the linear order of subordinate clauses relative to the matrix clause does not affect the salience status of the entities,
b) entities evoked in subordinate clauses are available as potential links between the current and previous or subsequent discourse,
c) topic shifts must be established in matrix clauses, and
d) backward anaphora in subordinate clauses is no longer "backward" as anaphors in subordinate clauses are processed before main clauses independent of their linear order.

Finally, we assume that anaphora across units obeys Centering's Pronoun Rule. However, we do not adopt the BFP algorithm for anaphora resolution across units. Instead, as suggested in section 3.3.1, the preferred antecedent for a pronominal in U_i is the highest ranked entity in U_{i-1} modulo agreement features.

The remainder of this section is organized as follows. First, we briefly review Kameyama's Tensed Adjunct Hypothesis, which states that subordinate clauses are independent processing units, and argue that on the basis of new empirical evidence the hypothesis cannot be maintained. Next, evidence is presented in support of the new definition of the update

unit. Data from English, Greek, and Japanese show that treating subordinate clauses as independent units yields a) counter-intuitive Centering transitions and b) violations of the Pronoun Rule.

3.5.2 The Centering Update Unit

Defining the update unit within the framework of the Centering model became central in very early work because Centering was adopted and modified mainly to account for anaphora resolution. Given that anaphoric elements occur in all types of clauses, it was crucial that the size of the unit was constrained to enable the handling of intrasentential anaphora. To a large extent, efforts to identify the appropriate unit were often dictated by needs specific to anaphora resolution algorithms.

Centering was not originally formulated as a model of anaphora resolution. For purposes of testing the suitability of the relevant unit in Centering, it would be desirable to derive a model which yields transitions that reflect our intuitions about perceived discourse coherence, as well as the degree of the processing load required by the hearer/reader at any given time in discourse processing. Reflecting degrees of continuity is not a concern for anaphora resolution algorithms.

Kameyama (1993, 1998) was concerned with the problem of intrasentential Centering and, in particular, the definition of the appropriate update unit when processing complex sentences. Kameyama suggested breaking up complex sentences according to the following hypotheses:

1. Conjoined and adjoined tensed clauses form independent units.

2. Tenseless subordinate clauses, report complements and relative clauses belong to the update unit containing the matrix clause.

With regard to her tensed adjunct hypothesis which treated tensed adjunct clauses (for reasons of convenience, we will henceforth use the term “subordinate” to refer to this class of clauses) as independent units, Kameyama brings support from backward anaphora. She

argues that the tensed adjunct hypothesis predicts that the pronoun in the fronted subordinate clause in (59c), for example, is anaphorically dependent to an entity already introduced in the immediate discourse and not to the subject of the main clause it is attached to:

(59) a. Kern$_i$ began reading a lot about the history and philosophy of Communism

b. but never 0$_i$ felt there was anything he as an individual could do about it.

c. When he$_i$ attended the Christina Anti Communist Crusade school here about six months ago

d. Jim$_i$ became convinced that an individual can do something constructive in the ideological battle

e. and 0$_i$ set out to do it.

This view on backward anaphora, in fact, was strongly professed by Kuno (1972), who asserted that there was no *genuine* backward anaphora: the referent of an apparent cataphoric pronoun must appear in the previous discourse. Kameyama's argument (also Kuno's) is weak in two respects. First, it is not empirically tested that in cases of backward anaphora the antecedent is found in the immediate discourse. Carden (1982) and van Hoek (1997) provide empirical evidence of pronouns which are the first mention of their referent in discourse. More recently, Tanaka (2000) reports that in the cataphora data retrieved from the Anaphoric Treebank, out of 133 total occurrences of personal pronouns encoded as "cataphoric", 47 (35.3%) are "first mentioned". Among the 47 cases of "first mention" cataphora, 6 instances are discourse initial.[14]

Secondly, this account leaves the use of a full NP in Kameyama's main clause (59d) unexplained (*Kern* and *Jim* have the same referent). Full NPs and proper names occurring

[14]The Anaphoric Treebank is a corpus of news reports, annotated, among other things, with type of anaphoric relations. The Anaphoric Treebank is developed by UCREL (Unit for Computer Research on the English Language) in Lancaster University, collaborating with IBM T.J. Watson Research Center, Yorktown Heights, New York.

in Continue transitions have been observed to signify a segment boundary, e.g., (Passonneau & Litman, 1993). Assuming that segment boundaries do not occur between a main clause and a subordinate clause associated with it, the use of a full NP in (59d) remains puzzling.

Empirical evidence in support of Kameyama's hypothesis that tensed subordinate clauses should be treated as independent processing units was brought forth by Di Eugenio (1990, 1998). Di Eugenio, reporting on Centering studies in Italian, proposes that the alternation of null and overt pronominal subjects in Italian can be explained in terms of Centering transitions. Typically, a null subject signals a Continue, and a strong pronoun a Retain or a Shift.[15]

Following (Kameyama, 1993), she treats subordinate clauses as independent update units. Her motivation for doing so comes from the following example where the use of a strong pronoun in the main clause cannot be explained if the preceding adjunct is not treated as an independent update unit. The translation, taken from (Di Eugenio, 1998), is literal but not word for word. For the utterance preceding (60), the Cb(U_{i-1})=$vicina_j$ (neighbor-fem) and Cf(U_{i-1})=$vicina_j$.

(60) a. Prima che i pigroni-i siano seduti a tavola a far colazione,
'Before the lazy ones-i sit down to have breakfast,'

b. 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.'

In Chapter 4, we discuss the results of a Centering study in Greek. One of the surprising findings in this study was that a few strong pronouns appeared in Continue transitions. The result was surprising because the overall distribution of nominal and pronominal forms isthat weak pronouns are most common in Continue transitions whereas strong pronouns, full noun phrases, and proper nouns are associated with Rough-Shift transitions. On closer

[15]Di Eugenio collapsed the distinction between Smooth and Rough Shifts. However, the reader is referred to (Miltsakaki & Kukich, 2000a, 2000b), and also Chapter 6 in this thesis for a discussion of the significance of Rough-Shifts in the evaluation of text coherence.

inspection, it was observed that, in 6 out of the 8 instances of strong pronouns in Continue transitions, the referent of the strong pronoun is contrasted on the basis of some property with some other entity belonging to a previously evoked set of entities.[16] Although the sample is too small to draw any definitive conclusions, we can at least entertain the hypothesis that strong pronouns in Italian serve a similar function. If this is true, then an alternative explanation is available for Di Eugenio's data: in (60b), *she*, the most salient entity in the current discourse, is contrasted with the *lazy ones*, in (60a), on the property of 'laziness'. It turns out that the hypothesis that the strong pronoun does not signal a Rough-Shift transition is confirmed by the preceding discourse, where the 'vicina' appears as the most salient entity, realized with multiple dropped subjects. The discourse immediately preceding (60) is shown in (61).[17][18]

(61) a. NULL$_j$ e' una donna non solo graziosa ma anche energica e dotata di spirito pratico;
'and not only is she$_j$ pretty but also energetic and endowed with a pragmatic spirit;'

b. NULL$_i$ e la combinazione di tutto cio' e', a dir poco, efficace.
'and the combination of all these qualities is effective, to say the least.'

c. NULL$_j$ si alza all'alba per sovrintendere a che si dia da mangiare alle bestie, si faccia il burro, si mandi via il latte che deve essere venduto; una quantita' di cose fatte mentre il piu' della gente se la dorme della grossa,
'She$_j$ gets up at dawn to supervise that the cows are fed, that the butter is made, that the milk to be sold is sent away; a lot of things done while most people sleep soundly. '

[16] One further instance of a strong pronominal in a Continue transition was ignored. In that case, the strong pronominal headed a relative clause and its use was determined by the grammar.

[17] Many thanks to Barbara Di Eugenio (personal communication) for providing us with the extra data in (61).

[18] We presume that Di Eugenio's coding of the null realization in (61b) is based on the inferable information that the noun phrase 'la combinazione di tutto ció' refers to *her* $_j$ qualities.

We now turn to English and Greek to show that treating subordinate clauses as independent Centering units yields counter-intuitive topic transitions. First, consider the constructed example from English shown in (62).

(62) Sequence:
main-subordinate-main

a. John had a terrible headache.
Cb= ?
Cf= John>headache
Transition=none

b. When the meeting was over,
Cb=none
Cf= meeting
Transition=Rough-Shift

c. he rushed to the pharmacy store.
Cb=none
Cf=John
Transition=Rough-Shift

(63) Sequence:
Main-main-subordinate

a. John had a terrible headache.
Cb=?
Cf=John>headache
Transition=none

b. He rushed to the pharmacy store
Cb=John
Cf=John>pharmacy store
Transitions=Continue

c. when the meeting was over.
Cb=none
Cf=meeting
Transition=Rough-Shift

Allowing the subordinate clause to function as a single update unit yields a sequence of two Rough-Shifts, which is diagnostic of a highly discontinuous discourse. Further, if indeed there are two Rough-Shift transitions in this discourse the use of the pronominal in the third unit is puzzling. A sequence of two Rough-Shift transitions in this short discourse is counterintuitive and unexpected given that of all Centering transitions, Rough-Shifts in particular have been shown to a) disfavor pronominal reference, among others, (Walker et al., 1994b), (Di Eugenio, 1998), (Miltsakaki, 1999), b) be rare in corpora, to the extent that the transition has been ignored by some researchers, among others (Di Eugenio, 1998), (Hurewitz, 1998), and c) be reliable measures of low coherence in student

essays (Chapter 6). In addition, simply reversing the order of the clauses, shown in (63), causes an unexpected improvement with one Rough-Shift transition being replaced with a Continue. Assuming that the two discourses demonstrate a similar degree of continuity in the topic structure (they are both *about* "John"), we would expect the transitions to reflect this similarity when, in fact, they do not.

Presumably, the introduction of a new discourse entity, "meeting", in the time-clause does not interfere with discourse continuity, nor does it project a preference for a shift of topic, as the Cp normally does when it instantiates an entity different from the current Cb. Notice that if we process the subordinate clause in the same unit as the relevant main clause, we compute a Continue transition independently of the linear position of the subordinate clause as the entities introduced in the main clause rank higher than the entities introduced in the subordinate clause. The computation is shown in (64).

(64) a. John had a terrible headache.
Cb=?
Cf=John>headache
Transition=none

b. When the meeting was over, he rushed to the pharmacy store.
Cb=John
Cf=John>pharmacy store>meeting
Transitions=Continue

Similar examples were identified in data collected from a short story in Greek (Chapter 4). Example (65), shown below, is representative.

(65) a. Ki epeza me tis bukles mu.
and I-was-playing with the curls my
'And I was playing with my hair.'
Cb=I, Cp=I, Tr=Continue

b. Eno ekini pethenan apo to krio,
while they were-dying from the cold

'While they were dying from the cold,'

Cb=none, Cp=THEY, Tr=Rough-Shift

c. ego voltariza stin paralia,
I was-strolling on-the beach
'I was strolling on the beach,'

Cb=NONE, Cp=I, Tr=Rough-Shift

d. ki i eforia pu esthanomun den ihe to teri tis
and the euphoria that I-was feeling not have the partner its
'and the euphoria that I was feeling was unequaled.'

Cb=I, Cp=EUPHORIA, Tr=Rough-Shift

Again, processing the while-clause in (65b) as an independent unit yields three Rough-Shift transitions in the subsequent discourse, reflecting a highly discontinuous discourse. When (65b) and (65c) are processed as a single unit, the resulting sequence of transitions for the entire discourse is a much improved Continue-Continue-Retain.

Further evidence in support of the proposed definition of the update unit comes from cross-linguistic observations on anaphora resolution. The most striking examples come from Japanese.[19] In Japanese, topics and subjects are lexically marked (wa and ga respectively) and null subjects are allowed. Note that subordinate clauses must precede the main clause. Consider the Japanese discourse (66). Crucially, the referent of the null subject in the second main clause resolves to the topic marked subject of the first main clause, ignoring the subject-marked subject of the intermediate subordinate clause.

(66) a. Taroo wa tyotto okotteiru youdesu
Taroo TOP a-little upset look
'Taroo looks a little upset.'

b. Jiroo ga rippana osiro o tukutteiru node
Jiroo SUB great castle OBJ is-making because
'Since Jiroo is making a great castle,'

[19]Thanks to Kimiko Nakanishi for providing us with the data. In a Centering study she conducted in Japanese (personal communication) she also concluded that treating subordinate clauses as independent units would yield a highly incoherent Japanese discourse.

c. ZERO urayamasiino desu
ZERO jealous is
'(He-Taroo) is jealous.'

In Section 3.2, a similar case was also identified in English. It is repeated here as (67d). Again, the referent of *he* in (67d) is co-specified with *ex-convict*, the subject of the previous main clause. If the because-clause were processed independently then the most salient referent available for the interpretation of the anaphoric in (67d) should be *Dodge*. Manipulating the semantics in the second main clause to make resolution to *Dodge* the most plausible choice does not seem sufficient to warrant felicitous pronominalization, as has been shown experimentally in (Suri et al., 1999), demonstrated here in (68). In (68), *he* is not the preferred form for reference to *Dodge* despite the fact that *Dodge* is the most natural referent for the argument of the predicate *screaming for help* in this context.

(67) a. Dodge was robbed by an ex-convict.

b. The ex-convict tied him up

c. because he wasn't cooperating.

d. Then he took all the money and ran.

(68) a. Dodge was robbed by an ex-convict the other night.

b. The ex-convict tied him up because he wasn't cooperating.

c. #Then he started screaming for help.

The low salience of subordinate clause entities is further confirmed in the experimental results reported in (Suri et al., 1999). In their experiment, the participants in the experiment judged that a natural way to refer to *Dodge* in (69c) is by name repetition.

(69) a. Dodge was robbed by an ex-convict the other night.

b. The ex-convict tied him up because he wasn't cooperating.

c. Then Dodge started screaming for help.

Finally, defining the main clause and its associated subordinate clauses as a single unit

points to interesting new directions in understanding backward anaphora. With the exception of a few modal contexts shown in (72),[20] backward anaphora is most commonly found in preposed subordinate clauses, (70), and not in sequences of main clauses, (71). From the proposed unit definition, it follows that surface backward anaphora is no longer "backward" once the Cf list is constructed and ranked. The referent of the pronoun in such cases appears lower in the Cf list ranking and, in fact, looks backwards for an antecedent as any other normal pronoun would. To illustrate the point, the Cf list for (70) contains *John>shower>he-referent*. The pronoun looks back for an antecedent, intrasententially, and resolves to the only compatible antecedent available, *John*.

(70) As soon as he arrived, John jumped into the shower.

(71) #He arrived and John jumped into the shower.

(72) He$_i$ couldn't have imagined it at the time but John Smith$_i$ turned out to be elected President in less than 3 years.

3.5.3 Discourse Salience vs. Information Structure

In the previous section, we suggested that the linear position of the subordinate clause does not affect topic continuity. This position leads itself to another question: if the linear position of subordinate clauses does not improve topic continuity, then what is the function of clause order variation?

Let us, briefly, turn our attention to the surface word order within a single clause. It is commonly assumed that for each language there is an underlying canonical order of the basic constituents. In an SVO language like Greek, the canonical order of the verb and its arguments is subject-verb-object. This, of course, is not always the attested surface order. In syntactic theories, it is commonly assumed that surface word order is derived by various movement operations. Some movement operations are dictated by the syntax of each

[20]Thanks to Ellen Prince for pointing out this example. Similar examples appear also in (Matthiessen & Thompson, 1988).

language and are necessary to yield grammatical sentences. However, it is also common, especially in free word order languages, for movement to be syntactically optional and for the surface word order to be used to satisfy information packaging needs (for example to arrange the information into old-new, or ground-focus, to mark open propositions etc.). Note that when this happens, it is only the surface word order that is altered and not the basic relation of the arguments to the predicate. To give an example from English, in (73) the internal argument of the verb (the object) has been fronted but its original relation to the verb has remained the same.

(73) Chocolate Mary hates.

Moving to the sentential level, we entertain the hypothesis that the same principle dictates the position of the clauses relative to each other. Each dependent clause stands in a specific relation to the main clause and this relation is not altered by the order in which the clause appears on the surface. In discourse grammars, this insight is captured in the discourse LTAG treatment of subordinate conjunctions. In discourse LTAGs, subordinate conjunctions are treated as predicates, anchoring initial trees containing the main and the subordinate clause as arguments. Each subordinate conjunction may anchor a family of trees to reflect variations of the surface order of the substituted argument clauses but the predicate argument relation remains the same (Webber & Joshi, 1998; Webber, Knott, Stone, & Joshi, 1999a, 1999b; Forbes, Miltsakaki, Prasad, Sarkar, Joshi, & Webber, 2001).

The above discussion relates to the definition of the Centering update unit in the following way. The Centering model keeps track of center continuations and center shifts. In other words it keeps track of discourse salience. If we dissociate salience from information structure the relevant unit for computing salience is at the sentence level, which we can visualize as a horizontal level (see Figure 1). The relative order of independent/dependent clauses is determined by information structuring, a process possibly orthogonal to the computing of salience. Subordinate links are not relevant to the salience mechanism. Salience is computed paratactically. A natural consequence of this model is that referents can be introduced on the vertical level without affecting the status of the salient entity on

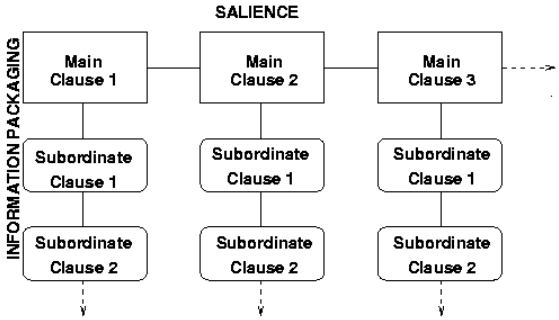

Figure 3.1: Salience model

the horizontal level. It follows that changes of topic must be established at the horizontal level. Such a conception of the salience structure suggests that text processing is not strictly incremental as commonly assumed. While it is possible that the Cf list is constructed incrementally, the final ranking is determined only after the sentence is complete.

Admittedly, the distinction between discourse salience and information packaging is hard to establish due to the inevitable overlap between information status and salience: attention centers, for example, tend to be discourse old. Still, there are other aspects of information packaging pertaining to clause order (e.g. temporal or logical sequences, open proposition frames inherited from previous discourse etc.) that do not necessarily relate to the salience of the participating entities. While a lot more work is required to understand the precise nature of the interaction between salience and information structure, we believe that we obtain a significant gain in keeping the two processes distinct.

Chapter 4

Adverbial Clauses

4.1 Introduction

In this chapter we focus on the reference patterns of pronouns in adverbial clauses. In controlled experimental conditions, with the main clause predicate held constant, we establish that the interpretation of subject pronouns in English and Greek adverbial clauses varies, in this case according to preferences projected by the semantics of the subordinate conjunction. Conversely, main clause subjects are consistently interpreted as the subject of the preceding main clause, thus confirming our hypothesis that intersententially anaphors opt for structurally salient entities. Previous work by Cooreman and Sanford (1996) is also reported which shows that the pronoun subject of a main clause following another main clause and a dependent adverbial clause is interpreted as the subject of the preceding main clause independently of its surface order with respect to the dependent adverbial clause. These results are also supportive of our hypothesis that entities in adjunct subordinate clauses are of lower salience than the entities of main clauses. The conclusions from the experimental studies are then confirmed by a Greek corpus study which shows that pronouns in main clauses resolve to the highest ranked entity in the preceding sentence. The interpretation of pronouns in subordinate clauses, however, varies with approximately 50% of pronouns resolving to an entity other than the highest ranked

entity in the preceding discourse. In the Greek corpus study, we included only tokens of sequences of main-main and main-subordinate clauses which contained at least two competing antecedents in the first main clause for the third person anaphor in the second main or adverbial clause. The search for relevant tokens in the Greek corpus was facilitated by the fact that Greek is a subject-drop language whose verb morphology marks number but not gender. We were not able to replicate the study for English because English third person pronouns are marked for gender and it is much harder to identify a sizable amount of tokens containing at least two competing antecedents for a third person pronoun contained in the subsequent main or adverbial clause.

This chapter is organized as follows. Section 4.2 reports the results of two experimental studies in English and related experimental work conducted by Cooreman and Sanford (1996). Section 4.3 reports the results of an experimental and a corpus study in Greek. Section 4.3 also includes the results of a preliminary corpus study in Greek which establishes that intersententially dropped subjects and weak pronouns are used for reference to topical entities and correlate with Centering's Continue transition. When hierarchical structure is taken into account, dropped subjects and weak pronouns are also shown to refer to a topical entity evoked in a higher segment. Strong pronouns and full NPs on the other hand are associated with reference to a non-topical entity and to signal a contrastive relation to members of a salient set of entities evoked in the previous discourse. The results of the preliminary study in Greek form the basis for the design of the experimental and corpus studies in the same language. Our conclusions from this chapter are discussed in Section 4.4.

4.2 English

In this section, we report two experimental studies in English. In both experiments, the interpretation of a subject pronoun was quantified in two conditions: a) the pronoun was located in a main clause following another main clause, and b) the pronoun was located in

an adverbial clause following a main clause. In experiment 1, we compared and contrasted structural and semantic effects on pronoun resolution via a close semantic match of subordinate and main clause adverbial connectives. In the main-main condition the second main clause was modified by an adverbial whose meaning approximated the meaning of the subordinate conjunctions. Experiment 2 also involved a set of connectives for both the main-main and main-subordinate conditions but the effort for one-to-one mapping of subordinate conjunctions and clause adverbials was abandoned for reasons discussed in Section 4.1.2.3.

4.2.1 Experiment 1

4.2.1.1 Materials and design

The method for this experiment was a sentence completion task. Participants were asked to read sets of two clauses. Each set of clauses consisted of a main clause followed by either a subordinate conjunction introducing an adverbial subordinate clause or by a period and a second main clause modified by a semantically matched sentence adverbial in initial position. In both conditions, the connective (main clause adverbial or subordinate conjunction) was followed by a subject pronoun. Participants were asked to complete the second clause in a natural way. Crucially, the first main clause contained two male or two female referents, one in the subject position and one in the object position. The referent of the subject pronoun in the second clause could be interpreted as either the subject or the object of the preceding main clause. The same gender referents were instantiated as role-NPs (e.g. groom, best man, witch, monk, etc.).[1] The main clause contained an action verb involving physical contact (e.g. hit, kick, hug, kiss, etc.). The subject of the verb was assigned the agent role and the object of the verb the patient role.

Both the main clause adverbials and the subordinate conjunctions were selected from two semantic classes: TIME and CONTRAST. The TIME class included the subordinate

[1]We opted for role NPs instead of individual names in order to minimize referent ambiguity in the participants' continuations.

connective *when* and the adverbial *then*. The CONTRAST class included the subordinate conjunction *although* and the adverbial *however*.

A sample stimulus set is shown in Figure 4.1. The experiment followed a 2X2 design. The factors were type of clause (main or subordinate) and semantic type (time or contrast). There were 12 target items which were combined with 24 fillers. All target items appeared an equal number of times in each condition but only once for each participant. Sixteen adult, native speakers of English volunteered to participate.

(74)	The groom hit the best man. However, he...
(75)	The beggar pushed the gentleman although he...
(76)	The boxer kicked the referee. Then, he...
(77)	The policeman shot the burglar when he...

Figure 4.1: Experiment 1: Sample of target items

4.2.1.2 Results

On average there were two ambiguous continuations per experimental set. In these cases, participants were asked to identify explicitly their interpretation of the pronoun immediately after the end of the experimental session.

The interpretation of the subject pronoun as the subject of the preceding main clause was quantified and converted to percentages. The scores were then submitted to a two-way ANOVA analysis. The results of the ANOVA showed a strong main effect for type of clause ($F(1,15)=25.6$, $p<0.0001$) and a marginal effect for semantic type ($F(1,15)=4.5$, $p<0.049$.

Figure 4.2 shows the percentages of reference to the subject of the first main clause by type of clause (main or subordinate) and semantic type (time or contrast). The percentages for each category show that, when the second clause was a main clause, the subject pronoun was more frequently interpreted as the subject of the main clause. On the other hand, when the second clause was a subordinate clause, the subject pronoun showed a much weaker tendency to be interpreted as the subject of the preceding main clause.

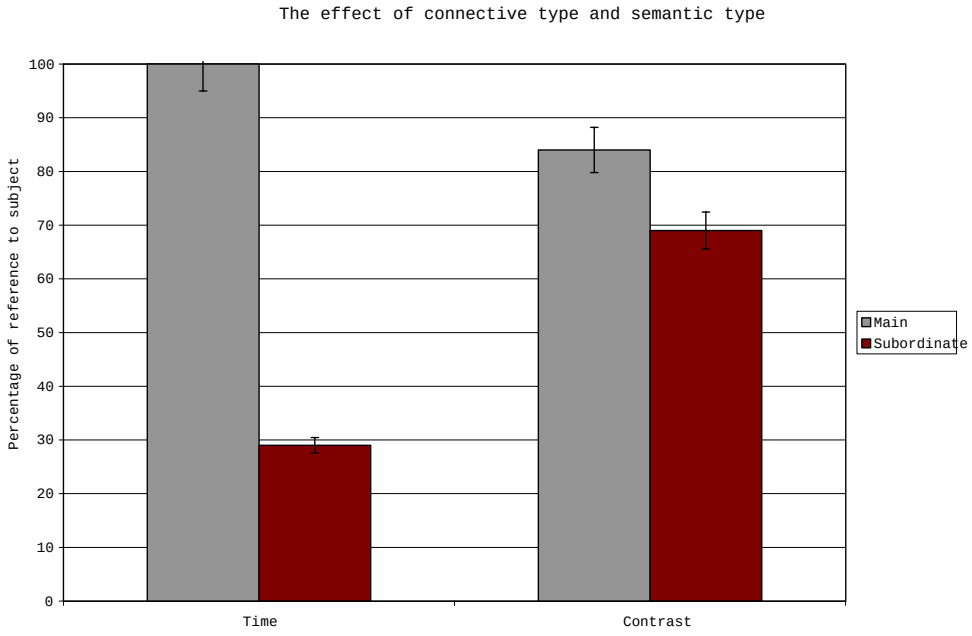

Figure 4.2: Percentage of reference to subject in English

4.2.1.3 Discussion

In this experiment we contrasted semantic type, time, and contrast, with type of clause, main and subordinate. The results show that the type of clause affects the interpretation of the pronoun it contains. In the main-main condition participants showed a significantly stronger tendency to interpret the subject pronoun as the subject of the preceding main clause than in the main-subordinate condition. This preference was demonstrated both for main clauses modified by the temporal adverb *then* and the contrastive adverb *however*. Conversely, in the main-subordinate condition, the subject pronoun was often interpreted as the object of the previous clause. The marginal effect of the semantic type shows that the effect of structural focusing in the main-main condition overrides the effect of semantic

focusing.

A word of caution is in order here. The comparison by semantic type was problematic for the temporal group containing *then* and *when*. In many cases, the *when*-clause continuations of the participants established a *causal* link between the events of the main clause and the *when*-clause. For example, in the continuation shown in (78), the fact that the son was lying on the ground seems to have *caused* the event in the main clause.

(78) The father shook the son vigorously when he saw him lying on the ground.

In fact, Moens and Steedman (1988) have argued that there is no true "temporal" interpretation for *when*-clauses. They argue that, in all cases, *when*-clauses predicate more than "temporal coincidence". They claim that *when*-clauses predicate some *contingency* relation such as a *causal link* or an *enablement relation* between the two events expressed in the main and subordinate clauses. This *causal link* link that the *when*-clause predicates is not what we normally understand as formally causal in that *when* seems to predicate an intransitive relation. For example, from (79a) and (79b) we cannot conclude (79c).

(79) a. When John left, Sue cried.

b. When Sue cried, her mother got upset.

c. When John left, Sue's mother got upset.

On the other hand, we believe that *because* seems to behave in a similar way (80), suggesting that causality in discourse processing is a more complex phenomenon than formal causality.

(80) a. Because John left, Sue cried.

b. Because Sue cried, her mother got upset.

c. Because John left, Sue's mother got upset.

Similar complications are likely to arise with respect to the semantics and pragmatics of other connectives, making it hard to validate comparisons by semantic type. For this

reason, we did not pursue any further analyses of the semantic typefactor. Instead, we redesigned the experiment dropping the semantic type factor and introducing, instead, a bigger number of subordinate conjunctions. Adding a bigger number of subordinate conjunctions allows for generalizations for the entire class of adverbial clauses versus main clauses independently of the semantics of connectives.

4.2.2 Experiment 2

4.2.2.1 Materials and design

The method for this experiment was a sentence completion task. As in experiment 1, participants were asked to read sets of two clauses. Each set either contained a sequence of two main clauses (main-main condition) or a sequence of a main and a subordinate clause (main-subordinate condition). The second clause contained a subject pronoun and participants were asked to complete the sentences in a natural way. The critical items in this experiment had the same structure as in experiment 1. Five subordinate conjunctions and five clausal adverbials were included. Both the subordinate conjunctions and the adverbials were chosen from a variety of semantic classes. Figure 4.3 contains the complete list of connectives included in this experiment.

Main clause adverbials:
however, then, period, as a result, what is more
Subordinate conjunctions:
although, because, while, when, so that

Figure 4.3: Experiment 2: Set of English connectives

Sample critical items are shown in Figure 4.4.

Each experimental set contained 30 critical items combined with 90 fillers. The fillers were also sentence completions with a different structure. Each condition (main-main or main-subordinate) appeared in fifteen versions : fifteen subordinate continuations and fifteen main clause continuations. Each connective appeared in three items in each complete experimental set.

(81)	The groom hit the best man. Moreover, he...
(82)	The beggar pushed the gentleman so that he...
(83)	The boxer kicked the referee. As a result, he...
(84)	The policeman shot the burglar because he...

Figure 4.4: Sample items from experiment 2

Twenty participants, native speakers of English, undergraduate students at the University of Pennsylvania, took part in the experiment in exchange for course credit. On average, participation time ranged from thirty to forty-five minutes.

The interpretation of the subject pronoun as the referent of the subject in the preceding main clause was first quantified and converted into percentages. As in experiment 1, ambiguous continuations were disambiguated by the participants immediately after the completion of the experimental session.

4.2.2.2 Results

The scores were submitted to an ANOVA analysis. The results of the ANOVA showed a strong main effect of the type of the clause type ($F(1,19)=79.33$, $p<0.000$)).

Figure 4.5 shows the percentages of reference to the subject of the first main clause in each condition. The results of this experiment confirm the results of experiment 1. The percentages for each category show that when the second clause was subordinate, the subject pronoun showed a much weaker tendeny to refer to the subject of the preceding main clause. Reference to the subject of the preceding main clause was strongly preferred when the subject pronoun appeared in a main clause.

4.2.2.3 Discussion

The purpose of this experiment was to test if we can generalize across connectives the effect of the type of clause obtained in experiment 1. The results of experiment 2 confirm this finding for a larger number of connectives, five subordinate conjunctions in the main-subordinate condition and five adverbials in the main-main condition . In the main-main

condition, the pronoun was interpreted as the subject of the previous main clause across all adverbials, confirming that structural focusing in this condition is the primary factor determining pronominal interpretation. If semantic focusing was the primary determinant of salience in this condition, we would expect to see a varied pattern depending on the semantics of the connective. In the main-subordinate condition, on the other hand, the percentage of reference to the subject of the previous main clause is significantly lower indicating that other factors override structural focusing.

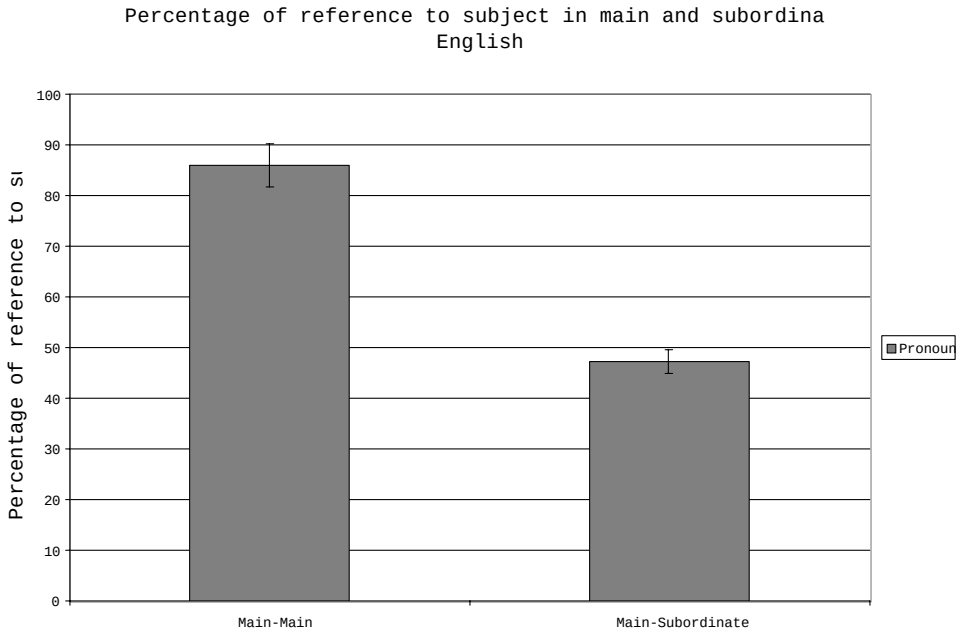

Figure 4.5: Percentage of reference to subject

4.2.3 Cooreman and Sanford

Cooreman and Sanford (1996) independently studied the effect that the main-subordinate distinction may have on the processing of subsequent discourse. Specifically, they investigated the interpretation of a subject pronoun following a main and an adverbial clause, each introducing a same gender referent. In a sentence completion task, they presented the participants with a complex sentence containing a main and an adverbial clause. Then, participants were prompted to start a continuation with a pronoun which would refer either to the entity introduced in the main clause or the same gender entity introduced in the adverbial clause. To check for clause order effects, the adverbial clause appeared both after and before the main clause. Three sets of subordinate conjunctions were used: *after/before*, *when/while*, and *because/since*. A sample set of items in Figure 4.6.

