CONNECTIVE POSITION, ARGUMENT ORDER AND INFORMATION STRUCTURE
OF DISCOURSE CONNECTIVES IN WRITTEN TURKISH TEXTS

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF INFORMATICS

OF

THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

IŞIN DEMİRŞAHİN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

THE DEPARTMENT OF COGNITIVE SCIENCE

SEPTEMBER 2008

Approval of the Graduate School of Informatics

_____

Prof. Dr. Nazife Baykal
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

_____

Prof. Dr. Deniz Zeyrek
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

_____

Assoc. Prof. Dr. Cem Bozşahin
Supervisor

Examining Committee Members

Prof. Dr. Deniz Zeyrek   (METU, COGS)   _____

Assoc. Prof. Dr. Cem Bozşahin   (METU, CENG)   _____

Assist. Prof. Dr. Annette Hohenberger(METU, COGS)   _____

Assist. Prof. Dr. Bilge Say   (METU, COGS)   _____

Dr. Ceyhan Temürcü   (METU, COGS)   _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this wok.

Name, Last name: Işın Demirşahin

Signature   :   _____

ABSTRACT

WORD ORDER AND INFORMATION STRUCTURE OF DISCOURSE
CONNECTIVES IN WRITTEN TURKISH TEXTS

Demirşahin, Işın

MS, Department of Cognitive Sciences

Supervisor: Assoc. Prof. Dr. Cem Bozşahin

September 2008, 103 pages

A text is a linguistic structure that is more than a random collection of sentences. A text is cohesive (Halliday & Hasan, 1976) and coherent (Mann & Thompson, 1987, 1988). Mainly ignored in the field of linguistics until recently, the text and the discourse structure have been inquired from various points of view (Asher, 1993; Asher & Lascarides, 1998; Grosz & Sidner, 1986; Mann & Thompson, 1987, 1988; Webber, 2004).

D-LTAG is a discourse grammar work that extends a lexicalized sentence level grammar LTAG (Joshi, 1987) to low-level discourse (Webber, 2004; Webber &

Joshi, 1998). In this framework, discourse connectives such as coordinating conjunctions, subordinating conjunctions, parallel connectives and discourse adverbials are predicates of discourse structure that take text spans that can be interpreted as abstract objects (Asher, 1993).

Turkish has a flexible word order in comparison to languages like English. In English, the discourse adverbials are noted for their ability to occupy positions unavailable to other discourse connectives. In Turkish, word order of other discourse connectives, coordinators and subordinators are not expected to be as restricted.

This thesis examines the connective position, argument order and the information structure of five Turkish discourse connectives in their eleven uses. The analyses show that the examined features of discourse connectives are related to the syntactic group the connective belongs to. Discourse connectives of the same syntactic groups exploit similar connective position and argument order possibilities, and they tend to be included in similar information units.

Keywords: Discourse, Discourse Connective, Information Structure, Argument Order, Connective Position, Word Order, Turkish

ÖZ

TÜRKÇE YAZILI METİNLERDE SÖYLEM BAĞLAÇLARININ BAĞLAÇ KONUMU, ÖGE DİZİLİMİ VE BİLGİ YAPISI

Demirşahin, Işın

Yüksek Lisans, Bilişsel Bilimler Bölümü

Tez Yöneticisi: Doç. Dr. Cem Bozşahin

Eylül 2008, 103 sayfa

Metin, rastlantısal bir şekilde bir araya gelmiş cümleler topluluğundan daha fazlasına sahip olan bir dilbilimsel yapıdır. Metin uyumlu (Halliday & Hasan, 1976) ve tutarlıdır (Mann & Thompson, 1987, 1988). Son zamanlara kadar dilbilim alanında genel olarak ihmal edilmiş olan metin ve söylem yapısı, daha sonra çeşitli bakış açılarından incelenmiştir (Asher, 1993; Asher & Lascarides, 1998; Grosz & Sidner, 1986; Mann & Thompson, 1987, 1988; Webber, 2004).

