Effects of Subordination on Referential Form and Interpretation

Eleni Miltsakaki

Abstract

The aim of this paper is to investigate the distribution of the structural and semantic focusing effect (e.g., Stevenson et al (1994) and Grosz, Joshi and Weinstein (1995) respectively) on pronominal intepretation, and determine the conditions under which one prevails over the other. We propose that the syntactic locality created by subordinate clauses defines a reference domain where the interpretation of a pronoun is primarily determined by semantic focusing. Across structurally independent units, i.e., main clauses, pronominal interpretation is guided primarily by structural focusing. The proposal was tested in three experiments: two in English and one in Greek. The data showed that there is a strong preference to interpret ambiguous pronouns in main clauses as the structurally most salient entity in the previous main clause, in our case, the subject. In subordinate clauses, the interpretation of the ambiguous pronouns varies, a result consistent with semantic focusing accounts.

1 Introduction

For the past thirty years, the issue of anaphoric interpretation has been the subject of extensive research across disciplines, most notably in computational linguistics and psycholinguistics. While approaches to anaphoric interpretation differ in the way they model anaphoric intepretation, the bulk of research converges to the following two uncontroversial claims: a) pronouns refer to the most salient entities and b) a number of factors may independently and/or in collaboration affect anaphoric interpretation. Assuming these two claims, this paper focuses on two central focusing factors, namely structural and semantic, and sets as its goal to investigate the conditions under which one predominates over the other in resolving pronominal reference.

Relatedly, Stevenson, Knott, Oberlander, and McDonald (2000), in recent work, investigated the effects of focusing accounts and coherence relations on pronominal interpretation. Specifically, three experimental studies were designed to evaluate the focusing and relational hypotheses. According to the relational hypothesis, (Hobbs, 1979), the interpretation of a pronoun is determined by the coherence relation that holds between two discourse segments and not by what is in focus. For example, a *result* relation is about the entity which instantiates the thematic role associated with the endpoint of the event. The choice of a *result* continuation leads the hearer to interpret the pronoun referring to the individual associated with the endpoint. So, in the case of the ambiguous connective *so*, which can signal either a 'result' or a 'purpose' relation, the interpretation of the subject pronoun is determined by the type of relation. In the 'result' version shown in (1), the subject pronoun will be interpreted as the entity instantiating the thematic role associated with the consequence, i.e., the patient. In the 'purpose' version, shown in (2), the subject pronoun is interpreted as the entity associated with the agent of the event.

(1) $Bill_i$ handed the plate to Bob_j so he_j balanced his glass on it.

(2) $Bill_i$ handed the plate to Bob_j so he_i could refill his glass.

The focusing hypothesis covers two types of focusing factors: structural focusing (Grosz, Joshi, & Weinstein, 1983) and semantic focusing (Stevenson, Crawley, & Kleinman, 1994). According to structural focusing (developed in the Centering framework (Grosz, Joshi, & Weinstein, 1995)), the form of referring expression in a discourse affects the resource demands made on the hearer and consequently the perceived coherence of the discourse. The center of attention (Centering's focused entity) is an entity that realizes a link between two utterances and the most likely referent of an upcoming pronominal form. Each utterance evokes a set of potential centers which are ranked according to their salience. For English it has been argued that a major factor determining the salience of entities in an utterance is grammarical role, in particular subjecthood (e.g., Kameyama (1985), Brennan, Walker-Friedman, and Pollard (1987)).

On the other hand, semantic focusing Stevenson et al. (1994) assumes that both verbs and connectives have focusing properties. The focusing properties of the verbs are articulated in terms of thematic roles and the endpoint of an event. For example, the endpoint of an action verb, requiring an agent and a patient as arguments, is associated with the patient-argument, making it the most salient entity independently of its grammatical role. The focusing property of a connective depends on its meaning. A connective such as 'so' directs attention to the consequence of an event, thus reinforcing the salience of the patient. A connective such as 'because' directs attention to the cause of an even, thus weakening the focus on the patient.