After the tenor opened his music store the conductor sneezed three times. He...
The conductor sneezed three times after the tenor opened his music score. He...

Figure 4.6: Sample items from Cooreman and Sanford's experiment

Their results revealed that for all three sets of connectors the main clause referent was the preferred choice for the interpretation of the pronoun in the continuation: 92.9% for *after/before*, 80.3% for *when/while*, and 79.8% for *because/since*. The order in which the main and adverbial clauses were presented did not make a difference except for the subordinate conjunction *because*: the main clause referent was the preferred choice for the interpretation of the pronoun in the continuation 75.2% in the main-subordinate order versus 85.4% in the subordinate-main order. No such effect was shown for any other subordinate conjunction, including *since*.

The experiment by Cooreman and Sanford directly addresses the question we have posed regarding the effect of main-subordinate syntax on topic continuity. In their experimental conditions, the complex sentence meets the definition of the center update unit that we have proposed. According to the model that we proposed in Chapter 3, we would also predict that the subsequent pronominal subject would be interpreted as the highest ranked

entity in the complex sentence (the main clause subject), i.e., the entity introduced as the most likely topic of the subsequent discourse. If, as we have proposed, the highest ranked entity in the complex clause is the subject of the main clause independently of the surface order of the subordinate clause, we would also predict that order of the subordinate clause would not affect the interpretation of the subsequent subject pronoun. As shown in the results of Cooreman and Sanford's experiment, the predictions of the proposed model are borne out.[2]

4.3 Greek

4.3.1 The Pronominal System in Greek

The pronominal system in Greek consists of two paradigms: strong pronouns and weak pronouns. Greek also allows null subjects, which are classified in the weak paradigm. Both strong and weak pronouns are subject to syntactic constraints which we present below.

Greek is a subject-drop language, so null pronouns are only allowed in subject position. Weak pronouns are used for direct and indirect objects, which are in fact clitics immediately preceding the verb. The order of clitics when both direct and indirect objects are present is also dictated by the grammar, with the indirect object pronoun always preceding the direct object pronoun. Strong pronominals are obligatory in prepositional phrases

[2]In a follow-up experiment, Cooreman and Sanford evaluated the effect of the main-subordinate distinction in a self-paced reading experiment. In this experiment, participants read a complex sentence and then a following target sentence which would cohere propositionally either with the main clause or with the subordinate clause. Then, they analyzed the reading times in each condition. The results of this experiment showed that reading times were faster when the target sentence cohered with the main clause for the temporal connectives (after, before, when and while). However, there was no significant difference in the reading times or the target sentence when the preceding complex sentence included causal connectives (because and since) indicating that the propositional content of a causal subordinate is equally "accessible" as the propositional content of the main clause.

and also when heading a relative clause. Both weak and strong pronouns are morphologically marked for case, number and gender. Greek has three genders, masculine, feminine and neuter. Nouns representing human referents are normally marked as male of female depending on the referent's sex (except for infants and kids). However, other animate and all inanimate objects can be masculine, feminine, or neuter.

Strong and weak forms are also available in possessive NPs. Weak possessive NPs consist of the head noun followed by a weak form in genitive, shown in (85). Strong possessive NPs are constructed with full NPs in the possessor and possessee roles, as in (87), or with with the emphatic form *dikos mu* "my own", preceding the possessive and marked with the same case as the head noun, shown in (86). Finally, we have classified the anaphoric *o idhios* in the strong paradigm. The anaphoric *o idhios* "self" is also morphologically marked for gender, number, and case. An example of *idhios* is shown in (88).

(85) I mitera mu
the mother my
'My mother.'

(86) I diki mu mitera
the own my mother.
'MY mother.'

(87) I mitera tis Marias.
the mother the-gen Maria-gen
'Maria's mother.'

(88) I idhia ostoso ihe apoliti sinesthisi.
the herself however had absolute awareness.
'(She) herself however was fully aware.'

4.3.2 Salience Ranking in Greek Main Clauses

The salience status of an entity is determined by a number of factors which may vary cross-linguistically. This is because languages may choose different linguistic strategies

and/or encoding to mark entities as more or less salient. As mentioned in Chapter 2, in English, for example, it has been proposed (among others, Kameyama (1985) and Brennan et al. (1987)) that the Cf list is partially determined by the grammatical role, with subjects ranking higher than objects. For German, Rambow (1993) has claimed that the salience of entities appearing between the finite and non-finite verbs (Mittelfeld) is determined by word order and used a diagnostic to test this claim.

One would expect that possibly in free word order languages, in general, the relevant salience of entities would be reflected by choices in word order. Word order effects on salience have been shown, for example by Kaiser (2003) for some anaphoric expressions in Finnish. Greek is also a free word order language. As a preliminary test for ranking entities in Greek, we have used Rambow's diangostic to contrast the effect of grammatical role versus word order. As discussed in some detail in Section 3.3.2, preliminary evidence from applying Rambow's diagnostic to Greek data indicates that salience ranking in Greek main clauses is primarily determined by grammatical functions, subjects ranking higher than objects independently of their surface order. We have tentatively assumed, then, that the ranking rule for Greek main clauses is: Subject>Object>Other.

In Greek, as in Turkish (Turan, 1995), in clauses with non-agentive psychological verbs, experiencer objects seem to rank higher than the theme subjects. Turan observed that the experiencer object is the highest ranked entity because it is the empathy locus in Turkish, in some respects analogous to *giving* and *receiving* verbs in Japanese, in which *empathy* ranks higher than subjects (Walker et al., 1994a). Turan (1995) also pointed out that in Turkish quantified indefinite subjects (qis) and impersonal plural pros rank very low. Again, the same observation appears to hold for Greek (also for Italian, (Di Eugenio, 1998)). In light of these observations we adopt for Greek the amended ranking rule proposed in (Turan, 1995), shown in Figure 4.7.

Empathy>Subject>Indirect object>Direct object>Others>qis, pro-arb

Figure 4.7: Salience ranking for Greek

In the next section, we present an exploratory corpus study on the distribution of weak and strong pronouns in Greek. We performed a Centering analysis of a Greek corpus and examined the distribution of referring expressions with respect to the four Centering transitions. The results of the study are supportive of the hypothesis that null subjects and weak pronouns are preferred in topic continuations. Null subjects and weak pronouns may also be used to signal a return to the topic of a higher discourse segment. Strong pronouns may also appear in topic continuations but only when a contrastive relation is expressed between the referent of the strong pronoun and some other referent evoked in the preceding discourse.

4.3.3 Preliminary Study on Weak and Strong Pronouns

The corpus in this study comprises a short story of approximately 6,000 words. The short story is an excerpt of the collection titled "I won't do this favor for you" (our translation), authored by the modern Greek novelist C.A. Chomenides. The text was chosen for its richness of nominal and pronominal expressions as the story involves multiple characters. A Centering analysis of the text was performed. For the Centering analysis of the text, we assumed the ranking rule shown in Figure 4.7 and the definition of center update unit proposed in Chapter 3, i.e., main and subordinate clauses were processed as a single center update unit. The computation of Centering transitions was made disregarding discourse segment boundaries.

A total of 467 sentences were identified, containing 371 weak forms (null subjects and weak pronominals) and 96 strong forms (full NPs and strong pronominals) in the highest ranked position, in most cases the subject of the main clause of the unit. The referring expression representing the highest ranked element of the Cf list of each unit, i.e., the Cp, was coded according to the coding schema shown in Table 4.1.

For every two consecutive units, Centering transitions were computed according to the Table 3.1, repeated here as Table 4.2 for convenience. The results of this annotation are shown in Table 4.3.

null	null subjects
weak	weak pronouns, weak possessives, quantified indefinite phrases realized as null and quantified indefinite phrases realized with a weak pronoun
full	full noun phrases (including proper names)
strong	strong pronouns, strong possessives, epithets, the anaphor *o idhios*

Table 4.1: Coding of Greek referring expressions

	$Cb(U_i)=Cb(U_{i-1})$	$Cb(U_i)\neq Cb(U_{i-1})$
$Cb(U_i)=Cp$	Continue	Smooth-shift
$Cb(U_i)\neq Cp$	Retain	Rough-shift

Table 4.2: Table of Centering transitions

What stands out from Table 4.3 is the strong tendency of Cps to be expressed with a null subject or a weak pronoun to appear in primarily Continue and secondarily Smooth-Shift transitions. Continue and Smooth-shift transitions have one important property in common: in both types of transitions the Cp is also the Backward-Looking Center of the current unit. With respect to the distribution of null and weak forms it is also noteworthy the unexpectedly high occurrence of these forms with Rough-Shift transitions. We will now turn to these cases for closer inspection.

On closer inspection of the instances of null and weak forms in Rough-Shift transitions we observed that their distribution patterns fell under the classification presented in Table 4.4.

Following Grosz and Sidner (1986), as "focus pops" were classified cases with a null

	Continue	Retain	Smooth-Shift	Rough-Shift
null	203	22	52	29
weak	44	1	9	10
total	243	23	61	39
full	6	8	3	54
strong	1	2	6	6
total	7	20	9	60

Table 4.3: Distribution of weak and strong pronouns in Greek

Focus pops	11
Mode switches	13
Missing arguments	6
deictic links	2
other	4

Table 4.4: Distribution of Rough-Shifts in Greek

or weak form appearing as the Cp of the unit immediately following a parenthetical or other embedded segment, in our cases descriptions of the setting of a scene. Such parenthetical interruptions or scene descriptions or interruptions halt temporarily the flow of the narrative and are sometimes used to give background information for a new setting in the narrative. For illustration, an example of an instance classified as a focus-pop is given in (90). The immediately preceding discourse is given in translation in (89). The discourse spanning over (90a) and (90b) temporarily freezes the narrative to provide additional information about the hotel and then (90c) resumes the narrative and temporally returns to the discourse in (89), immediately preceding the interruption. It is likely that the use of a null subject in (90c) to refer to the topic of (89) serves as a cue that the embedded segment is closed off and the narrative resumes continuing on the same topic that was established before the interruption.

(89) *I took him to a hotel for lovers in Victoria Square, where I used to go at the time of my relationship with Elias, the only boyfriend I ever had who didn't have a vacation house or at least a car.*

(90) a. Mesa se okto hronia o enikiazomenos peristerionas tu erota ihe
in to eight years the rentable pigeon-loft of love had
ekmondernisti
been-modernized.
'Within eight year the rentable pigeon-loft of love had been modernized.'

b. Ihane vali tileorasis sta domatia ke sistima exaerismou.
had-they put TVS in-the rooms and system of-air-condition.
'They had installed TVs and air-conditioning.'

c. Akinitopiisa to asanser anamesa ston proto kai ston
immobilized-I the elevator between to-the first(floor) and to-the
deftero.
second(floor).

'I stopped the elevator between the first and second floor.'

Similarly, we classified as "mode switches" instances where a null or weak form appeared in the unit immediately following a switch from narrative to direct speech and vise versa. Instances of "mode switches" are presented as a separate category in Table 4.4. However, given a hierarchical structure point of view of the discourse, mode switches can also be classified as focus-pops, in the sense that once a segment containing quoted speech is closed off the narrative resumes.

Moving to the next category, we classified as "missing arguments" cases where an argument was realized in the discourse implicitly. Modelling implicit arguments is a thorny issue both in theoretical and computational linguistics. While it seems intuitively obvious that in some cases implicit arguments can serve as links between discourse units, what their status is in a model of discourse representation is less obvious.

We classified as "deictic links" cases where the link between two units was established by discourse deixis, i.e. the use of a demonstrative pronoun like *afto* "this" to refer to a textual segment. Discourse deixis and the formulation of its contribution to discourse coherence as well as its interaction with entity-based coherence accounts is again an open research area. Finally, the category "other" included two character scenes represented by a dialogue between two characters containing first and second person references.

Turning to the distribution of full NPs and strong pronominals shown in Table (4.3), we see a high percentage of full NPs in Rough-Shift transitions, as expected. What is surprising in Table (4.3) is the number of occurrences of strong pronouns in Continue transitions. On closer inspection of these occurrences, Table 4.5, we observe that in 6 of the 7 occurrences of strong pronouns in a Continue transition, a "contrastive" relationship held between the referent of the strong pronoun and some other entity evoked in the preceding discourse. In fact this type of "contrastive" relationship appears under a "poset"

(partially ordered set) classification following (Prince, 1981b). Prince (1981b) argues that "contrast" is not a primitive notion. A "contrast" relation arises "when alternate members of some salient set are evoked and, most importantly, when there is felt to be a salient opposition of what is predicated of them" (Prince, 1998).

	poset (contrast)	relative	
strong	6	1	

Table 4.5: Strong forms in Continue transitions

A representative example of this function of strong pronouns is shown in (91). Taking into consideration the prior context, the propositional opposition in (91) is inferred between the referent of *them* trying to console the referent of *she* thinking that *she* was suffering when, in fact, *she* was experiencing pleasure from killing without being caught. In the remaining case, the use of a strong pronoun was obligatory by the grammar. The strong pronoun served as the head of a relative clause. In Greek, heads of relative clauses cannot be null subjects or weak pronouns.

(91) a. ke agonizondan na me parigorisun.
and were-trying-they subjun-prt me console-they
'and they were trying to console me.(Smooth-Shift)'

b. Omos ego iha epitelus vri ton eafto mu...
however I had finally found the self my...
'However, I had found myself... (Continue)'

c. O dikos tis iroikos thanatos den ihe tosi simasia oso i diki mu tapinosi.
the own her heroic death not had that-mush importance as the own my humiliation.
'HER heroic death was not as important (to her) as MY humiliation. (CONTINUE)'

This function of strong pronouns is in addition to another function of strong pronouns in Greek which has been pointed out by Dimitriadis (1996). Dimitriadis argues that strong

pronominals in Greek are used to indicate that the antecedent is *not* the Cp of the previous unit. Support of this claim is offered by a corpus study that he conducted in Greek. Dimitriadis does not recognize the "poset" function of strong pronouns. However, it is important to keep track of both functions of strong pronouns because, in fact, a strong pronoun *can* pick the Cp of the previous unit as its antecedent precisely in those cases that a poset relation holds between the Cp and some other entity evoked in the preceding discourse. A naturally occurring example of such a discourse is given in (92) (from the Greek newspaper *Eleftherotypia*).

(92) To idio kani ke i N.D$_i$.
the same does and the N.D.
'N.D.$_i$ (our note: Greek opposition political party) do the same.

(93) Null$_i$ gnorizi alla den null$_i$ lei.
null knows but not say.
'She$_i$ know but she$_i$ doesn't say.'

(94) Aoristos null$_i$ iposhete oti **afti**$_i$ tha diahiristi kalitera tin meta ONE epohi
Vaguely null promises that she will manage better the after ONE era
me to epihirima oti null-i ine to kat' exohin evropaiko komma.
with the argument that null$_i$ is the pre dominantly European party.
'She$_i$ vaguely promises that SHE$_i$ (our note: as opposed to the governing party) will manage the after ONE (European Currency Unification) era with the argument that she$_i$ is the predominantly European political party'

4.3.3.1 Summary

In this section we conducted an exploratory corpus study of the distribution of weak and strong pronominal forms in Greek. The purpose of this study was to confirm the hypothesis held for Greek that the preferred referring expression for reference to a topical entity is a weak form. Indeed, we found that weak forms are strongly associated with both Continue and Smooth-Shift transitions in which the highest ranked entity in the unit is also the Backward-Looking center (topic) of the same unit. As Dimitriadis (1996) has shown, strong pronouns signal reference to a non-Cp. In addition, as our study has shown, strong

pronouns in Greek may serve a second function of signalling a contrastive relation between entities belonging to a salient set of entities evoked in the discourse. In such cases, strong pronouns may, in fact, be used for reference to the highest ranked entity of the preceding discourse.

4.3.4 Experiment 3

In this section, we present the design and the results of an experiment conducted in Greek. As in Experiment 2, which was conducted in English, the purpose of this experiment is to evaluate the hypothesis that subject pronouns in main clauses are determined by the topic structure of the discourse whereas the interpretation of subject pronouns in subordinate clauses, in this case adverbial clauses, is determined by other factors, most likely the semantics of verbs and connectives and can therefore vary accordingly. The design of the Greek experiment was slightly modified due to the fact that established topics in Greek may be referenced with a dropped subject which, obviously, cannot be used as a prompt in a sentence completion task.

4.3.4.1 Materials and design

The Greek version of the experiment 2 in English was modified in the following way. A rating questionnaire was designed to elicit off-line judgments about naturalness. Participants were asked to read two versions of the same set of sentences. In version (1), the anaphoric element following the connective was a dropped subject. In version (2), the anaphoric element following the connective was the strong pronoun *ekinos* "that", marked with number and gender features. In both versions, the continuations following the dropped subject or strong pronoun were identical. The semantics of the second clause were controlled so that the referent of the anaphoric element would be unambiguously co-referent with the *object* of the preceding main clause. A sample stimulus set is shown in Figure 4.8. We quantified over the percentage of times that the participants judged the use of the strong pronoun as the most natural choice for reference to the object of the

preceding main clause.

(95) a. O astinomikos pirovolise ton lopoditi astrapiea **etsi oste**
the policeman shot-at the thief quickly **so that**
na min prolavi na apodrasi.
0 to not have-time to escape
'The policeman shot at the thief quickly so that he wouldn't escape.'

b. O astinomikos pirovolise ton lopoditi astrapiea **etsi oste**
the policeman shot-at the thief quickly **so that**
ekinos na min prolavi na apodrasi.
HE to not have-time to escape
'The policeman shot at the thief quickly so that HE wouldn't escape.'

(96) a. O raftis metrise ton kirio leptomeros. **Epipleon**
The tailor measured the gentleman with-detail. **Moreover**
0 stathike telios akinitos oso o raftis
0 stood completely still for-as-long-as the tailor
eperne metra.
was-taking measures
'The tailor measure the gentlemen in detail. Moreover he stood completely still while the tailor was taking measures.'

b. O raftis metrise ton kirio leptomeros. **Epipleon**
The tailor measured the gentleman with-detail. **Moreover**
ekinos stathike telios akinitos oso o
HE stood completely still for-as-long-as the
raftis eperne metra.
tailor was-taking measures
'The tailor measure the gentlemen in detail. Moreover HE stood completely still while the tailor was taking measures.'

Figure 4.8: Experiment 3: Sample items

As in the English experiments 1 and 2, the main clause contained two male or two female referents and the main clause verb was an action verb involving physical contact. Following the design of experiment 2, five subordinate conjunctions and five clause adverbials were selected for the continuations. Figure 4.9 shows the complete set of connectives.

There were 30 critical items combined with 90 fillers. The fillers consisted of pairs of

Main clause adverbials:
omos 'however' telia 'period' etsi 'so' epipleon 'moreover' epita 'then'
Subordinate conjunctions:
an ke 'although' yati 'because' eno 'while' otan 'when' etsi oste 'so that'

Figure 4.9: Experiment 3: Set of Greek connectives

sentences with a different variable to judge for naturalness, for example variations in word order in the continuations or variations in the use of a perfective or non-perfective form. Each condition (main-main or main-subordinate) appeared in fifteen versions: fifteen subordinate continuations and fifteen main clause continuations. Each connective appeared three times in each complete experimental set. Twenty adult participants, all native speakers of Greek, volunteered to take part in the experiment. On average participation time was 20-30 minutes.

4.3.4.2 Results

The number of times the strong pronoun was judged more natural for reference to the object of the preceding main clause was first converted to percentages and then the scores were submitted to an ANOVA analysis. The results of the ANOVA showed a strong main effect of the type of clausal connection ($F(1,18)=52.78$, $p<0.00$)).

Figure 4.10 shows the percentages of felicitous reference to the object of the preceding main clause using a strong pronoun. The percentages for each category show that strong forms are required for reference to the previous object across main clauses. When the anaphoric appears in a subordinate clause, reference to the object of the previous clause with a null subject is significantly facilitated.

4.3.5 Summary and Discussion of Experimental Studies

Experiment 1 for English provided preliminary evidence for the effect of clause type (main versus subordinate) on anaphora resolution. In their continuations, participants tended

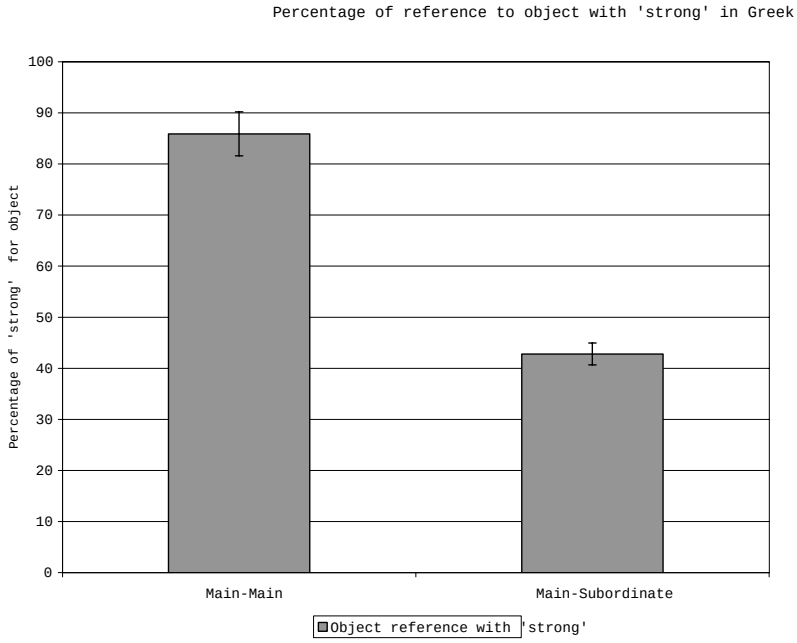

Figure 4.10: Experiment 3: Percentage of preference for “strong”.

to interpret the pronominal in the main clause condition as the subject of the previous main clause. No such pattern was identified in the main-subordinate condition where the interpretation of the pronominal varied across two types of subordinate clauses (time and contrast). In the same experiment, the semantic type had only a marginal effect on the interpretation of the pronominal.

In experiment 2 for English, a larger set of subordinate conjunctions was selected for the materials. The strong effect on type of clause was retained, confirming the preliminary results of experiment 1. Over a set of a total of ten connectives, 5 subordinate conjunctions and five clause adverbials from a variety of semantic classes, the preferred interpretation of the pronominal was consistently assigned to the subject of the preceding main clause when the pronominal appeared in a main clause. A varied pattern of interpretation was

observed when the pronominal appeared in a subordinate clause.

Experiment 3 tested the same conditions in Greek. The aim of experiment 3 was to investigate whether the effect of subordination on anaphoric interpretation applies to a language other than English. The results of experiment 3 showed that a strong pronoun, normally used for reference to an entity other than the most salient one in the previous discourse, was consistently judged more natural for reference to the object of the preceding of the main clause. On the other hand, in the main-subordinate condition the strong pronoun of the subordinate clause was not consistently judged more natural for reference to the object of the previous clause.

The results of all three experiments in English and Greek confirm the hypothesis that intrasententially subject pronouns in English and dropped subjects in Greek, are interpreted as the structurally most salient entity in the preceding main clause, the position that is often used to host topical entities. The results also confirm the second part of the hypothesis which suggests that, intrasententially, the interpretation of a subject pronoun in English and a dropped subject in Greek is determined by other factors which appear to facilitate interpreting the subject pronoun or dropped subject of a subordinate clause as the object of the preceding main clause.

Looking closer at the distribution of anaphoric interpretation per connective in English and Greek, shown here in Figures 4.11 and 4.12, respectively, we observe significant variation among subordinate connectives. Given the experimental design for both English and Greek, with the main clause predicate held constant, we tentatively conclude that the variation was largely due to the semantics of the subordinate connectives, although in some cases the subordinate connective could establish more than one type of relation between the propositions expressed in the main and subordinate clauses. Such could have been the case, for example, for the subordinate conjunction *when* which expresses either a temporal or a causal relation, and possibly the subordinate conjunction *while* which can express either a temporal or a contrastive relation.

The variation among adverbial connectives was much smaller but noticable revealing

a potentially interesting interaction between structural and semantic factors. The adverbials *then* and '*moreover* in both languages appear to enhance the salience of the previous subject, as the semantics of both connectives seem encourage continuation on the same topic. In fact, these two adverbials (and the period in Greek) show a ceiling effect in the diagram. On the other hand, *as a result* which, semantically, is predicted to shift attention to the object (the patient) of the preceding clause shows a lower percentage of reference to the subject of the preceding main clause by comparison to other clause adverbials but still higher than the subordinate conjunctions *because* or *so that*, which would also predicted by semantic accounts to shift attention to the patient. We take this as an indication of a tension revealed intersententially for continuing reference to the topic of the preceding clause despite semantically driven expectations for reference to the, as yet, non-topical entity. More research on the semantics of subordinate conjunctions and clausal adverbials is clearly needed to further illuminate the nature of the potential interaction between factors in various syntactic, semantic and pragmatic configurations.

4.3.6 Corpus Study

The central aim of this study is to evaluate the extend to which the results of the controlled experimental studies reported in the previous sections were also reflected in naturally occurring data. As mentioned earlier, Greek allows dropped subjects yielding more frequent referential ambiguity than pronominal references do in English. We were, therefore, able to collect a reasonable number of tokens fulfilling conditions similar to the experimental study. As in the experimental study, we wanted to compare and contrast the interpretation of anaphoric expressions in a main clause with the interpretation of anaphoric expressions in a subordinate clause. Unlike the experimental study, however, the search of anaphoric expressions in main and subordinate clauses was not restricted to subject pronouns.

The corpus used in this study contained approximately 800,000 words and contained primarily newspaper articles of the Greek newspapers *Eleftherotipia* and *To Vima*, available on-line at http://www.enet.gr and http://www.dolnet.gr, respectively.

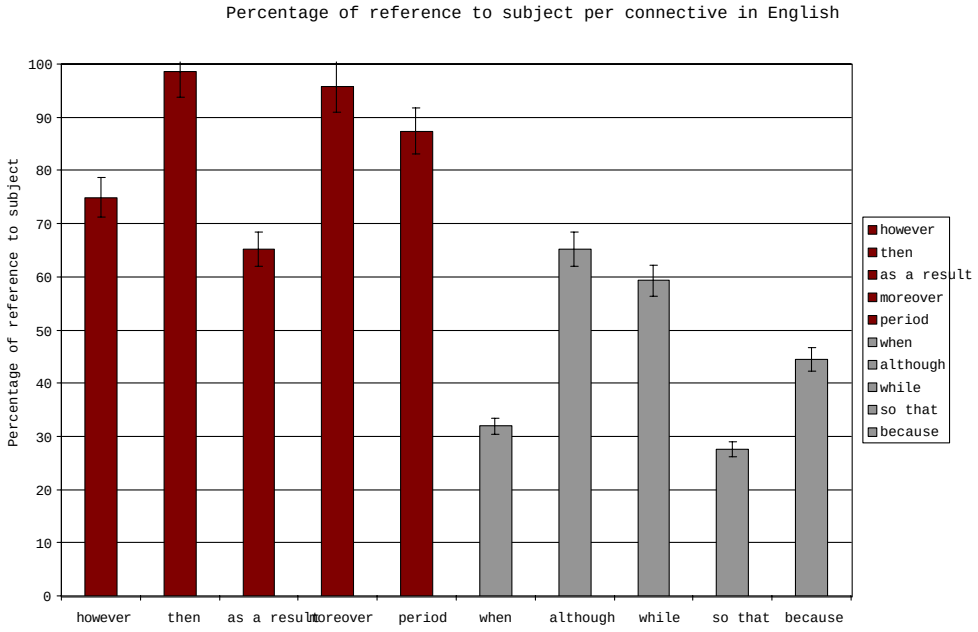

Figure 4.11: Percentage of reference to subject per connective

4.3.6.1 Data Collection and Coding

The requirements set for the dataset of this study were the following: a) the subordinate clause or second main clause contains a third person pro-dropped subject or weak pronominal, b) the preceding main clause or any of its other associated subordinate clauses contains at least two competing antecedents. A competing antecedent is defined as a full noun phrase, dropped subject or weak pronominal that agrees in gender (and in cases of adjectival predicates) number with the anaphoric expression.

For anaphoric reference in main-main and main-subordinate sequences, we would ideally like to include only those tokens where the second main or subordinate clause under investigation was preceded by a unit containing only a main clause. However, this extra constraint would invalidate a large number of the already limited number of tokens, so

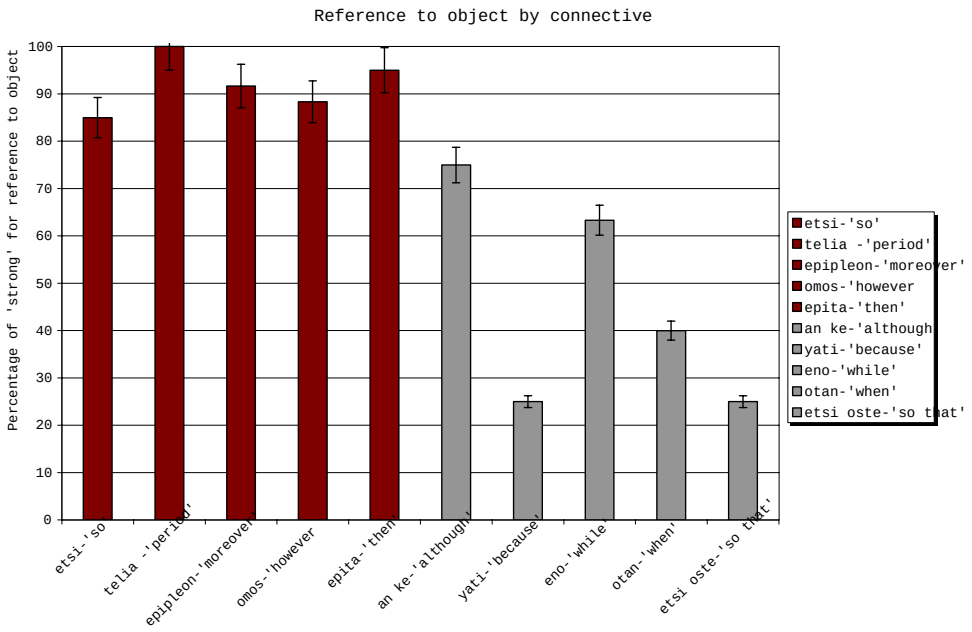

Figure 4.12: Percentage of reference to object with a strong pronoun per connective

we decided to relax it. Although a second pass at the data, with the purpose of studying the clausal location of these antecedents, would be useful, for the purposes of this study it was not crucial. On the contrary, a consistent pattern of reference in main-main sequences including cases with competing antecedents in intervening subordinate clauses provides further evidence that entities introduced in subordinate clauses do not override the salience of the main clause entities. We will provide an example to illustrate this point in the next section.

For the dataset of main-subordinate sequences, we searched for three types of subordinate clauses introduced by the following subordinate conjunctions: *otan* (‘when’), *yati* (‘because’), *oste* (‘so that’). The final dataset included only tokens which fulfilled the requirements described above.

For the dataset with main-main sequences, we randomly selected files from the corpus subdirectories and included tokens that fulfilled the requirements described above. The selection process was terminated when the number of qualifying tokens approximated one hundred.

Two coders, both native speakers of Greek, marked on the dataset the antecedent of the anaphoric expressions. One of the coders was the author and the other was a naive, non-linguist speaker of Greek. As expected for the simple task of identifying antecedents of third person anaphoric expressions (excluding discourse deixis), inter-coder agreement was high, 98%. The few cases of disagreement either involved instances perceived ambiguous by the coders or abstract complex NPs where there was disagreement as to whether the antecedent was the possessor or the possessee. Such cases were excluded from the final dataset.

All in all the final dataset included 88 instances of main-main sequences and 108 instances of main-subordinate sequences broken up as follows: 48 *otan*-clauses, 17 *yati*-clauses and 43 *oste*-clauses.

4.3.6.2 Ranking antecedents and coding

The competing antecedents were ranked according to the following rule given in Section 4.3.2, repeated below for convenience:

Ranking rule for Greek
Empathy>Subject>Indirect Object>Direct object>pro-arb, qis

Under *Empathy* were classified dative subjects of *psych* verbs. Such verbs are easily identified from a normally short exhaustive list that can be enumerated for each language. In our data, we only encountered the verb *like* from this verb category. [3]

[3]By way of demonstration, the expression "I like John" in Greek is glossed as 'me-genitive like-3rd singular John-nominative'. In the Greek example the experiencer of the *psych* verb is analyzed as *subject* despite its genitive marking. Such subjects are known as dative subjects. Modern Greek has lost the dative case whose function is now performed with the genitive case.

