

Toward an Aposynthesis of Topic Continuity and Intrasentential Anaphora

Eleni Miltsakaki*

University of Pennsylvania

The problem of proposing referents for anaphoric expressions has been extensively researched in the literature and significant insights have been gained through 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 various languages we propose an aposynthetic¹ model of discourse in which topic continuity, computed across units, and focusing preferences internal to these units are subject to different mechanisms. The observed focusing preferences across the units (i.e., intersententially) are best modeled structurally, along the lines suggested in centering theory. The focusing 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. We show that this distinction not only overcomes important problems in anaphora resolution but also reconciles seemingly contradictory experimental results reported in the literature. We specify a model of anaphora resolution that interleaves the two mechanisms. We test the central hypotheses of the proposed model with an experimental study in English and a corpus-based study in Greek.

1. The Problem

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, and Pollard 1987). The proposed algorithm (henceforth the BFP algorithm) gives the correct interpretation for the pronominal *he* in example (1), stating a preference to resolve the pronominal to *Max* rather than *Fred*.

- (1) a. Max is waiting for Fred.
b. He invited him for dinner.

It was soon observed, however, that the BFP algorithm was not capable of handling cases of intrasentential anaphora such as in (2) (adapted from Suri, McCoy, and De-Cristofaro [1999]).

* Institute of Research in Cognitive Science, Philadelphia, PA 19104. E-mail: elenimi@unagi.cis.upenn.edu

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

2 Although a significant amount of research in anaphora resolution has been carried out in statistical approaches, reviewing such approaches is well beyond the scope of the current article.

- (2) 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.

The centering-based BFP algorithm would have a preference to resolve *he* in (2d) 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 (3), 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*.

- (3) 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 (4), where the preferred interpretation for *he* is *John*, that is, the structural subject, independent of semantic/pragmatic factors.³ In such discourses it seems that a structural account is at play (in the sense of Grosz and Sidner [1986]).

- (4) a. John criticized Bill.
 b. Next, he insulted Susan.

This article sets out to explicate the behavior of pronominals demonstrated in the above examples. Gaining significant insights from current research in anaphora resolution, we reconcile what seem to be contradictory findings 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, topic continuity and the internal structure of the sentence.⁴ We conclude that intersentential anaphora is subject to structural constraints, whereas intrasentential anaphora is subject to grammatical as well as semantic/pragmatic constraints. We define the notion of discourse unit and propose a two-level approach to anaphora resolution. Within the unit, anaphora resolution is performed locally and is constrained by the grammar and semantic properties of the predicates and the subordinate conjunctions. This process outputs unresolved anaphoric expressions for which potential referents are picked from a centering-style ranked list of entities constructed in the previous unit.

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

⁴ We use the term "topic" to describe a centered entity, that is, 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. Elsewhere we have introduced the term "attention structure in discourse" to describe mechanisms, linguistic or nonlinguistic, that language users employ to navigate the hearer's attention in discourse. Topic continuity is derivative of attention structure in discourse. In this article, however, we have opted for the more transparent term "topic continuity," as it describes the phenomenon we are mostly concerned with in pretheoretical terms.

The article is organized as follows. In Section 2 we give a brief overview of centering-based models of anaphora resolution, discuss their shortcomings, and contrast them with the semantic-focusing account suggested in Stevenson et al. (2000). In Section 3, we present the discourse model we adopt and the specifications we propose for anaphora resolution across and within centering update units. In Section 4, we test the central hypotheses of the proposed model in two studies: an experimental study in English and a corpus-based study in Greek. We conclude with a general discussion in Section 5.

2. Issues and Insights in Anaphora Resolution

2.1 The BFP Algorithm

Brennan, Walker-Friedman, and Pollard (1987) were the first to use the centering model as the basis for an anaphora resolution algorithm. The centering model (Grosz and Sidner 1986; Grosz, Joshi, and Weinstein 1983) makes the following assumptions:

1. A discourse segment consists of a sequence of utterances, U_1, \dots, U_n .
2. For each utterance, a ranked list of evoked discourse entities is constructed, designated as the Cf list.
3. The highest-ranked element of the Cf list is called the preferred center (Cp).
4. 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 1). The distinction between a Smooth Shift and a Rough Shift is due to Brennan, Walker-Friedman, and Pollard (1987), who observed that the centering model generates ambiguity in cases such as (5):

- (5) a. Brennan drives an Alfa Romeo.
- b. She drives too fast.
- c. Friedman races her on weekends.
- d. She often beats her.

Adding weight to the status of the Cp in (5c) makes it possible to resolve the pronominal *she* in (5d) successfully to *Friedman*. We will return to the issue of ambiguity shortly.

The BFP algorithm consists of three basic steps:

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

Table 1
Centering transitions.

	$Cb(U_i) = Cb(U_{i-1})$	$Cb(U_i) \neq Cb(U_{i-1})$
$Cb(U_i) = Cp(U_i)$	Continue	Smooth Shift
$Cb(U_i) \neq Cp(U_i)$	Retain	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. Prasad and Strube's observations are made with respect to Hindi but hold in English and Modern Greek, as shown in (6) and (7), respectively.

The first of these errors occurs in cases in which $Cb(U_{i-1})$ is different from $Cp(U_{i-1})$. In such cases, the preference for a Continue transition is responsible for the pronominal in U_i being resolved to the $Cb(U_{i-1})$ and not to the $Cp(U_{i-1})$.

- (6) 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.
- (7) a. I Elen_i ide ti Maria_j sto sholio.
 the Elen_i saw the Maria_j at-the school.
 Elen_i saw Maria_j at school.
 b. I Maria_j den tis_i milise.
 the Maria_j 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_j took the friends her and NULL_j left.
 She_j took her_j friends and left.

There is an important observation to be made here, which we present 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 (6c) and (7c). On the other hand, anaphora aside, the topic transition identified by centering is intuitively correct. In (6) and (7), the discourse is initiated with *Ellen* as the current topic, *Maria* is introduced as an entity related to the current topic, and then the discourse shifts to *Maria* to elaborate on her doings. The shift is in fact anticipated by the promotion of *Maria* from the object position in (6a) and (7a) to the subject position in (6b) and (7b).

The second error observed by Prasad and Strube (2000) is that the BFP algorithm generates ambiguity when U_{i-1} is discourse initial. Example (8) is given as illustration.

- (8) 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 (8a) 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.

We conclude from this section that although centering transitions successfully identify topic continuity in discourse, in anaphora resolution the most useful centering notion is not the transitions themselves but the Cf list ranking in combination with the pronoun rule.

2.2 Functional Centering

Strube and Hahn (1996, 1999) elaborate on the nature of the Cf list and propose a centering-based model of anaphora resolution in which the Cf ranking is based not on grammatical function but on functional information status. They recast centering notions in terms of Daneš'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 that 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} . Although 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 using the same principles. They argue against Walker, Iida, and Cote (1994), in which the Cf ranking is viewed as a language-specific parameter that needs to be set.

In what follows we will remain agnostic to the suitability of the functional centering framework for German. We will show, however, that functional centering is not the appropriate framework for all free-word-order languages, much less for languages universally. We bring in evidence from Modern Greek, a free-word-order language.

To identify the factors determining Cf ranking in Greek, we employ Rambow's (1993) diagnostic⁵ to test whether surface word order or grammatical function is the most reliable indicator of salience. The relevant examples for the Greek version of

⁵ Rambow suggests that the order of entities in the position between finite and nonfinite verbs in German (Mittelfeld) affects their salience. Gender in German is grammaticized, so Rambow 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. "Fem" indicates that the noun phrases are gender marked "feminine."