Stevenson et al. (2000) in contrasting the focusing and the relational hypotheses on pronominal interpretation designed experiments which primarily tested the predictions of semantic focusing and the relational approach. The experiments showed that overall focusing approaches provide a better explanation for anaphoric interpretation than relational approaches. In their experiments, they also observed that there was a clear tendency to maintain consistency among structural focusing, verb semantics and coherence relation when it was possible. Interestingly, within the focusing hypothesis, (Stevenson et al., 2000) identified three patterns of pronominal interpretation: a) semantic focusing predominated over structural focusing, b) structural focusing enhanced semantic focusing and c) structural focusing predominated semantic focusing. Crucially, the conditions determining whether structural or semantic focusing will have a stronger effect on pronominal intepretation were unclear.

Filling this research gap, this paper investigates the conditions which determine the domain of applicability of structural and semantic focusing. Based on previous observations (Miltsakaki, 2002), we formulate and test the hypothesis that semantic focusing is predominant in resolving pronominal reference when the pronoun appears in subordinate clauses (intra-sententially) whereas structural focusing is predominant across sentences. This work is part of a larger project aiming to investigate the effects of all types of (tensed) subordinate clauses (adverbial, complement and relative clauses) on anaphoric interpretation. Here, the type of subordinate clause under investigation is 'subordinate adverbial', i.e. clauses introduced with subordinate conjunctions such as *when*, *while*, *although*, *because* etc. For simplicity, the term 'subordinate' in what follows is used to refer to subordinate adverbial clauses.

Sections 2 and 3 report two experiments in English. In section 3, a third experiment is reported for the Greek language. A brief discussion and conclusions follow in section 4.

2 Experiment 1: Structural and semantic focusing in English

In all three experiments reported in this study, the interpretation of a subject pronoun is quantified in two conditions: a) when the pronoun is located in a main clause following a main clause, and b) when the pronoun is located in a subordinate clause (adverbial) following a main clause. In experiment 1, we compared and contrasted structural and semantic focusing by a close semantic match of connectives. In the main-main condition the second main clause was modified by an adverbial whose meaning approximated the meaning of one of the subordinate conjunctions used in the experiment. Experiments 2 and 3 also involve a set of connectives for both the main-main and main-subordinate conditions but the effort for one-to-one mapping of subordinate conjunctions and clause adverbials was abandoned for reasons to be discussed in section

2.3.

2.1 Materials and design

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

Both the main clause adverbials and the subordinate conjunctions were selected from two semantic classes: TIME and CONTRAST. The TIME class included the subordinate connective *when* and the adverbial *then*. The CONTRAST class included the subordinate conjunction *although* and the adverbial *however*.

A sample stimulus set is shown in Figure (1). The experiment followed a 2X2 design. The factors were type of clause (main or subordinate) and semantic type (time or contrast). There were 12 target items which were combined with 24 fillers. All items appeared an equal number of times in each condition but only once for a given participant. Sixteen adult, native speakers of English volunteered to participate.

2.2 Results

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

The interpretation of the subject pronoun as the subject of the preceding main clause was quantified and converted to percentages. The scores were

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

(3)	The groom hit the best man. However, he...
(4)	The beggar pushed the gentleman although he...
(5)	The boxer kicked the referee. Then, he...
(6)	The policeman shot the burglar when he...

Figure 1: Experiment 1: Sample items

then submitted to a two-way ANOVA analysis. The results of the ANOVA showed a strong main effect for type of clause ($F(1,15)=25.6$, $p<0.0001$) and a marginal effect for semantic type ($F(1,15)=4.5$, $p<0.049$.

Figure (2) shows the percentages of reference to the subject of the first main clause by type of clause (main or subordinate) and semantic type (time or contrast). The percentages for each category show that when the second clause is a main clause the subject pronoun is more frequently interpreted as the subject of the main clause. On the other hand, when the second clause is a subordinate clause the subject pronoun shows a weaker tendency to be interpreted as the subject of the preceding main clause.