All entities introduced in subordinate clauses associated with the main clause are ranked by the same rule but lower than the main clause entities. So, for example, if the evoked entities are main subject, main object, and subordinate subject, the Cf list is ranked as follows: main subject>main object>subordinate subject. It is not clear what the ranking would be in cases of multiple subordinate clauses but this extra ranking specification was not crucial for the current study.

What was crucial for the study was the ranking of entities evoked within Complex NPs. Greek Complex NPs are normally constructed with two nouns: the "possessor" marked with genitive, and the "possessee" marked with nominative, accusative, or, more rarely, genitive, depending on its grammatical role. The "possessee" always precedes the "possessor".[4] Noun-noun modification is not allowed in Greek. In complex NPs animate referents rank higher than inanimates. In all other cases, "possessor" ranks higher than "possessee". For clarification we present an example below, followed by the ranking of the evoked entities.

(97) I mitera tis Marias ipe sto Yani oti o Giorgos den tha erhotan.
the mother of-the Maria said to-the John that the George not would come
'Maria's mother told John that George would not come.'

Maria>mother>John>George

The salience ranking as specified above was then used for a second pass of coding done by the author. For each set of candidate antecedents, the intended referent was marked as either "preferred antecedent", designated as Ap, or "non-preferred antecedent", designated as Anp. The referent of an anaphoric expression was marked as preferred antecedent when it was the highest ranked entity in the set of competing antecedents. The referent of an anaphoric expression was marked as non-preferred antecedent when it was not the highest

[4]We use the terms "possessor" and "possessee" here for convenience, to label the structural position of nouns in complex NPs. However, these terms do not always describe the semantic relationship between two nouns. For example, in "John's participation", "John" can hardly be characterized as the "possessor" but in Greek "participation" would always precede "John" and would be case-marked "genitive".

ranked entity in the set of competing antecedents. In most cases, the set of candidate antecedents included only two candidates so sub-categorizing non-preferred antecedents was not crucial.

Example (98) is demonstrative of cases with the referent of the anaphoric expression in the second main clause marked as Ap. The competing antecedents in (98a) are *ta opla* and *anthropous* because they are the same person and number as the anaphoric *tus* in (98b). The NP *ta opla* is marked as the preferred antecedent because it is the highest ranked element in the list of potential antecedents and is the intended referent of the anaphoric. We report this particular example for the additional reason that it shows that, outside complex NPs, animacy may not be a crucial factor in determining the ranking of entities even in cases where the semantics of the verb taking the referent of the anaphoric as an argument favors the human, in this case, antecedent.[5] Example (99), also, demonstrates a case of reference to Ap. Here, the competing antecedents are both male characters and semantically plausible subjects of the verb *egrafe* "wrote". Notice that the assumed ranking receives further support with this example, since the anaphoric resolves to the subject of the previous clause and not to the most recent, equally plausible, entity.

(98) a. [Ta opla]$_i$ ine kataskevasmena ya na skotonun [anthropus]$_j$.
the guns are made in-order to kill people
'Guns$_i$ are made to kill people$_j$.'

b. Aftos ine o skopos [tus]$_i$.
This is the purpose their
'This is their$_i$ purpose.'

(99) a. [O Turen]$_i$ vriskete apo filosofiki apopsi ston antipoda [tu
the Turen is-placed from philosophical view at-the opposite-side of-the
Popper]$_j$.
Popper
'From a philosophical point of view Tourraine$_i$ is the very opposite of Popper$_j$.'

b. Prosfata [0]$_i$ egrafe oti iparhun dio idon dianoumeni...
recently 0 wrote that there-are two types of-intellectuals

[5]However, see Chapter 5 for a potential counterexample.

'Recently, he$_i$ wrote that there are two types of intellectuals.'

In (100) the referent of the dropped subject in (100b) is marked as Anp. The list of competing antecedents in (100a) contains *PAOK*, the name of a football team, and *Pikulin Ortith*, the name of a player, both being singular and masculine. Number agreement with the verb is sufficient to create ambiguity because Greek verbs are marked for number but not gender. Also, in Greek, subject collective nouns marked singular always take a singular verb. The intended referent is marked as Anp because it is ranked lower than the subject *PAOK*.

(100) a. Ya mia sira praxeon [o PAOK]$_i$ kali [ton Pikulin Ortith]$_j$ na
For a series of-deeds the PAOK summons the Pikulin Ortith to
apologithi amesa,
confess immediately
'PAOK is asking Pikulin Ortith to confess immediately for a series of things,'

b. yati [0]$_j$ ehi prokalesi megisti agonistiki ke ithiki zimia.
because 0 has caused enormous competitive and moral damage
'because [he]$_j$ has caused enormous damage (to the team) both morally and in the championship.'

Finally, example (101) demonstrates that competing antecedents in dependent clauses do not override the salience of main cause antecedents. Note that *i kinonikes igesies* is a plausible candidate for the subject of the verb *pistevun*.

(101) a. [I esiodoxi]$_i$ pistevun oti ehun dimiurgithi [i kinonikes igesies]$_j$ pu
the ambitious believe that have created the social leaderships which
mporun na antiparatethun stin katestimeni exusia.
can to object to-the established leadership
'$[Theambitiousones]_i$ believe that there have been formed
$[socialauthorities]_j$ which can stand up to the established leadership/political power.'

b. $[NULL]_i$ pistevun oti o agonas tus den ehi akrivos kerdithi alla oti
NULL believe that the fight their not has exactly been-won but that
NULL vriskete se "dromo horis epistrofi".
NULL is-found in "road without return"

'$[They]_i$ believe that their fight has not exactly been won but that it is at a point with no return.'

In the next section we present the results of the analysis of the distribution of anaphoric references based on the values Ap and Anp.

4.3.6.3 Results and discussion

Table 4.6 shows the distribution of anaphoric reference. The first column shows the number of times the anaphoric expression resolves to the preferred antecedent, Ap. The second column shows the number of times the anaphoric expression resolves to a non-preferred antecedent, Anp, and the third column summarizes the total number of tokens per condition.

	Ap	Anp	Total
Main-Main	81 (92%)	7 (7%)	88
Main-Subordinate	55(51%)	53(49%)	108

Table 4.6: Reference in Main and Subordinate Clauses

The corpus-based results support the hypothesis that anaphora in main and subordinate clauses does not obey the same rules.[6] Clearly, the preferred antecedent as defined structurally is a strong predictor of the referent of main clause anaphoric expressions whereas the picture appears more complicated in subordinate clauses.

In the main-main condition, the Anp instances have interesting properties in common. Four out of the seven Anp cases involved complex NPs where both competing antecedents belonged to the same complex NP construction. It turned out that the ranking we assumed for complex NPs did not always predict the intended referent correctly. For example, in (102), the ranking of the complex NP *i simetohi tu k. Avramopulu* "Mr Avramopulos's participation" is AVRAMOPULOS>SIMETOHI because *Avramopulos* (current Mayor of Athens) is animate and ranks higher. However, the intended referent of the dropped

[6]Not surprisingly, chi-square gives a highly significant $p < 0.0005$.

subject in the coordinated clause is *simetohi*. An alternative plausible analysis would be to treat (102) as a case of VP coordination in which case the two VPs would share the same subject.

(102) a. Apo afto prokiptei oti [i simetohi]$_i$ [tu k. Avramopulu]$_j$ stis
From this concludes that the participation of-the Mr. Avramopulos at-the
prosehis ekloges epireazi apofasistika tin tihi tis ND
next elections affects decisively the fate of-the ND-(name of political party)
'From this it is concluded that Mr Avramopoulos' participation at the next elections decisively affects the fate of ND

b. ke [0]$_i$ evnoi antistihos to PaSoK.
and 0 favors correspondingly the PaSoK-(name of political party)
'and [it]$_i$ favors PaSoK at the same time.'

However, the same phenomenon was observed in cases in which a VP-coordination analysis cannot be maintained, shown in (103). Again, in this case the anaphoric resolves to *koma* "political party" and not to *Avramopulos* as would be expected. A possible explanation here is that the political party has been erroneously characterized as "inanimate" because, in fact, it denotes a particular group of people. If we treat the political party as "animate" the ranking works as expected.

(103) a. [To koma]$_i$ [tu Avramopulu]$_i$ emfanizete se thesi na anadihthi se
the party of-the Avramopulos appears in position to promote to
paragonta pu tha tropopiisi tus orus tu politiku pehnidiu.
factor which will change the terms of-the political game
'Avramopoulos's political party appears to be in a position to get promoted to a factor what will change the terms of the political game.'

b. Me to 14,7% pu pistonete os 'prothesi psifu' [0]$_i$ katagrafi
with the 14,7% which gets-credited as 'intention of-vote 0 records
axiologi apihisi protu kan anadihthun ta politika haraktiristika tu.
significant appeal before even get-revealed the political characteristics its
'With the 14,7% which gets recorded as 'vote intention' [it]$_i$ records a significant appeal even before its political characteristics are shown.'

The remaining cases of reference to Anp involved complex discourses which either required inferencing or the referent was placed in an adverbial located in the same clause as the anaphoric itself. The following example contains both cases.

(104) a. Legete oti [o 'Mihanismos']$_i$ metaferotan stin Arhea Romi
it-is-said that the 'Mechanism' was-being-transfered to-the Ancient Rome
gia na epidihthi ston Kikerona
in-order that be-shown to-the Kikerona
'It is said that the 'Mechanism' was being transfered to Ancient Rome in order to be shown to Cicero.'

b. ala to plio vithistike exo apo ta Kithira.
but the boat sank outside of the Kithira
'but the boat sank off (the coast of) Kithira.

c. To navagio entopistike stis arhes tu eona
the shipwreck was-located at-the beginning of the century
'The shipwreck was found at the beginning of the century'

d. ke meta tin anelkisi mathimatiki ke arheologi [ton]$_i$
and after the hoisting mathematicians and archaeologists it
anasinthesan.
they-reconstructed
'and after the hoisting, mathematicians and archaeologists reconstructed it$_i$.'

The pronoun in (104d) resolves to *Mihanismos* in (104a). The entity *Mihanismos* is evoked much more recently in (104d) via inference -the hoisting of the Mechanism- and it appears in the same clause as the anaphoric itself. Such complex cases are extremely rare and generally very hard to resolve with a structure-based algorithm.

To complete the analysis of the data, we further broke down the distribution of reference to Ap and Anp for each subordinate clause. The results are shown in Table 4.7. Chi-square shows no significant differences among the three types of subordinate clauses ($p < 0.182$).

These results indicate that the focusing preference of the connectives do not by themselves predict the interpretation of the anaphoric expressions. They are however consistent with Stevenson et al.'s (2000) conclusions that the effect of the connective on the interpretation of pronominals depends on the event structure of the preceding clause, either

	Ap	Anp	Total
Main-when(otan)	23(48%)	25(52%)	48
Main-because(yati)	6(35%)	11(65%)	17
Main-so that(oste)	26(60%)	17(40%)	43

Table 4.7: Distribution of Ap/Anp

reinforcing or reducing the effect of the verb focusing projections. Lack of correlations between subordinate type and anaphora resolution is not surprising since the data included various types of verbs.

4.4 Conclusions

In this chapter we have compared the interpretation of subject pronouns in main and adverbial clauses. Our primary goal was to evaluate the hypothesis that reference inter- and intrasententially is subject to different mechanisms. Based on the proposed model of topic continuity presented in Chapter 3, intersententially, we predicted that subject pronouns would resolve to the highest ranked entity in the preceding unit. This prediction was borne out in three experiments designed to test the interpretation of a subject pronoun in main and adverbial clauses in English and Greek. A corpus study also confirmed the same prediction for naturally occurring data in Greek. With regard to the role of adverbial clauses in topic continuity a study independently conducted by Cooreman and Sanford provide support for the hypothesis that entities evoked in adverbial clauses do not override the salience status of main clause entities, specifically, main clause subjects.

Chapter 5

Relative Clauses

5.1 Introduction

The purpose of this chapter is twofold: a) to investigate the topical status of entities evoked in relative clauses compared with main clauses and b) to test the hypothesis that relative clauses do not form independent processing units in the computation of topic structure in discourse.

To evaluate the topical status of entities evoked in relative clauses compared with entities evoked in the main clause, we perform a group of studies in English and Greek which we will call the *reference test*. For the reference test, we have extracted 300 tokens of *who*-, *which*- and *that*-relatives for English, and 200 tokens for the corresponding *opios*- and *pu*-relatives in Greek. The total of 500 tokens were annotated with the following set of features.

- Subsequent reference to the entity evoked by the head noun
- Type of referring expression used for reference to the head noun entity
- Subsequent reference to "other" entities evoked in the relative clause
- Type of referring expression used for reference to "other" entities evoked in the relative clause

- Restrictive/non-restrictive

This annotation can be analyzed at various levels. First, we wanted to see to what extent entities evoked in relative clauses are featured in the subsequent discourse. We predict that references to the head noun entity will be more frequent than other entities evoked in the relative clause. This is because the head noun entity is evoked in the main clause and therefore is higher ranked than other entities evoked in the relative clause. We also predict that pronouns will be featured as the most preferred type of referring expression for reference to the head noun if the head noun is the highest ranked in the main clause, in most cases, if it is the subject of the main clause. References with full noun phrases are predicted to be more frequent in tokens with a non-subject head noun unless, of course, the subsequent discourse already contains pronominal reference to the highest ranked entity in the main clause. For "other" entities evoked in the relative clause we predict low frequency of subsequent references. In all cases of reference to an "other" entity, references with full noun phrases are expected unless the discourse already contains pronominal references to higher ranked entities. Finally, we have included the feature restrictive/non-restrictive to see if there are any differences in the reference patterns between the two types of relatives. In the syntactic literature, non-restrictive relative clauses are sometimes analyzed on a par with main clauses.

To evaluate the hypothesis that relative clauses do not form independent units in the computation of topic structure, we perform a second group of studies, which we will call the *coherence test*. For this group of studies we have extracted non-restrictive relative clauses in sentence final position. In sentence final position, entities evoked in the relative clause are more recent than entities evoked in the main clause. We have chosen to include only non-restrictive relative clauses because, between the two types of relative clauses, non-restrictives have been claimed to behave more like main clauses. For a total of 200 tokens of English and Greek sentence-final non-restrictive relative clauses, we have computed Centering transitions in two conditions. Consistent with our hypothesis, in the first condition the main and the relative clause are processed as a single unit. In the second

condition, both the main and the relative clause are processd as single units. We predict that in the first condition we will obtain more "coherent" Centering transitions compared to the second condition. The definition of "more coherent" is based on the ranking rule for Centering transitions: Continue>Retain>Smooth-Shift>Rough-Shift.

5.2 English

5.2.1 Background

The study of English relatives includes relatives introduced by the relative pronouns *who* and *which*, and the complementizer *that*. Reduced relatives, free relatives, and relatives with the null complementizer have been excluded. As mentioned in the introduction, the studies of relative clauses includes restrictives and non-restrictives. Informally, a restrictive relative clause helps to identify the referent of the word that it modifies as in (105).

(105) The man that you see is my cousin.

A non-restrictive relative clause is a relative clause which does not aid in the identification of the referent of its head noun, but only provides information about it, as in (106).

(106) John, who passed the test, was elated.

English (also Greek) is not a language in which restrictive and non-restrictive relative clauses can reliably be distinguished syntactically or lexically (i.e., by the form of the relative pronoun). Style guides for English recommend that *who* and *which* are reserved for non-restrictives and *that* for restrictives, and that a comma precede a non-restrictive relative clause, but such conventions are not followed consistently and therefore cannot be used reliably to classify relative clauses.

A formal criterion for distinguishing between restrictive and non-restrictive clauses is offered by (McCawley, 1981), who notes that with a restrictive clause, as in (107a), the

antecedent of the elided material in (107) must include the relative clause if it includes the head noun, whereas with a non-restrictive clause, as in (107b), it cannot include the relative clause. Example (107a) implies that Sam's cat once belonged to Fred, whereas (107b) does not imply that Sam's violin once belonged to Heifetz. However, this criterion can only be applied selectively when coding naturally occurring data.

(107) a. Tom has two cats that once belonged to Fred, and Sam has one.

b. Tom has two violins, which once belonged to Heifetz, and Sam has one.

In the studies that we report in this chapter, we have coded relative clauses as *restricting* versus *non-restricting* depending on their function. *Restricting* relative clauses are necessary to identify the referent of the head noun. In Heim's file card terms, (Heim, 1983), a card may be selected for update only after the relative clause has been processed. *Non-restricting* relative clauses provide additional information about the referent of the head noun, so a card can be selected for update before the relative clause is processed. Intuitively, the two categories correspond closely to the restrictive/non-restrictive divide mentioned above. The terms restricting/non-restricting have been adopted for added clarification and to remind the reader that this classification is not based on formal properties.

5.2.2 The Reference Test

The corpus used for the studies in the following three sections contains naturally occurring data from the following sources:

1) The Brown corpus, available from LDC (http://www.ldc.upenn.edu). Size 1,000,000 words.

2) The Wall Street Journal Corpus, available from LDC. Size: approximately 2,600,000.

3) The Switchboard corpus, available at LDC. The Switchboard corpus consists of telephone conversations between strangers on a pre-assigned topic. Size: 130,650 words. (Total size: 3,000,000 words).

4) “The Adventures of Sherlock Holmes,”, by Arthur Conan Doyle, available from the Project Gutenberg corpus (http://promo.net/pg). Size: 104,693 words.

5) “The Discovery and Settlement of Kentucke,” by John Filson, available from the Project Gutenberg corpus. Size: 8,843 words.

6) “Increasing Human Efficiency in Buisiness,” by Walter Dill Scott, available from the Gutenberg corpus. Size: 61,608 words.

7) “The Agrarian Crusade, A Chronicle of the Farmer in Politics,” by Solon J. Buck, available from the Project Gutenberg corpus. Size: 43,850

The set of features used for the annotation of the data for this group of studies is shown in Table 5.1. For the sake of completeness, we have included in the group of labels characterizing the type of referring expression the categories “NP-assoc” and “implicit”. The category “NP-assoc” has been used for NPs that are anaphorically related to a previously evoked entity via association (e.g., “the house”-“the door”). The category “implicit” has been applied to tokens whose interpretation includes an entity anaphorically related to a previously evoked entity but which is not lexically realized in the subsequent discourse. We do not pursue any further analyses of these two types of reference.

5.2.2.1 *Who*-Relatives

For this study, 100 tokens of *who*-relatives were extracted from “The Adventures of Sherlock Holmes” and the Brown corpus. Each token was annotated with the set of features shown in Table (5.1).

The results of the reference annotation are summarized in Table (5.2). The column “Ref. to the head noun” shows how many times the head noun referent was referenced in the subsequent discourse. For *who*-relatives, we observe that the head noun was subsequently referenced in almost 50% of the tokens. Reference to the head noun referent with a pronoun occurred 14 times, that is, 29% of the total number of references to the head

Features	**Feature values**
Reference to head noun entity	Yes/No
Referring expression for reference to head noun entity	Non-applicable (N/A) NP NP-associative Pronoun/Zero Implicit Other
Reference to "other" entities in subsequent sentence	Yes/No/Non-applicable(N/A)
Referring expression for reference to "other" entities	Non-applicable (N/A) NP NP-associative Pronoun/Zero Implicit Other
Restricting function of relative clause	Yes/No

Table 5.1: Set of annotation features for *who*-, *which*- and *that*-relatives

noun. For 7 out of the 14 instances of pronominal reference to the head noun, the head noun was the highest ranked entity of the sentence, Centering's preferred center (Cp) and therefore the most likely topic of the subsequent discourse (Centering's backward-looking center, Cb). A typical example is given in (108).

(108) a. Barber$_i$, who$_i$ is in his 13^{th} year as a legislator, said there are "some members$_j$ of our congregational delegation in Washington who$_j$ would like to see it (the resolution) passed."

b. But he$_i$ added that none of Georgia's congressmen specifically asked him to offer a resolution.

From the remaining 7 tokens of pronominal reference to the head noun, in 2 cases the reference was in the same sentence, in 1 case the subject of the main clause was already pronominalized, and in the remaining 4 cases there was no competing antecedent in the main clause and syntactic constraints made the realization of the head referent in subject position either impossible or awkard. A typical example of this last type is given in (109)

below.

(109) a. A special presentation was made to Mrs. Geraldine Thompson of Red Bank, who is stepping down after 35 years in the committee.

b. She was also the original GOP national committeewoman from New Jersey in the early 1920s following adoption of the women's suffrage amendment.

For 40% of the remaining references, the preferred referring expression was a full NP. Closer inspection of the relevant tokens reveals that a full NP was used primarily when the head referent was a non-subject in the clause that it was evoked. This was the case for 16 out of the 19 instances of NP reference to the head noun. In 2 of these 16 cases, the head noun was further embedded in a complement clause. A representative example is given in (110). Using a full NP to refer to the head noun in a *subject* position in the subsequent discourse, as is the case in (110), is probably an indication of the writer's intention to promote the head noun referent to topic in subsequent discourse.[1] From the remaining 3 cases of NP reference to the head noun, in two instances the NP expression was across a paragraph boundary, and one instance involved NP reference in a parenthetical *say*-phrase shown in (111).

(110) a. Two tax revision bills were passed.

b. One, by Sen. Louis Crump of San Saba, would aid more than 17,000 retailers$_i$ who$_i$ pay a group of miscellaneous taxes by eliminating the requirement that each return be notarized.

c. Instead, retailers$_i$ would sign a certificate of correctness, violation of which would carry a penalty of one to five years in prison, plus a $1,000 fine.

[1]This strategy has been observed by Turan (1995) for Turkish. In Turkish, Turan has observed that an entity evoked in a non-subject position can be established as a new topic in the subsequent discourse by placing it in the subject position with a full NP. It can, then, be referenced with a pronoun (or null). This strategy is spelled out in the "center promotion rule" which possibly holds for English as well.

Ref. to head noun		Referring expression	
Yes	47	N/A	52
No	52	NP	19
		NP-assoc.	2
		1st/2nd pers.	4
		Pronoun/null	14
		Poss	2
		Other	6 (1 Quant, 5 rel.pron)
Ref. to "other"		Referring expression	
Yes	7	N/A	92
No	88	NP	5
N/A	4	Pronoun	0
		Other Rel pron?	2 (rel. pron+pred NF)

Table 5.2: Reference in *who*-relatives

(111) a. County Supervisor Weldon R. Sheets$_i$, who$_i$ is a candidate for the Democratic gubernatorial nomination, today called for an end to paper ballots in those counties in the state which still use them.

b. The proposal, Sheets$_i$ said, represents part of his$_i$ program for election reforms necessary to make democracy in New Jersey more than a "lip service word".

The column "Ref. to 'other' " shows how many times an entity evoked in the relative clause, other than the head noun, was referenced in the subsequent discourse. As can be seen in Table 5.2, "other" entities were rarely referenced. Reference to an "other" entity occurred in 7 of the 95 tokens. (The N/A category shows how many times the relative clause did not contain any "other" entity.) For *who*-relatives, there were no instances of pronominal reference to an "other" entity. In 5 of the 7 cases of reference to an "other" entity, the referring expression used was a full NP. A representative example is given in (112). Example (112) is especially interesting because it contains two male referents. The first one is introduced as the main clause subject. The other one is introduced as the main clause object and is modified by the relative clause. The head noun referent is the only third person male entity in the relative clause, which is then referenced in the subject

position of the subsequent sentence with a full NP. Crucially, if we replaced the full NP "Mr. Breeden" with the pronoun "he" in (112c), the pronoun would be interpreted as "Mr. Brady", who is the subject of the main clause. It would be hard to interpret a pronoun in that position as "Mr. Breeden" despite the fact that "Mr. Breeden" is the subject of the relative clause and, additionally, has already been pronominalized in the complement clause in (115).

(112) a. In testimony to the Senate securities subcommittee, **Mr. Brady** disputed the view of SEC Chairrman **Richard Breeden**,

b. who told a House panel Wednesday that he doesn't want the ability to halt the market.

c. **Mr. Breeden** contended that discretionary power could have an impact on the markets if rumors were to circulate about when the exchanges might be closed.

d. **He** added that the president already has the power to close the markets in an emergency.

5.2.2.2 *Which*-Relatives

The same set of annotation features and corpus used for the annotation of *who*-relatives were also used for the annotation of *which*-relatives. The results of the reference annotation of *which*-relatives are summarized in Table 5.3. The colunn "Ref. to head noun" shows how many times the entity evoked by the head noun was referenced in the subsequent sentence. In *which*-relatives the head noun referent is referenced in the subsequent discourse in 17% of the tokens. Comparing these results with the results of *who*-relatives, we observe that reference to the head noun is much less frequent for *which*-relatives. Given that *who*-relatives modify NPs that evoke human referents and *which*-relatives modify NPs that evoke non-human referents, this lower frequency probably reflects a tendency for human referents to be more salient in most discourses. The column "Referring expression" shows the type of referring expression used for reference to the head noun. A

pronoun was used for reference to the head noun in 7 of the 17 cases. Looking closer at these 7 tokens, we observe the following distribution. In 4 cases the head noun was the highest ranked entity in the sentence. A typical example is (113), where the antecedent of the pronoun *it* in (113b) is the subject of the preceding sentence (113a). In another case, the pronoun appeared in the same sentence containing the relative clause. The remaining two pronominal references were harder to analyze. In one case, the pronoun appeared in an elliptical utterance and, in the other, it occurred in a discourse containing complement clauses and switches from indirect to quoted speech. The relevant example is given in (114). Analyzing this pronominal use requires a better understanding of topic management and its interaction with discourse structure in quoted speech. We will leave this question open for future research.

Ref. to head noun		Referring expression	
Yes	17	N/A	82
No	82	NP	8
		Pronoun	7
		Poss	1
		Disc. deixis	1
Ref. to "other'		Referring expression	
Yes	39	N/A	65
No	58	NP	13
N/A 6		NP-assoc	1
		1st/2nd pers.	11
		Pronoun	9 (2 poss)

Table 5.3: Reference in *which*-relatives

(113) a. **The road**$_i$ in **which**$_i$ we found ourselves as we turned round the corner from the retired **Saxe-Coburg Square**$_j$ presented as great a contrast to **it**$_j$ as the front of a picture does to the back.

b. **It**$_i$ was one of the main arteries which conveyed the traffic of the City to the north and west.

(114) a. The only day they “have a chance to compete with large supermarkets is on Sunday,” the council’s resolution said.

b. The small shops “must be retained, for they provide essential service to the community,” according to **the resolution**$_i$, **which**$_i$ added that they “ also are the source of livelihood for thousands of our neighbors”.

c. **It**$_i$ declares that Sunday sales licenses provide “great revenue” to the local government .

A full NP was used for reference to the head noun in 8 tokens. In 7 of these 8 tokens the head referent was not the highest ranked entity in the sentence containing the relative clause (6 non-subjects and 1 subject in a complement clause). The remaining token, shown in (115) should probably best be classified as NP-assoc. The head referent is a complex NP containing noun-noun modification.[2]

(115) a. A House Committee which heard his local option proposal is expected to give it a favorable report, although the resolution faces hard sledding later.

b. The house passed finally, and sent to the Senate, a bil extending the State Health Department’s authority to give planning assistance to cities.

The column “ ‘other’ entities” shows how many times an entity other than the head noun is subsequently referenced. The row “N/A” under “Ref. to ‘other’ ” gives the count

[2]Although we have not studied the status of entities evoked in noun-noun modifications, it seems that in such NPs a subsequent pronoun can only refer to the entity evoked by the entire NP and not to any of the individual entities evoked by each constituent noun even, certainly not when the individual entities are competing antecedents.

(1) a. The garden table is set up.

b. It$_{thegardentable}$/*It$_{thegarden}$ ...

(2) a. The Clinton administration was very popular.

b. He$_{clinton}$ was happy.

of relative clauses that did not evoke any "other" entities. An "other" entity was subsequently referenced in 39 cases. The column "Referring expression" shows the type of referring expression used for reference to "other" entities. An "other" entity was subsequently referenced in 35 tokens. A pronoun was used in 9 tokens, of which in 7 cases the "other" entity was already pronominalized in the relative clause. In all of these cases, the discourse contained other first person pronominal references and no competing antecedents. A typical example of this category is shown in (116).

(116) a. Indeed, apart from the nature of the investigation which **my friend** had on hand, there was something in **his** masterly grasp of a situation, and **his** keen, incisive reasoning,
which made it a pleasure to me to study **his** system of work, and to follow the quick, subtle methods
by which **he** disentangled the most inextricable mysteries.

b. So accustomed was I to **his** invariable success that the very possibility of **his** failing had ceased to enter into my head.

From the remaining 2 cases, in one the pronoun appeared *intra*-sententially (shown in (118)) and the other is shown in (117). In this case, the main clause in (117a) contains a *there*-construction and the main clause as a whole provides the setting against which the main character in the story is presented. Note that the only entities evoked in the main clause are furniture items. In earlier work on narratives, Labov (1972) made a distinction between narrative and non-narrative clauses and mapped that distinction to main and subordinate clauses. According to Labov, subordinate clauses by definition are not narrative clauses because their order relative to the main clause can be reversed without disturbing the "temporal sequence of the original semantic interpretation". Later, Reinhart (1984) defined "foreground" as the sequence of narratives as defined by Labov, and suggested that "a powerful means for marking background is the use of syntactic embedding". However, both Reinhart and later Thompson (1987) acknowledged that writers can manipulate the foreground-background relations such that a narrative clause can function as background

and a subordinate clause as foreground. Reinhart views the use of a subordinate clause for foregrounding as a stylistic option of the writers while Thompson attempts to show that when this happens an additional discourse function is performed. Whatever the correct analysis of the relationship between foregrounding/backgrounding and the choice of syntactic form, the above discussion leads to the possibility that the main character introduced in the relative clause in (117a) receives topical status because the preceding main clause is simply processed as a description of the room in which the main character was located. An expectation that the *small man* is going to be central in the subsequent discourse is facilitated by the fact that no other character is introduced in the sentence and is probably cued by the non-canonical post-verbal position of the subject in the relative clause.[3]

(117) a. There was nothing in the office but a couple of wooden chairs and a deal table, behind which sat **a small man**$_i$ with a head that was even redder than mine.

b. **He**$_i$ said a few words to each candidate as he came up, ...

(118) **The law**$_i$ **which**$_i$ governs **home rule charter petitions**$_j$ states that **they**$_j$ must be referred to the chairman of the board of canvassers for verification of the signatures within 10 days.

In the remaining 13 cases of reference to an "other" entity, the referring expression used was a full NP. A typical example is shown in (119).

(119) a. A difference of opinion arose between Mr. Martinelli and John P. Bourcier, town solicitor, over the exact manner in which **the vote**$_i$ is handled.

b. Mr. Martinelli has, in recent weeks, been of the opinion that a special town meeting would be called for **the vote**$_i$, while Mr. Bourcier said that a special election might be called instead.

For *which*-relatives we also performed a supplementary study using data from the Switchboard corpus which consists of spoken dialogues. In the previous study, 63 of

[3]It is, of course, possible that the post-verbal position of the subject is due to the heaviness of the NP.

the 100 tokens were extracted from *The Adventures of Sherlock Holmes*. Some instances of *which*-relatives in this corpus sounded pedantic and somewhat obsolete. For this reason we decided to check if there are any different patterns in colloquial English. For reasons that will become obvious shortly, the results do not address directly the questions that concern this thesis. However, they reveal an interesting pattern of the use of *which*-relatives in spoken English, which we will discuss briefly below.

The Switchboard corpus contained 68 tokens of *which*- relatives. Table 5.4 shows the set of annotation tags that turned out to be revealing in this corpus and the results of the annotation. The head noun column shows the syntactic type of the head noun. Examples of the categories "NP", "PP", and "clause" are given in (120), (121), and (122) respectively. The category "other" includes cases of dysfluencies which made it hard to identify the grammatical category of the head noun (e.g., (126) and (127)). The column "RC Pred. Type" characterizes the predicates the relative clauses. As can be seen in the table, *which*-relatives contain either a regular verb, (e.g., (123) and (128)), or the verb *-to be* followed by either an adjective or a predicational NP (e.g., (124)). The category "dysfluency" contains cases where the relative clause was started but was interrupted before completion, e.g., (125).

Head noun	Tokens	RC Pred. Type	Tokens
NP	32	Verb	19
PP	3	Copula	33
Clause	29	Dysfluency	16
Other	4	0	

Table 5.4: Switchboard data for *which*-relatives

Interestingly, we observed that in all cases of a nominal head referent the predicate in the relative clause was the verb *to be* followed by an adjective or a predicational NP and the relative clause contained no other referents. On the other hand, in all the cases that the predicate of the *which*-relative was a regular verb, the relative clause was preceded by a comma, and it modified a clausal constituent. Finally, *which*-clauses very often appeared in dysfluent speech (interrupted abruptly after the relative pronoun and then repaired).