Rambow's diagnostic are shown in (9) and (10). The null pronominal in (9b) and (10b) resolves to the subject irrespective of its surface position.⁶ The relevant indicator of salience in the Cf list appears to be grammatical function, at least subjecthood.⁷

- (9) a. I profsati diefthetisi_i veltioni tin ikonomiki politikij?
 the recent arrangement improve the economic policy?
 Does the recent arrangement_i improve the economic policy_j?
- b. Ohi, (null)_i ine aneparkis.
 No, (it) is inadequate.
 No, it_i is inadequate.
- (10) a. Tin ikonomiki politikij t_j veltioni i profsati diefthetisi?
 the economic policy it-(clitic) improve the recent arrangement?
 Does the recent arrangement_i improve the economic policy_j?
- b. Ohi, (null)_i ine aneparkis.
 No, (it) is inadequate.
 No, it_i is inadequate.

Further evidence for the role of grammatical function in Greek comes from syntactic objects.⁸ In Greek (and also in Turkish), a strong pronominal or a full noun phrase (NP) must be used to promote the object of U_{i-1} to the subject position of U_i .⁹ As the infelicitous interpretations (indicated by the pound sign) show in (11b), reference to the object *Yorgo* becomes felicitous only with the use of name repetition or a strong

-
- (1) a. Glauben Sie, dass [eine solche Massnahme]_i [der russischen Wirtschaft]_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.

6 Gender and lexical considerations are controlled. Both *economical policy* and *arrangement* are feminine and they can both be *inadequate*. Also, we have presented the diagnostic test and confirmed the judgment with a sizable group of native speakers of Greek attending the 15th International Symposium on Theoretical and Applied Linguistics (Miltakaki 2001).

7 It is interesting that in Turkish, another free-word-order language, it has also been shown (Turan 1998) that the strongest indicator of salience is subjecthood.

8 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 Miltakaki (2000), we argue 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 three-way distinction in their nominal system (i.e., full noun phrases and weak and strong pronominals) use strong pronominals to signal reference to an entity previously evoked in discourse that is not, however, the most salient entity. This use of strong pronominals is equivalent to certain prosodic effects in English. As noted by an anonymous reviewer, there is extensive literature on the effects of prosody in pronominal interpretation. For example, it has been observed that prominent stress on the pronominals in (1) yields cospecification 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 should be trivial.

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

pronominal, shown in (11c) and (11d).¹⁰ We take this as further evidence that objects rank lower than subjects in Greek.

- (11) O Yannis_i proskalese ton Yorgo_j.
 the John invited the Yorgo.
 John_i invited George_j.
- a. null_i tu_j profere ena poto.
 he him offered a drink.
 He_i offered him_j a drink.
- b. #null_j #tu_i profere ena poto.
 he him offered a drink.
 He_j offered him_i a drink.
- c. O Yorgos_i tu_i profere ena poto.
 the George him offered a drink.
 George offered him_i a drink.
- d. Ekinos_j tu_i profere 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, we replace the definite subject in (9) with an indefinite noun phrase. As shown in (12), the subject is an indefinite noun phrase representing new (or hearer-new) information and the object is a definite phrase, encoding old (or hearer-old) information. The null pronominal in (12b) resolves to the subject of (12a), and the information status of the potential antecedents is disregarded.

- (12) a. Mia kainourgia diefthetisi_i tha veltiosi tin ikonomiki politiki_j?
 a new arrangement will improve the economic policy?
 Will a new arrangement_i improve the economic policy_j?
- b. Ohi, (null_i) tha ine aneparkis.
 No, (it) will be 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 and Strube (2000) for Hindi. In both of these languages, the relevant factor for the ranking of elements in the Cf list is grammatical function.

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

¹⁰ Empirical evidence for the use of strong pronominals to signal reference to nonsalient entities in Greek is provided in Dimitriadis (1996). Further functions of strong pronominals in Greek are identified in Miltsakaki (1999, 2001).

2.3 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 with the ordering among elements of what he calls the S-list, that is, the list of salient discourse entities. The S-list ranking criteria define a preference for hearer-old over hearer-new discourse entities and are 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 lookup in the S-list. The elements of the S-list are looked up for compatibility in the order determined by their information status ranking (hearer-old entities are looked up before hearer-new entities). When the analysis of the utterance is finished (processed left to right), the discourse entities that 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.

Although 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 in which 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 (13), 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).¹¹

- (13) 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 (13a) and (13b) the pronominal resolves to *Geoff*, the verb argument with the **stimulus** role. The ordering in the S-list in (13a), however, 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 that attempts to handle intrasentential anaphora according to a fixed ranking of entities in a salience list.¹²

Apparently, certain discourse algorithms relying on a fixed ordering of potential antecedents are not capable of resolving anaphora successfully. In sections 4 and 5, we argue that such cases are most commonly identified intrasententially.

2.4 Revised Algorithms for Focus Tracking/Revised Algorithms for Pronoun Resolution

Based on previous work (Suri and McCoy 1994), Suri, McCoy, and DeCristofaro (1999) propose a methodology of extending their Revised Algorithms for Focus Tracking/Revised Algorithms for Pronoun Resolution (RAFT/RAPR) 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

¹¹ This strategy was also adopted by Prasad and Strube (2000) in the implementation of the S-list algorithm for Hindi.

¹² It is conceivable that a discourse can be constructed in which the semantics will force a similar pattern of resolution intersententially. Hudson-D'Zmura and Tanenhaus (1998), however, report experimental results that show that in such cases, sentence processing is slowed down.

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.

- (14) (S1) Dodge was robbed by an ex-convict the other night.
 (15) (S2) The ex-convict tied him up because he wasn't cooperating.
 (16) (S3) Then he took all the money and ran.

Suri, McCoy, and DeCristofaro's findings indicate that the pronoun resolution facts within S2 are consistent with the expectations of both centering and RAFT/RAPR. On completing the processing of the SY clause, however, 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.

Although the prefer-SX hypothesis repairs the algorithm with respect to the construction in question, it seems to be missing a generalization regarding inconsistencies observed within versus across sentences. We return to this issue in Section 4.

2.5 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, 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, and connectives like *so* direct attention to the consequences of the event. Thus in a sentence like (17), 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.

- (17) 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 on the entity associated with the endpoint of the event, namely, the patient, independent 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 (18a) 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 there as an object, as in (18b).

- (18) 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 is observed with state verbs, shown in (19), where *he* in the continuation is interpreted as the experiencer of the event independent of its structural position.

- (19) 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.

Hudson-D'Zmura and Tanenhaus (1998), however, report experimental results that, at first blush, contradict this view. They conducted a similar experiment to test whether subject-object or stimulus-experiencer is the crucial distinction for pronominal interpretation. Subjects were given sentence (20) followed by the continuations (20a)–(20b) and were asked to judge the continuations for naturalness.

- (20) Max despises Ross.
 a. He always gives Ross a hard time.
 b. He always gives Max a hard time.

Hudson-D'Zmura and Tanenhaus's results show that there is a strong preference for the subject interpretation independent of the thematic role.

What are we to conclude from these inconsistent results? The results show that the same type of verb (i.e., a state verb) in some cases projects its own focus preference (e.g., experiencer), but in other cases it does not. In order to account for the facts, 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 in which we divide the labor of anaphoric interpretation between the two mechanisms and define the domains of their applicability.

3. The Proposal: Aposynthesis

3.1 Outline of the Discourse Model

We assume that the discourse is organized hierarchically in linear and embedded segments as specified in Grosz and 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, is left unspecified. Because transitions are computed for each utterance, we will rename utterances as **centering update units** and argue that a centering update unit consists of a matrix clause and all the dependent clauses associated with it. For each update unit a list of forward-looking centers is constructed and ranked according to the salience of each. Consistent with the proposed definition of unit, we argue that entities evoked in subordinate clauses are of lower salience than entities evoked in the matrix clause and are ranked accordingly. The proposed centering specifications have the following corollaries:

1. The linear order of subordinate clauses relative to the matrix clause does not affect the salience status of the entities.