2.3 Discussion

In this experiment we contrasted semantic type, i.e. time and contrast, with type of clause, i.e. main and subordinate. The results show that the type of clause affects the interpretation of the pronoun it contains. In the main-main condition participants showed a significantly stronger tendency to interpret the subject pronoun as the subject of the preceding main clause than in the main-subordinate condition. This preference was demonstrated in both adverbial connectives, namely 'then' and 'however'. Conversely, in the main-subordinate condition, the subject pronoun was often interpreted as the object of the previous clause. The marginal effect of the semantic type shows that the pressure for structural focusing in the main-main condition overrides the pressure from semantic focusing. However, a word of caution is in order here. The semantic parallelism between the pairs of connectives, then-when and however-although broke for the temporal connectives. In many cases, the continuations in the when-clauses indicated a causal and not a temporal interpretation of the relation between main and subordinate clause, thus infecting the semantic uniformity of the semantic class TIME. An example from this category of continuation is shown in (7).

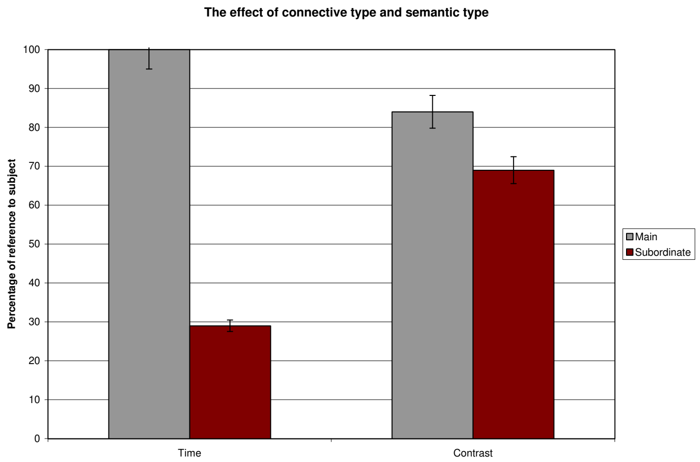

Figure 2: Percentage of reference to subject in English

(7) The actress kissed the hostess enthusiastically when she complimented her on her performance.

Experiment 2 was designed to fix this problem. In a sentence continuation task, it is very hard to control the interpretation of connectives by the participants. For this reason, in experiment 2 the semantic class condition was left out. Instead, a large number of connectives, subordinate and adverbial, was added in order to analyze the effect of the type of clause and draw a general and reliable conclusion about the main-subordinate distinction .

3 Experiment 2: Focusing in main and subordinate clauses in English

3.1 Materials and design

The method for this experiment was a sentence completion task. As in experiment 1, participants were asked to read sets of two clauses. Each set either contained a sequence of two main clauses (main-main condition) or a sequence of a main and a subordinate clause (main-subordinate condition). The second clause contained a subject pronoun and participants were asked to complete the sentences in a natural way. The critical items in this experiment had the same structure as in experiment 1. Five subordinate conjunctions and five clausal adverbials were included. Both the subordinate conjunctions and the adverbials were chosen from a variety of semantic classes. Figure (3) contains the complete list of connectives included in this experiment.

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

Figure 3: Experiment 2: Set of English connectives

Sample critical items are shown in Figure (4).

(8) The groom hit the best man. Moreover, he...

(9) The beggar pushed the gentleman so that he...

(10) The boxer kicked the referee. As a result, he...

(11) The policeman shot the burglar because he...

Figure 4: Sample items from experiment 2

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

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

3.2 Results

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

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

Figure (5) shows the percentages of reference to the subject of the first main clause in each condition. The results of this experiment confirm the results of experiment 1. The percentages for each category show that when the second clause is a subordinate clause, the subject pronoun shows a much weaker tendeny to refer to the subject of the preceding main clause. Reference to the subject is the preceding main clause, however, is strongly preferred when the subject pronoun appears in a main clause.

3.3 Discussion

The purpose of this experiment was to test if we can generalize across connectives the type of clause effect in experiment 1. The results of experiment 2 confirm this finding across a reasonably large number of connectives (five subordinate conjunctions and five adverbials). In the main-main condition, the pronoun was interpreted as the subject of the previous main clause across all adverbials, confirming that structural focusing in this condition is the primary factor determining pronominal interpretation. If semantic focusing was the primary determinant of salience in this condition we would expect to see a varied pattern depending on the semantics of the connective. In the main-subordinate condition, on the other hand, the percentage of reference to the subject of the previous main clause is significantly lower indicating that other factors override structural focusing.