While these patterns are interesting by themselves we did not pursue the "reference test" annotation because in cases of NP modification there were no "other" entities evoked in the relative clause and cases of clausal modification are hard to evaluate in the framework of this thesis. Clausal referents are referenced with discourse deixis whose nature in a model of anaphora and/or topic continuity is unclear.

(120) B.13: [Laughter] Well, well, I can, uh, I can understand that weather. We're having, uh, we are, i-, we had a late, um, um, an ice storm here, about two weeks ago, <REL REF=2 XP=NOMINAL RC=DYSFLUENCY> which </REL> is, you know, um, and, and it's, and it's, they're calling it the worst ice storm in like the last hundred years. And, um, and then to the point where about three hundred thousand people in our, in our area lost power and,

A.14: Isn't it a little late in the season for that type of ice storm though?

(121) Um, we had, uh, um, my wife and I own our own home, and we have a very big willow tree in our backyard, that, um, that ninety percent of <REL REF=3 XP=PREP RC=VERB> which </REL> came down in one night. Things like that. It was a, it, it definitely is very, um, it wa-, it was-, it wasn't supposed to be as bad. They, they have, I, I guess, apparently, you know, ice storms here occasionally, but never anything this bad this late in the season.

(122) B.31: Oh, sheesh, yeah, um, um, it's right about now when they're talking about hitting thirty-five and forty degree days, we're thinking heat wave. Because, occasion-, what'll, what'll happen here is, with the wind chill it can go below zero quite often, <REL REF=4 XP=CLAUSAL RC=COPULA> which </REL> is not, not fun. I personally hate it. I'm only here for school.

(123) A.19: And the judge was extremely concerned for our welfare, if we were well, if we were comfortable, and, uh, things of that nature <REL REF=7 XP=CLAUSAL RC=VERB> which </REL> made us feel good.

(124) A.25: That's right. About two months ago OMNI MAGAZINE, <REL REF=9

XP=NOMINAL RC=COPULA> which </REL> is a, a science magazine – B.26: Uh-huh.

A.27: – are you familiar with that?

B.28: Yeah.

A.29: Ran an article, um, really restating some conclusions of some, some, uh, programs or some investigations or research, whatever you want to call it, into the, the actual physical differences between men and women, and questioning whether those physical differences, uh, accounted [cough] in background for some of the differences in ways women handled stress as [cough] in background opposed to men.

(125) B.24: [Laughter] Yeah, and I, I thought that was particularly interesting in the, the Gulf War, that there were pieces of information that, that were apparently, uh, leaked just as a, as, as, uh, a ploy.

A.25: Uh-huh.

B.26: <REL REF=34 XP=CLAUSAL RC=DYSFLUENCY> Which </REL> was, uh, I, I find that fascinating that, uh,

A.27: Yeah, I do, too. Do you ever listen to the radio or any,

B.28: I listen to K R L D, and, uh, uh, K L I F, the news talk radio.

A.29: Uh-huh.

(126) A.135: Uh, but you could fill a whole bunch of, uh, holes with these things. I used to, I used to advertise buying wheat pennies. Um, I'd give a dollar a roll, <REL REF=57 XP=OTHER RC=DYSFLUENCY> which </REL> two cents apiece <REL REF=58 XP=CLAUSAL RC=COPULA> which </REL> is basically overpriced.

(127) B.64: Well, I know my parents like to camp a lot, and they, uh, they've been going to Gulf Shores, Alabama.

A.65: Oh, uh-huh.

B.66: And, uh, <REL REF=66 XP=OTHER RC=COPULA NOTE= AMBIGUOUS NOTE=PRONOUN> which </REL> is really, really neat. They said they've got the white beaches and the sand,

A.67: Yes.

B.68: and it's not real populated so they can, they feel like they're in the outdoors and still close to the ocean.

(128) A.19: And the judge was extremely concerned for our welfare, if we were well, if we were comfortable, and, uh, things of that nature <REL REF=7 XP=CLAUSAL RC=VERB> which </REL> made us feel good.

5.2.2.3 *That*-Relatives

The corpus used for the annotation of *that*-relatives includes *The Adventures of Sherlock Holmes* by Arthur Conan Doyle, *Increasing Human Efficiency in Business* by Walter Dill Scott, *From the Discovery and Settlement of Kentucke* by John Filson, *Minnesota Historical Society* by Solon J. Buck, and parts from the Brown Corpus and the Wall Street Journal. The same set of features used for the annotation of *who*- and *which*-relatives was also used for the annotation of *that*-relatives. The results of this annotation are presented in Table 5.5.

The column "Ref. to head noun" shows how many times the head noun referent was referenced in the subsequent sentence. As with *which*-relatives, reference to the head noun of a *that*-relatively is low, (18%). The head noun was subsequently reference in 18 tokens. An anaphoric expression other than a NP was used 3 times. In two cases, the anaphoric was the null subject of a participial form occurring in the same sentence that contained the relative, e.g., (129), and one case of "one" intra-clausal, associative anaphora shown in (130). So, all 3 cases were therefore instances of anaphora intra-*clausaly*. There were no pronominal references to the head noun in the subsequent sentence.

(129) ..., there being no more forts of white men in the country, except at the Falls, a considerable distance from these, and all taken collectively, were but a handful

Ref. to head noun		Referring expression	
Yes	18	NA	81
No	81	NP-assoc	7
		NP	5
		Pronoun/zero/one	3
		Disc. deixis	1
Ref. to "other"		Referring expression	
Yes	27	N/A	71
No	55	NP-assoc	14
N/A	17	NP	9
		Pronoun	1
		Rel. pron	1
		Zero	1

Table 5.5: Reference in *that*-relatives

to **the numerous warriors**$_i$ that were every where dispersed through the country, **null**$_i$ intent upon doing all the mischief that savage barbarity could invent.

(130) Frequently it is not **the team**$_i$ with the greater muscular development or speed of foot **that**$_i$ wins the victory, but **the one**$_{i-assoc}$ with the more grit and perseverance.

The most frequent expression for reference to the head noun was a full NP. Consistent with our earlier results on *who*- and *which*-relatives, a full NP was used when the head noun had a non-subject grammatical role in the main clause, as in (131). This was the case for all 5 instances of NP reference to the head noun.

(131) a. Mr. Doherty kept only those muscles tense that were used in the game.

b. The muscles especially necessary for tennis were also, so far as possible, kept lax except at the instant for making the stroke.

Turning to "other" referents in *that*-relatives, again we observe that the reference pattern for "other" referents is similar to that of *which*-relatives with approximately 32% reference to "other" in subsequent discourse (27 tokens). From those cases, one instance

contained pronominal reference to the head noun intrasententially, (132), another contained zero anaphora intraclausaly and a third case included anaphora to the head noun with a relative pronoun (with no further reference in the following sentence), again intrasententially. So, as in the case of the head noun, there were no pronominal references to an "other" entity in the subsequent sentence. Excluding "NP-assoc" from the analysis, in the remaining 9 tokens that an "other" entity was referenced, the referring expression was a full NP.

(132) And being informed, by **two of their number**$_i$ **that**$_i$ went to their **town**$_j$, that the Indians had entirely evacuated **it**$_j$, we proceeded no further and ... .

5.2.2.4 Restricting versus non-restricting relative clauses

Restricting and non-restricting relative clauses demonstrate a significant difference with respect to the discourse model of entity update. The head noun of a non-restricting relative clause represents an entity that is introduced in the discourse before the relative clause is processed. It follows that the entity represented by the head noun of a non-restricting clause is available for subsequent reference as soon as the head noun is processed (or even earlier if the head noun had already been introduced in the discourse). In restricting relative clauses, on the other hand, the head noun represents an entity whose referent can be identified only after the relative clause is processed. To evaluate if the different functions performed by restricting and non-restricting relative clauses affects the salience of the entity represented by the head noun, we looked at the reference patterns associated with the head nouns of restricting and non-restricting relative clauses across *who-*, *which-* and *that*-relatives. Because subject head nouns evoke entities that are already salient by virtue of their grammatical role in the main clause, we focused only on object head nouns, including head nouns appearing in prepositional phrases that served as an argument to the verb (e.g., verbs that take a locative argument and indirect objects in ditransitive constructions). For each token of an object relative clause, we summarized the frequency of subsequent reference to the entity represented by the head noun and the type of referring

expression. The results are shown in Table 5.6.

	Ref. to head noun	Referring Expression
Non-restricting	Yes=10 (32%) No= 21 (68%) Total=31	NP=5 (50%) Pron.=5 (50%)
Restricting	Yes=9 (17%) No=44 (81%) Total=53	NP=9 (100%) Pron.=0 (0%)

Table 5.6: Reference to head nouns of restricting versus non-restricting relative clauses

Two differences between restricting and non-restricting relative clauses emerge from Table 5.6. First, head noun objects of non-restricting relative clauses are more likely to be referenced than head noun objects of restricting relative clauses. Although the number of tokens in our corpus was too small to draw any definitive conclusions, they do however show a tendency for the head noun of a non-restricting clause to be more frequently referenced. Given the discourse status of the head noun in these cases (an entity that enters the discourse model before processing the relative clause), the non-restricting relative clause is already an elaboration on the entity evoked by the head noun, albeit within a single topic update unit. The second difference between restricting and non-restricting relative clauses is related to the referring expression used for reference to the head noun. As Table (5.6) shows, a pronoun was used for reference to the head noun of non-restricting clauses in the subsequent discourse half of the time whereas no pronominal references were found for reference to the head noun of the restricting clauses. Again, the numbers are too low for definitive conclusions but they reveal a tendency to disprefer a pronoun for reference to the object head noun of a restricting relative clause. A question tightly related to our main question on the topical status of entities is to what extend the pronominal references to the object head noun are reflections of topicality.

Looking closer at the five pronominal references, we observe that in one case the relative clause is contained in a complement clause, in another the pronominal reference

occurs in reported speech, and in a third case the pronominal reference appears intrasententially. In the fourth instance, the interpretation of the pronoun is ambiguous between the referent of the head noun and another entity evoked earlier in the discourse. The relevant example is given in (133). In this example, the pronoun *it* is ambiguous between referring to *one* of the two tax revision bills evoked in (133b) or the head noun *a certificate of correctness* evoked in (133c). We find that the most likely interpretation is *it* is *one* of the tax revision bills evoked earlier (such long-distance antecedents have also been discussed in the literature as antecedents whose referents are in *global focus*, e.g., Grosz and Sidner (1986), Hitzeman and Poesio (1998)). If, this is the intended interpretation, then this example calls for a better understanding of the effect of the hierarchical structure of discourse on pronominal interpretation along the lines suggested by Grosz and Sidner (1986) (reviewed in Chapter 3). In this case, the sentence shown in (133a) creates the expectation that two sub-discourses will follow, one centering one of the two revision bills that were passed and one centering the other. These expectations are in fact met. If *one of the two bills* is interpreted as the topic of the first sub-discourse then the use of the pronoun *it* in (133d) is not surprising. However, the ambiguity created by the introduction of the competing head noun in the preceding sentence is indicative of the tension that is created when local and global centers of attention are in competition.

(133) a. Two tax revision bills were passed.

b. **One**, by Sen. Louis Crump of San Saba , would aid more than 17,000 retailers who pay a group of miscellaneous excise taxes by eliminating the requirement that each return be notarized.

c. Instead, retailers would sign **a certificate of correctness**, violation of which would carry a penalty of one to five years in prison, plus a $1,000 fine.

d. **It**$_i$ was one of a series of recommendations by the Texas Research League.

e. The other bill...

In the remaining case of pronominal reference, the object head noun appears to represent the topical entity of the sentence although it is not introduced in subject position. As discussed in Section 5.2.2.1, this case, repeated here as (134), introducing the head noun in a subject position would be awkard and rare from a syntactic point of view, especially if the relative clause is to remain in sentence-final position. Another consideration is the information status of the head noun. Despite the definite description, the referent of *Mrs. Geraldine Thompson* is discourse new and possibly hearer new, too, as indicated by the prepositional modifier *of the Red Bank* (Prince's *brand new anchored* type of entity, or, for some readers *unused*). As suggested in Prince (1981a), it is possible that the post-verbal position of the head referent is simply a reflection of the tendency for new entities to avoid the subject position. The topical status of the head noun referent, on the other hand, can be established by the fact that the head noun referent is the only human referent and that the agent of the main predicate is suppressed by passivization.

(134) a. A special presentation was made to **Mrs. Geraldine Thompson of Red Bank**$_i$, **who**$_i$ is stepping down after 35 years on the committee.

b. **She**$_i$ also was the original GOP national committee woman from New Jersey in the early 1920s following adoption of the women's suffrage amendment.

In light of these observations, we tentatively conclude that the type of relative clause affects the likelihood for subsequent reference to the head noun entity. However, the type of relative clause does not appear to have a significant effect on the topical status of the head noun referent. In both restricting and non-restricting relative clauses, the head noun referent does not acquire topical status in the processing unit that it is first evoked, and, therefore, does not license the use of a pronoun for reference in the subsequent sentence. With only few exceptions, referents of head nouns in both restricting and non-restricting relative clauses need to be established as topics in the subsequent main clause before they can be referenced with a pronoun intersententially (except, of course, for cases where the established topic is already pronominalized).

5.2.3 The Coherence Test

5.2.3.1 Design and results

For the *coherence test*, 100 tokens of non-restrictive relative clauses were extracted from the Wall Street Journal corpus. The tokens included in this study were selected according to the following criteria:

1. The relative clause is preceded by a comma.

2. The sentence following the relative clause includes reference to at least one entity evoked in the sentence containing the relative clause, either in the main clause or in the relative clause.

3. The relative clause is the sentence-final position.

For each token, Centering transitions were computed in two versions. In version A, two Centering transitions were computed: one for the sentence containing the relative clause and one for the sentence following the relative clause. In version B, three Centering transitions were computed. One for the first sentence excluding the relative clause, one for the relative clause and one for the sentence following the relative clause. The results of the computation of Centering transitions in the two conditions are shown in Table (5.7). The column "more 'coherent' transition" contains the number of cases where a more "coherent" transition was computed in the final sentence in version B. The column "less 'coherent' transition" shows how many times a less "coherent" transition was computed in the final sentence in version B, and, finally, the column "no effect" shows how many times the same transition was computed in both versions A and B. The relevant degree of coherence is specified according to the Centering transitions rule: Continue>Retain>Smooth-Shfit>Rough-Shift.

A typical example of the category "less 'coherent' transition" is given in (135) and (136). In this case, version B, (136) yields a Rough-Shift transition, which is ranked less "coherent" than the Continue transition computed in version A, (135) for the same last

More "coherent" transition	Less "coherent" transition	No effect	Total
13	46	41	100

Table 5.7: Effect of English non-restrictive relatives on Centering transitions

sentence. A typical example of the category "more 'coherent" transition" is given in (139) and (140). In this case, version B, (140) yields a Continue transition, which is ranked more "coherent" than the Smooth-Shift transition computed in version A, (139) for the same last sentence.

(135) VERSION A

(A disaffected, hard-drinking, nearly-30 **hero**$_i$ sets off for snow country in search of an elusive sheep with a star on its back at the behest of a sinister, erudite mobster with a Stanford degree.)

a. **He**$_i$ has in tow his prescient girlfriend, whose sassy retorts mark her as anything but a docile butterfly.

Cb=hero

Cp=hero

Tr=Continue

b. Along the way, **he**$_i$ meets a solicitous Christian chauffeur who offers the hero God's phone number;

Cb=hero

Cp=hero

Tr=Continue

(136) VERSION B

(A disaffected, hard-drinking, nearly-30 **hero**$_i$ sets off for snow country in search of an elusive sheep with a star on its back at the behest of a sinister, erudite mobster with a Stanford degree.)

a. **He**$_i$ has in tow his prescient **girlfriend**$_j$,

Cb=hero

Cp=hero

Tr=Continue

b. **whose**$_j$ sassy retorts mark her as anything but a docile butterfly.

Cb=girlfriend

Cp=girlfriend

Tr=Smooth-Shift

c. Along the way, **he**$_i$ meets a solicitous Christian chauffeur who offers the hero God's phone number;

Cb=none

Cp=hero

Tr=Rough-Shift

Examples such as the above, repeated as (137) below for ease of reference, provide strong evidence that the relative clause does not behave as an independent unit in the computation of topic continuity. If we processed the relative clause as an independent unit, we would be faced with three problems. First, we would process the *girlfriend* as the most likely topic of the subsequent discourse when, in fact, she is not even mentioned in the following sentence. Second, we would be left with a sequence of two processing units, the relative clause and the following sentence, that have no links. Such discourses are predicted to be hard to process as they place on the hearer the extra burden to infer and establish a link in order to incorporate the meaning of the current unit to the preceding one. Third, we would have to revise substantially our theories of pronominalization. If the most salient entity after processing the relative clause is the *girlfriend* then we have no explanation as to why a pronoun was used for reference to the *hero* of the preceding discourse. If, however, the relative clause indeed belongs to the same unit as the main clause, then none of the above problems arise. The highest ranked entity in that unit, the *hero*, is processed as the most likely topic of the discourse, an expectation that is met with the use of a pronoun in subject position in the subsequent sentence. The same entity

also serves as a link between the two processing units. Indeed, the results from this study challenge the view that the relative clause is processed as an independent unit as such an assumption yields a more "coherent" transition only in 13 of the 100 cases that we looked at, whereas for 41 cases, it yields a less "coherent" transition. Let us look more closely at instances of a more "coherent" transition.

(137) (A disaffected, hard-drinking, nearly-30 **hero**$_i$ sets off for snow country in search of an elusive sheep with a star on its back at the behest of a sinister, erudite mobster with a Stanford degree.) **He**$_i$ has in tow his prescient **girlfriend**$_j$, **whose**$_j$ sassy retorts mark her as anything but a docile butterfly. Along the way, **he**$_i$ meets a solicitous Christian chauffeur who offers the hero God's phone number;

Let us now turn our attention to cases where processing the relative clause as an independent unit yields a more "coherent" transition. For ease of reading, we, also, present the relevant discourse in prose style below in (138). In this discourse, processing the relative clause as an independent unit yields a Continue transition, more "coherent" than the Smooth-Shift transition computed for the subsequent sentence when the relative is processed as a single unit with the main clause. On closer inspection, however, we observe that the head noun, which appears in subject position in the relative clause, is referenced in the subsequent discourse with a full NP, a strategy that, as we have noted earlier, is possibly used to signal switch to a new topic. If, indeed, the relative clause was independent, then the referent of the head noun would have been established as the current topic in the relative clause. In that case, we would expect to see a pronominal reference to this entity in the subsequent discourse, which would appropriately reflect a Continue on the same topic. Note that the full NP does not provide any further information for the referent and the sentence that contains it does not cross a paragraph boundary. These would be two potential reasons for using an NP when a pronoun would otherwise be appropriate. A third possibility would be that a full NP is associated with a time change in discourse structure (McCoy & Strube, 1999). However, if any of the above possibilities applied in this case, we would still expect to be able to replace the full NP with a pronoun and get the

same interpretation. This is not the case. According to the judgment of the native speakers that we consulted, the preferred interpretation for a subject pronoun in the last sentence would be *Wilson Taylor*, i.e., the subject of the main clause in the preceding discourse. We conclude, therefore, that even the small number of cases identified under the more "coherent" condition are not necessarily counterevidence to the hypothesis that relative clauses do not form independent processing units in the computation of topic continuity.

(138) **Wilson H. Taylor**$_i$, president and chief executive officer of this insurance and financial services concern, was elected to the additional post of chairman. **Mr. Taylor**$_i$, 45 years old, succeeds **Robert D. Kilpatrick**$_j$, 64, **who**$_j$ is retiring, as reported earlier. **Mr. Kilpatrick**$_j$ will remain a director.

(139) VERSION A

a. **Wilson H. Taylor**$_i$, president and chief executive officer of this insurance and financial services concern, was elected to the additional post of chairman.
Mr. Taylor$_i$, 45 years old, succeeds **Robert D. Kilpatrick**$_j$, 64, who is retiring, as reported earlier.
Cb=Taylor
Cp=Taylor
Tr=Continue
Mr. Kilpatrick$_j$ will remain a director.
Cb=Kilpatrick
Cp=Kilpatrick
Tr=Smooth-Shift

(140) VERSION B

a. **Wilson H. Taylor**$_i$, president and chief executive officer of this insurance and financial services concern, was elected to the additional post of chairman.
Mr. Taylor$_i$, 45 years old, succeeds **Robert D. Kilpatrick**$_j$, 64,
Cb=Taylor

Cp=Taylor

Tr=Continue **who**$_j$ is retiring, as reported earlier.

Cb=Kilpatrick

Cp=Kilpatrick

Tr=Smooth-Shift **Mr. Kilpatrick**$_j$ will remain a director.

Cb=Kilpatrick

Cp=Kilpatrick

Tr=Continue

5.3 Greek

5.3.1 Background

In Greek, relative clauses can be introduced either by the relative pronoun *o opios* or by the complementizer *pu* (null complementizers are not allowed). The expression *o opios* must agree in gender and number with the noun phrase it modifies and it must be in the case appropriate to its grammatical role in the relative clause. In (141), for example, *i opii* is plural masculine agreeing with the head noun *anthropous* and it is in the nominative case because it is the subject of the relative clause.[4] It is also possible for the relative pronoun to be the noun phrase complement of a prepositional phrase. In such cases the preposition is always followed by the *o opios*-paradigm. A preposition cannot combine with *pu*. However, a relative clause introduced by *pu* may *stand* for a prepositional phrase, as in (142). This construction is more common in spoken Greek than in written text.

(141) Null katalaveno panta tus anthropus i opii lene psemata.
Null understand always the people-ACC the who-NOM tell lies
'I can always tell people who lie.'

[4]The examples and the grammatical information presented in this section are based on Holton, Mackridge, and Philippaki-Warburton (1997).

(142) To kuti pu null evale to doro.
The box that null put the present
'The box in which s/he put the present.'

The relative clauses introduced by the relative complementizer *pu* may contain no NPs referring to the head noun if it is the subject, object, or indirect object but, with a direct or indirect object, it is possible to use the corresponding clitic pronoun in the relative clause. In fact, there is a strong preference to do so with indirect objects. The relevant examples are shown below.

(143) O nearos pu ton idame stin taverna htes
the young-man that him saw at-the tavern yesterday
'The young man whom we saw in the tavern yesterday.'

(144) To pedi pu tu dosame to vivlio
the child that it gave the book
'The child to whom we gave the book.'

In Greek, as in English, relative clauses can be restrictive or non-restrictive. The function of a restrictive clause is to further specify a noun in such a way so as to enable the hearer to identify its referent. A non-restrictive relative clause is used to add some additional piece of information about the noun. That information is not crucial for the identification of the referent.

(145) a. O kathigitis pu mas ekane istoria itan poli kalos.
the professor that of-us did history was very good.
'The teacher who taught us history was very good.'

b. O kenurgios mas kathigitis, pu spudase sto Reading, ine poli kalos.
the new our professor, that studied in-the Reading, is very good.
'Our new professor, who studied at Reading, is very good.

In general, the relative clause immediately follows the head noun, as in (145), but it is also possible for both restrictive and non-restrictive relative clauses to be separated from the head noun by a verb. In such cases the restrictive relative clause is more likely to be

introduced by *pu* than *o opios*, e.g., (146a) while the non-restrictive is much more likely to be introduced by *o opios* than *pu*, e.g., (146b).

(146) a. I kopela efige pu null ithele na su milisi.
the girl left that null wanted to you talk
'The girl who wanted to talk to you left.'

b. Irthe ki o Gianis na ti di, o opios den ine poli filos tis.
came and the Gianis to her see, the who not is very friend her.
'John came to see her, too, who is not a close friend of hers.'

This separation of the relative clause from its antecedent is more likely to occur when the relative clause is long. Especially when the relative clause is separated by the head noun, but in other cases too, it is possible to use the expression *o opios* not as a pronominal but as a determiner followed by a repetition of the head noun, as in (147). Such constructions are possible only with non-restrictive relative clauses.

(147) Null skorpai ta lefta tu edo ki eki, ta opia lefta ta vgazi
Null throws-around the money his here and there, the which money them makes
me megalo kopo.
with big effort.
'He throws his money around, which money he makes with a lot of toil.'

Non-restrictive relative clauses tend to be introduced with *o opios* more often than with *pu*. Also, non-restrictive relative clauses introduced with either *o opios* or *pu* tend to be preceded by a comma. However, none of the two conventions are strictly followed as is clear from the studies reported in the following sections. The presence of a comma is a strong indicator that the relative clause is a non-restrictive one but the absence of a comma is not consistently associated with a restrictive relative clause.

For lack of a formal criterion of distinguishing between restrictive and non-restrictive relative clauses, in the studies that follow we use the terms *restricting* and *non-restricting* to distinguish between relative clauses which are crucial for the identification of the head noun referent and relative clauses which provide additional information for a referent

already known to the hearer/reader. As in English, reduced and free relatives have been excluded for the studies.

5.3.2 The Reference Test

The corpus used for the studies in the following two sections contains data from the following sources:

> 1) *To fantasma* by John Dickson Carr. Name of the Greek translator unavailable in the source. The European Corpus Initiative Multilingual Corpus, available at LDC, http://www.ldc.upenn.edu.
>
> 2) *The ladies of Missalongeei*, by Colleen McCullough. Translated to Greek by Philippos Letzis. Greek title: *I ginekes tu Messologiou*. The European Corpus Initiative Multilingual Coprus, available at LDC, http://www.ldc.upenn.edu.
>
> 3) *Dio haraktires*. Author and translator information not available at the source. The European Corpus Initiative Multilingual Corpus, available at LDC, http://www.ldc.upenn.edu.
>
> 4) To Vima, Greek newspaper, available at http://tovima.dolnet.gr

A set of features similar to the one used for the annotation of the English relative clauses has also been used for the annotation of Greek relative clauses and is shown in Table 5.8. In this set, the type of referring expression "strong pronoun" has been added, which, as argued earlier, is used to mark reference to an entity other than the most salient entity in the preceding discourse. For this set of data, the presence or absence of a comma before the relative clause is, also, an added feature whose purpose is to evaluate if the presence of a comma, a feature associated with non-restrictive relatives, correlates with our coding of non-restricting relatives.

Features	Feature values
Reference to head noun entity	Yes/No
Referring expression for reference to head noun entity	Non-applicable (N/A) NP NP-associative Strong pronoun Null/Pronoun Implicit Other
Presence of "other" entities in relative clause	Yes/No
Reference to "other" entities in subsequent sentence	Yes/No
Referring expression for reference to "other" entities	Non-applicable (N/A) NP NP-associative Strong pronoun Null/Pronoun Implicit Other
Restricting function of relative clause	Yes/No
Presence of comma	Yes/No

Table 5.8: Set of annotation features for *o opios*- and *pou*-relatives

5.3.2.1 *O opios*-relatives

For the study of *o opios*-relatives in Greek, 100 tokens were extracted from "To fantasma", "I ginekes tu Messologiou", "Dio haraktires" and "To Vima". Each token was annotated with the set of features shown in Table 5.8. The results of the annotation are presented in Table 5.9.

The column "Ref. to the head noun" shows how many times the head noun referent was referenced in the subsequent sentence. The head noun was subsequently referenced in almost 50% of the tokens, a pattern similar to the *who*-relative data in English. Reference to the head noun with a pronoun/null, occurred 11 times, again a number similar to *who*-relatives in English. An example of pronominal reference to the head noun is given in (148). In this case, the head noun is in a prepositional phrase. In (148c), the possessive

Ref. to head noun		Referring expression	
Yes	50	N/A	49
No	49	NP	20
		NP-assoc	11
		1st/2nd pers.	1
		Strong pron.	5
		Null/pronoun	11
		Implicit	2
Ref. to "other"		Referring expression	
Yes	46	N/A	53
No	46	NP	24
N/A	7	NP-assoc	1
		1st/2nd pers.	3
		Strong pron.	1
		Null/pronoun	15 (12old)
		Implicit	3

Table 5.9: Reference in *o opios*-relatives

tis 'her" is used to refer to the head noun referent *Brigitte Roussellen*. The head noun referent is not the highest ranked entity in the main clause that contains the head referent as it was often the case in the analogous studies in English. However, the two other entities evoked in (148a), both higher ranked than *Brigitte Roussellen* have already been pronominalized in (148c). The most crucial evidence for the non-topical status of the head noun referent, however, is the interpretation of the dropped subject in (148c) (3^{rd} person singular, morphologically marked on the verb). Both *Leonardo* and *Brigitte* are possible antecedents for the dropped subject in (148c). *Brigitte* is the highest ranked entity in the relative clause and *Leonardo* is the highest ranked entity (with compatible features) in the main clause. If the relative clause was processed as an independent unit then the highest ranked entity would be *Brigitte*. In fact, the antecedent of the dropped subject is *Leonardo*, who is the highest ranked compatible entity in the main clause.

(148) a. Ne simera ta mesanihta, giro sti mia to proi estila ton
yes today the midnight, around at-the one the morning sent(1st-p) the

Leonardo$_i$ na erevnisi gia tin Mprizit Ruselen$_j$,
Leonardo to research for the Brigitte Roussellen,
'Yes, today at midnight, about one in the morning, I sent **Leonard**$_i$ to do some research for **Brigitte Roussellen**$_j$,

b. tis opias$_j$ tin kordela iha prosexi htes to vradi.
of-the who the ribbon had(1st-p) noticed yesterday the evening.
'**whose**$_j$ ribbon I had noticed last evening.'

c. Otan null$_i$ girise, mu null$_i$ ipe pos diafori tipi
when null(nom) returned(3rd-p), me(gen) null(nom) told that various guys
periferontan giro apo to spiti tis.
were-strolling around from the house her$_j$.
' When **he**$_i$ returned, **he**$_i$ told me that various guys were strolling around **her**$_j$ house.'

From the remaining 10 instances of pronominal reference to the head noun, in 6 cases the head noun had already been pronominalized as in the case shown in (148), or it was the highest ranked entity in the main clause. In another case, the subject of the main clause was a discourse deictic expression. The remaining three cases are discussed below. In one, given in (149), the head referent is not the highest ranked entity but there is no competing antecedent and the subject of the main clause is an inanimate noun phrase, which could be a relevant factor for the salience ranking in Greek.[5] The last two cases are most probably exceptions to the pattern that we have seen so far. An example of this reference pattern is given in (150). Note that in this case, the dropped subject is known to be masculine due to the obligatory gender agreement between subject and the adjectival predicate. However, until the adjective is processed there are two competing antecedents, *Misi* and *Sir William Harlingford*. After the adjective is processed this example is similar in behavior to (149).

(149) a. Ta hirokrotimata$_i$ proerhontan apo ton kirio Lathrop$_j$,
the clappings came-from from the Mr Lathrope
'**The clapping**$_i$ came from **Mr Lathrope**$_j$'

[5]Turan (1995) has found that after subjecthood, animacy is an important salience factor in Turkish.

b. o opios$_j$ ihe mpi ke kathisi aparatiritos ligo markia tus
the who had come-in and sat unobserved a-little far their
'**who**$_j$ had come in and had sat, without anyone noticing, at some distance from them.'

c. Ke tin ekdilosi tu enthusiasmu tu$_j$ akoluthise, ligo pio irema,
and the demonstration of-the enthusiasm his followed, a-bit more calmly,
o giatros Artser$_k$
the doctor Archer-NOM
'And the expression of **his**$_j$ enthusiasm was seconded, a bit more calmly by **Dr Archer**$_k$.'

(150) a. afti i ekentriki astiki onomatologia ofilotan ston propapo$_i$
this the eccentric bourgeois nomenclature was-due to-the great-grandfather
tis Misi$_j$, ston Ser Wuiliam Harlingrford$_i$
her, to-the Sir William Harlingford
'This eccentric bourgeois nomenclature was due to **Misi**$_j$'s **great-grandfather**$_i$, **Sir William Harlingford**$_i$,'

b. o opios$_i$ ihe vasisi tin idrisi tis polis sto vivlio X
the who had based the founding of-the town to-the book X
'**who**$_i$ had based the founding of the town on book X.'

c. null$_i$ itan malista toso ikanopiimenos giati null ihe anakalipsi ena toso
null was in-fact so satisfied-MASC that null$_i$ had discovered one such
megalo logotehniko ergo pu ...
big literary work that ...
'**He**$_i$ was in fact so pleased because **he**$_i$ had discovered such a big literary piece that ...'

A full NP was used for reference to the head noun in 20 tokens. Of the 20 tokens of NP reference to the head noun, in 16 cases the head noun was a non-subject, in two cases it was the subject of an embedded clause, and in the remaining two the head noun was a main clause subject. In one case, though, the NP reference in the subsequent discourse was obligatory because it appeared after a preposition and in the other one, the NP reference was contained in speech directly quoted in part of the subsequent discourse. So, on closer inspection, the data show that when a choice was possible, a full NP was used to refer to

the head noun when the head noun was not the highest ranked entity in the main clause.

The column "Ref. to 'other' " shows how many times an entity evoked in the relative clause, other than the head noun, was referenced in the subsequent sentence. As can be seen in Table 5.9, reference to "other" entities in Greek relative clauses was much more frequent than in their English counterparts, occurring in 46 tokens. A pronoun or null was used for reference to the head noun in 15 tokens. In 12 of those 15 instances the highest ranked entity in the main clause was already pronominalized but appeared as an "other" referent in the relative clause. A typical example is shown in (151).