2. Entities evoked in subordinate clauses are available as potential links between the current and previous or subsequent discourse.
3. Topic shifts must be established in matrix clauses.
4. 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 2.1, 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, we present evidence in support of a new definition of the update unit. Data from English, Greek, and Japanese show that treating subordinate clauses as independent units yields counterintuitive centering transitions and violations of the pronoun rule.

3.1.1 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 update unit be constrained to enable the handling of intrasentential anaphora. To a large extent, efforts to identify the appropriate size for the unit were often dictated by needs specific to anaphora resolution algorithms.

Centering was not originally formulated, however, 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 that 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 complex sentences are processed. 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 in support from backward anaphora. She argues that the tensed adjunct hypothesis predicts that the pronoun in the fronted subordinate clause in (21c), for example, is anaphorically dependent on an entity already introduced in the immediate discourse and not on the subject of the

main clause to which it is attached:

- (21) 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_j 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 of 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 (also Kuno's) argument 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 that are the first mention of their referent in discourse. Most recently, Tanaka (2000) reported that in the cataphora data retrieved from the Anaphoric Treebank, out of 133 total occurrences of personal pronouns encoded as "cataphoric," 47 (35.3%) were "first mentioned." Among the 47 cases of first-mention cataphora, 6 instances were discourse initial.¹³

Secondly, this account leaves the use of a full NP in Kameyama's main clause (21d) unexplained (*Kern* and *Jim* have the same referent). Full NPs and proper names occurring in Continue transitions have been observed to signify a segment boundary. 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 (21d) 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), who carried out centering studies in Italian. Di Eugenio (1990) proposed that the alternation of null and overt pronominal subjects in Italian could be explained in terms of centering transitions. Typically, a null subject signals a Continue and a strong pronoun a Retain or a Shift.¹⁴

Following Kameyama (1993), Di Eugenio treats subordinate clauses as independent update units. Her motivation for doing so comes from the following example, in which 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 (22) the $Cb(U_{i-1}) = vicina_j$ (neighbor-fem) and $Cf(U_{i-1}) = vicina_j$.

13 The Anaphoric Treebank is a corpus of a collection of news reports, annotated with, among other things, type of anaphoric relations. It was developed by UCREL (Unit for Computer Research on the English Language) at Lancaster University, collaborating with the IBM T.J. Watson Research Center.

14 Di Eugenio collapsed the distinction between Smooth and Rough Shifts. The reader is referred, however, to Miltsakaki and Kukich (2000a, 2000b) for a discussion of the significance of Rough Shifts in the evaluation of text coherence.

- (22) 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 Miltsakaki (2001), we report 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 revealed that weak pronouns were most common in Continue transitions, whereas strong pronouns, full noun phrases, and proper nouns were associated with Rough Shift transitions. On closer inspection, we observed that in six out of the eight instances of strong pronouns in Continue transitions, the referent of the strong pronoun was contrasted on some property with some other entity belonging to a previously evoked set of entities.¹⁵ 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 (22b), *she*, the most salient entity in the current discourse, is contrasted with the *lazy ones*, in (22a), 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, in which the "vicina" appears as the most salient entity, realized with multiple dropped subjects. The discourse immediately preceding (22) is shown in (23).¹⁶

- (23) 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_j 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.

We now turn to English and Greek to show that treating subordinate clauses as independent centering units yields counterintuitive topic transitions. First, consider the constructed example from English shown in (24).

¹⁵ We ignored one further instance of a strong pronominal in a Continue transition, as in that case the strong pronominal headed a relative clause and its use was forced by the grammar.

¹⁶ Many thanks to Barbara Di Eugenio (personal communication) for providing me with the extra data in (23). I presume that Di Eugenio's coding of the null realization in (23b) is based on the inferable information that the noun phrase 'la combinazione di tutto cio'' refers to *her_j* qualities.

- | | |
|---|---|
| <p>(24) Sequence:
main-subordinate-main</p> <p>a. John had a terrible headache.
Cb = ?
Cf = John > headache
Transition = none</p> <p>b. When the meeting was over,
Cb = none
Cf = meeting
Transition = Rough Shift</p> <p>c. he rushed to the pharmacy
store.
Cb = none
Cf = John
Transition = Rough Shift</p> | <p>(25) Sequence:
main-main-subordinate</p> <p>a. John had a terrible headache.
Cb = ?
Cf = John > headache
Transition = none</p> <p>b. He rushed to the pharmacy
store
Cb = John
Cf = John > pharmacy store
Transition = Continue</p> <p>c. when the meeting was over.
Cb = none
Cf = meeting
Transition = Rough Shift</p> |
|---|---|

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 (Walker, Iida, and Cote 1994; Di Eugenio 1998; and Miltsakaki 1999, among others), (b) be rare in corpora, to the extent that the transition has been ignored by some researchers (Di Eugenio 1998 and Hurewitz 1998, among others), and (c) be reliable measures of low coherence in student essays (Miltsakaki and Kukich 2000a). In addition, simply reversing the order of the clauses, shown in (25), 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 C_p normally does when it instantiates an entity different from the current C_b . Notice that if we process the subordinate clause in the same unit as the relevant main clause, we compute a Continue transition independent 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 (26).

- (26) 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
Transition = Continue

Similar examples were identified in data collected from a short story in Greek (Miltsakaki 2001, 1999). Example (27), shown below, is indicative.

- (27) 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 (27b) as an independent unit yields three Rough Shift transitions in the subsequent discourse, reflecting a highly discontinuous discourse. When (27b) and (27c) 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.¹⁷ 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 (28). 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.

- (28) 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,
- c. ZERO urayamasiino desu
ZERO jealous is
(he-Taroo) is jealous.

¹⁷ Thanks to Kimiko Nakanishi for providing me with the data. In a centering study she conducted in Japanese, she also concluded that treating subordinate clauses as independent units would yield a highly incoherent Japanese discourse.

In section 2.4, a similar case was also identified in English. It is repeated here as (29d). Again, the referent of *he* in (29d) is cospecified 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* (29d) 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, McCoy, and DeCristofaro (1999), demonstrated here in (30). In (30), *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.

- (29) 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.
- (30) 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, McCoy, and DeCristofaro (1999). In their experiment, participants judge that a natural way to refer to *Dodge* in (31c) is by name repetition.

- (31) 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 (34),¹⁸ backward anaphora is most commonly found in preposed subordinate clauses (32) and not in sequences of main clauses (33). From the unit definition we propose, 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 backward for an antecedent, as any other normal pronoun would. To illustrate the point, the Cf list for (32) contains *John* > *shower* > *he-referent*. The pronoun looks back for an antecedent, intrasententially, and resolves to the only compatible antecedent available, *John*.

- (32) As soon as he arrived, John jumped into the shower.
 (33) #He arrived and John jumped into the shower.
 (34) He-i couldn't have imagined it at the time but John Smith-i turned out to be elected President in less than three years.

¹⁸ Thanks to Ellen Prince for pointing out this example.

3.1.2 Discourse Saliency versus Information Structure. In the previous section, we argued that the linear position of the subordinate clause does not affect topic continuity. This position leads 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 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 language and are necessary to yield grammatical sentences. It is also common, however, especially in free-word-order languages, that movement is syntactically optional and the surface word order is used to satisfy information-packaging needs (for example, to arrange the information into old-new or ground-focus or mark open propositions). 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 (35) the internal argument of the verb (the object) has been fronted, but its original relation to the verb has remained the same.

(35) 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 Lexicalized Tree Adjoining Grammar (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 in the surface order of the substituted argument clauses, but the predicate-argument relation remains the same (Webber and Joshi 1998; Webber et al., 1999a, 1999b).

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 saliency. If we dissociate saliency from information structure, the relevant unit for computing saliency 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 saliency. Subordinate links are not relevant to the saliency mechanism. Saliency is computed paratactically.