D-LTAG tümce düzeyinde sözcükçeleştirilmiş bir dilbilgisi olan LTAG'in (Joshi, 1987) alt-düzey söylem için genişletilmesiyle ortaya çıkmış olan bir söylem dilbilgisidir (Webber, 2004; Webber & Joshi, 1998). Bu çerçevede, eş dizimli bağlaçlar, art dizimli bağlaçlar, koşut bağlaçlar ve söylem zarfları gibi söylem bağlaçları, soyut nesneler (Asher, 1993) olarak yorumlanabilecek metin aralıklarını öğe alarak alan yüklemler olarak görev yapmaktadır.

Türkçe, İngilizce gibi dillere göre esnek bir sözcük dizilimine sahiptir. İngilizce'de söylem zarfları, diğer söylem bağlaçlarının erişimi olmayan konumlara gelebilmeleriyle dikkat çekerler. Türkçe'de diğer söylem bağlaçlarının, yani eş dizimli ve art dizimli yapısal söylem bağlaçlarının bu şekilde bir kısıtlamanın etkisinde olması beklenmemektedir.

Bu tezde, beş Türkçe söylem zarfı, on bir değişik kullanım şeklinde, bağlaç yeri, öge dizilimi ve bilgi yapısı açılarından incelenmektedir. Analizler, incelenen özelliklerin zarfın ait olduğu sözdizimsel sınıfla ilişkili olduğunu göstermektedir. Aynı sözdizimsel sınıftan olan bağlaçlar benzer bağlaç yeri ve öge dizilimi olanaklarını kullanmakta ve benzer bilgi birimlerine dahil olma eğilimi göstermektedirler.

Anahtar Kelimeler: Söylem, Söylem Bağlacı, Bilgi Yapısı, Öge Dizilimi, Bağlaç Yeri, Sözcük Dizilimi, Türkçe

To all my friends, real and imaginary.

ACKNOWLEDGMENTS

BKFT, where there are too many names and too many things to thanks for. I can only name a few: I am grateful to Alp Yürüm for everything, especially for being such a nice crow; Meltem Cemre Üstünkaya, Sevcan Hakyemez and Leyla Önal for sharing my unusual interests; Başar Tuna, Çetiners, Okumuşs, Malkoçs, Beyazgüns and soon-to-be Soykals for comprising such an implausibly comfortable community.

I want to express my special gratitude to dear Algan Uskarcı for providing me with food, shelter, warmth and affection. Without him, I would fall apart.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

B:	Background
CAO:	Connective-Argument Order
D-LTAG:	Lexicalized Tree Adjoining Grammar for Discourse
LF:	Logical form
K:	Kontrast
MTC:	METU Turkish Corpus
N:	Nucleus (RST framework)
NP:	Noun phrase
PDTB:	Penn Discourse TreeBank
R:	Rheme (Information unit)
	Reader (RST framework)
RST:	Rhetorical Structure Theory
S:	Sentence (Syntnactic representation)
	Satellite (RST framework)
T:	Theme
T-K:	Theme-Kontrast
VP:	Verbal phrase
W:	Writer (RST framework)
β:	Background
θ:	Theme
ι:	Intermediate phrase
Φ:	Complete phrase
ρ:	Rheme

η: Unaccented form

CHAPTER 1

INTRODUCTION

For years, discourse has been an academically less studied area of linguistic research. Syntax has been associated with sentence level structure. Recently, researchers started to pay attention to the linguistic environment surrounding the sentence. Many of them attributed structure to its organization and finally, inquired how language progresses from sentence level structure to discourse level structure.

An utterance composed of the same words can convey different information structures depending on the context of utterance. A sequence of utterances and the way they are linked together provides a structured background for the interpretation of the subsequent utterances. Discourse studies focus on the structure of language beyond isolated sentences, how utterances are linked together to build a sequential. The aim of this study is to investigate the relation between sentence level information structure and discourse structure. This study will make use of lexicalized grammars to analyze utterances and discourse.