(151) a. Amesos null$_i$ sigrotise mia omada apo dodeka aristokrates$_j$ tis
Immediately null put-together a team from twelve aristokrats of-the
kakias oras, ... ,
bad time, ... ,
'Immediately **he**$_i$ put together a team of **twelve pitiful aristokrats**$_j$,'

b. stus opius$_j$ null$_i$ apokalipse ena meros mono tis ipothesis.
to whom null revealed a part only of-the case
'to **whom**$_j$ **he**$_i$ revealed only a part of the case.'

c. tus$_j$ null$_j$ metamorfose se dinus sinomotes, ikanus gia kathe
them null transformed to skillful conspirators, capable for every
vromodulia.
dirty-trick
'**He**$_i$ transformed **them**$_j$ into skillful conspirators, capable for any dirty trick.'

One of the remaining three cases of pronominal reference to an "other" entity that was not the highest ranked entity of the main clause is given (152), which simply combines properties of the pronominal use for reference to "other" that we have seen previously. To comprehend this example, some background information is needed. "ND" stands for "New Democracy", which is the name of the right wing party in the Greek parliament. "PASOK" is the name of the socialist party, which is, at the time of writing, the government party. "Mr. Avramopoulos" used to be a member of ND, who (recently) formed his own political party. "Mr. Avramopoulos" is evoked for the first time in the relative

clause but is then referenced with a pronoun. In this case the interpretation of the pronoun in (152c) is facilitated by the fact that the preceding sentence does not contain any competing antecedents. Also, as in the other 12 cases, "ND" which is the highest ranked entity of the main clause is already pronominalized (*meros tis vasis tis* "part of **its** base"). The full NP reference to "ND" in the prepositional phrase *gia ti ND* is obligatory due to grammar constraints. As mentioned earlier a full NP or strong pronoun are the only available choices for the nominal complement of a preposition. Assuming that "ND" has topical status, the use of a strong pronoun in this case would be inappropriate, as it would indicate that the referent is an entity other than the most salient entity in the preceding discourse.

(152) a. Ostoso i kinisi tis ND pros to 'kentro' kathistate disheris
However, the move of-the ND towards the 'center' is-established difficult
ke os-ek-tutu periorizete i ikanonita tis na epofelithi apo ti
and therefore is-limited the ability her to benefit from the
fthora tu PASOK
wearing-out of PASOK
'However, ND's move towards the 'center' is now becoming difficult (to achieve) and therefore its ability to benefit from PASOK's attrition has been limited.'

b. to megalitero meros tis opias isprati o k. Avramopoulos.
the biggest part of-the which receives the Mr. Avramopoulos
'the biggest part of which (of the attrition) benefits Mr. Avramopoulos.'

c. Prostheti disheria gia ti ND apoteli to oti meros tis vasis tis
Additional difficulty for the ND constitutes the that part of-the base her
ton eklamvani os melontiko simaho tis ke sinepos i "polosi" me
him take as future ally her and therefore the 'polarization; eith
to neo koma isos apodihthei epizimia gia ton k. Karamanli.
the new party perhaps prov-to-be harmful for Mr. Karamanlis.
'An additional problem for ND is the fact that part of its base perceives him as its future ally and therefore the 'polarization' with the new party may prove to be harmful for Mr. Karamanlis.'

From the remaining two cases of pronominal reference to an "other" entity, in one case

the subject of the main clause was a discourse deictic expression and in the other the text explicitly sets the "other" referent as the topic of the subsequent discourse, shown in (153). Excluding "NP-assoc.", "1st/2nd pers.", and 'implicit" from the analysis, the remaining references to an "other" entity were either with a full NP (24 tokens) or a strong pronoun (1 token). The use of a strong pronoun was again obligatory because it occurred in the quantified phrase *ola afta* "all these". In fact it is possible that the strong pronoun in this token was a case of discourse deixis. The relevant example is shown in (154).

(153) a. Su grafo gia na su po dio logia gia ti Misi Rait$_i$, mia
to-you write-1st-pers for to to-you say two words about the Misi Rait, a
ftohi ke shetika mikri gerontokori$_i$,
poor and relatively young spinster,
'I'm writing to you to say a few things about **Missi Right**$_i$, **a poor and rather young spinster**$_i$'

b. i opia$_i$ epathe tulahiston mia krisi$_j$.
the who had at-least one crisis
'**who**$_i$ had at least **one crisis**$_j$.'

c. Null$_j$ ekdilothike me pono sto stithos ke dispnia meta apo mia kurastiki ke viastiki pezoporia.
null was-manifest with pain at-the chest and laborious-breathing after from a tiring and rushed marching
'**It**$_j$ got manifest with a chest pain and laborious breathing after an exhausting and fast walk.'

(154) a. Tin apantisi edose o Huper, o opios diigithike ta gegonota$_i$ opos ta
the answer gave the Huper, the who narrated the facts as them
ihe zisi null.
had lived null
'Hooper gave the answer, who narrated **the facts**$_i$ as he had experienced them.'

b. Idate sta alithia ola afta$_i$; epemene o giatros
saw-2nd-pers in-the truth all these? insisted the doctor
' "Did you really see **all these**$_i$?", the doctor insisted.'

5.3.2.2 *Pou*-relatives

In this study, 100 tokens of *pou*-relatives were extracted from "To fantasma". Each token was annotated with the set of features shown in Table 5.8.

The results of the reference annotation are summarized in Table 5.10. The column "Ref. to the head noun" shows how many times the head noun referent was referenced in the subsequent sentence. By comparison to *o opios*-relatives, we observe that reference to the head noun is less frequent in *pou*-relatives. A head noun referent was referenced in 24 tokens (24%). The head noun was subsequently referenced with a null/pronoun in 15 cases. An example of pronominal reference to the head noun is given in (155). As can be seen in this example the head noun referent is the subject of the main clause and the highest ranked entity in the main clause. That was the case for 10 of the 15 cases of pronominal reference to the head noun. In another 2 cases, the head referent was the highest ranked *animate* entity with compatible features in the main clause and the main clause did not contain any other animate referents. An example is given in (156). In the remaining 3 cases, one involved a switch from direct speech to narrative to direct speech, one was in an elliptical clause and the last one was ambiguous between referring to the head noun or a clausal constituent. A full NP was used for reference to the head noun twice. In both cases, the head referent was a non-subject, as shown in (157).

(155) a. Null$_i$ ine o pio paraxenos anthropos pu$_i$ eho sinantisi.
null is the most strange man that have-1st-pers met.
'**He**$_i$ is the strangest guy that I have met.'

b. pes tu$_i$ oli tin ipothesi.
tell him all the situation
'Tell *him*$_i$ everything about the situation.'

(156) a. Xafnika mia foni$_i$ vrahni ke enrini, vgenontas apo to megafono,
suddenly a voice coarse and nasal, coming-out of-the speaker, seemed
fanike na xipna tus ligus taxidiotes$_j$
to wake-up the few travellers
'Suddenly, **a coarse and nasal voice**$_i$, coming out of the speaker seemed to wake up **the few travellers**$_j$'

Ref. to head noun		Referring expression	
Yes	24	N/A	76
No	76	NP	2
		NP-assoc	3
		1st/2nd pers.	0
		Strong pron.	1
		Null/pronoun	15
		Implicit	1
		Disc. deixis	1
		Other	1
Ref. to "other" entities		Referring expression	
Yes	28	N/A	72
No	61	NP	3
N/A	11	NP-assoc	0
		1st/2nd pers.	0
		Null/pronoun	23 (22 old)
		Implicit	2
		Other	0

Table 5.10: Reference in *pou*-relatives

b. pu$_j$ kathontan me ifos sovaro ke katavevlimeno stin ethusa
that were-sitting with expression serious and worn-out in-the room
anamonis.
of-waiting

'**that**$_j$ were sitting in the waiting room with a serious and worn out expression (on their face)'

c. Null$_j$ sikothikan ke vgikan arga eno ...
null got-up and went-out slowly whiel ...

'**They**$_j$ got up and went out slowly while ... '

(157) a. O Max ... anevike stin kentriki skala gia na pai sto saloni$_i$ pu$_i$ vriskotan sti gefira A.

The Max ... went-up to-the central staircase in-order to gp to-the sitting-room that was-located to-the bridge A

'Max ... went up the central stairs to go to **the sitting area**$_i$ **that**$_i$ was at

Bridge A.'

b. To domatio$_i$ itan gemato polithrones ke ...
the room was full-of armchairs and ...
'**The room**$_i$ was full of armchairs and ... '

The column "Ref. to 'other' " shows how many times an entity evoked in the relative clause, other than the head noun, was referenced in the subsequent discourse. As can be seen in Table 5.10, an "other" entity was subsequently referenced in 28 tokens (28%). A null/pronoun was used for reference to an "other" entity in 23 cases. Of these 23 cases, in 22 the "other" referent was already pronominalized or was already evoked as the highest ranked entity in the main clause. A typical example is shown in (158).

(158) a. O Max Mathius$_i$ hamilose to vlema tu sto tsigaro$_j$
Max Matthews lowered the gaze his to-the cigarette
'**Max Matthews**$_i$ lowered his eyes to **the cigarette**$_j$'

b. pu$_j$ null$_i$ kratuse anamesa sta dahtila tu
that null was-holding between the fingers his
'**that**$_j$ **he**$_i$ was holding between his fingers'

c. ke null$_i$ thimithike pos vriskotan pano se ena plio pou metefere
and null remembered that was on to ea ship that was-carrying
ekriktika
explosives
'and **he**$_i$ remembered that he was on a ship that was carrying explosives.'

d. Null$_i$ anarotithike an epitrepotan to kapnisma sti gefira
null wondered if was-allowed the smoking on-the bridge
'**He**$_i$ wondered if smoking was allowed on the bridge.'

The "other" referent in the remaining case was an instance of what has been discussed in the literature, as an entity being in *global focus* (Grosz & Sidner, 1986), (Hitzeman & Poesio, 1998). The relevant example with the preceding discourse is given in (159). As can be seen in the discourse preceding the relative clause, the "doctor" was already introduced earlier as the topic and intuitively he remains the topic of the entire segment. The sentence interfering between the last mention of "doctor" and the reference in the

relative clause gives a description of the setting including the "doctor" as well as other characters introduced in a previous segment in the narrative. Note that reference to those characters is with a null although they were not mentioned in the preceding sentence. It seems, then, that in this case, a higher-level discourse structure, which we have not captured, is responsible for the attested reference patterns. Excluding "implicit" from the analysis, the remaining 2 references to an "other" entity was with a full NP.

(159) a. O giatros$_i$ edihne kosmikos tipos. Ntimenos ligaki exezitimena, null$_i$
The doctor looked cosmopolitan type. Dressed a-little eccentric, null
ihe monima sto prosopo tu$_i$ mia ekfrasi epiikias, ena
had permanently to-the face his an expression of-forbearance, a
elafro hamogelo ke null$_i$ den eperne pote to logo an o
slight smile and null not took ever the speech-turn if the
sinomilitis tu$_i$ den ihe oloklirosi.
co-conversant his not had finished

'**The doctor**$_i$ seemed the cosmopolitan type. Dressed a little eccentric, **he**$_i$ always had on **his**$_i$ face an expression of forbearance, a slight smile and **he**$_i$ never took (his) turn if his co-conversant had not finished (what he was saying).'

b. Null$_j$ kathismeni-pl oli giro apo to tzaki, null$_j$ fotizontan-3rd-pl apo
null seated all around from the fireplace, null lit by
lampes pano apo ta kefalia tus$_j$ -giati mono enas mikros arithmos
lamps above from the heads their -because only a small number
apo finistrinia itan anihta- ke to fos pu epefte loxa pano
of scuttles were open- and the light that was-falling sideways above
apo sto giatro$_i$ tonize tis ametrites mikres ritides giro apo
from to-the doctor highlighted the numerous small wrinkles around from
ta matia tu$_i$.
the eyes his

'Seated around the fireplace, **they**$_j$ were lit by the lamps above **their**$_j$ heads -since only a small number of scuttles were open- and the light that was falling sideways on **the doctor**$_i$ highlighted the numerous, small wrinkles around **his**$_i$ eyes.'

c. Null$_i$ itan sigura pio megalos apo oso null$_i$ fenoten se proti opsi.
null was surely more old than what null looked at first sight

'**He**$_i$ surely was older than what **he**$_i$ looked at first sight.'

5.3.3 The Coherence Test

5.3.3.1 Design and results

The corpus for the coherence test consists mainly of newspaper articles, available on line at http://www.enet.gr. The size of the corpus is approximately 800,000 words. For the study of coherence, we have included 100 tokens of Greek final non-restrictive relative clauses. As in the analogous English study, we included only non-restrictives. Also, we included only sentence-final relative clauses because in such a position the entities evoked in the relative clauses are linearly more recent than the entities introduced in the main clause. The selection criteria are specified below:

1. The relative clause is preceded by a comma.

2. The sentence following the relative clause includes reference to at least one entity evoked in the sentence containing the relative clause, either in the main clause or in the relative clause.

3. The relative clause is in sentence-final position.

For each token, Centering transitions were computed in two versions. In version A, two Centering transitions were computed: one for the sentence containing the relative clause and one for the sentence following the relative clause. In version B, three Centering transitions were computed. One for the first sentence excluding the relative clause, one for the relative clause and one for the sentence following the relative clause. The results of the computation of Centering transitions in the two conditions are shown in Table (5.11). The column "more 'coherent' transition" contains the number of cases where a more "coherent" transition was computed in the final sentence in version B. The column "less 'coherent' transition" shows how many times a less "coherent" transition was computed in the final sentence in version B, and, finally, the column "no effect" shows how many times

the same transition was computed in both versions A and B. The relevant degree of coherence is specified according to the Centering transitions rule: Continue>Retain>Smooth-Shfit>Rough-Shift.

More coherent transition	Less coherent transition	No effect	Total
8	44	48	100

Table 5.11: Effect of Greek non-restrictive relatives on Centering transitions

Before discussing the results, let us first look at some examples from the data. In what follows, we first present examples from the Greek corpus glossed and translated in English. For ease of reading we, then, present the two coded versions with respect to the English translations. The English translations accurately reflect the grammatical role of the crucial entities, sometimes at the expense of naturalness in English. Examples (160) and (161) demonstrate a case coded as "less coherent". In this case, version B, yields a Rough-Shift transition, which is ranked less "coherent" than the Continue transition computed in version A for the same last sentence. What is important to note in this example is that in version B, where the relative clause is processed as an independent unit, there is no link between the relative clause and the subsequent utterance. Note that in the last sentence we have coded *Kostas Karamanlis* as the Cp although he is not the subject. We did so because the subject a negated indefinite NP. As mentioned earlier, a possibility that we have not fully investigated is that *Kostas Karamanlis* outranks the subject due to animacy. However, even if we code the subject as the Cp, we still get a less coherent transition in version B.

(160) a. I prothesi tu Kosta Karamanli$_i$ itan safestati htes sti
The intention of Kosta Karamanli was most-clear yesterday at-the
sindedriasi$_j$ tis ektelestikis epitropis tu komatos, apo tin opia$_j$
meeting of-the executive committee of-the party, from the which
apusiazan i Mil. Evert ke Ntora Mpakogiani.
were-missing the Mil. Evert and Ntora Mpakogiani.

'**Kosta Karamanli's**$_i$ intention was very clear yesterday at **the meeting**$_j$ of the party's executive committee, from **which**$_j$ (meeting) Mil. Evert and Ntora Mpakogiani were missing.'

b. Omos apo tin topothetisi tu proedru$_i$ tis ND den apusiase
However from the position of-the president of-the ND not was-missing
ki enos idus ehmi gia to dimarho tis Athinas ke to endehomeno
and of-one kind dig for the mayor of-the Athens and the possibility
dimiurgias neu komatos apo afton.
of-founding of-new party by HIM
'However, in the position of **ND's president**$_i$ a kind of dig was present against the mayor of Athens and the possibility of the founding of a new party by HIM.

(161) VERSION A

a. **Kosta Karamanli's**$_i$ intention was very clear yesterday at **the meeting**$_j$ of the party's executive committee, from **which**$_j$ (meeting) Mil. Evert and Ntora Mpakogiani were missing.

Cb=none

Cp=Karamanlis

Tr=none

b. However, in the position of **ND's president**$_i$ a kind of dig was present against the mayor of Athens and the possibility of the founding of a new party by HIM.

Cb=Karamanlis

Cp=Karamanlis

Tr=Continue

(162) VERSION B

a. **Kosta Karamanli's**$_i$ intention was very clear yesterday at **the meeting**$_j$ of the party's executive committee,

Cb=none

Cp=Karamanlis

Tr=None

b. from **which**$_j$ (meeting) Mil. Evert and Ntora Mpakogiani were missing.

Cb=meeting

Cp=Evert and Mpakogiani

Tr=Retain

c. However, in the position of **ND's president**$_i$ a kind of dig was present against the mayor of Athens and the possibility of the founding of a new party by HIM.

Cb=none

Cp=Karamanlis (ND's president)

Tr=Rough-Shift

An example of the category "more 'coherent' transition" is given in (164) and (165) glossed in (163). In this example, version B yields a Continue transition, which is ranked more "coherent" than the Smooth-Shift transition computed in version A, (139) for the same last sentence. In this example the last sentence contains a full NP reference to the *Patriarch*. However, the choice of a full NP in this case is not helpful in deciding if we indeed have the establishment of a new topic in the last sentence (as we has seen in English) because the NP contains the adjectival modifier *new* which necessitates the use of NP reference to *Patriarch*. It is also possibly relevant here that neither *Patriarch* nor *new Patriarch* are referring to an entity in the real world. Strictly speaking, they are not co-referential. The entity evoked by the first mention of *Patriarch* appears in the intentional context created by the verb *epithimi* "desire" in the first main clause. It is possible that pronominal reference is harder outside the scope of the intentional context. We discuss further examples from this category shortly.

(163) a. To Israil$_i$, ..., dihni malon na epithimi Patriarhi$_j$, o opios$_j$ de tha
The Israel, ..., seems rather to want Patriarch, the who not will
dimiurgi provlimata stis dikes tu epidioxis ke sta dika tu simferonta.
create problmes to-the own its plans and to-the own its interests

'**Israel**$_i$, ..., seems to probably want **a Patriarch**$_j$, **who**$_j$ will not create problems to its own plans and interests.'

b. Ek ton pragmaton provali ke o exisoropistikos rolos tu neu
From the facts emerges and the balancing role of-the new
prokathimenu$_j$.
position-holder
'Pragmatically, the balancing role of **the new patriarch**$_j$ is also emerging.'

(164) VERSION A

a. **Israel**$_i$, ..., seems to most likely want **a Patriarch**$_j$, **who**$_j$ will not create problems to its own plans and interests.

Cb=Israel

Cp=Israel

Tr=Continue

b. Pragmatically, the balancing role of **the new patriarch**$_j$ is also emerging.

Cb=Patriarch

Cp=Patriarch

Tr=Smooth-Shift

(165) VERSION B

a. **Israel**$_i$, ..., seems to most likely want **a Patriarch**$_j$,

Cb=Israel

Cp=Israel

Tr=Continue

b. **who**$_j$ will not create problems to its own plans and interests.

Cb=Patriarch

Cp=Patriarch

Tr=Smooth-Shift

c. Pragmatically, the balancing role of **the new patriarch**$_j$ is also emerging.

Cb=Patriarch

Cp=Patriarch

Tr=Continue

Looking back at the table of results in (5.11), we see that for 44% of the tokens, processing the relative clause as an independent unit yields a "less coherent" Centering transition in the subsequent sentence, whereas a "more coherent" transition is was computed for only 8% of the tokens. For the remaining 48%, there was no effect. If we, now look closer at the 8 tokens of the "more coherent" transition we observe that our coding schema actually failed to capture implicit links present in the relevant discourses. This was the case for 7 of the 8 examples (the eighth case was given as an example of the "more coherent" type above). A typical example is given in (167) and (168) - glossed in (166).

(166) Sta vretanika egrafa$_i$ tu Iuliu 1970 anaferete oti o Papadopulos
In-the british documents of-the July 1970 is-mentioned that the Papadopulos
ihe plirofories gia dolofoniki apopira enantion tu Pipineli.
had information for murder attempt against of-the Pipineli
'In the **British**$_i$ documents of July 1970 it is mentioned that Papadopulos had information about a murder attempt against Pipineli.'

a. O Vretanos presveftis$_j$ epikalite 'to gnosto kikloma sizigon', ..., me
The British ambassador appeals 'the known circle of-wives$_{kl}$', ..., with
tis opies$_k$ o ser Maikl Stiuart$_j$ diatiruse filikes shesis.
the whom$_{kl}$ the sir Michael Steward maintained friendly relationships
'**The British Ambassador**$_j$ appeals to 'the well-known **circle of wives**$_{kl}$', ..., with **whom**$_{kl}$ Sir Michael Steward had friendly relationships.'

b. I sizigos$_l$ tu Pipineli diohetefse tin pliroforia afti ston Stiuart ke
the wife of-the Pipineli released the information this to-the Steward and
...
...
'**Pipineli's wife**$_l$ released this information to **Steward**$_j$ and ... '

(167) VERSION A

In **the British documents**$_i$ of July 1970 it is mentioned that Papadopulos had information about a murder attempt against Pipineli.

a. **The British Ambassador**$_j$ appeals to 'the well-known **circle of wives**$_{kl}$', ..., with **whom**$_k$ **Sir Michael Steward**$_j$ had friendly relationships., ...,'
Cb=none
Cp=British Ambassador
Tr=Rough-Shift

b. **Pipineli's wife**$_l$ released this information to **Steward**$_j$ and ...
Cb=British Ambassador
Cp=wife
Tr=Rough-Shift

(168) VERSION B

In the British documents of July 1970 it is mentioned that Papadopulos had information about a murder attempt against Pipineli.

a. **The British Ambassador**$_j$ appeals to 'the well-known **circle of wives**$_{kl}$', ...,'
Cb=none
Cp=British Ambassador
Tr=Rough-Shift

b. with **whom**$_k$ **Sir Michael Steward**$_k$ had friendly relationships.
Cb=British Ambassador
Cp=British Ambassador
Tr=Smooth-Shift

c. **Pipineli's wife**$_l$ released this information to **Steward**$_j$ and ...
Cb=British Ambassador
Cp=wife
Tr=Retain

In the example above, *Pipineli's wife* was coded as the Cb in (166b) by virtue of its set-membership relation to the entity evoked by the *circle of wives*. We indicate that by marking the set *circle of wives* with two subscripts, *k* and *l*. The latter is then retained for

Pipineli's wife. Note that in both versions A and B, a Rough-Shift transition was computed in (167a) and (168a), respectively. This is because none of the entities introduced in this sentence is coreferent with any of the entities in the preceding discourse. However, the (167a) and (168a) *are* linked to the preceding discourse via functional dependency between *british documents* and the *British Ambassador*. The British Ambassador or his associates are most likely the authors or otherwise responsible for the contents of the documents mentioned at the beginning of the discourse. The *British Ambassador* is associated with the *british documents*, even though they are not in a co-referring relationship. We believe that this is the link between the two sentences, although we have not been able to capture it in our computation of Centering transitions. In this particular example, the association could be identified by lexical similarity but in a computational framework the complete range of relevant associations would be hard to capture on lexical grounds. Centering acknowledges that *realization* of an entity may combine syntactic, semantic, discourse, and intentional factors and recognizes that the center of an utterance may be functionally dependent on an entity evoked in the set of forward-looking centers in the preceding discourse. However, what semantic theory is needed to best capture such dependencies and their interactions with other discourse processes such as inferencing awaits further research.

For our purposes, it suffices to point out that recognizing functional dependencies crucially affects the representation of discourse coherence. If we were able to capture the association between *british documents* and the *British Ambassador* we would obtain a different result, shown in (169) and (170). In (170) we observe that processing the relative clause as an independent unit no longer improves the transition computed in the sentence following it. As shown in (169) and (170), both versions A and B yield a Retain transition for the last sentence.

(169) VERSION A

In **the British documents**$_{i}j$ of July 1970 it is mentioned that Papadopulos had information about a murder attempt against Pipineli.

a. **The British Ambassador**$_j$ appeals to ‘the well-known **circle of wives**$_{kl}$’, ..., with **whom**$_k$ **Sir Michael Steward**$_j$ had friendly relationships., ...,’
Cb=British Ambassador
Cp=British Ambassador
Tr=Continue

b. **Pipineli’s wife**$_l$ released this information to **Steward**$_j$ and ...
Cb=British Ambassador
Cp=wife
Tr=Retain

(170) VERSION B

In **the British documents**$_ij$ of July 1970 it is mentioned that Papadopulos had information about a murder attempt against Pipineli.

a. **The British Ambassador**$_j$ appeals to ‘the well-known **circle of wives**$_{kl}$’, ...,’
Cb=British Ambassador
Cp=British Ambassador
Tr=Continue

b. with **whom**$_k$ **Sir Michael Steward**$_j$ had friendly relationships.
Cb=British Ambassador
Cp=British Ambassador
Tr=Continue

c. **Pipineli’s wife**$_l$ released this information to **Steward**$_j$ and ...
Cb=British Ambassador
Cp=wife
Tr=Retain

5.4 Conclusions

In this chapter we tested the following two hypotheses: a) entities evoked in relative clauses do not override the topical status of entities evoked in main clauses and b) relative clauses do not form independent topic update units. To test the two hypotheses we performed a series of corpus annotations in both English and Greek. With respect to the first hypothesis the results of the corpus studies showed that entities evoked in relative clauses are not subsequently referenced with a pronoun unless a) they are already pronominalized in the relative clause or b) the highest ranked entity in the main clause is also pronominalized. The observed patterns of reference are supportive of the first hypothesis. With respect to the second hypothesis, we performed a Centering analysis of written text in two conditions. In the first condition we computed Centering transitions assuming that the relative clause consists a single unit with the main clause. In the second condition we assumed that the relative clause is processed as an independent unit. We predicted that more coherent Centering transitions would be computed in the first condition under the assumption that topic discontinuities are avoided in carefully planned text. The prediction was borne out, thus supporting our hypothesis that in the computation of topic structure relative clauses are not processed as independent center update units.

Chapter 6

Implications for Textual Coherence: An Evaluation Model

6.1 Introduction

Existing software systems for automated essay scoring can provide language researchers with opportunities to test certain theoretical hypotheses. In this experiment, we employ the Educational Testing Service's *e-rater* essay scoring system to examine whether local discourse coherence, which we define by a measure of Centering Theory's Rough-Shift transitions, might be a significant contributor to the accuracy of computer-generated essay scores. In accordance with the topic continuity view of Centering of this thesis, Rough-Shifts within students' paragraphs are generated by topics that are short-lived and unconnected, and are therefore indicative of poor topic development. The results of the *e-rater* Centering study reported in this chapter show that adding the Rough-Shift based metric to the system improves its performance significantly, better approximating human scores and providing the capability of valuable instructional feedback to the student. This result indicates that Rough-Shifts do indeed capture a source of incoherence, one that was not closely examined in the Centering literature. They not only justify Rough-Shifts as a valid transition type, but they also support the original formulation of Centering as a measure

of discourse continuity even in pronominal-free text.

This chapter is organized as follows. Section 6.2 briefly introduces the concepts of automated essay scoring. Section 6.4 describes the central tenets of the Centering Model as employed for the *e-rater* study. Section 6.5 focuses specifically on the role of *Rough-Shift* transitions. Sections 6.6, 6.6.6, and (6.6.7) describe the *e-rater* Centering study and present its results followed by discussion. Finally, in Section 6.6.8 we discuss related open issues.

6.2 The *E-rater* Essay Scoring System

Approaches to essay scoring vary in their use of NLP techniques and other methods to assess the writing ability exhibited in an essay. Very early work by Page (1966, 1968) and Page and Peterson (1995) demonstrated that computing the fourth root of the number of words in an essay provides a highly accurate technique for predicting human-generated essay scores. Such measures of essay length have two main weaknesses which render them impractical for writing evaluation. First, scoring criteria based on a superficial word count make the automated system susceptible to deception. Furthermore, due to their lack of explanatory power, such measures cannot be translated into instructional feedback to the student. To improve the efficiency of automated writing evaluation systems, we need to build models which more closely represent the criteria that human experts use to evaluate essays.

Two more recent approaches have attempted to define computational techniques based on these criteria. Both of these approaches are able to predict human scores with at least as much accuracy as length-based approaches. One of these systems, the Intelligent Essay Assessor (Landauer, 1998; Foltz, Kintsch, & Landauer, 1998; Schreiner, Rehder, Landauer, & Laham, 1997), employs a technique called Latent Semantic Analysis (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990) as a measure of the degree to

which the vocabulary patterns found in an essay reflect the writer's semantic and linguistic competence. Another system, the Electronic Essay Rater, *e-rater* (Burstein, Kukich, Wolff, Chodorow, Braden-Harder, Harris, & Lu, 1998), employs a variety of NLP techniques, including sentence parsing, discourse structure evaluation, and vocabulary assessment techniques to derive values for over fifty writing features.

The writing features that *e-rater* evaluates were specifically chosen to reflect scoring criteria defined by Educational Testing Service (ETS) writing evaluation experts for the essay portion of the Graduate Management Admissions Test (GMAT). The GMAT test is one of several criteria used by most U.S. graduate business schools to evaluate applicants. Over 200,000 GMAT tests are administered each year. Fully computerized, the GMAT test includes both a multiple choice section and an essay writing section. In the essay section, each examinee must compose two essays on general business-related topics randomly chosen by computer from a large pool of topics. Examinees are allowed 30 minutes to compose each essay and the average length of the essays is about 250 words. Essays are scored on a scale of 1 to 6 points, where a score of 1 indicates an extremely poor essay and a score of 6 indicates an excellent essay. Until recently, each essay was first scored by two trained writing evaluation experts. For those essays whose first two scores differ by more than one point (about ten percent), additional experts' scores are solicited. Starting in 1999, scores generated by *e-rater* were used in place of one of the first two experts, yielding a similar ten percent disagreement rate. The procedure for invoking additional experts as needed remains the same.

The essay scoring criteria used by GMAT writing evaluation experts, including, among others, syntactic variety, argument development, logical organization and clear transitions, are fully articulated in GMAT test preparation and scoring materials, which can be found at http://www.gmat.org. In the *e-rater* system, syntactic variety is represented by features that quantify occurrences of clause types. Logical organization and clear transitions are represented by features that quantify cue words in certain syntactic constructions. The existence of main and supporting points is represented by features that detect where new

points begin and where they are developed. *E-rater* also includes features that quantify the appropriateness of the vocabulary content of an essay.

One feature of writing valued by writing experts that is not explicitly represented in the current version of *e-rater* is coherence.

6.3 An Ideal Corpus for Centering

E-rater evaluates the quality of student essays. A set of students' essays evaluated by *e-rater* as well as human raters is used as the corpus for the experiment reported in this chapter. Such a corpus is invaluable in that it provides a unique opportunity to test the strengths of Centering Theory as a model of discourse coherence. Unlike all the corpora in previous work on Centering, in this corpus textual coherence is not assumed. In previous work Centering Theory was used to make predictions about referential form and interpretation based on the crucial assumption that text is *maximally* coherent. For one thing, it is questionable whether *maximal* coherence should be assumed for any text. Most importantly, though, corpora assumed to be maximally coherent are not suitable for testing what, we believe, Centering Theory is best suited for, namely, evaluation of textual coherence with respect to topic continuity.

If the Centering Model is indeed capable of making correct judgments on textual coherence, then it is plausible to hypothesize that it can enhance the *e-rater*'s performance by adding a coherence feature to its evaluation criteria. To gain some initial insight, we first performed a preliminary study on a small sample of GMAT essays. We applied the Centering algorithm manually to a set of 32 essays, 8 from each of the top four levels, 6, 5, 4, and 3, counting the number of occurrences of each of the four types of Centering transitions (*Continue, Retain, Smooth-Shift* and *Rough-Shift*) in each essay. We observed that essays that received higher scores by writing experts tended to have significantly lower percentages of Centering Theory's *Rough-Shift* transitions than essays with lower scores.[1]

[1]For this preliminary study, we omitted the two lowest levels because essays scored 1 and 2 were not

Specifically, for 15 of the 16 essays scored 5 or 6, less than 25 percent of the total number of transitions were *Rough-Shifts*, while the percentage of *Rough-Shift* transitions was greater than 40 percent for almost all of the essays scored 3 or 4. None of the other three Centering transition types showed either a positive or negative pattern across essay scores. A detailed account of the essay scores and transition counts can be found in Table A.1 in Appendix A.

This observation encouraged us to undertake a fuller study to explore the hypothesis that the Centering Model provides a reasonable measure of coherence (or lack of) reflecting the evaluation performed by writing experts. Specifically, in the study described here, we investigate the effect of adding a *Rough-Shift* percentage feature to *e-rater*'s existing array of features.

6.4 Specifications of the Centering Model

In this section, we, briefly, repeat the basic definitions and assumptions in Centering reviewed in Chapter 2 and then specify and motivate the ones that were made in this study.