A natural consequence of this model is that one can introduce referents on the vertical level without affecting the status of the salient entity on the horizontal level. It follows that changes of topic must be established at the horizontal level. Such a conception of the saliency structure suggests that text processing is not strictly incremental, as commonly assumed. Although 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 saliency and information packaging is hard to establish because of the inevitable overlap between information status and saliency: 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) that do not

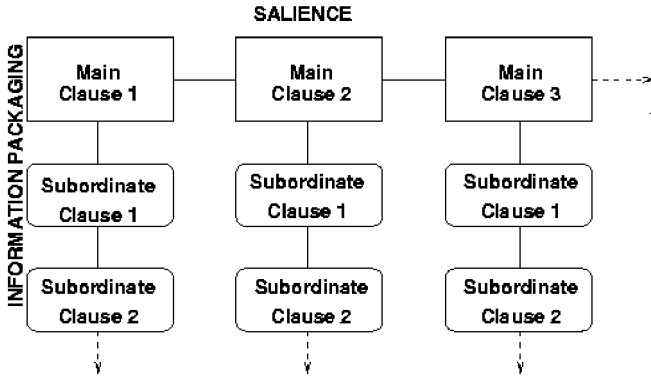


Figure 1
Saliency model.

necessarily relate to the saliency of the participating entities. Although a great deal of additional work is required to understand the precise nature of the interaction between saliency and information structure, we believe that we obtain a significant gain in keeping the two processes distinct.

3.2 Outline of the Anaphora Resolution Model

In Sections 2 and 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 reconciled. We are now able to offer an explanation. The basic idea is that topic continuity and intrasentential anaphora are handled by two distinct mechanisms. Topic continuity is computed across centering update units. Anaphoric reference spanning across update units relates to topic continuity and is therefore determined structurally in accordance with centering rules and constraints. Within the unit, anaphora is constrained by focusing 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 we have seen so far. 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 (1998) 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, McCoy, and DeCristofaro’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.

The remainder of this section is organized as follows. First, we provide definitions for the basic tenets of the model we propose and describe the basic steps required for combining the two mechanisms in a single anaphora resolution algorithm. Next, we

discuss some remaining issues raised by the English connective *so* and certain types of preposed subordinate clauses.

3.2.1 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. Dependent clauses are of three types: sentential complements of verbs, relative clauses, and subordinate clauses. Sentential complements of verbs and relative clauses are identified syntactically. Subordinate clauses are introduced with subordinate conjunctions. To identify subordinate conjunctions, we apply the **reversibility test**: a tensed clause is introduced by a subordinator when the clause it introduces can be preposed.¹⁹ For example, in (36), *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 (38) is not classified as a subordinator, because preposing the clause with which it is associated yields ungrammaticality.

- (36) John traveled by air although he is afraid of flying.
 (37) Although he is afraid of flying, John traveled by air.
 (38) John traveled by air. However, he is afraid of flying.
 (39) #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 complex NPs, we assume left-to-right ranking of entities, as suggested in Walker and Prince (1996).²⁰

Given the above input for N units, $U_{i=1 \dots N}$, the anaphora resolution algorithm starts at the last identified unit. The basic steps are specified below. Some of the steps require 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 the focusing preferences of verbs and connectives as well as identifying thematic roles will, of course, await further research.

¹⁹ “Reversibility” is identified as a characteristic of subordinate clauses in Quirk et al. (1972).

²⁰ Complex NPs are noun phrases containing multiple nouns, for example, *John’s father*.

0. Start at the last identified unit U_i with $i = N$.
1. Identify pronominal expressions in the rightmost subordinate clause.
2. Input antecedents from the Cf list.
3. Apply grammar-driven constraints (number and gender agreement, contra-indexing, etc.) to reduce list of potential antecedents.
4. Resolve from right to left to the first available antecedent inside the subordinate clause. Output unresolved pronominals.
5. Using the Cf list resolve pronominals according to semantic focusing constraints. Output unresolved pronominals.
6. If there is another subordinate clause to process go to step 1.
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.
8. Input Cf list of potential antecedents from previous unit.
9. Apply grammar-driven constraints to reduce list of potential antecedents.
10. Resolve pronominals starting from the leftmost to the highest-ranked element of the list of available antecedents.
11. If an antecedent is found go to step 13.
12. If the list of potential antecedents is empty and there is a unit to process go to step 8, else mark unknown.
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 (2), repeated here in (40)–(42).

(40) Dodge was robbed by an ex-convict.

(41) The ex-convict tied him-3 up because he-2 wasn't cooperating.

(42) 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 > DODGE.
- 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 because of semantic focusing. HE-2 resolves 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.

3.3 Comparison with Related Algorithms

The crucial difference between our approach and related anaphora resolution algorithms is in the treatment of subordinate clauses. Whereas steps 7–10 are similar to other approaches that 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 set forth in Section 3.2.1, (a) subordinate clauses are part of the same unit containing the main clause with which they are associated, 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.²¹ Also, intrasentential anaphora is preferred in the cases of anaphoric elements occurring in subordinate clauses but not in main clauses (assuming grammatical filtering), again irrespective of their linear order.

We will now demonstrate these differences with respect to Lappin and Leass’s (1994) and Hobbs’s (1978) algorithms, which are conceptually the closest to our approach. 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 and rewards parallel syntactic positions and intrasentential antecedents (sentence recency).

As we have already mentioned, backward anaphora need not receive any special treatment in our approach. Lappin and Leass penalize cases of backward anaphora severely, which seems to work well on empirical grounds, presumably because backward anaphora is rather rare. In absence of an explicit method of identifying real cases of backward anaphora, however, the system is likely to miss such cases. In our

²¹ Assuming that backward anaphora is restricted to subordinate clauses. Special treatment is required for the *but* clauses discussed in section 3.1.1, example (34).

approach, 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 Lappin and Leass involve cases of intersentential anaphora such as the following:

- (43) 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.

The RAP algorithm resolves the pronoun *it* in (43b) to the *controller* in (43a). This is because, in RAP, the subject of the main as well as the when clause in (43a) are of equal salience. In this case, *the controller* wins because it is more recent. In our approach, *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.

Hobbs'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 clause is reached.

Intersententially, Hobbs's syntactic algorithm favors subjects over objects, as subjects are higher up in the parse tree than objects. Intersententially, our approach and Hobbs's algorithm would opt for the same type of antecedent. As Lappin and Leass (1994) have pointed out, however, 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 in which salience is determined by other factors (e.g., information status, as has been argued for German [Strube 1998]), Hobbs's 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 undermined by nonsyntactic 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 and impersonal pronouns). The flexibility of constructing lists of entities according to salience both optimizes the capabilities of an anaphora resolution algorithm and is best suited to accommodate the multiplicity of factors that may have to be taken into account in determining reference salience.

Hobbs's algorithm is, in effect, similar to our approach in the treatment of subordinate clauses. Subordinate clauses belong to the same parse tree as the main clause to which they are subordinate. This is equivalent to our claim that subordinate clauses are not independent processing units. With respect to backward anaphora, in particular, Hobbs's 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 (44), the pronoun would correctly resolve to *Susan*. In a case like (45), however, Hobbs's 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.

- (44) After she phoned Barbara, Susan went out for dinner.
 (45) Susan criticized Barbara because she was lazy.

3.4 Some Remaining Issues

As mentioned above, the proposed model for anaphora resolution accounts comfortably for the results reported in Stevenson et al. (2000) except, however, for the experiment involving the connective *so*.

According to the reversibility test, *so* is classified as a subordinate conjunction depending on its interpretation. In English, *so* denotes two relations: consequence and purpose. The examples below indicate that only purpose-*so* behaves as a subordinate conjunction.²²

- (46) I had to give up my job so I could be happy again.
 (47) So I could be happy again, I had to give up my job.
 (48) I had just been to the bank, so I had money.
 (49) #So I had money, I had just been to the bank.