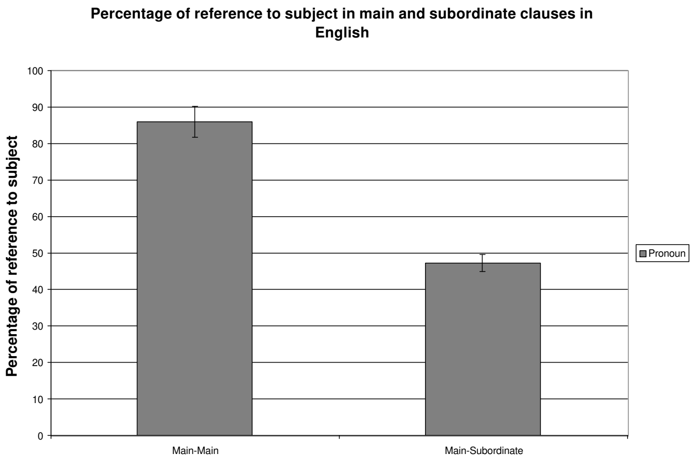

Figure 5: Percentage of reference to subject

4 Experiment 3: Focusing in main and subordinate clauses in Greek

4.1 The pronominal system in Greek

Greek is a subject-drop language. Its pronominal system consists of two paradigms: strong and weak. The weak paradigm includes null subjects and a set of weak pronominal forms marked for gender, number and case, cliticized to the verb in object position. The strong paradigm includes two sets of strong forms: the demonstrative/deictic paradigms *aftos-masc, afti-fem, afto-neut* 'this-one' and *ekinos-masc, ekini-fem, ekino-neut* 'that one'. It has been shown for Greek that one of the functions of the strong forms is to signal reference to a less salient entity, at least in cases when two competing antecedents are available ((Dimitriadis, 1996), (Miltsakaki, 1999), (Miltsakaki, 2001)). Null subjects and weak forms are used to refer to the most salient entities. This property of strong pronouns in Greek was utilized in the design of the Greek experiment as specified in the next section.

4.2 Materials and design

The Greek version of the experiment was modified in the following way. The method in this case was a rating questionnaire that elicited off-line judgments about naturalness. Participants were asked to read two versions of the same set of sentences. In version (1), the anaphoric element following the connective was a dropped subject. In version (2), the anaphoric element following the connective was the strong pronoun *ekinos, ekini, ekino*. In both versions, the continuations were identical. The semantics of the second clause were such that the referent of the anaphoric element was identified as the *object* of the preceding main clause. A sample stimulus set is shown in Figure (7). We quantified over the percentage of times that the participants judged the use of the strong pronoun to be the most natural way to refer to the object of the preceding main clause.

As in experiment 1 and 2, the main clause contained two male or two female referents and the main clause verb was an action verb involving physical contact. Following the design of experiment 2, five subordinate conjunctions and five clause adverbials were selected for the continuations. Figure (6) shows the complete set of connectives.

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

Figure 6: Experiment 3: Set of Greek connectives

There were 30 critical items combined with 90 fillers. The fillers consisted of pairs of sentences with a different variable to judge for naturalness, for example variations in word order in the continuations or variations in the use of a perfective or non-perfective form. Each condition (main-main or main-subordinate) appeared in fifteen versions: fifteen subordinate continuations and fifteen main clause continuations. Each connective appeared three times in each complete experimental set. Twenty adult participants, native speakers of Greek, volunteered to take part in the experiment. On average participation time was 20-30 minutes.

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

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

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

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

Figure 7: Experiment 3: Sample items

4.3 Results

The number of times the strong pronoun was judged more natural for reference to the object of the preceding main clause was first converted to percentages and then the scores were submitted to an ANOVA analysis. The results of the ANOVA showed a strong main effect of the type of clausal connection ($F(1,18)=52.78$, $p<0.000$)).

Figure (8) shows the percentages of felicitous reference to the subject of the preceding main clause using a strong pronoun. The percentages for each category show that strong forms are required for reference to the previous object across main clauses. When the anaphoric appears in a subordinate clause, reference to the object of the previous clause with a null subject is significantly facilitated.