Discourse consists of a sequence of textual segments and each segment consists of a sequence of utterances. In Centering Theory, utterances are designated by $U_i - U_n$. Each utterance U_i evokes a *set* of discourse entities, the Forward-looking Centers, designated by $Cf(U_i)$. The members of the Cf set are ranked according to discourse salience. (Ranking is described in Section 4.4.) The highest-ranked member of the Cf set is the Preferred Center, Cp. A Backward-looking Center, Cb, is also identified for utterance U_i. The highest ranked entity in the previous utterance, $Cf(U_{i-1})$, that is *realized* in the current utterance, U_i, is its designated Backward-looking Center, Cb. The Backward-looking Center is a special member of the Cf set because it represents the discourse entity that U_i is about, what in the literature is often called the "topic" (Reinhart, 1981; Horn, 1986).

The Cp for a given utterance may be identical with its Cb, but not necessarily so.

only infrequent but they typically contained very little text.

Depending on the identity relations among Cb's and Cp's in subsequent utterances, four different types of transitions are defined, Continues, Retains, Smooth-Shifts, and Rough-Shifts. This distinction between looking back in the discourse with the Cb and projecting preferences for interpretations in the subsequent discourse with the Cp provides the key element in computing local coherence in discourse.

6.4.1 Discourse Segments

Segment boundaries are extremely hard to identify in an accurate and principled way. Furthermore, existing segmentation algorithms (Morris & Hirst, 1991; Youmans, 1991; Hearst, 1994; Kozima, 1993; Reynar, 1994; Passonneau & Litman, 1997; Passonneau, 1998) rely heavily on the assumption of textual coherence. The same is true for work done in the Centering framework. Passonneau (1998), for example, implemented Centering to detect segment boundaries. The rationale of her approach was that assuming coherent texts, Rough-Shifts can be used to locate segment boundaries. As explained above, in this corpus, textual coherence cannot be assumed. Given that text organization is also part of the evaluation of the essays, we decided to use the students' paragraph breaks to locate segment boundaries. The Rough-Shift based metric that we propose evaluates textual coherence *within* each paragraph in an essay. The final score is summative, adding up the coherence evaluation of each paragraph. In other words, first the degree of coherence within each segment is computed and then a single score is produced for all the segments in an essay. The proposed metric does not compute textual coherence *across* segments.

6.4.2 Centering Transitions

For convenience, we include here the four types of Centering transitions, reflecting four degrees of coherence. Their order of precedence is shown in transition ordering rule (1).

(1) Transition ordering rule: Continue is preferred to Retain, which is preferred to Smooth-Shift, which is preferred to Rough-Shift.

Centering's Pronoun rule will be discussed in detail in Section 6.5.

Table 6.1: Centering transitions

	$Cb(U_i)=Cb(U_{i-1})$	$Cb(U_i)\neq Cb(U_{i-1})$
$Cb(U_i)=Cp$	Continue	Smooth-Shift
$Cb(U_i)\neq Cp$	Retain	Rough-Shift

For a segment initial U_{i-1} with Cb=none, we assume $Cb(U_i)=Cb(U_{i-1})$ for the computation of the Centering transition in U_i. For a segment medial U_{i-1} with Cb=none, we assume $Cb(U_i)\neq Cb(U_{i-1})$ for the computation of the Centering transition in U_i.

6.4.3 Utterance

Based on the discourse model proposed in Chapter (3) and the experimental and corpus results in Chapters 4 and 5, the "utterance", a single center update unit, consists of one main clause and all its associated dependent clauses.

6.4.4 Cf Ranking

As mentioned earlier, the Preferred Center of an utterance is defined as the highest ranked member of the Cf set. The ranking of the Cf members is determined by the salience status of the entities in the utterance and may vary across language. Kameyama (1985) and Brennan et al. (1987) proposed that the Cf ranking for English is determined by grammatical function as follows:

(2) Rule for ranking of forward-looking centers:

SUBJ>IND. OBJ>OBJ>OTHERS

Later crosslinguistic studies based on empirical work (Di Eugenio, 1998; Turan, 1995; Kameyama, 1985) determined the following detailed ranking, with QIS standing for quantified indefinite subjects (people, everyone etc.) and PRO-ARB (we, you) for arbitrary plural pronominals.

(3) Revised rule for the ranking of forward-looking centers:

SUBJ>IND. OBJ>OBJ>OTHERS>QIS, PRO-ARB.

We assumed the Cf ranking given in (3). The content and the ranking of the Cf list may also vary across different types of essays within the same language. Indeed, we have made a few modifications to reflect the properties of the type of essay under investigation. We will turn to those shortly. Overall, though, the Cf ranking in (3) worked well for the GMAT essays. This is because text coherence in students' paragraphs was often achieved by centering a certain individual or concept as shown in (171).

(171) Another example of an individual who has achieved success in the business world through the use of conventional methods is **Oprah Winfrey**. One may not think of her as a "businesswoman", however **she** has managed to install her own production company, all done through hard work and perseverance. Indeed, perseverance is a time honored method of gaining success. **She** has indeed been able to persevere through all the obstacles which she had to face throughout her career. It is because of this hard work and perseverance (again, conventional practices), that **she** has been able to attain her success.

To construct the ranking of the Cf list under the assumption that the "utterance" contains both a main clause and its subordinate clauses, we assume the augmented Cf **ranking rule** shown below. The "M" prefix stands for main clause and the "S_n" prefix stands for the nth subordinate clause. The relevant ranking of the various types of subordinate clauses is currently left unspecified. In our study, the relevant ranking of subordinate clauses was never crucial. In our study, the "S" in the *augmented ranking rule*, stands for any tensed dependent clause.

Augmented ranking rule
M-Subject $>$ M-indirect object $>$ M-direct object $>$ M-other $>$ S1-subject $>$ S1-indirect object $>$ S1-direct object $>$ S1-other $>$ S2-subject $>$ ...

The **augmented ranking rule** is insensitive to the linear order of the subordinate clauses. While no corpus study has yet been conducted to specifically test whether the insensitivity of the rule to linear order is justified, there is accumulating evidence pointing to this direction across languages (discussed in Chapters 3 and 4).

Returning to the Cf ranking in the *e-rater* study, a modification we made involved the status of the pronominal *I*. [2] We observed that in low-scored essays the first person pronominal *I* was used extensively, normally presenting personal narratives. However, personal narratives were unsuited to this essay writing task and were assigned lower scores by expert readers. The extensive use of *I* in the subject position produced an unwanted effect of high coherence. We prescriptively decided to penalize the use of *I*'s in order to better reflect the coherence demands made by the particular writing task. The way to penalize was to omit *I*'s. As a result, coherence was measured with respect to the treatment of the remaining entities in the *I*-containing utterances. This modification yielded the desired result of distinguishing those *I*-containing utterances which made coherent transitions with respect to the entities they were talking about and those that did not.

A further modification made to the Cf ranking involved constructions containing the verb *to be*. In these constructions (e.g., Another company would be Gerber..., There is more promise ...), the noun phrase following the verb *to be* is ranked higher than its structural subject. The rationale for this modification is as follows.

The verb *to be* appears in two types of constructions: specificational and predicational. The modification is relevant only for the specificational cases. The predicational *be* in, for example, the sentence *John is happy/a doctor/the President of the United States*, does not make any semantic contribution. The post verbal nominal phrase forms the predicate of the sentence and assigns a property holding of *John*. It does not introduce another entity distinct from *John*.

[2] In fact, a similar modification has been proposed by Hurewitz (1998) and Walker (1998) observed that the use of *I* in sentences such as "I believe that...", "I think that..." does not affect the focus structure of the text.

The specificational *be*, as in *The cause of his illness is this virus here*, is a predicate of identity or equation (Heycock & Kroch, 1997). It is in these cases that the post verbal nominal is ranked higher than the subject. In (172), for example, *Oprah Winfrey* is the highest ranked entity in the Cf list because the verb *to be* is specificational.

(172) Another example of an individual who has achieved success in the business world through the use of conventional methods is Oprah Winfrey.

Finally, expletives do not evoke discourse entities and therefore do not participate in the Cf list. In (173), for example, the highest ranked entity is *success*.[3]

(173) It is possible to achieve real success in business by following conventional methods.

6.4.5 Complex NPs

In the case of complex NPs, which have the property of evoking multiple discourse entities (e.g. his mother, software industry), the working hypothesis commonly assumed (e.g., Walker and Prince (1996)) is ordering from left to right.[4] With respect to complex NPs containing *possession* relationships the following clarification is in order. English has two types of *possessive* constructions. The first construction is the *genitive* construction realized with an apostrophe plus the letter *s* at the end of the noun. In this construction, the *possessor* is to the left of the *possessee*, for example *Mary's father*. The second construction contains the preposition *of*. In this case, the *possessor* is to the right of the *possessee*. To maintain uniformity for the ranking of the complex NP, we assume linearization of the complex NP according to the genitive construction and then rank from left to right. In (174b), for example, *TLP* ranks higher than both *success* and the *secret*. The ranking is easy to see if we linearize *The secret of TLP's success* to *TLP's success's secret*.

[3] In accordance with the Cf ranking rule (3), the subject of the infinitival construction *to achieve* is ranked low because it is a non-referential indefinite noun phrase.

[4] But see also Di Eugenio (1998) for the treatment of complex NPs in Italian.

(174) a. Trade & Leisure Publications is a successful publishing house in Russia, with two market-leading monthly consumer magazines.

b. The secret of TLP's success, however, is not based on developing or exploiting some new technology or business strategy.

c. Rather, TLP follows a business strategy that has been known since business began.

6.5 The Significance of Rough-Shift Transitions

To date most Centering-related research has focused on its applicability to the problem of pronoun resolution. As already mentioned, the Centering model includes the Pronoun Rule. The Pronoun Rule reflects the intuition that pronominals are felicitously used to refer to discourse-salient entities. As a result, Cbs are often pronominalized, or even deleted (if the grammar allows it). The Pronoun Rule predicts that if there is only one pronoun in an utterance, this pronoun must realize the Cb. The Pronoun Rule and the distribution of forms (definite/indefinite NPs and pronominals) over transition types plays a significant role in the development of anaphora resolution algorithms in NLP.

Note that the utility of the Pronoun Rule and the Centering transitions in anaphora resolution algorithms relies heavily on the assumption that the texts under consideration are maximally coherent. In maximally coherent texts, however, Rough-Shifts transitions are rare, and even in less than maximally coherent texts they occur infrequently. For this reason the distinction between Smooth-Shifts and Rough-Shifts was collapsed in previous work (Di Eugenio, 1998; Hurewitz, 1998). The status of Rough-Shift transitions in the Centering model was therefore unclear, receiving only negative evidence: Rough-Shifts are valid because they are found to be rare in coherent discourse.

In this study we gain insights pertaining to the nature of the Rough-Shifts precisely because we are forced to drop the coherence assumption. After we applied the Centering

algorithm and computed a Rough-Shift coherence measure for 100 student essays as described in detail in the next section, we observed a crucial pattern. Namely, in the students' essays, the incoherence detected by the Rough-Shift measure is *not* due to violations of Centering's Pronominal Rule or infelicitous use of pronominal forms in general.

Table 6.2 shows the distribution of nominal forms over Rough-Shift transitions. Out of the 211 Rough-Shift transitions found in the set of 100 essays, in 195 instances, the preferred center or Cp as indicated in the rules in Table 6.1 was a nominal phrase, either definite or indefinite.

Pronominals occurred in only 16 instances, of which 6 cases instantiated the pronominals *we* or *you* in their generic sense. These findings strongly indicate that the incoherence found in student essays is not due to the processing load imposed on the reader to resolve anaphoric references. Instead, the incoherence in the essays is apparently due to discontinuities caused by introducing too many undeveloped topics within what should be a conceptually uniform segment, i.e., the paragraph. This is, in fact, what the Rough-Shift measure picked up. In the next section it is shown that Rough-Shift transitions provide a reliable measure of *incoherence*, correlating well with scores provided by writing experts.

Table 6.2: Distribution of nominal forms over Rough-Shifts

	Def. Phr.	Indef. Phr.	Prons	Total
Rough-Shifts	75	120	16	211
Total	195		16	211

These results not only justify Rough-Shifts as a valid transition type but they also support the original formulation of Centering as a measure of discourse continuity even when anaphora resolution is not an issue. It seems that Rough-Shifts are capturing a source of incoherence that has been overlooked in the Centering literature. The processing load in the Rough-Shift cases reported here is not increased by the effort required to resolve anaphoric reference (i.e., the use of pronouns for entities the readers are not attending to). Instead, the processing load is increased by the effort required to find the relevant topic connections when readers' attention is required to jump from one entity to another, in a

discourse bombarded with a rapid succession of multiple entities. That is, Rough-Shifts are the result of absent or extremely short-lived Cbs. We interpret the Rough-Shift transitions in this context as a reflection of the incoherence perceived by the reader when s/he is unable to identify the topic structure of the discourse.

6.6 The *E-rater* Centering Experiment

In this experiment, we test the hypothesis that a predictor variable derived from Centering can significantly improve the performance of *e-rater*. Since we are in fact proposing Centering's ROUGH-SHIFTs as a predictor variable, the model, strictly speaking, measures *incoherence*. The data consist of student essays whose degree of coherence is under evaluation and therefore cannot be assumed.

The corpus for the experiment came from a pool of essays written by students taking the GMAT test.[5] We randomly selected a total of 100 essays (the same set of 100 essays also mentioned in Section 5) covering the full range of the scoring scale, where 1 is lowest and 6 is highest, as shown in Table A.2 and A.3 in Appendix A. Using students' paragraph marking as segment boundaries (for reasons specified in Section 6.4.1), we applied the Centering algorithm to all 100 essays, calculated the percentage of ROUGH-SHIFTs in each essay and then ran multiple regression to evaluate the contribution of the proposed variable to the *e-rater*'s performance. Although the ROUGH-SHIFT measure itself is simple, its automatic computation raises some interesting research challenges which are discussed here.

6.6.1 Implementation

For this study, we decided to manually tag coreferring expressions despite the availability of coreference software. This decision was made because a poor performance of the coreference software would give distorted results and make it impossible to test our hypothesis.

[5]Many thanks to Jill Burstein who provided the essay set and human and *e-rater* scores.

Similarly, we manually tagged Preferred Centers (as Cp's) for the same reason. The difficulties that can arise with regard to manual annotation and inter-annotator agreement are well-known, and we address this issue in the next section. We also manually tagged other entities in utterances, but we only needed to mark them as OTHER, since this information is sufficient for the automatic computation of the Cb and all of the transitions indicated in Table 6.1. From a natural language engineering perspective, this work highlights the need for more research and development toward reliable named-entity recognizers, coreference resolvers, and software needed to determine Cf ranking, for example syntactic parsers and semantic role identifiers.

Discourse segmentation and the implementation of the Centering algorithm for the computation of the transitions were automated. Segment boundaries were automatically marked at paragraph breaks, and transitions were computed according to the rules given in Table 6.1. As output, the system computed the percentage of Rough-Shifts for each essay. The percentage of Rough-Shifts was calculated as the number of Rough-Shifts over the total number of identified transitions in the essay.[6]

6.6.2 Inter-Annotator Agreement

Manually annotating corpora for specific linguistic features is known to be fraught with difficulties. See Poesio and Vieira (1998) for an account of the issues regarding annotating for definite descriptions. As mentioned in the previous section, we chose to annotate essays manually to identify co-referring expressions and Cp's because truly robust and accurate software for these tasks does not yet exist. We believed that manual tagging would produce more reliable data, especially since the Cp is a well-defined concept, and we did not expect high disagreement. As a reality check for this belief, we performed a small inter-annotator agreement study. We randomly extracted five essays from each of the six scoring levels in the experimental set of 100 essays. We used this set of thirty

[6]We are grateful to Ramin Hemat from the NLP group at the Educational Testing Service for providing the code for the computation of Centering Transitions and the percentage of Rough-Shifts per essay.

essays to compare inter-annotator agreement. A second annotator independently tagged only the Cp in each utterance of these thirty essays in accordance with the Cf ranking rule given in Section 6.4.4.[7] The thirty essays of this inter-annotation set contained 444 utterances.

For the total of 444 annotated Cps, the two annotators were in agreement in 405 cases, that is in 91% of all utterances. In 39 cases the two annotators marked a different noun phrase as the Cp. To examine the effect of the Cp mismatch, I looked at those cases to check if the transition change involved Rough-Shifts. For 31 of the 39 cases of Cp mismatch, choosing a different Cp did not affect the computation of the transition. This is because in most of these cases no Cb was identified in the subsequent utterance, so the Cp of the current utterance did not matter. For 7 of the 8 cases where the Cp mismatch would change the transition, the change involved Continue, Retain and Smooth-Shift transitions (for example, changing a Continue to a Retain or Smooth-Shift and so on). In only one case would the transition change from a Smooth-Shift to a Rough-Shift, thus affecting the value of the Rough-Shift metric for that essay. The results of the inter-annotator study and the close inspection of the effect of the mismatches were very encouraging. In effect, only one case out of the 444 would affect the value of the Rough-Shift metric. To further validate the use of manual tagging, we computed the Kappa statistic for our small study. In the following section, we discuss the computation of the Kappa statistic.

6.6.3 The Kappa Statistic

The Kappa statistic (Cohen, 1960; Kraemer, 1982), introduced to NLP by Carletta (1996) for corpus annotation, has been widely used in the field as a measure of inter-annotator agreement. The Kappa calculation provides a statistical method to correct for chance agreement among annotators. For Kappa > 0.8 annotation is considered reliable. For $Kappa < 0.68$, annotation is considered unreliable. Values in between may allow some tentative conclusions to be drawn (Poesio & Vieira, 1998).

[7] My deepest thanks go to Karen Kukich for volunteering to do the annotation of the set of thirty essays.

The usefulness of the Kappa statistic to quantify levels of agreement has been questioned, however (Maclure & Willett, 1988; Guggenmoos-Holzmann, 1993). The criticism is that the Kappa computation is reliable only in cases where the statistical independence of raters is guaranteed, and raters are by definition dependent because they all rate the same cases according to a pre-specified rule. Critics point out that "Lacking an explicit model of decision-making, it is not clear how chance affects the decisions of actual raters and how one might correct for it."[8] Keeping these concerns in mind, we find it useful to compute the Kappa statistic as a means to compare with Kappa statistics that have been reported in other inter-annotator studies.

The formula for the computation of Kappa is given below:

$$K = \frac{P(A) - P(E)}{1 - P(E)},$$

where P(A) is the proportion of times the annotators agree and P(E) is the proportion of times that we would expect the annotators to agree by chance.[9] To compute the P(E), Poesio and Vieira (1998) give the formula:

$$P(E) = (\frac{\text{number of instances of classification category}}{\text{total number of classification judgments}})^2.$$

To compute P(E) in this case, we observed that the probability of an annotator correctly tagging the Cp is the probability of picking the correct NP out of all the NPs in an utterance. So we computed the average number of NPs for each utterance (by dividing the total number of NPs by the total number of utterances). The average number of NPs per utterance is 4.83. The chance probability of two annotators tagging the same NP as the Cp is $(1/4.83)^2$. P(A) is the percentage agreement for all descriptions, 0.91 in our case. The final computation is given below.

$$K = \frac{(P(A) - P(E))}{(1 - P(E))} = \frac{(0.91 - 0.04)}{(1 - 0.04)} = \frac{0.87}{0.96} = 0.91.$$

[8] http://ourworld.compuserve.com/homepages/jsuebersax/kappa.htm and references therein.

[9] For the details of the formula, its description, and its computation we have consulted (and replicated) the excellent presentation of the Kappa statistic in Poesio and Vieira (1998).

A Kappa of .91 indicates very good inter-annotator reliability, as we expected for this relatively simple task.

This simple study was perhaps even more useful in that it helped us identify causes of disagreement that can be used to further refine a future algorithm for the identification of a Cp. We found that the disagreement instances fell in two main groups. The first group contained instances where there was some apparent confusion as to the ranking of phrases such as *a person, people, impersonal "we" and "they"*, etc. with respect to other indefinite phrases. For example, in (176), one annotator picked *they* as the referent because it was the subject of the sentence. The other picked *rich or lasting success* because *they* referred to *the person*, which is impersonal.

(175) However, real success can be measured depending on what the person wants out of life.

(176) How they define rich or lasting success.

The second group contained cases with *I* as one of the potential Cps. Apparently, it was unclear whether all *I's* were to be ignored, or just the *I's* in the constructions *I think, I believe, I agree*, etc. For example, in (177), one annotator picked *I* as the Cp and the other picked *the service*.

(177) I do not do so because the service has unconventional way of couriering documents.

6.6.4 An Example of *Coherent* Text

What follows is a small excerpt (a paragraph) of a student essay scored 6.[10][11] For each utterance, enclosed in the <UT-n> and </UT> tags, the Preferred Center and OTHER entities are tagged as <CP> and <OTHER> respectively. Each entity is assigned a unique ID number, REF. Following each utterance, the Cb, Cp and transition type are identified. The following paragraph demonstrates an example of a maximally coherent text, centering the company "Famous name's Baby Food" and continuing with the same center through the entire paragraph.

<UT-1> Yet another company that strives for the "big bucks" through conventional thinking is <CP REF='3'>Famous name's Baby Food</CP>.</UT> Cb=none Cp=3 Tr=none
<UT-2><CP REF='3'>This company</CP> does not go beyond the norm in their product line, product packaging or advertising.</UT> Cb=3 Cp=3 Tr=*Continue*
<UT-3>If they opted for an extreme market-place, <CP REF='3'>they</CP> would be ousted.</UT> Cb=3 Cp=3 Tr=*Continue*
<UT-4>Just look who <CP REF='3'>their</CP> market is!</UT> Cb=3 Cp=3 Tr=*Continue*
<UT-5>As new parents, <CP REF='3'>the Famous name</CP> customer wants tradition, quality and trust in their product of choice.</UT> Cb=3 Cp=3 Tr=*Continue*
<UT-6><CP REF='3'>Famous name</CP> knows this and gives it to them by focusing on "all natural" ingredients, packaging that shows the happiest baby in the world and feel good commercials the exude great family values.</UT> Cb=3 Cp=3 Tr=*Continue*
<UT-7><CP REF='3'>Famous name</CP> has really stuck to the typical ways of doing things and in return has been awarded with a healthy bottom line.</UT> Cb=3 Cp=3

[10] Only proper names have been changed for privacy protection. Spelling and other typographical errors have been corrected, also for privacy reasons.

[11] In this and the following example, the identified transitions evaluate the degree of (in)coherence in the quoted paragraphs. This evaluation may not reflect the final score of the essay. The final (in)coherence score for the essay as a whole is based on the sum of the scores of all the paragraphs contained in that essay.

Tr=*Continue*

In the first utterance, the *Famous name's Baby Food* is marked as the Cp because it appears in a main clause, after the verb *to be* in a specificational construction (see Section 4.4). In the second utterance, *this company* is marked as the Cp because it is the subject of the main clause. Similarly, in the third utterance, the referent of *they* is the Cp because it is the subject of the main clause. In the fourth utterance, the implicit subject of imperative form, the impersonal *you*, is ignored, so the referent of *their* is the Cp because it is the highest ranked entity in the complex NP *their market*, following the rule for ranking entities in complex NPs from left to right as explained in Section (6.4.5). In the fifth utterance, the first entity in the complex NP in the subject role, *the Famous name*, is the Cp following the left-to-right ranking of entities in complex NPs. In the sixth and seventh utterances, *Famous name* is the Cp because in both cases it realizes the subject of the main clause.

6.6.5 An Example of *Incoherent* Text

Following the same mark-up conventions, we demonstrate text incoherence with an excerpt (a paragraph again) of a student essay scored 4. In this case, repeated *Rough-Shift* transitions are identified. Several entities are centered, *opinion*, *success*, and *conventional practices*, none of which is linked to the previous or following discourse. This discontinuity, created by the very short lived Cbs, makes it hard to identify the topic of this paragraph, and at the same time it captures the fact that the introduced centers are poorly developed.

<UT-8>I disagree with <CP REF='1'>the opinion</CP> stated above.</UT> Cb=none Cp=1 Tr=none
<UT-9>In order to achieve <CP REF='4'>real and lasting success</CP> <OTHER REF='2'>a person</OTHER> does not have to be a billionaire.</UT> Cb=none Cp=4 Tr=*Rough-Shift*

<UT-10>And also because <CP REF='3'>conventional practices and ways of thinking </CP> can help a person to become rich.</UT> Cb=2 Cp=3 Tr=*Rough-Shift*

In utterance 8, the referent of *I* is ignored and the only other entity realized in the utterance is marked as the Cp. In utterance 9, there is only a main clause, as the infinitive *in order to achieve* is not a tensed clause and therefore does not count as a separate subordinate clause according to our definition. The subject of the main clause, *a person*, ranks lower than the other entities in the utterance because it is an indefinite, non specific, non-referential NP. Furthermore, the verb *to be* in the main clause is predicational and therefore the NP *a billionaire* does not evoke an entity. The subject of the infinitive, the impersonal *you*, is not retrieved. The remaining NP *real and lasting success* is marked as the Cp. In utterance 10, the only available subject *conventional practices* is marked as the Cp.

6.6.6 Results

A summary of the results of applying the Centering algorithm to 100 GMAT essays is shown in Table 6.3. The first column in Table 6.3, labeled HUM, indicates the score level of the essays as graded by human raters. The second column, labeled E-R, gives the average *e-rater* score for all essays at each (human) score level. There were twenty essays each for score levels 6, 5, 4, and 3, and ten essays each for score levels 2 and 1, totaling 100 essays. The third column, labeled ROUGH, shows the average *Rough-Shift* measure at each score level. The full details of the human scores, *e-rater* scores and Rough-Shift measure for each of the 100 essays are shown in Table A.2 and Table A.3 in Appendix A.

Comparing columns HUM and ROUGH in Table 6.3, we observe that essays with scores from the higher end of the scale tend to have lower percentages of *Rough-Shifts* than those from the lower end, repeating the same pattern observed in the preliminary study of 32 essays. To statistically evaluate whether this observation can be used to improve *e-rater*'s performance, we regressed the variable X=ROUGH (the predictor) by

HUM	E-R	ROUGH
6	5.25	22.7
5	4.8	24.95
4	3.6	43.25
3	3	54.37
2	2.33	55.44
1	1.6	55.40

Table 6.3: Summary table with average E-R and ROUGH scores for each essay score

Y=HUM. As expected, the regression yielded a negative coefficient (ROUGH=0.013) for the ROUGH predictor, thus penalizing occurrences of *Rough-Shifts* in the essays. It also yielded a highly significant p-value ($p<0.0013$) on the t-test for ROUGH for these 100 essays, suggesting that adding the variable ROUGH to the *e-rater* model can contribute to the accuracy of the model. [12] The magnitude of the contribution indicated by this regression is approximately 0.5 point, a reasonably sizable effect given the scoring scale (1-6). The full details of the regression output are shown in Table 6.4.

Additional work is needed to precisely quantify the contribution of ROUGH. Ideally, we would incorporate the variable ROUGH into the building of a new *e-rater* scoring model and compare the results of the new model to the original *e-rater* model. Because we could not modify the original *e-rater* model directly, we used a standard statistical technique known as jackknifing (Becker & Chambers, 1984; Mosteller & Tukey, 1977) to simulate the effect of incorporating the ROUGH variable into an *e-rater* model.[13] Jackknifing calls for repeatedly using a random portion of a data set to predict values for the unused portion and averaging over all subset predictions to estimate a whole set prediction. I performed 100 tests with ERATER as the sole variable, leaving out one essay each time, and recorded the prediction of the model for that essay. Then we repeated the procedure

[12] The t ratio is formed by first finding the difference between the estimate and the hypothesized value and then dividing that quantity by its standard error. A significant t ratio indicates that for the tested variable the null hypothesis must be rejected. In our case, the t ratio indicates that the ROUGH variable is significant.

[13] Many thanks to Henry Brown who kindly discussed some statistical issues.

Lack of Fit Source	DF	Sum of Squares	Mean Square	F-Ratio
Lack of Fit	71	53.55	0.75	1.30
Pure Error	24	13.83	0.57	Prob>F
Total Error	95	67.38	0.23	
Max RSq				
0.94				
Parameter Estimates	Estimate	Std Error	t-Ratio	Prob> $\|t\|$
Term				
Intercept	1.46	0.37	3.92	0.0002
E-RATER	0.80	0.06	11.91	<.0001
ROUGH	-0.013	0.0041	-3.32	0.0013
Effect Test	DF	Sum of Squares	F-Ratio	Prob> F
E-RATER	1	100.56	141.77	<.0001
ROUGH	1	7.81	11.01	0.0013

Table 6.4: Regression for the ROUGH variable

using both the ERATER and ROUGH variables. This procedure enabled us to estimate the scores predicted by both *e-rater* alone and *e-rater* enhanced with a Rough-Shift measure.

The predicted values for ERATER alone and ERATER+ROUGH are shown in columns E(PRED) and E+R(PRED) in Table 6.5.

HUM	E(PRED)	E+R(PRED)
6	5.29	5.36
5	4.89	4.98
4	3.78	3.75
3	3.24	3.12
2	2.63	2.59
1	1.97	2.03

Table 6.5: Summary table with E(PRED) and E+R(PRED) scores for each essay score level

As can be seen by comparing the columns E(PRED) and E+R(PRED), the addition of

the *Rough-Shift* measure moved the e-rater score closer to the human score for levels 6, 5, 4, and 2. By examining the detailed comparisons of predictions for each of the 100 essays, shown in Tables A.2 and Table A.3 in Appendix A, we observe that, indeed, 57 % of the predicted values shown in the E+R(PRED) column are better approximations of the human scores, especially in the cases where the *e-rater* score differs by 2 or more points from the human score. In all these cases, the E+R(PRED) value unmistakably tilts the predicted score in the right direction. In summary, the results clearly indicate a greater agreement with human expert scores using a Rough-Shift enhanced version of *e-rater*.

6.6.7 Discussion

The positive finding of this experiment, namely that Centering Theory's measure of relative proportion of Rough-Shift transitions is indeed a significant contributor to the accuracy of computer-generated essay scores, has several practical and theoretical implications. Clearly, it indicates that adding a local coherence feature to *e-rater* could significantly improve *e-rater*'s scoring accuracy. Note, however, that overall scores and coherence scores need not be strongly correlated. Indeed, the essay corpus contains several examples of essays with high coherence scores, i.e., low percentages of Rough-Shifts, but low overall scores and vice versa.

In collaboration with Karen Kukich, director of the NLP group at ETS at the time of the study and leader of the *e-rater* project, we briefly reviewed these cases with several ETS writing assessment experts to gain their insights into the value of pursuing this work further.[14] In an effort to maximize the use of their time with us, we carefully selected three pairs of essays to elicit specific information. One pair included two high-scoring (6) essays, one with a high coherence score and the other with a low coherence score. Another pair included two essays with low coherence scores but differing overall scores (a 5 and a 6). A final pair was carefully chosen to include one essay with an overall score of 3

[14]Many thanks to Mary Fowles, Peter Cooper, and Seth Weiner who provided the valuable insights of their writing assessment expertise.

that made several main points but did not develop them fully or coherently, and another essay with an overall score of 4 that made only one main point but did develop it fully and coherently.

After briefly describing the Rough-Shift coherence measure and without revealing either the overall scores or the coherence scores of the essay pairs, we asked the ETS writing experts for their comments on the overall scores and coherence of the essays. In all cases, the ETS writing experts precisely identified the scores the essays had been given. In the first case, they agreed with the high Centering coherence measure, but one expert disagreed with the low Centering coherence measure. For that essay, one expert noted that ”coherence comes and goes” while another found coherence in a ”chronological organization of examples” (a notion beyond the domain of Centering Theory). In the second case, the experts’ judgments confirmed the Rough-Shift coherence measure. In the third case, the ETS experts specifically identified both the coherence and the development aspects as determinants of the essays’ scores.

Overall, the evaluation of coherence by the ETS writing experts precisely reflected the evaluation of the Centering-based measure of coherence (in its applicable domains) providing extra support of the basic insight we gained from this experiment, namely that Centering under the proposed specification of its center update unit correctly evaluates those aspects of discourse coherence that relate to topic continuity in discourse even in cases in which Centering’s Pronoun rule cannot be evoked for testability.

6.6.8 Remaining Issues

From a practical point of view, the *Rough-Shift* algorithm relies heavily on the efficiency of automated coreference systems. Discourse deictic expressions and nominalizations are especially hard both from a practical and a theoretical point of view and they suggest a number of interesting research projects. We discuss these issues below.

Discourse deixis describes the phenomenon whereby speakers use demonstrative expressions such as *this* and *that* to refer to propositions or in general lengthier parts of the

preceding discourse. Webber (1991) argues that referents for discourse deixis are provided by discourse segments on the right frontier of a formal tree structure. However, what the status of such entities is within the Centering framework remains unclear. Further research is required to determine the effect that the use of such expressions has on textual coherence, compared with simpler entities such as *John* or *the newspaper*.[15] In addition to discourse deixis, the status of nominalizations of verbs or verb phrases is also unclear. The issue of anaphoric nominalizations (essentially, another form of discourse deixis) raises itself in cases where a coherence link could arguably be established between the verb of one utterance and a nominalized version of it, occurring in the subsequent utterance. To give an example, it is possible that in (178) and (179) below the coherence link is established by the semantics of the verb *changes* and the noun *change*.