The anaphora resolution model we propose predicts that the interpretation of pronouns in consequence-*so* sentences is determined structurally. This prediction was not borne out. Stevenson et al. (2000) report a main effect of semantic focusing in consequence-*so* continuations.

There are two options available to explain the data. First, we may hypothesize that subordination is determined on structural grounds, in which case it is likely that languages may 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 a hypothesis on empirical grounds. Alternatively, we may hypothesize that the crucial factor in characterizing subordination is by its semantic properties, that is, 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. It runs into the following problem, however.

In Modern Greek, the equivalent conjunction for the English *so* is *etsi* or *ki etsi* (= 'and so'), which is not polysemous and is not a subordinate conjunction. Greek *etsi* links clauses paratactically (i.e., links sequences of main clauses). The examples below show that Greek behaves differently from English in the *so* cases.

- (50) #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.
 (51) 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.

²² Although in many cases preposing a purpose-*so* clause seems unnatural, at least for some native speakers of English preposing of a purpose-*so* clause is possible, given the appropriate context. For all native speakers we have consulted, there is a marked difference in the acceptability of (47) and (49).

The Modern Greek data show that the null subject in the *so* clause cannot be interpreted as the object of the previous clause. If subordination was to be defined on semantic grounds, then we should not expect focusing differences between the two languages, but in fact we do.

Finally, we notice that Greek is much like English when the second clause is linked through other types of subordination, as shown in (52)–(53):

- (52) 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
- (53) 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 was being naughty.

The reason for 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. In this article, we do not claim to understand the intricacies of subordination any better. In the next section, it is shown that the distinction between main and subordinate clauses is in the right direction. It is not yet clear, however, what property of subordination—structural, semantic or other—is responsible for the pattern we observe. We will leave this issue open for future work.

Another issue that requires special attention in the proposed account pertains to some special cases of preposed subordinate clauses. Example (54) 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.

- (54) After Susan phoned Barbara, she went out for dinner.

The ranking in the Cf list for (54) is she-referent > Susan > Barbara. In effect, what we are faced with here is analogous to backward anaphora. In its current form, however, 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 be fixed locally by forcing resolution to the highest entity in the current unit (i.e., *Susan*). Also, as a reviewer has suggested, the algorithm presented in Section 3.2.1 could be modified so that in step 2 the Cf list includes all possible antecedents from the current utterance U_i . With this modification, (54) would be processed correctly, but as the same reviewer points out, this does not explain the contrast in (55):

- (55) 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 (55a) is an instance of intersentential anaphora, and there is a subject reference for the pronoun as predicted. Example (55b) is a case of intrasentential

anaphora, and there is no clear subject reference. Example (55c) is another instance of intrasentential anaphora, but in this case the subject preference is clearly on a par with the intersentential case in (55a). Whatever required modification to the algorithm will prove to be more useful, the fact remains that the similarity between (55a) and (55c) remains unexplained in purely structural terms. We suspect that the difference between (55b) and (55c) and the similarity between (55a) and (55c) 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 “nonnatural order” with respect to the event(s) of the 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 (55a) and (55c) is due to the fact that both sequences of clauses reflect the linear succession of events. The proposed 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.

4. Empirical Studies

In this section we report the results of two empirical studies designed specifically to test the central hypotheses of the proposed model: (a) that anaphoric reference that spans across centering update units is determined structurally (as specified in Section 3.2.1) and (b) that subordinate clauses do not form independent processing units. In the first study, in Section 4.1, we report the results of a sentence completion experiment in English. We quantify over the interpretation of a subject pronoun to one of two ambiguous antecedents evoked in the preceding clause. Two factors were analyzed, type of clause and semantic type of connection, in four conditions. The results show a strong main effect for type of clause. In the main-clause conditions, the subject pronoun resolved to the subject antecedent of the previous sentence. In the subordinate-clause conditions, the interpretation of the subject pronoun was varied. In the second study, we tested the same hypotheses on a Greek corpus. We selected three types of subordinate clauses under conditions similar to the ones set in the experimental study. We contrasted anaphoric interpretation in subordinate clauses with anaphoric interpretation in main clauses. The results provide strong evidence for the accuracy of the centering-based algorithm proposed for anaphora resolution across units. The results also confirm that anaphoric interpretation in subordinate clauses is not determined structurally.

4.1 Experimental Data

The aim of the experiment discussed in this section is to investigate the hypothesis that subject pronouns in main clauses follow a different pattern of interpretation from subject pronouns in subordinate clauses. The participants in this experiment read sentences containing sequences of two clauses. In two of the four conditions of the experiment, the sequence consisted of one main and one subordinate clause. In the other two, the sequence consisted of two main clauses. In all the conditions, the first

clause introduced two individuals of the same gender. The second clause contained a subject pronoun of the same gender as the individuals introduced in the first main clause. Participants were asked to complete the second clause. We were interested in the interpretation of the subject pronominal in the second clause.

For this experiment, two subordinate conjunctions were selected from two semantic classes, namely, *time* and *contrast*. Also, two adverbial conjunctions, *then* and *however*, were selected from the same semantic classes to introduce the second clause in the main-main conditions.

We predicted that the pronominal in the main clause would consistently be interpreted as the subject of the preceding main clause. We also predicted that the interpretation of the subject pronoun in the subordinate clause would vary. In all critical items the verb of the main clause belonged to the same category (see below), so any variation pattern observed across the two subordinate conjunctions would be attributed to the focusing preferences projected by the semantics of the subordinate conjunctions.

The grouping of connectives into semantic types, namely, *then*, *when* for time and *however*, *although* for contrast and structural types, *then*, *however* for main clauses and *when*, *although* for subordinate clauses, enables us to compare and contrast the effect of two factors: semantic type and type of clause. A main effect of semantic type would mean that the interpretation of the subject pronoun is primarily determined by the semantics of the connectives. Conversely, a main effect of type of clause would mean that the interpretation of the subject pronominal is primarily determined by the type of clause.

4.1.1 Study and Results. Sixteen adults, native speakers of English, participated in the experiment. They were asked to perform a sentence completion task as described in the previous section. Each participant received a form containing 36 fillers and 12 critical items, 6 per type of clause and 6 per semantic type. The example below demonstrates a complete set of conditions; each form contained three such sets interspersed with fillers. Each main clause appeared in all conditions across participants, and each condition appeared three times within participants. The order of appearance of critical items was varied in two LISTS (two variations in the order of appearance of critical items).

- | | |
|---|------------------------------|
| a. The groom hit the best man violently. However, he... | (contrast, main-main) |
| b. The groom hit the best man violently
although he... | (contrast, main-subordinate) |
| c. The groom hit the best man violently. Then, he... | (time, main-main) |
| d. The groom hit the best man violently when he... | (time, main-subordinate) |

All main clauses contained an action verb with two human arguments of the same gender. An adverbial phrase was added at the end of the main clause to achieve naturalness. The arguments of the verbs were always realized as role NPs to minimize referential ambiguity in the continuations. In general, sentence continuation involving role NPs tends to disambiguate the intended reference. Nevertheless, an average of two ambiguous continuations were identified per participant. On completing the experiment, the participants were asked to identify explicitly the intended referent of the ambiguous continuations.