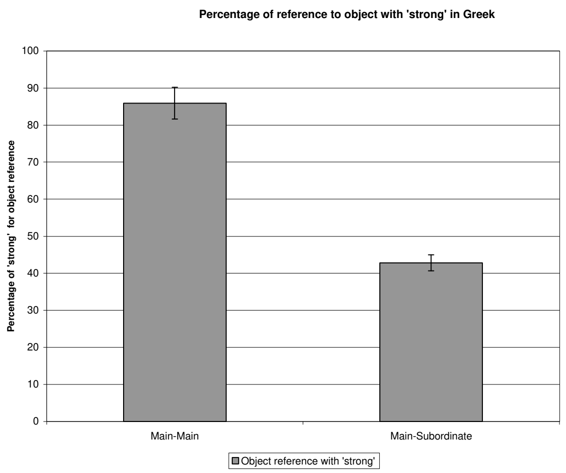

Figure 8: Experiment 3-Percentage of preference for ’strong’

5 General discussion

The basic question explored in this paper is whether syntactic subordination is one of the factors determining the strength of the effect of structural versus semantic focusing in pronominal interpretation.

Experiment 1 provided preliminary evidence for the effect of clause type (main versus subordinate) on anaphora resolution. In their continuations, participants tended to interpret the pronominal in the main clause condition as the subject of the previous main clause. No such pattern was identified in the main-subordinate condition where the interpretation of the pronominal varied across two types of subordinate clauses (time and contrast). In the same experiment, the semantic type had only a marginal effect on the intepretation of the pronominal. The result on semantic type was questionable. This was because connectives such as ’when’, prototypically temporal, may be used with alternative intepretations, making it hard to directly contrast syntactic clause types with semantic relations. For example, in experiment 1, the continuations on 'when' clauses would often establish a causal rather than temporal relation between the main clause and the subordinate clause.

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

Experiment 3 tested the same conditions in Greek. The aim of experiment 3 was to investigate whether the effect of subordination on anaphoric interpretation extends to other languages or whether it is a language-specific property of attentional organization in the English discourse. The results of experiment 3 show that a strong pronoun, normaly reserved for reference to an entity other than the most salient one in the previous discourse, is consistently judged more natural for reference to the object of the preceding of the main clause. On the other hand, in the main-subordinate condition the strong pronoun of the subordinate clause was not always necesary for reference to the object of the previous clause.

We conclude so far that structural focusing is predominant across main clauses whereas syntactic subordination creates a locality where other focusing factors are at work. The distribution of anaphoric interpretation per connective in English indicates that within this locality semantic focusing is most probably the predominant focusing factor. Preliminary examination of figure (4) in the appendix shows the percentage of pronominal reference to the subject of the preceding main clause per connective in English. In the experimental design, the main clause predicate was held constant, so the tendencies exhibited by each connective are due to the focusing properties of the connective combined with the focusing properties of action verbs. Figure (4) in the appendix also reveals an interesting tension between structural and semantic focusing. The adverbials 'then' and 'moreover' enhance the salience of the previous subject because they indicate continuation on the same topic reached. These two adverbials show a ceiling effect in the diagram. On the other hand, 'as a result' which, semantically, is predicted to shift the focus to the object (the patient) of the preceding clause shows a lower percentage of reference to the subject of the preceding main clause by comparison to other clause adverbials but still higher than the subordinate 'because' or 'so that'. Taking a closer look at the subordinate conjunctions, they, too, 'because' , 'so that' and 'although' show clear tendencies whereas it appears that some noice was introduced in the cases of the polysemous 'when' (temporal and causal) and 'while' (temporal and concessive).

Figure (10) shows a similar pattern for the Greek experiment. As in the case of English, the only subordinate conjunctions showing weak resolution tendencies are *eno* 'while' and *otan* 'when' which share the same semantic ambiguity as their English counterparts.

Acknowledgments

I am grateful to Jesse Snedeker, Felicia Hurewitz, John Trueswell, Lila Gleitman and the Gleitmans' CHEESE seminar participants for sharing their expertise in experimental design and analysis. Special thanks to Alistair Knott for very useful discussions on structural and semantic focusing. Last but not least, I would like to thank the members of the Institute of Language and Speech Processing in Athens, Greece, who participated in the Greek experiment and Maria Gavriilidou for her invaluable help in organizing the experimental session.

References

Brennan, S., Walker-Friedman, M., & Pollard, C. (1987). A Centering Approach to Pronouns. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics, Stanford, CA*, pp. 155–162.