(178) Many software companies changed their policy.

(179) This change brought about a series of new problems.

Within the Centering framework it is possible to treat these as cases where an utterance has no Cb as, indeed, there are utterances with no Cb. This is always the case, for example, when the utterance is discourse initial, but utterances with no Cb may also be found segment medially (Poesio, Cheng, Henschel, Hitzeman, Kibble, & Stevenson, 2000). In these cases, the Centering literature is unclear as to what the effect of 'No Cb' is on the computation of transitions. We have considered a discourse medial utterance with 'No Cb' equivalent to an utterance whose Cb is different from the previous utterance and the Cb of the current utterance is different from the Cp of the current utterance. This means that a discourse medial utterance with no Cb yields a Rough-Shift transition because there is no link to establish coherence between two consecutive utterances, either by continuing on the previous center or promoting a new center.

On the other hand, discourse deixis and nominalizations are qualitatively different and we would therefore like to distinguish them from cases with no Cb. Unlike cases with no

[15]The judgments required to establish even a working hypothesis were too fine to make and so we decided to omit the utterances including discourse deictic expressions.

Cb to establish a link between two utterances, discourse deictic expressions and anaphoric nominalizations do establish a coherence link between the current utterance and the previous discourse. One problem in integrating this intuition into the current model is that it is not obvious how to represent verb meanings in the Cf set and what the relevant ranking of such entities would be. In the original formulation of the Centering Model, discourse centers are defined as discourse constructs that establish a link between the current, previous, and subsequent discourse. Discourse centers are semantic objects, not "words, phrases, or syntactic forms" (Grosz et al., 1995). It was later shown that in most cases discourse centers can conveniently be mapped to syntactic forms, Brennan et al. (1987), Kameyama (1985), inter alia, but as we see in the case of anaphoric nominalizations, for example, this mapping is not always trivial. To return to the system of this experiment, even if we forced it to detect these cases by comparing the verbs and nouns on a lexico-morphological level, we would still miss cases where the link is based on synonymy or more complex inferencing. Since this issue currently remains unsolved, those potential links were simply missed. Fortunately, such cases were rare. In the essay corpus, there were only three such instances.

Chapter 7

Implications for Anaphora Resolution: A New Algorithm

7.1 Outline of a New Anaphora Resolution Model

In Chapter 3, we discussed a number of challenging cases for anaphora resolution, including some puzzling experimental data. We raised the question of how the data are to be handled. The explanation proposed was based on the hypothesis that topic continuity and intrasentential anaphora are handled by two distinct mechanisms. Topic continuity is computed across center update units. Anaphoric reference across update units relates to topic continuity and is determined structurally in accordance with Centering rules and constraints. Within the unit, anaphora is constrained by resolution preferences projected by the matrix predicate and the extended arguments of the predicate that can be locally realized through subordination.

This basic outline is sufficient to explain most of the data discussed in Section 3.2. The experiments reported in (Stevenson et al., 2000), which show a main effect of thematic focusing, involve the interpretation of anaphoric expressions in subordinate clauses. On the other hand, Hudson-D'Zmura and Tanenhaus's (1998) experiments on similar types of verbs show a main effect of structural focusing. The difference between the two sets

of experiments is that Hudson-D'Zmura and Tanenhaus's experiments involve sequences of main clauses whereas in (Stevenson et al., 2000) the relevant experiments involve subordinate clauses. Furthermore, Stevenson et al. (2000) report results on a different set of experiments showing a main effect of structural focusing and these are precisely the experiments containing sequences of main clauses. Further, Suri et al.'s (1999) "SX because SY" construction indicates that the referent appearing in the subordinate clause is not the preferred focus in the subsequent discourse, whereas resolution to the subject of the main SX clause yields the desired interpretation.

In this chapter we propose a model for anaphora resolution which articulates and integrates the two anaphora resolution mechanisms. The remainder of this chapter is organized as follows. First, in Section 7.2 we provide definitions for the basic tenets of the anaphora resolution model we propose and describe the basic steps required for combining the two mechanisms in a single anaphora resolution algorithm. Next, in Section 7.3 we compare the proposed algorithm with related algorithms and discuss some issues raised by the English connective *so* and certain types of preposed subordinate clauses in Section 7.4. Finally, we conclude in Section 7.5.

7.2 Algorithm and Model Specifications

Discourse consists of a sequence of segments. Each segment consists of a sequence of Centering update units. A single Centering update unit consists of one main clause and all its associated dependent clauses. The three basic types of tensed dependent clauses include: sentential complements of verbs, relative clauses, and adverbial clauses. Sentential complements of verbs and relative clauses are identified syntactically. Adverbial clauses are introduced with subordinate conjunctions. The **reversibility test** is applied to identify subordinate conjunctions: A tensed clause is introduced by a subordinator when

the clause it introduces can be preposed.[1] For example, in (180), *although* is classified as a subordinator and the although-clause is classified as a subordinate clause because placing the although-clause before the main clause retains grammaticality. Conversely, *however* in (182) is not classified as a subordinator because preposing the clause it is associated with yields ungrammaticality.

(180) John traveled by air although he is afraid of flying.

(181) Although he is afraid of flying, John traveled by air.

(182) John traveled by air. However, he is afraid of flying.

(183) # However, he is afraid of flying. John traveled by air.

Update units are identified and numbered. For each identified update unit the list of forward-looking centers is constructed and its members are ranked according to the **ranking rule for English**. The 'M' prefix stands for main clause and the 'S' prefix stands for subordinate clause. The relevant ranking of the various types of dependent clauses is currently left unspecified.

Ranking rule for English
M-Subject $>$ M-indirect object $>$ M-direct object $>$ M-other $>$ S1-subject $>$ S1-indirect object $>$ S1-direct object $>$ S1-other $>$ S2-subject $>$ ...

For the ranking of entities within a clause we assume the following ranking rule: Subject>Indirect Object>Direct Object>qis, pro-arb. For Complex NPs, we assume left-to-right ranking of entities in Complex NPs, as suggested in (Walker & Prince, 1996) and further specified in Chapter 6.

Given the above input for N units $U_{i=1..N}$, the anaphora resolution algorithm starts at the last identified unit.[2] The basic steps are specified below. Some of the steps require

[1] 'Reversibility' is identified as a characteristic of subordinate clauses in (Quirk, Greenbaum, Leech, & Svartvik, 1972).

[2] Starting at the last identified unit is merely a choice made to simplify the implementation of the algorithm. The proposed resolution model can apply starting at the beginning of the discourse and proceed

information that is obtainable by currently available natural language systems: syntactic parsers, morphological analyzers, automated proper name identification, and electronic lexical databases such as Wordnet (to check animacy, for example, as would be necessary for the ranking of entities in Greek). Others, such as understanding and modeling the focusing preferences of verbs and connectives as well as identifying thematic roles, await further research.

ALGORITHM

STEP 0: Start at the last identified unit U_i with $i = N$.

STEP 1: Identify pronominal expressions in the rightmost subordinate clause.

STEP 2: Input antecedents from the Cf list.

STEP 3: Apply grammar-driven constraints (number and gender agreement, contra-indexing etc.) to reduce list of potential antecedents.

STEP 4: Resolve from right-to-left to the first available antecedent inside the subordinate clause. Output unresolved pronominals.

STEP 5: Using the Cf list, resolve pronominals according to semantic focusing constraints. Output unresolved pronominals.

STEP 6: If there is another subordinate clause to process, go to STEP 1.

STEP 7: Identify pronominals in the main clause. Apply grammar-driven constraints (number and gender agreement, contra-indexing etc) to reduce list of potential antecedents. Resolve from right-to-left to the first available antecedent inside the current clause. Output unresolved pronominals.

---

processing one unit at a time, thus more closely resembling human processing.

STEP 8: Input Cf list of potential antecedents from previous unit.

STEP 9: Apply grammar-driven constraints to reduce list of potential antecedents.

STEP 10: Resolve pronominals starting from the leftmost to the highest ranked element of the list of available antecedents.

STEP 11: If an antecedent is found, go to STEP 13.

STEP 12: If the list of potential antecedents is empty and there is a unit to process, go to STEP 8, else mark UNKNOWN.

STEP 13: If U_i is the first unit U_1, terminate, else start processing U_{i-1} and go to STEP 1.

By way of demonstration, we apply the algorithm to resolve the anaphoric expressions in discourse (184)-(186).

(184) Dodge was robbed by an ex-convict.

(185) The ex-convict tied him-3 up because he-2 wasn't cooperating.

(186) Then he-1 took all the money and ran.

- Step 0 applies. Move to Step 1.
- No subordinate clause is identified. Jump to Step 7.
- Step 7 applies. The pronoun he-1 is identified. There is no potential antecedent in the current clause. Move to Step 8.
- Step 8 applies. The Cf list from the previous unit contains EX-CONVICT>HIM-3>HIM-2.
- Step 9 applies. Grammar constraints do not reduce the list of potential antecedents.
- Step 10 applies. HE-1 resolves to the EX-CONVICT.
- Step 13 applies. Move to Step 1.

- Step 1 applies. The pronoun HE-2 is identified.
- Step 2 applies. The Cf list is empty (it contains only the unresolved pronoun he-2).
- Steps 3 and 4 apply vacuously. There are no potential antecedents in the current clause.
- Step 5 applies. The Cf list contains HIM-3>EX-CONVICT due to semantic focusing. HE-2 resolve to HIM-3.
- Step 7 applies. HIM-3 is identified. Grammar constraints apply and contra-index EX-CONVICT with HIM-3.
- Step 8 applies. The Cf list from the previous unit contains DODGE>EX-CONVICT.
- Step 9 applies. Grammar constraints do not reduce the list of potential antecedents.
- Step 10 applies. HIM-3 resolves to DODGE.
- Steps 11-13 apply. The algorithm terminates.

7.3 Comparison with Related Algorithms

The crucial difference between the proposed approach and related anaphora resolution algorithms is in the treatment of subordinate clauses. While steps 7-10 are similar to other approaches which opt to resolve a pronoun to the highest ranked element of the Cf list of the previous clause, the resolution process described in steps 0-7 and the Cf ranking assumptions described earlier are not. As indicated in the ranking rule for English in the previous section, a) subordinate clauses are part of the same unit containing the main clause they are associated with, and b) there is a single Cf ranking list for both the main and the subordinate clauses. Because the entities in the subordinate clauses rank lower than the entities in the main clause, the linear position of the subordinate clause does not affect the resolution process. We have seen that this "restoring" of a basic clause order

results in virtually eliminating backward anaphora which in other approaches requires special treatment.[3] Also, intrasentential anaphora is preferred in the cases of anaphoric elements occurring in subordinate clauses but not in main clauses (assuming grammatical filtering), again irrespectively of their linear order.

We will, now, demonstrate these differences with respect to Lappin and Leass's (1994) and Hobbs' (1978) algorithms, which are conceptually the closest to our approach.

7.3.1 Lappin and Leass 1994

Lappin and Leass's RAP (Resolution of Anaphora Procedure) algorithm applies to the output of McCord's (1990) Slot Grammar parser and utilizes measures of salience derived from syntactic structure and a simple model of attentional state. Potential anaphor antecedents receive a salience score on which they are evaluated. The scoring system penalizes backward anaphora while it rewards parallel syntactic positions and intrasentential antecedents (sentence recency).

As we have already mentioned, backward anaphora need not receive any special treatment in the algorithm proposed here. Lappin and Leass' penalizing of cases of backward anaphora seems to work well on empirical grounds presumably because backward anaphora is rather rare. However, in the absence of an explicit method of identifying real cases of backward anaphora, the system is likely to miss those. In the proposed algorithm, this is not a problem because the Cf ranking of the processing unit implicitly identifies all real cases of backward anaphora and converts them into forward anaphora.

Further, some of the limitations of the system discussed by the authors involve cases of intersentential anaphora such as the following.

(187) a. This green indicator is lit when the controller is on.

b. It shows that the DC power supply voltages are at the correct levels.

[3] Assuming that backward anaphora is restricted to subordinate clauses. Special treatment is required for the *but*-clauses discussed in Section 3.5.2, example (72).

The RAP algorithm resolves the pronoun *it* in (187b) to the *controller* in (187a). This is because, in RAP, both the subject of the main as well as that of *when*-clause in (187a) are of equal salience. In this case, *the controller* wins because it is more recent. In the algorithm that we propose, *it* would resolve to the highest ranked entity of the previous unit, which in this case is correctly identified as the *green indicator*. This is because the *when*-clause is not treated as an independent unit. The entities evoked in the *when*-clause are *linearly* but not *structurally* more recent.

7.3.2 Hobbs 1978

Hobb's (1978) syntactic algorithm is based on a well-defined search procedure (left-to-right in most cases, breadth-first) applied on the surface parse tree. The algorithm has three main components. The first component treats reflexive pronouns by constraining the search procedure with special configurational requirements. The second component takes over when the antecedent of an anaphor is to be found in previous sentences and the third component searches subparts of the parse tree in cycles until the highest S node is reached.

Intersententially, Hobbs' syntactic algorithm favors subjects over objects as subjects are higher up in the parse tree than objects. In such cases, our approach and Hobbs' algorithm would opt for the same type of antecedent. However, as Lappin and Leass (1994) have pointed out, the syntactic search procedure seems to work pretty well in English because grammatical order corresponds to phrase order. For other languages either free word order languages like Greek or languages where salience is determined by other factors (e.g., information status, as it has been argued for German, (Strube, 1998)), Hobbs' search procedure would fail because it is too rigid to accommodate linguistic variation in marking salience. Even for languages like English, the relevant salience of entities may be determined by non-syntactic factors. As has already been suggested by Turan (1998), among others, certain types of NPs are less salient than others independent of their grammatical function (e.g. indefinite quantified expressions, impersonal pronouns etc). The flexibility of constructing lists of entities according to salience both optimizes

the capabilities of an anaphora resolution algorithm and is better suited to accommodate the multiplicity of factors that may have to be taken into account in determining reference salience.

Hobbs' algorithm is, in effect, similar to our approach in the treatment of subordinate clauses. Subordinate clauses belong to the same parse tree with a main clause. This is equivalent to our claim that subordinate clauses are not independent processing units. With respect to backward anaphora, in particular, Hobbs' use of the "command" relation achieves the same result as our lower-ranking of entities appearing in subordinate clauses. The subject of a subordinate clause would be lower in the parse tree than the subject of the main clause independent of the linear position of either. So, for example, in (188), the pronoun would correctly resolve to *Susan*. However, in a case like (189), Hobbs' algorithm would always resolve the pronoun to *Susan* since the search procedure has no way of making a distinction between different types of subordinate connectives (or verbs) and their effect on reference salience.

(188) After she phoned Barbara, Susan went out for dinner.

(189) Susan criticized Barbara because she was lazy.

7.4 Remaining Issues

As mentioned above, the proposed model for anaphora resolution accounts comfortably for the results reported in (Stevenson et al., 2000) with one exception: the experiment involving the connective *so*.

In English, *so* denotes two relations: consequence and purpose. Consequence-*so* is a clause-modifying adverbial. Purpose-*so* is a subordinate conjunction, possibly having dropped the subordinator *that*.[4] The anaphora resolution model we propose predicts

[4]Interestingly, preposed purpose-*so* clauses sound unnatural despite the fact that they are subordinate.

(1) I gave up my job so I could be happy again.

(2) # So I could be happy again, I had to give up my job.

that the interpretation of pronouns in consequence-*so* sentences is determined structurally. This prediction is not borne out. Stevenson et al. (2000) report a main effect of semantic focusing in consequence-*so* continuations.

There are two possibilities available in order to explain the data. First, we may hypothesize that languages arbitrarily characterize their set of subordinate conjunctions. Under this option, we may hypothesize that *so* in English is uniformly a subordinate conjunction and then set out to investigate the implications of such hypothesis on empirical grounds. Alternatively, we may hypothesize that the crucial factor in characterizing subordination is given by its semantic properties, i.e., the type of relation it establishes with the proposition denoted in the main clause. This second option seems intuitively appealing and more promising in explaining this otherwise puzzling linguistic phenomenon, namely, the structural distinction between main and subordinate clauses. However, it runs into the following problem.

In Modern Greek, the equivalent conjunction for the English *so* is *etsi* or *ki etsi* 'and so', which is not polysemous and not a subordinate conjunction. Greek *etsi* links clauses paratactically (i.e., it links sequences of main clauses). The examples below show that Greek behaves differently from English in the *so* cases.

(190) #I Maria-i htipise tin Eleni-j, ki etsi NULL-j evale ta klamata.
the Maria hit the Eleni and so she put the tears.
'Maria-i hit Eleni-j and so she-j started crying.'

(191) I Maria-i xilokopithike apo tin Eleni-j ki etsi NULL-i evale ta klamata.
the Maria was-hit by the Eleni and so she put the tears.
'Maria-i was hit by Eleni-j and so she-i started crying.'

The Modern Greek data show that the null subject in the *so*-clause cannot be interpreted as the object of the previous clause. This is in contrast with the English data reported

(3) I had just been to the bank, so I had money.

(4) # So I had money, I had just been to the bank.

in Stevenson et al, where in the equivalent examples a subject pronoun in the *so*-clause is interpreted as the object of the preceding clause. If subordination was to be defined on semantic grounds then we should not expect focusing differences between the two languages but in fact such differences exist.

Finally, in other instances of subordinate clauses, Greek is much like English as shown in (192)-(193).

(192) I Maria-i htipise tin Eleni-j giati NULL-j ekane ataxies.
the Maria hit the Eleni because she did naughty-things
'Maria-i hit Eleni-j because she-j was being naughty.

(193) I Eleni-j xilokopithike apo ti Maria-i giati NULL-j ekane ataxies.
the Eleni was-hit by the Maria because she did naughty-things.
'Eleni-j was hit by Maria-i because she-j being naughty.'

The difference between the two languages with respect to *so*-clauses is hard to explain. This difficulty in understanding the cross-linguistic variation is also telling of our fundamental lack of understanding subordination in languages. While we have shown that the distinction between main and subordinate clauses is in the right direction, it is not yet clear what property of subordination —structural, semantic or other— is responsible for the observed pattern.

Another issue that requires special attention in the proposed account pertains to some special cases of preposed subordinate clauses. Example (194) presents a problem for the proposed model because the antecedent of the subject pronoun in the matrix clause is the subject of the preposed subordinate clause.

(194) After Susan phoned Barbara, she went out for dinner.

The ranking in the Cf list for (194) is she-referent>Susan>Barbara. In effect, what we are faced with here is analogous to "backward anaphora". However, in its current form, the proposed algorithm would process the subordinate clause first and would then move to the matrix clause. The matrix clause contains a pronoun and no possible antecedent so, on completing the processing of the unit, the algorithm would output the unresolved

pronoun from the matrix clause and would continue searching for an antecedent in the previous unit. Such cases can be identified easily by even shallow parsing and can be fixed locally by forcing resolution to the highest entity in the current unit, i.e., *Susan*. Also, the algorithm presented in Section 7.2 could be modified so that in step 2 the Cf list includes all possible antecedents from the current utterance U_i. With this modification, (194) would be processed correctly but this modification would not explain the contrast in (195).

(195) a. Susan phoned Barbara. Then, she went out for dinner.

b. Susan phoned Barbara before she went out for dinner.

c. After Susan phoned Barbara, she went out for dinner.

Example (195a) is an instance of intersentential anaphora and there is a subject reference for the pronoun as predicted. Example (195b) is a case of intrasentential anaphora and there is no clear subject reference. Example (195c) is another instance of intrasentential anaphora but in this case the subject preference is clear, on a par with the intersentential case in (195a). Whichever required modification to the algorithm will prove to be more useful, the fact remains that the similarity between (195a) and (195b) remains unexplained in purely structural terms. We suspect that the difference between (195b) and (195c) and the similarity between (195a) and (195c) is the result of an interaction with a discourse function of subordinate clauses. Subordinate clauses normally convey background information and do not by themselves move the narrative forward. They also have the property of enabling information to appear in a "non-natural order" with respect to the event(s) of main clause. A "natural order" for temporal connections would be to express events in the order in which they happened. For causal connections, a "natural order" would be to express the 'cause' before the "effect". So, it seems plausible to hypothesize that subordinate structures can be used to introduce background (or presupposed) information and even discourse-new characters without disturbing the narrative structure of the discourse and the salience of the centers of attention already established in the narrative. If this line of thinking is on the right track, then it is possible that the similarity between

(195a) and (195b) is due to the fact that both sequences of clauses reflect the linear succession of events. The preposed *after*-clause does not disturb the natural temporal order of events both of which are predicated of the same center, which in this case is introduced in the subordinate clause. Further empirical work is clearly needed to evaluate this line of explanation.

7.5 Summary and Conclusions

The interpretation of anaphoric expressions in natural language processing is not a trivial problem. Extensive research in the past 30 years has made significant contributions to our understanding of the phenomenon, and a considerable amount of theoretically motivated and/or corpus-based anaphora resolution algorithms have been built with more or less success. However, the task remains a challenge and the slow rate of improvement in the performance of anaphora resolution systems is somewhat alarming.

The review of the relevant literature in Chapter 3 revealed that a lot of the complications and inconsistencies in anaphora resolution start when algorithms are faced with anaphoric elements in complex sentences. In particular, we saw that the interpretation of anaphoric expressions in certain types of clauses would defy any algorithm based on registers of NPs and a uniform look-up mechanism.

The main finding in this thesis, i.e., the importance of subordinate clauses in understanding the distinction between topic continuity and intrasentential anaphora, applies directly to anaphoric interpretation and justifies the specification of two systems that determine preferences for anaphoric interpretation. Contra some earlier views on the status of subordinate clauses, in this thesis we have argued that subordinate clauses do not constitute independent processing units. In fact, subordinate clauses can be seen as filling up extended argument positions required by the predicate of the matrix clause and, in this respect, intrasentential relationships which hold between predicates and participating entities should be expected to be closely determined on semantic grounds. We have identified

the boundaries of the basic discourse units with the boundaries of the unit containing a matrix clause and all its dependent clauses and suggested that anaphoric interpretation within this unit is determined semantically by the focusing properties of the verbs and connectives.

On the other hand, topic continuity, as evaluated in the Centering Model, requires rather arbitrary specifications of salience in order to facilitate discourse processing and efficient integration of meaning to previous discourse. Discourses grow enormous very quickly. Unrestricted semantic representations and the resulting inferencing load imposed by exploding semantic computations would considerably slow down discourse processing (Kohlhase & Koller, 2000). The notion of salience, in the sense of Centering (Joshi & Kuhn, 1979), is arguably crucial for efficient processing not only for NLP systems but also for humans. Topic continuity therefore is evaluated by a salience mechanism operating across processing units and we have showed that this mechanism is structural and best defined in Centering terms. We then argued that anaphoric reference which spans across units is also determined structurally.

Regarding Centering-based anaphora resolution algorithms, which seem the best candidates for anaphora across units, the technical problems discussed in Chapter (3) are easily fixed. The algorithm presented here selects as the preferred antecedent the highest ranked entity in the previous unit (see Chapter 3). This modification is, in fact, consistent with Centering's Pronoun Rule and at the same time does not rely on the assumption that text is maximally coherent.

The corpus-based study reported in Section 4.2.2 tests the hypothesis that two mechanisms are indeed at work and also evaluates the strengths of the modified Centering-based algorithm for resolving anaphoric reference across units. The results were robust despite the moderate sample size, suggesting a number of future projects in this direction, the most challenging of which will probably be further understanding the structural and semantic properties of subordination and its role in the organization, representation, and structure of discourse.

Chapter 8

Conclusions

The study of entity salience, topichood, and discourse coherence is plagued with the well-known and daunting complexity that the analysis of even the simplest of discourses readily reveals. Numerous factors have been identified over decades of research in discourse interpretation, some of them somewhat easier to control, such as the semantics of individual linguistic forms, and others much harder to model, such as the effect of speaker's knowledge, contextual information, and communicative intent. Recent years have seen an increased interest in models of interpretation based on the interaction of multiple constraints applying simultaneously and competing with one another in on-line interpretation. This approach is especially prevalent in the psycholinguistics literature. In the computational linguistics literature, modeling discourse interpretation is in its infancy. The most prevalent path of investigation is increasingly concerned with the contribution of overt lexical information at the expense of modeling principles of discourse organization and interpretation that may be inscrutable from overt lexicalizations such as the nature and amount of inferencing that is required and *expected* in discourse interpretation.[1]

The study of multiple constraints applying in interpreting discourse are based mainly on frequency information, speaker's intent, and knowledge store is undoubtedly useful and most likely to yield a wider coverage of the phenomena at hand. However, such an

[1]Notable exceptions are the contributions of B. Grosz, C. Sidner, A. Joshi and their collaborators.

approach yields comprehension and production models of such high computational complexity that it makes the task of explaining the speed and efficiency of language-mediated communication hard. While it is premature and possibly simply wrong to equate language processing as performed by human brains to the level of information processing that a machine is capable of, human brains and machines are nevertheless similar in that they are both constrained by bounded resources. Models based on unbounded resources may be successful in the tasks set for themselves but they may not reveal the nature and amount of computation that is *minimally* necessary to process language and effect communication.

This thesis is driven by an approach to the study of discourse interpretation that seeks to tease apart possibly autonomous or semi-autonomous mechanisms that apply at different levels of discourse interpretation. Ultimately, a complete model of discourse interpretation will have to provide a unified model of the interaction and collaboration of such subsystems and the effect of such a unified model in deriving discourse meaning. However, we see a significant advantage in identifying and studying sub-components of discourse organization and interpretation in isolation even if later integration of such sub-models will require significant revisions and readjustments of their properties and responsibilities.

This thesis has been concerned with that component of discourse organization that is responsible for topic management. Based on previous research, we have assumed that, in an imaginary discourse consisting of a sequence of main clauses only, each main clause would serve as an independent unit for the computation of topic structure. However, discourses are rarely constructed as a sequence of main clauses. Instead, we find that main clauses are often accompanied by one or more (tensed) subordinate clauses.

Our investigation of the role of tensed subordinate clauses in topic management has focused on those subordinate clauses that do not serve as arguments of the matrix verbs, namely, adjunct subordinate clauses. Chapters 4 and 5 were devoted to the analysis of adverbial and relative clauses, respectively. For both adverbial and relative clauses, our findings indicate that entities evoked in adjunct subordinate clauses are less salient than the entities evoked in main clauses. Crucially, this is true even for entities evoked in subject

position in the subordinate clause.

With respect to relative clauses, in particular, our conclusions were drawn from a number of studies that we conducted in English and Greek. As a first evaluation of the potential of a relative clause to host a topical entity, we counted how many times entities evoked in a relative clause in any position were subsequently referenced. English and Greek differed in the degree to which relative clause entities continued to be relevant in the subsequent discourse. In Greek, relative clause entities were referenced more frequently than their English counterparts. However, in both languages the numbers were low, especially so for entities that were not coreferent with the head noun, which was already evoked in the main clause. An analysis of the referring expressions used for reference to relative clause entities confirmed their low salience status. With only few exceptions a pronoun was used to refer to a relative clause entity only when that entity was co-referent with a subject head noun or when the subject of the main clause had already been pronominalized in the same sentence containing the pronominal reference to the relative clause entity.[2] To explicitly evaluate the contribution of relative clauses in topic continuity in discourse, we conducted a Centering-based study whose findings indicate that, in the computation of topic continuity, relative clauses are processed as a single unit with the main clause on which they depend. If we focus on sentence-final relative clauses, preceding another main clause, our findings show that processing relative clauses as independent topic update units creates topics discontinuities and raizes puzzles in the use of pronouns to refer to an entity that was evoked in the main clause but was not mentioned in the more recent relative clause itself. Configuring the relative clause as an atomic unit for updating topics yields a Centering topic transition of relatively low coherence. In contrast, when the same relative clause is processed together with its main clause, topic continuity is recovered and the use of a pronoun for reference to the topic of that unit (established in the main clause) is to be

[2]More precisely, a pronoun was used to refer to the head noun entity evoked in the main clause when that entity was the most salient entity of the main clause. In some cases, the most salient entity was not the subject, as, for example, when the subject represented an impersonal referent or was not referring at all.

expected as a signal of a discourse continuing on the same topic.

For the study of the salience of entities in adverbial clauses, we added two experimental studies in English and Greek which gave us additional insights into pronominal interpretation. In controlled experimental conditions we compared the most likely interpretation of a subject pronoun in main and adverbial clauses. The results of the experiments showed that subject pronouns in main clauses were consistently interpreted as the subject of the preceding main clauses. In contrast, the interpretation of the subject pronoun in an adverbial clause varied between the subject and the object of the preceding clauses. Given the strict experimental conditions whereby all first main clauses contained an action predicate and two same-gender referents, the variation was most likely due to the semantics of the subordinate conjunctions. The main-main condition also included adverbial phrases in clause-initial position of varying semantics but that did not significantly affect the overall tendency of the subsequent subject pronoun to be interpreted as the subject of the preceding main clause. Since the interpretation of a main clause subject pronoun is associated with a topical entity, whereas the interpretation of a subordinate subject pronoun is not, the results of these experiments support to the hypothesis that pronominal interpretation is subject to at least two distinct mechanisms. *Within* topic update units, defined by the syntactic locality formed by the main clause and its dependent clauses, pronouns are resolved locally and their interpretation is primarily determined by the semantics of verbs and connectives. Unresolved pronouns search for their antecedent across units. *Across* units, where topic continuity is computed, we expect pronoun interpretation to be affected by the strategies used by speakers for topic management. In English and Greek, topics are represented in structurally prominent positions, e.g., the subject of the main clause. So, if there is a single unresolved pronoun, we expect this pronoun to be interpreted as the subject of the main clause in the preceding unit, or, if the subject is not an acceptable option, the next available entity in the ranking. For the computation of topic continuity, therefore, the salience ranking of entities evoked in the topic update unit is crucial as it projects preferences for the interpretation of upcoming unresolved pronouns, which presumably serve

as links between the units. This distinction between pronominal interpretation within and across topic update units goes a long way in accounting for numerous anaphora resolution puzzles and for some seemingly contradictory results identified in the related literature and discussed in Chapter 3.

The findings of the experimental results were replicated in a corpus study in Greek, reported in Chapter 4, where, again, we found that a dropped subject or weak pronoun in Greek main clauses was co-referring with the most salient entity in the preceding sentence (with compatible morphological features), often skipping over other more recent competing antecedents. Consistent with the main-subordinate experimental condition, the interpretation of a dropped subject or weak pronoun in an adverbial clause varied, half of the time co-referring with an entity that ranked lower than other compatible entities evoked in the preceding discourse.

The topical status of entities evoked in adverbial clauses with respect to subsequent reference and the choice of referring expression was independently investigated by Cooreman and Sanford (1996). Their experimental findings are consistent with our corpus findings for relative clauses. They show that a subject pronoun in a main clause following a main and a subordinate clause, each introducing a same gender referent in subject position, is interpreted as the subject of the main clause independently of the surface order of the preceding main and adverbial clause. These findings confirm the hypothesis that the interpretation of pronouns in main clauses is subject to the mechanism responsible for computing topic structure in discourse and are consistent with the hypothesis that an atomic topic update unit includes both the main clause and its dependent subordinate clauses.

The main contribution of our studies in relative and adverbial clauses is disentangling two processes that are often confounded in the literature: topic continuity and anaphora resolution. With respect to topic continuity, in Chapter 3, we specified the details of a very simple model that we have proposed for the computation of topic structure. According to the proposed model, topic structure is a component of discourse representation that can be

computed independently of information structure and which is distinct from anaphora resolution. Information structure is relevant to the computation of topic continuity only for those languages which might use word order, for example, to determine entity salience. Topicality overlaps with information status (old versus given, for example), possibly cross-linguistically, only to the extent that topics by definition tend to be discourse-old entities. In the model proposed, the basic unit for the computation of topic structure is the sentence defined in syntactic terms. Topic continuity is computed across topic update units. For the computation of topic continuity, entities evoked in each unit are ranked according to a salience hierarchy which, as proposed by researchers working in the Centering paradigm, may vary cross-linguistically. For English and Greek, the salience hierarchy is determined primarily by grammatical function. In this hierarchy, entities evoked in subordinate clauses rank lower than entities evoked in main clauses, independently of their surface order. While entities evoked in subordinate clauses are available for subsequent reference, topic continuations and shifts to new topics are established in main clauses.

The proposed model of topic continuity is articulated in Centering terms. Centering provides all and only the necessary concepts for modeling topic continuity. In fact, we have argued extensively that, despite previous attempts to transform Centering from a model of local coherence to a model of anaphora resolution, Centering is best suited to modeling entity-based topic structure. The main contribution of this dissertation to the Centering model is the definition of the center update unit, Centering's previously undefined *utterance*, on an empirical basis. Centering's Pronoun Rule, when applicable, reflects precisely our finding that only a subset of pronouns, in fact a single pronoun per unit, is interpreted as the current topic of the local discourse.

With respect to anaphora resolution, in Chapter 7 we proposed an anaphora resolution algorithm based on the insight that the interpretation of pronouns is affected by two distinct mechanisms: the mechanism responsible for resolving the pronouns to topical entities and the mechanism responsible for assigning interpretations to pronouns appearing in the syntactic locality formed by the main clause and its dependent subordinate clauses.