Figure 2 summarizes the distribution of the data within participants. The *y*-axis shows the number of references to the main-clause subject. The lighter grey columns show the distribution of the interpretation of the main-clause subject pronoun, and the

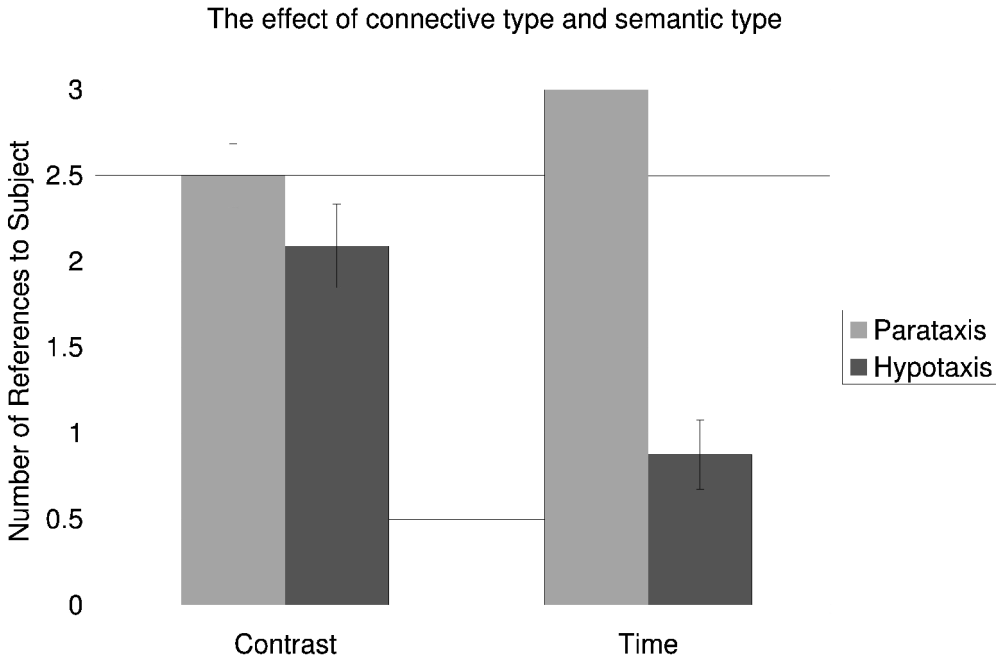


Figure 2
Experiment on anaphora.

darker grey columns show the distribution of the interpretation of the subordinate-clause subject. For convenience, we labeled the main-main clause condition *parataxis* and the main-subordinate clause condition *hypotaxis*.²³

Analyses of variance revealed a strong main effect of type of CLAUSE: the pronoun resolved to the subject of the previous clause significantly more often when it appeared in a main clause ($F = 74.16, df = 1, 15, p < .0001$). Semantic type showed only a marginal effect ($F = 4.59, df = 1, 15, p < .049$).

With regard to the comparison based on the semantic similarity between *however/although* and *then/when*, it turned out that in several cases when clauses were assigned a causal interpretation. For example, the *when* clause continuation in (56) is interpreted as giving the cause of the event described in the preceding main clause, rather than the time specification.

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

For this reason, it was hard to pursue the semantic-based comparison between the paratactic *then* and the hypotactic *when*, and we therefore did not make any further analyses within the semantic-type factor. Since it would be hard to control the interpretation of the connectives used in this type of experiment, in future experiments we plan to give up on the effort to pair connectives semantically. Instead, we plan to

²³ *Parataxis* and *hypotaxis* are terms borrowed from traditional grammars to describe the two types of connections. Main clauses are linked to each other through parataxis, whereas a subordinate clause is linked to its superordinate clause through hypotaxis.

include a larger number of subordinate conjunctions in order to be able to generalize more reliably over the whole class of subordinate conjunctions.

4.1.2 Discussion. The results of this experiment show that structural focusing is prominent across main clauses. The analyses of variance revealed a strong main effect for the type of clause despite the fact that, in the experiment, we included main-clause continuations introduced with *however*. Given the verb type in the first main clause, we in fact stretched structural focusing to its limits, as it would be reasonable to expect that the contrast relation established with *however* would shift attention to the second individual. For example:

(57) The Pope_i tapped the priest_j on the shoulder. However, he_j ignored him.

Still, the results clearly indicate that the type of connection affects the interpretation of the pronoun in the second main clause across, or despite, semantic types. Also, the variation of pronominal reference across types of subordinate conjunctions indicates that the interpretation of anaphoric expressions in subordinate clauses is determined by other, apparently nonstructural, factors.²⁴

Finally, the results are consistent with the hypothesis that structural focusing determines the interpretation of pronominals across, but not within, processing units. According to the discourse model suggested in Section 3.1, a main clause is an independent processing unit, so a sequence of main clauses constitutes a sequence of units. On the other hand, a subordinate clause does not form an independent unit, so a sequence containing a main and a subordinate clause is simply internal to the unit. The reference variation observed in the subordinate conditions is then expected from the model.

Unfortunately, in this experiment, we lost the contrast for the type of connection within semantic types because of the noise created by the causal interpretation of some when clauses. To enable the conclusions of this experiment to be generalized across the entire classes of paratactic and hypotactic connectives, this experiment will be repeated with a larger set of connectives.

An interesting extension of this work would be to investigate whether structural focusing is active within units, before it is eventually overridden by semantic focusing, or whether focusing preferences projected by the semantics of the verbs and connectives are immediately accessed during on-line processing.

4.2 Corpus-Based Data

The corpus-based study reported in this section is based on a Greek corpus built with text from Greek newspapers available on the Web. The corpus consists of approximately 800,000 words downloaded from the sites of the Greek newspapers *Eleftherotipia* and *To Vima*.

The aim of the corpus study is similar to that of the experimental study reported in the previous section. Greek allows dropped subjects, which yields higher referential ambiguity than English pronouns. We were therefore able to collect a reasonable number of tokens fulfilling the conditions of the experimental study and to test our hypotheses against naturally occurring data.

²⁴ It is understood that in the main-subordinate clause sequences, the subordinate clause is linked to the main clause in question and not to some subsequent main clause (as, for example, in cases in which the subordinate clause is preposed).

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 in the experimental study, however, the search of anaphoric expressions in main and subordinate clauses was not restricted to subject pronouns.

4.2.1 Data Collection and Coding. Greek has two pronominal systems: weak and strong. Weak pronominals include dropped subjects and object clitics. Clitics are marked with case, gender, and number features and attach to the verb. Direct-object clitics are case-marked “accusative,” and indirect or “dative” clitics are case marked “genitive.” Strong pronominals are also marked with case, gender, and number features, but syntactically, they behave as common nouns. Their functions are mentioned in Sections 2.2 and 3.1.1. In this study, only weak pronominals were included.

We established the following requirements for the data set of this study: (a) the subordinate clause or second main clause contains a third-person dropped subject or weak pronominal, and (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 number with the anaphoric expression.

For anaphoric reference in main-main and main-subordinate sequences, ideally we would have liked to include only those tokens in which the second main or subordinate clause under investigation was preceded by a unit containing only a main clause. Imposing this extra constraint, however, would have invalidated a large number of the already limited number of tokens, so we decided against doing so. Although conducting a second pass of the data, with the purpose of studying further the characteristics of these antecedents, would have been useful, for the purposes of this study it was not crucial. Further, a consistent pattern of reference in main-main sequences including cases in which competing antecedents are present 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 the point in the next section.

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

For the data set 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 data set the antecedent of the anaphoric expressions. One of the coders was the author and the other was a naive, nonlinguist speaker of Greek. Intercoder reliability was particularly high (98%). We attribute the high intercoder reliability to the fact that discourse-deictic expressions, known to lower intercoder reliability, were not included in this study.²⁵ The few cases of disagreement between the coders involved either instances perceived as ambiguous by the coders or abstract complex NPs about which there was disagreement as to

²⁵ Discourse-deictic expressions include demonstratives, such as *this* and *that*, used to refer to chunks of previous discourse. Discourse-deictic expressions in Greek are identical in form with neuter strong pronominals. Dropped subjects can refer to discourse-deictic expressions, but such cases were excluded.

whether the antecedent was the possessor or the possessee. Such cases were excluded from the final data set.

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.