Dimitriadis, A. (1996). When Pro-Drop Languages Don't: Overt Pronominal Subjects and Pragmatic Inference. In *Proceedings of the 32nd Meeting of the Chicago Linguistics Society, Chicago*, pp. 33–47.

Grosz, B., Joshi, A., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics, MIT Press, Cambridge*, pp. 44–50.

Grosz, B., Joshi, A., & Weinstein, S. (1995). Centering: A Framework for Modeling Local Coherence in Discourse. *Computational Linguistics*, *21*(2), 203–225.

Hobbs, J. (1979). Coherence and Coreference. *Cognitive Science*, *3*, 67–90.

Kameyama, M. (1985). *Zero Anaphora: The Case of Japanese*. Ph.D. thesis, Stanford University.

Miltsakaki, E. (1999). Locating Topics in Text Processing. In *Computational Linguistics in the Netherlands: Selected Papers from the Tenth CLIN Meeting, (CLIN '99), Utrecht*, pp. 127–138.

Miltsakaki, E. (2001). Centering in Greek. In *Proceedings of the 15th International Symposium on Theoretical and Applied Linguistics, Thessaloniki.*

Miltsakaki, E. (2002). Towards an Aposynthesis of Topic Continuity and Intra-Sentential Anaphora. *Computational Linguistics*, *28(3)*, 319–355.

Stevenson, R., Crawley, R., & Kleinman, D. (1994). Thematic Roles, Focusing and the Representation of Events. *Language and Cognitive Processes*, *9*, 519–548.

Stevenson, R., Knott, A., Oberlander, J., & McDonald, S. (2000). Interpreting Pronouns and Connectives: Interactions among Focusing, Thematic Roles and Coherence Relations. *Language and Cognitive Processes*, *15*(3), 225–262.

Appendix

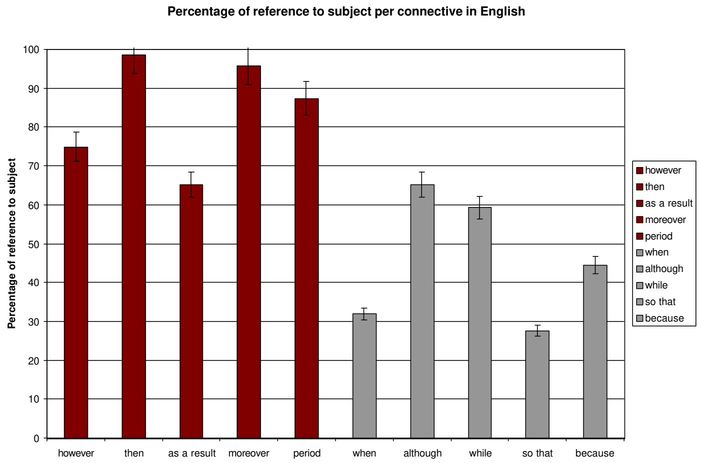

Figure 9: Experiment 2-Percentage of reference to subject per connective

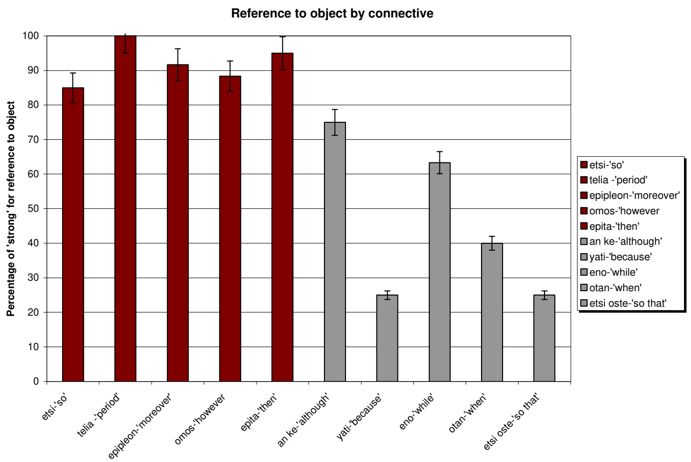

Figure 10: Experiment 3-Percentage of preference to ’strong’ per connective