The two mechanisms are interleaved in the algorithm. Roughly, the algorithm uses the proposed center update unit as the basic processing unit. Entities evoked in each unit are ranked according to a salience ranking rule whereby entities in subordinate clauses rank lower than entities in the main clause. Pronouns appearing intrasententially are first resolved locally according to the (as yet unknown) preferences projected by verb semantics and the semantics of subordinate conjunctions. Unresolved entities, most likely to be found in the main clause, then search for their antecedents across units, starting from the highest ranked entity of the preceding unit. We did not perform an evaluation of the algorithm as a significant component of the algorithm, i.e., the one responsible for the resolution of intrasentential entities, requires further research into the semantics of predicates and conjunctions.

Finally, in Chapter 6, we proposed a computational model for the evaluation of discourse coherence. The model is based on Centering and incorporates the basic findings of this dissertation with regard to topic continuity. For the evaluation of the proposed model we used a corpus of essays written by students taking the GMAT exam. For each essay we had two scores available: the score received by human raters and the score received by *e-rater*, an electronic essay rating system developed at ETS. Based on a preliminary examination of the correlation between centering transitions and essay scores we devised an evaluation metric based on Rough-Shift transitions. We then performed a statistical analysis of the contribution of the Rough-Shift metric to the performance of the *e-rater*. The purpose of the statistical analysis was to evaluate if the essay score predicted by the *e-rater* reinforced by the Rough-Shift metric of discourse coherence better approximated human scores. Our positive results support the validity of Centering as a model of local discourse coherence and the definition of the center update unit proposed in this dissertation, and open up new directions for the applicability of the Centering model to natural language processing. Further, the *e-rater* experiment provided substantial support for the validity of the previously unstable role of Rough-Shift transitions in the Centering model. One of the new insights obtained in the course of conducting the *e-rater* experiment was

the previously unnoticed potential of the Centering model to capture a source of low coherence that is not attributed to poor uses of pronouns. The Rough-Shift metric captured a source of incoherence that was due to extremely short-lived or absent topic links, which made it hard for the reader to integrate the meaning of the current unit to the preceding discourse. We showed that this source of incoherence was captured independently of Centering's Pronoun Rule, thus providing evidence for the validity of the model even in cases where the Pronoun Rule cannot apply as a reality test for the model.

This thesis has focused on the effect of adjunct subordinate clauses on entity salience and topic structure. For a complete model of topic structure in discourse, further work is required. Of high priority is work on the salience status of entities evoked in complement clauses. Complement clauses can be especially interesting because they serve as arguments of a limited type of verbs (e.g., *say*, *believe*, *claim*, etc), also known as attitude verbs. On an intuitive level, the higher clause evokes an entity whose attitude towards the proposition expressed in the complement clause is expressed in the higher clause. It is the complement clause that contains the actual proposition. With respect to topic continuity, however, it is unclear what the topic of such a unit is. It could be either the/an argument of the higher clause, or else some salient entity evoked in the complement clause on which the information in the complement clause is predicated. Following the complement clause, the subsequent discourse may elaborate either on entities introduced in the complement clause or continue on the entities introduced in the higher clauses. That both continuations are possible can easily be verified by eyeballing even small sections of the *Wall Street Journal*, which often reports what topical characters say or claim. Elaborations of an entity evoked in a complement clause may span several clauses. Still, it is possible for the discourse to return to the entity evoked in the higher clause. What is not clear is exactly how such a complex type of topic management is organized. In other words, it is not clear what strategies are used to help the reader recognize when the discourse opened under the complement clause is completed, the conditions under which it is possible to return to a topical entity introduced in the higher clause, and the kinds of

referring expressions that can be used when doing so. Another important research direction suggested by the findings of this dissertation includes the study of reference patterns when the discourse relation established by a subordinate conjunction is merely implicit in the discourse. A causal relationship, for example, established with the subordinate conjunction *because* may be inferred between two main clauses. Will the interpretation of the subject of the second main clause then be driven by the semantics of the causal relationship, or will its interpretation be driven by the mechanism of topic management that we have proposed? We leave the questions raised by complement clauses and by implicit subordinating relations open for future research.

Appendix A

E-rater Tables

Score	File	CONTINUE	RETAIN	SMOOTH SHIFT	ROUGH SHIFT	% OF ROUGH SHIFTS
6	1	5	4	1	3	23
6	2	6	4	-	1	9
6	3	5	1	2	1	10
6	4	5	2	3	-	0
6	5	7	-	1	2	20
6	6	3	2	1	4	40
6	7	3	-	-	-	0
6	8	7	4	1	3	20
						$< 25\%$
5	1	11	-	-	-	0
5	2	9	-	2	2	15
5	3	6	2	3	3	21
5	4	4	1	3	1	11
5	5	7	6	-	1	7
5	6	2	4	1	2	22
5	7	7	3	2	4	25
5	8	5	2	-	2	22
						$< 25\%$
4	1	1	2	3	4	40
4	2	3	-	2	4	44
4	3	1	-	1	2	50
4	4	1	-	-	5	83
4	5	1	1	1	2	40
4	6	-	-	1	2	66
4	7	4	4	4	1	7
4	8	9	-	2	-	0
						$> 40\%$
3	1	2	1	1	3	50
3	2	-	-	-	-	-
3	3	3	-	2	4	44
3	4	1	1	1	3	50
3	5	-	3	-	2	40
3	6	3	1	-	-	0
3	7	-	-	-	-	-
3	8	2	2	1	3	37
						$> 40\%$

Table A.1: Table with centering transitions of 32 GMAT essays

HUM	E-R	ROUGH	E(PRED)	E+R(PRED)
6	5	15	5.05	5.26
6	6	22	5.9921	5.9928
6	6	15	5.99	6.09
6	6	22	5.9921	5.9928
6	6	24	5.99	5.96
6	4	22	4.13	4.35
6	4	13	4.13	4.46
6	6	28	5.99	5.90
6	5	30	5.0577	5.0594
6	4	30	4.13	4.24
6	4	0	4.13	4.62
6	5	20	5.05	5.19
6	6	21	5.99	6.00
6	6	50	5.99	5.58
6	6	25	5.99	5.94
6	5	21	5.05	5.18
6	6	6	5.99	6.22
6	5	35	5.05	4.98
6	5	25	5.05	5.12
6	5	30	5.057	5.059
5	4	15	4.14	4.46
5	5	7	5.07	5.40
5	4	5	4.14	4.60
5	5	38	5.07	4.96
5	4	40	4.14	4.12
5	5	45	5.07	4.86
5	6	27	6.02	5.95
5	4	30	4.28	4.14
5	5	21	5.07	5.20
5	5	16	5.07	5.27
5	5	20	5.07	5.22
5	6	32	6.02	5.88
5	4	40	4.143	4.148
5	4	10	4.14	4.53
5	4	23	4.14	4.35
5	5	20	5.07	5.22
5	6	25	6.02	5.98
5	4	25	4.14	4.33
5	5	50	5.07	4.79
5	6	10	6.02	6.20
4	3	11	3.22	3.71
4	5	45	5.09	4.88
4	4	46	4.15	4.04
4	3	50	3.22	3.17
4	3	36	3.22	3.37
4	3	33	3.22	3.41
4	5	42	5.09	4.92
4	3	50	3.22	3.17
4	4	36	4.15	4.18
4	4	40	4.15	4.13

Table A.2: Table with the human scores (HUM), the *e-rater* scores (E-R), the Rough-Shift measure (ROUGH), the (jackknifed) predicted values using *e-rater* as the only variable, E(PRED), and the (jackknifed) predicted values using the *e-rater* and the added variable Rough-Shift, E+R(PRED). The ROUGH measure is the percentage of Rough-Shifts over the total number of identified transitions. The question mark appears where no transitions were identified.

HUM	E-R	ROUGH	E(PRED)	E+R(PRED)
4	3	11	3.22	3.71
4	3	75	3.22	2.79
4	4	38	4.15	4.16
4	3	62	3.22	3.00
4	4	12	4.15	4.53
4	4	40	4.15	4.13
4	5	48	5.09	4.84
4	3	9	3.22	3.74
4	3	81	3.22	2.69
4	3	100	3.22	2.34
3	3	55	3.24	3.11
3	4	30	4.16	4.28
3	4	81	4.16	3.59
3	4	42	4.16	4.11
3	3	50	3.24	3.18
3	3	66	3.24	2.96
3	3	42	3.24	3.30
3	2	40	2.30	2.50
3	3	75	3.24	2.83
3	3	40	3.24	3.33
3	3	78	3.24	2.78
3	3	62	3.24	3.02
3	2	55	2.30	2.29
3	2	30	2.30	2.64
3	3	?	3.29	?
3	5	45	5.11	4.91
3	3	80	3.24	2.75
3	2	37	2.30	2.54
3	3	75	3.24	2.83
3	2	50	2.30	2.36
2	2	67	2.32	2.14
2	2	67	2.32	2.14
2	4	78	4.17	3.68
2	3	67	3.25	2.97
2	3	41	3.25	3.33
2	2	?	2.32	?
2	1	67	1.37	1.30
2	2	20	2.32	2.84
2	2	42	2.32	2.50
2	2	50	2.32	2.39
1	2	50	2.35	2.41
1	2	0	2.35	3.29
1	1	67	1.42	1.35
1	3	71	3.26	2.95
1	3	57	3.26	3.12
1	0	100	0.44	-0.03
1	1	85	1.42	1.09
1	1	67	1.42	1.35
1	2	57	2.35	2.31
1	1	0	1.42	2.48

Table A.3: (Continued from Table 4) Table with the human scores (HUM), the *e-rater* scores (E-R), the Rough-Shift measure (ROUGH), the (jackknifed) predicted values using *e-rater* as the only variable, E(PRED), and the (jackknifed) predicted values using the *e-rater* and the added variable Rough-Shift, E+R(PRED). The ROUGH measure is the percentage of Rough-Shifts over the total number of identified transitions. The question mark appears where no transitions were identified.

Score	File	Continue	Retain	Smooth Shift	Rough Shift
6	1	10	3	4	3
6	2	4	1	3	5
6	3	6	4	4	4
6	4	9	2	5	3
6	5	8	3	1	3
6	6	8	1	5	4
6	7	13	1	3	5
6	8	5	2	2	4
6	9	10	3	6	8
6	10	15	2	5	7
6	11	23	1	3	2
6	12	10	0	3	7
6	13	4	6	2	0
6	14	6	3	2	3
6	15	11	2	3	4
6	16	4	0	9	13
6	17	5	1	0	2
6	18	10	3	5	4
6	19	7	0	5	4
6	20	14	0	2	7
5	21	12	3	0	1
5	22	7	5	7	2
5	23	3	3	2	5
5	24	5	1	4	6
5	25	6	1	5	10
5	26	4	2	2	3
5	27	6	0	3	4
5	28	10	1	0	2
5	29	10	1	2	3
5	30	5	1	5	3
5	31	7	3	2	3
5	32	13	1	3	8
5	33	9	0	0	1
5	34	4	3	3	3
5	35	5	0	3	2
5	36	8	6	5	5
5	37	3	0	0	1
5	38	3	1	4	3
5	39	12	11	4	3
5	40	7	0	1	6

Table A.4: Table with Centering transitions for essay scores 5 and 6

Score	File	Continue	Retain	Smooth Shift	Rough Shift
4	41	2	3	2	2
4	42	3	0	0	7
4	43	7	0	1	5
4	44	3	0	3	6
4	45	4	1	2	4
4	46	2	2	0	2
4	47	1	1	2	3
4	48	1	2	3	6
4	49	7	1	3	7
4	50	8	1	3	7
4	51	5	3	0	1
4	52	0	0	1	3
4	53	5	0	3	5
4	54	0	1	2	5
4	55	6	0	1	1
4	56	4	0	1	2
4	57	6	4	3	12
4	58	4	4	2	1
4	59	0	2	0	9
4	60	0	0	0	5
3	61	2	2	0	5
3	62	2	4	1	3
3	63	2	0	0	9
3	64	6	0	2	6
3	65	1	0	3	4
3	66	0	1	1	4
3	67	4	0	0	3
3	68	4	1	0	3
3	69	0	0	1	3
3	70	1	1	1	2
3	71	2	0	1	11
3	72	0	1	3	5
3	73	2	1	1	5
3	74	6	1	0	7
3	75	7	1	1	4
3	76	0	0	0	0
3	77	4	5	3	10
3	78	0	0	1	4
3	79	3	0	2	3
3	80	0	1	3	12

Table A.5: Table with Centering transitions for essay scores 3 and 4.

Score	File	Continue	Retain	Smooth Shift	Rough Shift
2	81	0	0	1	2
2	82	0	1	0	2
2	83	0	0	2	7
2	84	2	2	0	8
2	85	5	0	2	5
2	86	0	0	0	0
2	87	1	0	1	4
2	88	0	2	2	1
2	89	2	1	1	4
2	90	1	1	1	3
1	91	1	0	1	2
1	92	2	2	1	0
1	93	0	0	1	2
1	94	1	1	0	5
1	95	2	1	0	4
1	96	0	0	0	2
1	97	0	0	1	6
1	98	1	0	0	2
1	99	3	0	3	8
1	100	1	0	0	0

Table A.6: Table with Centering transitions for essay scores 1 and 2. Note that the counts in Tables (A.4), (A.5), and (A.6) are based on the earlier Cf ranking rule proposed in Brennan, Friedman and Pollard 1987.

Bibliography

Ariel, M. (1988). Referring and Accessibility. *Journal of Linguistics, 24*, 65–87.

Ariel, M. (1990). *Accessing Noun-phrase Antecedents*. London: Routledge.

Arnold, J. (1998). *Reference Form and Discourse Patterns*. Ph.D. thesis, Stanford University.

Becker, R., & Chambers, J. (1984). *An Interactive Environment for Data Analysis and Graphics*. Wadsworth, Inc., Belmont CA.

Brennan, S., Walker-Friedman, M., & Pollard, C. (1987). A Centering Approach to Pronouns. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics, Stanford, CA*, pp. 155–162.

Burstein, J., Kukich, K., Wolff, S., Chodorow, M., Braden-Harder, L., Harris, M. D., & Lu, C. (1998). Automated Essay Scoring Using a Hybrid Feature Identification Technique. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics, Montreal, Canada*.

Caramazza, A., & Gupta, S. (1979). The Roles of Topicalization, Parallel Function and Verb Semantics in the Interpretation of Pronouns. *Linguistics, 3*, 497–518.

Carden, G. (1982). Backwards Anaphora in Discourse Context. *Journal of Linguistics, 18*, 361–387.

Carletta, J. (1996). Assessing Agreement on Classification Tasks: The Kappa Statistic. *Computational Linguistics, 22*(2), 249–254.

Chafe, W. (1976). Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View. In Li, C. N. (Ed.), *Subject and Topic*, pp. 27–55. New York: Academic Press.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psycological Measurement, 20*, 37–46.

Cooreman, A., & Sanford, A. (1996). Focus and Syntactic Subordination in Discourse. Tech. rep., Human Communication Research Center.

Daneŝ, F. (1974). Functional Sentence Perspective and the Organization of the Text. In Daneŝ, F. (Ed.), *Papers on Functional Sentence Perspective*, pp. 106–128. Prague: Academia.

Deerwester, S., Dumais, S., Furnas, G., Landauer, T., & Harshman, R. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science, 41*, 391–407.

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

Di Eugenio, B. (1998). Centering in Italian. In Walker, M., Joshi, A., & Prince, E. (Eds.), *Centering Theory in Discourse*, pp. 115–137. Clarendon Press: Oxford.

Dimitriadis, A. (1996). When Pro-Drop Languages Don't: Overt Pronominal Subjects and Pragmatic Inference. In *Proceedings of the 32nd Meeting of the Chicago Linguistics Society, Chicago*, pp. 33–47.

Foltz, P., Kintsch, W., & Landauer, T. (1998). The Measurement of Textual Coherence with Latent Semantic Analysis. *Discourse Processes, 25*, 285–307.

Forbes, K., Miltsakaki, E., Prasad, R., Sarkar, A., Joshi, A., & Webber, B. (2001). D-LTAG System: Discourse Parsing with a Lexicalized Tree Adjoining Grammar. In *Proceedings of the ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*. (First four authors in alphabetical order).

Fox, B. (1987). *Discourse Structure and Anaphora*. Cambridge: Cambridge University Press.

Givón, T. (1983). *Topic Continuity in Discourse: A Quantitative Cross-Language Study*, chap. Topic Continuity in Discourse: An Introduction, pp. 1–42. Amsterdam: John Benjamins Publishing.

Grosz, B. (1977). The Representation and Use of Focus in Dialogue Underastanding. Tech. rep. No. 151, Menlo Park, Calif., SRI International.

Grosz, B., Joshi, A., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics, MIT Press, Cambridge*, pp. 44–50.

Grosz, B., Joshi, A., & Weinstein, S. (1995). Centering: A Framework for Modeling Local Coherence in Discourse. *Computational Linguistics*, *21*(2), 203–225.

Grosz, B., & Sidner, C. (1986). Attentions, Intentions and the Structure of Discourse. *Computational Linguistics*, *12*(3), 175–204.

Guggenmoos-Holzmann, I. (1993). How Reliable Are Chance-Corrected Measures of Agreement?. *Statistics in Medicine*.

Gundel, J. (1974). *The Role of Topic and Comment in Linguistic Theory*. Ph.D. thesis, University of Texas.

Gundel, J. (1985). 'Shared Knowldege' and Topicality. *Journal of Pragmatics*, *9*(1), 83–107.

Gundel, J., Hedberg, N., & Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language, 69*, 274–307.

Hajicova, E., & Sgall, P. (2001). Topic-focus and Salience. In *Meeting of the Association for Computational Linguistics*, pp. 268–273.

Halliday, M. (1967). Notes on Transitivity and Theme in English. *Journal of Linguistics, 3*, 199–244.

Hearst, M. (1994). Multiparagraph Segmentation of Expository Text. In *Proceedings of the 32nd Annual Meeting of the Accosiation for Computational Linguistics*.

Heim, I. (1983). File Change Semantics and the Familiriaty Theory of Definiteness. In Bauerle, R., Schwarze, C., & von Stechow, A. (Eds.), *Meaning, Use and Interpretation of Language*. De Gruyter, Berlin.

Heycock, C., & Kroch, A. (1997). Inversion and Equation in Copular Sentences. In *Workshop on (Pseudo)Clefts, ZAS, Berlin*.

Hitzeman, J., & Poesio, M. (1998). Long Distance Pronominalisation and Global Focus. In *Proceedings of ACL/COLING '98*.

Hobbs, J. (1978). Resolving Pronoun References. *Lingua, 44*, 311–338.

Hobbs, J. (1985). On the Coherence and Structure of Discourse. Tech. rep. CSLI-85-37, Center for the study of language and information, Stanford University.

Holton, D., Mackridge, P., & Philippaki-Warburton, I. (1997). *Greek: A Comprehensive Grammar of the Modern Language*. Routledge: London and New York.

Horn, L. (1986). Presupposition, Theme and Variations. In *Chicago Linguistics Society*, Vol. 22, pp. 168–192.

Hudson-D'Zmura, S., & Tanenhaus, M. (1998). Assigning Antecedents to Ambiguous Pronouns: The Role of the Center of Attention as a Default Assignment. In Walker,

M., Joshi, A., & Prince, E. (Eds.), *Centering Theory in Discourse*, pp. 199–226. Clarendon Press: Oxford.

Hurewitz, F. (1998). A quantitative look at discourse coherence. In Walker, M., Joshi, A., & Prince, E. (Eds.), *Centering Theory in Discourse*, chap. 14, pp. 273–291. Clarendon Press: Oxford.

Joshi, A., & Kuhn, S. (1979). Centered Logic: The Role of Entity Centered Sentence Representation in Natural Language Inferencing. In *6th International Joint Conference on Artificial Intelligence, Tokyo*, pp. 435–439.

Joshi, A., & Weinstein, S. (1981). Control of Inference: Role of Some Aspects of Discourse Structure: Centering. In *7th International Joint Conference on Artificial Intelligence*, pp. 385–387.

Kaiser, E. (2003). Differences in the Referential Properties of Finnish Pronouns and Demonstratives. In *Proceedings from the Main Session of the Chicago Linguistics Society's Thirty-Eighth Meeting*, Vol. 38.

Kameyama, M. (1985). *Zero Anaphora: The Case of Japanese*. Ph.D. thesis, Stanford University.

Kameyama, M. (1993). Intrasentential Centering. In *Proceedings of the Workshop on Centering*. University of Pennsylvania.

Kameyama, M. (1998). Intrasentential Centering: A Case Study. In Walker, M., Joshi, A., & Prince, E. (Eds.), *Centering Theory in Discourse*, pp. 89–112. Clarendon Press: Oxford.

Kohlhase, M., & Koller, A. (2000). Towards a Tableaux Machine for Language Understanding. In *Proceedings of the 2nd Workshop on Inference in Computational Semantics (ICOS-2), Dagstuhl*, pp. 57–88.

Kozima, H. (1993). Text Segmentation Based on Similarity Between Words. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics (Student Session)*, pp. 286–288.

Kraemer, H. C. (1982). Kappa Coefficient. In *Encyclopedia of Statistical Sciences*. New York: John Wiley & Sons.

Kuno, S. (1972). Functional Sentence Perspective: A Case Study from Japanese and English. *Linguistic Inquiry*, *3*, 269–320.

Labov, W. (1972). The Transformation of Experience in Narrative Syntax. In *Language in the Inner City*. Philadelphia: University of Pennsylvania Press.

Landauer, T. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, *25*, 259–284.

Lappin, S., & Leass, H. (1994). An Algorithm for Pronominal Anaphora Resolution. *Computational Linguistics*, *20*(4), 535–561.

Maclure, M., & Willett, W. (1988). Misinterpretation and Misuse of the Kappa Statistic. *American Journal of Epidemiology*, *126*(2), 161–169.

Matthiessen, C., & Thompson, S. (1988). *Clause Combining in Grammar and Discourse*, chap. The Structure of Discourse and "Subordination", pp. 275–329. Amsterdam: Benjamins.

McCawley, J. (1981). The Syntax and Semantics of English Relative Clauses. *Lingua*, *53*, 99–149.

McCord, M. (1990). Slot Grammar: A system for Simpler Construction of Practical Natural Language Grammars. In Studer, R. (Ed.), *Natural Language and Logic: International Scientific Symposium, edited by R. Studer*, pp. 118–145. Lecture Notes in Computer Science, Berlin: Springer Verlag.

McCoy, K., & Strube, M. (1999). Generating Anaphoric Expressions: Pronoun or Definite Description. In Cristea, D., Ide, N., & Marcu, D. (Eds.), *Proceedings of the Workshop on The Relation of Discourse/Dialogue Structure and Reference*, pp. 63–71. 38th Annual Meeting of the Association for Computational Linguistics, University of Maryland, College Park, MD.

McDonald, J., & MacWhinney, B. (1995). The Time Course of Anaphor Resolution: Effects of Implicit Causality and Gender. *Journal of Memory and Language, 34*, 543–566.

Miltsakaki, E. (1999). Locating Topics in Text Processing. In *Computational Linguistics in the Netherlands: Selected Papers from the Tenth CLIN Meeting, (CLIN '99), Utrecht*, pp. 127–138.

Miltsakaki, E. (2000). *Attention Structure in Discourse: A Cross-Linguistic Investigation*. Dissertation topic proposal, University of Pennsylvania.

Miltsakaki, E. (2001). Centering in Greek. In *Proceedings of the 15th International Symposium on Theoretical and Applied Linguistics, Thessaloniki.*

Miltsakaki, E., & Kukich, K. (2000a). Automated Evaluation of Coherence in Student Essays. In *Proceedings of the Workshop on Language Rescources and Tools in Educational Applications, LREC 2000, Athens*, pp. 7–14.

Miltsakaki, E., & Kukich, K. (2000b). The Role of Centering Theory's Rough Shift in the Teaching and Evaluation of Writing Skills. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics (ACL 2000), Hong-Kong*, pp. 408–415.

Moens, M., & Steedman, M. (1988). Temporal Ontology and Temporal Reference. *Computational Linguistics*, *14*(2), 15–28.

Morris, J., & Hirst, G. (1991). Lexical Cohesion Computed by Thesaural Relations as an Indicator of the Structure of the Text. *Computational Linguistics*, *17*, 21–28.

Mosteller, F., & Tukey, J. (1977). *Data Analysis and Regression, a Second Course in Statistics*. Addison-Wesley, Reading MA.

Page, E. (1966). The Imminence of Grading Essays by Computer. *Phi Delta Kappan*, *48*, 238–243.

Page, E. (1968). Analyzing Student Essays by Computer. *International Review of Education*, *14*, 210–225.

Page, E., & Peterson, N. (1995). The Computer Moves into Essay Grading: Updating the Ancient Test. *Phi Delta Kappan*, *March*, 561–565.

Passonneau, R. (1998). Interaction of Discourse Structure with Explicitness of Discourse Anaphoric Noun Phrases. In Walker, M., Joshi, A., & Prince, E. (Eds.), *Centering Theory in Discourse*, pp. 327–358. Clarendon Press: Oxford.

Passonneau, R., & Litman, D. (1993). Intention-Based Segmentation: Human Reliability and Correlation with Linguistic Cues. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics (ACL-93)*, Vol. 19, 3, pp. 148–155.

Passonneau, R., & Litman, D. (1997). Discourse Segmentation by Human and Automated Means. *Computational Linguistics*, *23*(1), 103–139.

Poesio, M., Cheng, H., Henschel, R., Hitzeman, J., Kibble, R., & Stevenson, R. (2000). Specifying the Parameters of Centering Theory: a Corpus-Based Evaluation using Text from Application-Oriented Domains. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics, Hong-Kong*.

Poesio, M., & Vieira, R. (1998). A Corpus-Based Investigation of Definite Description Use. *Computational Linguistics*, *24*(2), 183–216.

Prasad, R., & Strube, M. (2000). Pronoun Resolution in Hindi. In *Working Papers in Linguistics*, Vol. 6. University of Pennsylvania.

Prince, E. (1981a). *Radical Pragmatics*, chap. Toward a Taxonomy of Given-New Information, pp. 223–255. NY: Academic Press.

Prince, E. (1981b). Topicalization, Focus-Movement, and Yiddish-Movement: A Pragmatic Differentiation. In Alford, D. (Ed.), *Proceedings of the Seventh Annual Meeting of the Berkeley Linguistics Society*, pp. 249–264.

Prince, E. (1988). Yiddish Wh-clauses, Subject-Postposing, and Topicalization. In Powers, J., & de Jong, K. (Eds.), *ESCOL '88*. Columbus, OH: Ohio State University.

Prince, E. (1992). *Discourse Description: Diverse Analyses of a Fund Raising Text*, chap. The ZPG Letter: Subjects, Definiteness, and Information Status, pp. 295–325. Philadelphia/Amsterdam: John Benjamins B.V.

Prince, E. (1998). On the Limits of Syntax, with Reference to Left-Dislocation and Topicalization. In Culicover, P., & McNally, L. (Eds.), *The Limits of Syntax*, Vol. 29 of *Syntax and Semantics*. NY: Academic Press.

Prince, E. (1999). *How not to Mark Topics: 'Topicalization' in English and Yiddish*, chap. 8. Texas Linguistic Forum. Austin: University of Texas.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A Grammar of Contemporary English*. London:Longman.

Rambow, O. (1993). Pragmatic Aspects of Scrambling and Topicalization in German. In *Workshop on Centering Theory in Naturally Occurring Discourse*. Institute of Research in Cognitive Science, University of Pennslylvania.

Reinhart, T. (1981). Pragmatics and Linguistics: An Analysis of Sentence Topics. *Philosophica*, *27*(1), 53–93.

Reinhart, T. (1984). Principles of Gestalt Perception in the Temporal Organization of Narrative Texts. *Linguistics*, *22*, 779–809.

Reynar, J. (1994). An Automatic Method of Finding Topic Boundaries. In *Proceedings of the 32nd Annual Meeting of the Accosiation for Computational Linguistics (Student Session)*, pp. 331–333.

Romero, M., & Hardt, D. (2003). ACD and Discourse Relations. In *Proceedings of the 27th Penn Linguistics Colloquium*. University of Pennsylvania.

Schreiner, M., Rehder, B., Landauer, T., & Laham, D. (1997). How Latent Semantic Analysis (LSA) Represents Essay Semantic Content: Technical Issues and Analysis. In Shafto, M., & Langley, P. (Eds.), *Proceedings of the 19th Annual Meeting of the Cognitive Science Society*, p. 1041. Mawhwah, NJ: Erlbaum.

Sidner, C. (1979). *Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse*. Ph.D. thesis, MIT.

Stevenson, R., Crawley, R., & Kleinman, D. (1994). Thematic Roles, Focusing and the Representation of Events. *Language and Cognitive Processes*, *9*, 519–548.

Stevenson, R., Knott, A., Oberlander, J., & McDonald, S. (2000). Interpreting Pronouns and Connectives: Interactions among Focusing, Thematic Roles and Coherence Relations. *Language and Cognitive Processes*, *15*(3), 225–262.

Strube, M. (1998). Never Look Back: An Alternative to Centering. In *Proceedings of the 17th International Conference on Computational Linguistics and the 36th Annual Meeting of the Association for Computational Linguistics (COLING-ACL '98), Montreal*, pp. 1251–1257.

Strube, M., & Hahn, U. (1996). Functional Centering. In *Proceedings of the 34th Annual Conference of the Association for Computational Linguistics (ACL '96), Santa Cruz*, pp. 270–277.

Strube, M., & Hahn, U. (1999). Functional Centering: Grounding Referential Coherence in Information Structure. *Computational Linguistics*, *25*(3), 309–344.

Suri, L., & McCoy, K. (1994). RAFT/RAPR and Centering: A comparison and discussion of problems related to processing complex sentences. *Computational Linguistics*, *20*(2), 301–317.

Suri, L., McCoy, K., & DeCristofaro, J. (1999). A Methodology for Extending Focusing Frameworks. *Computational Linguistics*, *25*(2), 173–194.

Talmy, L. (2000). *Toward a Cognitive Semantics*, Vol. I&II. MIT Press, Cambridge, Massachessetts.

Tanaka, I. (2000). Cataphoric Personal Pronouns in English News Reportage. In *Proceedings of the 3rd International Conference of Discourse Anaphora and Anaphor Resolution Conference, DAARC 2000, Lancaster*, Vol. 12, pp. 108–117.

Thompson, S. (1987). *Coherence and Grounding in Discourse*, Vol. 11 of *Typological Studies in Language*, chap. "Subordination" and Narrative Event Structure, pp. 435–454. John Benjamins Publishing Company.

Turan, U. (1995). *Null vs. Overt Subjects in Turkish Discourse: A Centering Analysis*. Ph.D. thesis, University of Pennsylvania.

Turan, U. (1998). Ranking Forward-Looking Centers in Turkish: Universal and Language Specific Properties. In Walker, M., Joshi, A., & Prince, E. (Eds.), *Centering Theory in Discourse*, pp. 139–160. Clarendon Press: Oxford.

van Hoek, K. (1997). *Anaphora and Conceptual Structure*. Chicago: University of Chicago Press.

Walker, M., Joshi, A., & Prince, E. (Eds.). (1998). *Centering Theory in Discourse*. Clarendon Press: Oxford.

Walker, M. (1996). Limited Attention and discourse Structure. *Computational Linguistics*, *22*(2), 255–64.

Walker, M. (1998). Centering : Anaphora Resolution and Discourse Structure. In Walker, M., Joshi, A., & Prince, E. (Eds.), *Centering Theory in Discourse*, pp. 401–35. Clarendon Press: Oxford.

Walker, M., Iida, M., & Cote, S. (1994a). Japanese Discourse and the Process of Centering. *Computational Linguistics, 20*(2), 193–233.

Walker, M., Iida, M., & Cote, S. (1994b). Japanese Discourse and the Process of Centering. *Computational Linguistics, 20*(2), 193–233.

Walker, M., & Prince, E. (1996). A Bilateral Approach to Givenness: A Hearer-Status Algorithm and a Centering Algorithm. In Fretheim, T., & Gundel, J. (Eds.), *Reference and Referent Accessibility*, pp. 291–306. Amsterdam: John Benjamins.

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

Webber, B., & Joshi, A. (1998). Anchoring a Lexicalized Tree Adjoining Grammar for Discourse. In *ACL/COLING Workshop on Discourse Relations and Discourse Markers, Montreal*, pp. 8–92. Montreal, Canada.

Webber, B., Knott, A., Stone, M., & Joshi, A. (1999a). Discourse Relations: A Structural and Presuppositional Account Using Lexicalized TAG. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics, Maryland,* pp. 41–48. College Park MD.

Webber, B., Knott, A., Stone, M., & Joshi, A. (1999b). What are Little Texts Made of? A Structural and Presuppositional Account Using Lexicalized TAG. In *Proceedings of the International Workshop on Levels of Representation in Discourse (LORID '99), Edinburgh*, pp. 145–149.

Youmans, G. (1991). A New Tool for Discourse Analyis: The Vocabulary-Management Profile. *Language, 67*, 763–789.