4.2.2 Ranking Antecedents and Coding. Based on earlier work on the ranking of entities in Greek (Miltsakaki 1999), the competing antecedents were ranked according to the following rule:

Ranking Rule for Greek

Empathy > Subject > Indirect Object > Direct Object >
Indefinite Quantified NPs, Nonspecific Indefinites

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 encountered only the verb *like* from this verb category.²⁶

All entities introduced in subordinate clauses associated with the main clause are ranked using 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.²⁷ 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:

- (58) 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 by “Ap”) or “nonpreferred antecedent” (designated by “Anp”). The referent of an anaphoric expression was marked

²⁶ 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 by the genitive case.

²⁷ We use the terms *possessor* and *possessee* here for convenience to label the structural position of nouns in complex NPs. These terms, however, 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.

as a preferred antecedent when it was the highest-ranked entity in the set of competing antecedents. The referent of an anaphoric expression was marked as a nonpreferred antecedent when it was not the highest-ranked entity in the set of competing antecedents. In most cases, the set of candidate antecedents included only two candidates, so sub-categorizing nonpreferred antecedents was not crucial.

In what follows, example (59) is demonstrative of cases in which the referent of the anaphoric expression in the second main clause is marked “Ap.” The competing antecedents in (59a) are *ta oplá* and *anthropus*, because they have the same number and gender as the anaphoric *tus* in (59b). The NP *ta oplá* is 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 is not a factor in determining the ranking of entities, even in cases in which the semantics of the verb taking the referent of the anaphoric as an argument favors the human, in this case, antecedent. Example (60) 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’). Note 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.

- (59) a. [Ta oplá]_i ine kataskevasména 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]_j.
 This is the goal their.
 This is their_i goal.
- (60) a. [O Turen]_i vrisquete apo filosofiki apopsi ston antipoda.
 the Turen is-placed from philosophical view at-the opposite-side
 [tu Popper]_j.
 of-the 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.
 Recently, he_i wrote that there are two types of intellectuals.

In (61) the referent of the dropped subject in (61b) is marked “Anp.” The list of competing antecedents in (61a) contains *PAOK*, the name of a football team, and *Pikulín Ortith*, the name of a player, both being singular and masculine. In this case, 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 Anp, because it is ranked lower than the subject *PAOK*.

- (61) 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 (62) demonstrates that competing antecedents in dependent clauses do not override the salience of main-clause antecedents. Note that [*i kinonikes igesies*]_j is a perfectly natural candidate for the subject of the verb *pistevun*.

- (62) a. [I esiodoxi]_i pistevun oti ehun dimiurgithi [i kinonikes igesies]_j
 the ambitious believe that have been-created the social leaderships
 pu mporun na antiparatethun stin katestimeni exusia.
 which can to object to-the established leadership.
 [*The ambitious ones*]_i believe that there have been formed [*social
 authorities*]_j
 which can stand up to the established leadership/political power.
- b. [NULL]_i pistevun oti o agonas tus den ehi akrivos kerdithi alla
 NULL believe that the fight their not has exactly been-won but
 oti NULL vriskete se “dromo horis epistrofi.”
 that 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 of Ap and Anp.

4.3 Results and Discussion

Table 2 shows the distribution of anaphoric reference in the experiment described in Section 4.2. 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 nonpreferred antecedent (Anp), and the third column summarizes the total number of tokens per condition.

The corpus-based results support the hypothesis that anaphora does not obey the same rules in main and subordinate clauses.²⁸ 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 in which both competing antecedents belong to the 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 (63), the ranking of the complex NP *i simetohi tu k. Avramopulu* (Mr. Avramopoulos's participation) is *Avramopulos* > *Simetohi* because *Avramopulos* (the current mayor of Athens) is animate and ranks higher. The intended referent of the

²⁸ Not surprisingly, chi-square gives a highly significant $p < 0.0005$.

Table 2
Reference in main and subordinate clauses.

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

dropped subject in the co-ordinated clause, however, is *simetohi*. A possible explanation for this pattern is to analyze (63) as verb phrase (VP) coordination, in which case the two VPs share the same subject.

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

The same phenomenon was observed, however, in cases with no VP coordination, as shown in (64). 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 concept *political party* is not "inanimate" in the sense that it denotes a particular group of people. In this case, it would be an animate possessor, and the ranking would work as expected. In conclusion, although it seems possible that the ranking of complex NPs can be fixed reliably, taking into account special cases of VP coordination and animacy of collective nouns, the number of instances of such special cases is too small to draw any definitive conclusions.

- (64) a. [To koma]_i [tu Avramopulu]_j emfanizete se thesi na anadihthi
the party of-the Avramopulos appears in position to be-promoted
se paragonta pu tha tropopiisi tus orus tu politiku pehnidiu.
to 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 that 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
significant appeal before even get-revealed the political characteristics
tu.
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 in which either inferrable information was needed or the referent was placed in an adverbial

located in the same clause as the anaphoric itself. The following example contains instances of both cases:

- (65) a. Legete oti [o 'Mihanismos']_i metaferotan stin Arhea
 it-is-said that the 'Mechanism' was-being-transferred to-the Ancient
 Romi gia na epidihthi ston Kikerona
 Rome in-order that be-shown to-the Kikerona
 It is said that the 'Mechanism' was being transferred 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 (65d) resolves to *Mihanismos* in (65a). The entity *Mihanismos* is evoked much more recently as an inferrable entity in (65d), however—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 3. Chi-square shows no significant differences among the three types of subordinate clauses ($p < 0.182$).

These results indicate that the focusing preferences 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 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.

Table 3
 Distribution of Ap/Anp.

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

5. General Discussion

The interpretation of anaphoric expressions in natural language processing is not a trivial problem. Extensive research in past years has made significant contributions to our understanding of the phenomenon, and a considerable number of theoretically motivated and/or corpus-based anaphora resolution algorithms have been built with more or less success. The task remains a challenge, however, and the slow rate of improvement in the performance of anaphora resolution systems is somewhat alarming.

The detailed review of the literature provided in Section 2 revealed that many 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 lookup mechanism.

The main contribution of this article is precisely the distinction and specification of two systems that determine preferences for anaphoric interpretation. Contra earlier views on the status of subordinate clauses, we 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 that hold between predicates and participating entities should be expected to be closely determined on semantic grounds. We 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 to facilitate discourse processing and efficient integration of meaning to the 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 and Koller 2000). The notion of salience, in the sense of centering (Joshi and Kuhn 1979), is arguably crucial for efficient processing not only for natural language processing systems, but also for humans. Topic continuity therefore is evaluated using a salience mechanism operating across processing units, and we showed that this mechanism is structural and best defined in centering terms. We then argued that anaphoric reference that spans across units is also determined structurally.

Regarding centering-based anaphora resolution algorithms, which seem the best candidates for resolving anaphora across units, a few technical problems were discussed in Section 2. We suggested, however, that these problems can be fixed easily and proposed that the algorithm should select as the preferred antecedent the highest-ranked entity in the previous unit. 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 5.2 was designed to test the hypothesis that two mechanisms are indeed at work and also to evaluate the strengths of the modified centering-based algorithm for resolving anaphoric reference across units. The results were robust despite the moderate sample size, prescribing a route for a number of future projects in this direction, the most challenging of which will probably be understanding the structural and semantic properties of subordination and its role in the organization, representation, and structure of discourse.

Acknowledgments

I am most grateful to Ellen Prince and Aravind Joshi as well as to Robin Clark, Barbara Grosz, Michael Kohlhase, Alexander Koller, and Alistair Knott for very useful and stimulating discussions. I would also like to thank Felicia Hurewitz, Jesse Snedeker, and John Trueswell for their generous help in the design and analysis of the experiment as well as the other members of the Gleitmans' CHEESE seminar in the Psychology Department, University of Pennsylvania, for their continuous support and valuable feedback. Finally, I am grateful to three anonymous reviewers, whose thoughtful and constructive criticism helped me improve the quality of this article significantly. The research on which this article is based was supported by the Institute of Research in Cognitive Science, grant no. NSF-SBR 8920230, University of Pennsylvania.

References

- Brennan, Susan, Marilyn Walker-Friedman, and Carl Pollard. 1987. "A centering approach to pronouns." In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, pages 155–162, Stanford, California.
- Carden, Guy. 1982. Backwards anaphora in discourse context. *Journal of Linguistics* 18:361–387.
- Daneš, Frantisek. 1974. "Functional sentence perspective and the organization of the text." In F. Daneš, editor, *Papers on Functional Sentence Perspective*. Academia, Prague, pages 106–128.
- Di Eugenio, Barbara. 1990. "Centering theory and the Italian pronominal system." In *Proceedings of the 13th International Conference on Computational Linguistics (COLING 90)*, pages 270–275, Helsinki.
- Di Eugenio, Barbara. 1998. "Centering in Italian." In M. Walker, A. Joshi, and E. Prince, editors, *Centering Theory in Discourse*. Clarendon Press, Oxford, pages 115–137.
- Dimitriadis, Alexis. 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, pages 33–47.
- Grosz, Barbara, Aravind Joshi, and Scott Weinstein. 1983. "Providing a unified account of definite noun phrases in discourse." In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, pages 44–50, MIT Press, Cambridge, Massachusetts.
- Grosz, Barbara and Candace Sidner. 1986. "Intentions, intentions and the structure of discourse." *Computational Linguistics* 12(3):175–204.
- Hobbs, Jerry. 1978. Resolving pronoun references. *Lingua* 44:311–338.
- Hudson-D'Zmura, Susan and Michael Tanenhaus. 1998. "Assigning antecedents to ambiguous pronouns: The role of the center of attention as a default assignment." In M. Walker, A. Joshi, and E. Prince, editors, *Centering Theory in Discourse*. Clarendon Press, Oxford, pages 199–226.
- Hurewitz, Felicia. 1998. "A quantitative look at discourse coherence." In M. Walker, A. Joshi, and E. Prince, editors, *Centering Theory in Discourse*. Clarendon Press, Oxford, pages 273–291.
- Joshi, Aravind and Steven Kuhn. 1979. "Centered logic: The role of entity centered sentence representation in natural language inferencing." In *Sixth International Joint Conference on Artificial Intelligence*, pages 435–439, Tokyo.
- Kameyama, Megumi. 1993. "Intrasentential centering." In *Proceedings of the Workshop on Centering*, University of Pennsylvania, Philadelphia.
- Kameyama, Megumi. 1998. "Intrasentential centering: A case study." In M. Walker, A. Joshi, and E. Prince, editors, *Centering Theory in Discourse*. Clarendon Press, Oxford, pages 89–112.
- Kohlhase, Michael and Alexander Koller. 2000. "Towards a tableaux machine for language understanding." In *Proceedings of the Second Workshop on Inference in Computational Semantics (ICOS-2)*, pages 57–88, Dagstuhl, Germany.
- Kuno, Susumu. 1972. Functional sentence perspective: A case study from Japanese and English. *Linguistic Inquiry* 3:269–320.
- Lappin, Shalom and Herbert Leass. 1994. "An algorithm for pronominal anaphora resolution." *Computational Linguistics* 20(4):535–561.
- McCord, Michael. 1990. "Slot grammar: A system for simpler construction of practical natural language grammars." In R. Studer, editor, *Natural Language and Logic: International Scientific Symposium*. Lecture Notes in Computer Science. Berlin, Springer Verlag, pages 118–145.
- Miltsakaki, Eleni. 1999. "Locating topics in text processing." In *Computational Linguistics in the Netherlands: Selected Papers from the Tenth CLIN Meeting (CLIN '99)*, pages 127–138, Utrecht, Netherlands.
- Miltsakaki, Eleni. 2000. Attention Structure

- in Discourse: A Cross-Linguistic Investigation. Dissertation topic proposal, University of Pennsylvania, Philadelphia.
- Miltsakaki, Eleni. 2001. "Centering in Greek." In *Proceedings of the 15th International Symposium on Theoretical and Applied Linguistics*, Thessaloniki, Greece.
- Miltsakaki, Eleni and Karen Kukich. 2000a. "Automated evaluation of coherence in student essays." In *Proceedings of the Workshop on Language Resources and Tools in Educational Applications, LREC 2000*, pages 7–14, Athens.
- Miltsakaki, Eleni and Karen Kukich. 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)*, pages 408–415, Hong Kong.
- Prasad, Rashmi and Michael Strube. 2000. "Pronoun resolution in Hindi." In A. Williams and E. Kaiser, editors, *Working Papers in Linguistics*, volume 6. University of Pennsylvania, pages 189–208.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1972. *A Grammar of Contemporary English*. Longman, London.
- Rambow, Owen. 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 Pennsylvania.
- Stevenson, Rosemary, Rosalind Crawley, and David Kleinman. 1994. Thematic roles, focusing and the representation of events. *Language and Cognitive Processes* 9:519–548.
- Stevenson, Rosemary, Alistair Knott, Jon Oberlander, and Sharon McDonald. 2000. Interpreting pronouns and connectives: Interactions among focusing, thematic roles and coherence relations. *Language and Cognitive Processes* 15(3):225–262.
- Strube, Michael. 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)*, pages 1251–1257, Montreal.
- Strube, Michael and Udo Hahn. 1996. "Functional centering." In *Proceedings of the 34th Annual Conference of the Association for Computational Linguistics (ACL '96)*, pages 270–277, Santa Cruz.
- Strube, Michael and Udo Hahn. 1999. Functional centering: Grounding referential coherence in information structure. *Computational Linguistics* 25(3):309–344.
- Suri, Linda and Kathleen McCoy. 1994. RAFT/RAPR and centering: A comparison and discussion of problems related to processing complex sentences. *Computational Linguistics* 20(2):301–317.
- Suri, Linda, Kathleen McCoy, and Jon DeCristofaro. 1999. A methodology for extending focusing frameworks. *Computational Linguistics* 25(2):173–194.
- Tanaka, Izumi. 2000. "Cataphoric personal pronouns in English news reportage." In *Proceedings of the Third International Discourse Anaphora and Anaphor Resolution Conference, DAARC 2000*, volume 12, pages 108–117, Lancaster, England.
- Turan, Umit. 1998. "Ranking forward-looking centers in Turkish: Universal and language specific properties." In M. Walker, A. Joshi, and E. Prince, editors, *Centering Theory in Discourse*. Clarendon Press, Oxford, pages 139–160.
- van Hoek, Karen. 1997. *Anaphora and Conceptual Structure*. University of Chicago Press, Chicago.
- Walker, Marilyn, Masayo Iida, and Sharon Cote. 1994. Japanese discourse and the process of centering. *Computational Linguistics* 20(2):193–233.
- Walker, Marilyn and Ellen Prince. 1996. "A bilateral approach to givenness: A hearer-status algorithm and a centering algorithm." In T. Fretheim and J. Gundel, editors, *Reference and Referent Accessibility*. John Benjamins, Amsterdam, pages 291–306.
- Webber, Bonnie and Aravind Joshi. 1998. "Anchoring a lexicalized tree adjoining grammar for discourse." In *ACL/COLING Workshop on Discourse Relations and Discourse Markers*, pages 8–92, Montreal, Quebec, Canada.
- Webber, Bonnie, Alistair Knott, Matthew Stone, and Aravind Joshi. 1999a. "Discourse relations: A structural and presuppositional account using lexicalized TAG." In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, pages 41–48, College Park, Maryland.
- Webber, Bonnie, Alistair Knott, Matthew Stone, and Aravind Joshi. 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)*, pages 145–149, Edinburgh.