Steedman (2000) states that Information Structure defines how Logical Form relates to the discourse structure. He proposes integration of syntactic, semantic and

information structural properties in the Logical Form within the framework of Combinatory Categorial Grammar. Information Structural properties of Turkish have been studied at sentential level in the same framework (Hoffman, 1995; Özge, 2003).Following a lexicalized approach to discourse (Webber, 2004; Webber & Joshi, 1998), this study aims at investigating how syntactic structure extends to cover discourse, and more specifically, how the syntactic positioning of discourse elements affect the information structure.

This thesis is organized as follows: In chapter 2, general information about corpus based linguistic studies and the METU Turkish Corpus will be introduced and the discourse level annotation project will be presented.

In chapter 3, discourse structure in general will be discussed. Various approaches to discourse will be examined and compared. D-LTAG will be discussed in greater detail. The types of discourse connectives introduced by the theory, and the types of discourse connectives that exist in Turkish will be compared.

In chapter 4 the word order and information structure of Turkish will be presented. Previous work on Turkish word order and the main factors affecting the Turkish word order will be analyzed. Information structure of Turkish and previous work about the issue will be discussed. The interaction of intonation and word order will be presented. The effect of intonation on written texts will be inquired.

In chapter 5, the connective-argument order of Turkish discourse connectives will be presented through examples from the METU Turkish corpus and their variants. The word order, connective-argument order and the information structure will be cross-examined in order to discover the interactions between them.

In chapter 6, chapters 3 and 4 will be discussed upon the findings from chapter 5. How Turkish word order and information structure affect connective-argument order will be questioned.

CHAPTER 2

CORPUS STUDY AND THE MTC

A corpus is a compilation of sample texts that can be taken to be representative for a language (Tognini-Bonelli, 2001). Linguistic corpora provide the linguistic studies with a large sample of naturally generated language for analysis. The study may arise from the analysis of the corpus itself by observing the data, building hypotheses, making generalizations and achieving statements, or it may use the corpus to test or exemplify aspects of a language or a theory. Either case, the corpus gives the linguist an empiric ground to put their feet on.

2.1   METU Turkish Corpus

METU Turkish Corpus (MTC) is a natural written language source of 2 million words from multiple genres (Say, Zeyrek, Oflazer, & Özge, 2002). MTC sources are all post-1990 written Turkish.

MTC is composed of samples that are approximately 2000 words (the samples end where the last sentence ends). A variety of Turkish resources by Turkish authors were used as allowed by copyright agreements. At most three samples were taken

from a single source and there are no more than five publications or ten samples for each author. The genres in the corpus include, but are not limited to, novels, short stories, essays, research monographs, interviews, memoirs and news. Such a diverse corpus offers a wide variety of linguistic use.

The examples in this thesis are all natural generations of written language from the MTC, or their word order variants and questions constructed to accommodate the information structure of such examples, unless explicitly stated otherwise.

2.2   MTC as a Written Discourse Resource

There is an ongoing project to annotate the discourse connectives in the MTC (Zeyrek, Turan, & Demirşahin, 2008; Zeyrek & Webber, 2008). The annotations are intended to be as theory free as possible in order to allow the generation of an unbiased linguistic source.

The 2000-word samples of the MTC provide sufficient context to make decisions about discourse and the information flow most of the time. When it does not, there are usually two cases. Either real world knowledge is necessary to interpret the text, or the analyzed part is at, or too close to, the boundaries of the sample. Either way, these cases are very rare.

The annotation scheme used in this thesis, following earlier representations for both English and Turkish (Webber et al., 2005; Zeyrek et al., 2008), will be as follows: the connective will be in bold typeface and will be underlined. The argument syntactically hosting the connective, Arg2, will be in bold typeface. The other argument of the connective, Arg1, will be in italics.

Discourse connective, and its modifier: **<u>Boldface and underlined</u>**

The argument syntactically hosting the connective: **Boldface**

The other argument of the connective: *Italics*

For annotation samples, see Appendix A.

CHAPTER 3

DISCOURSE STRUCTURE AND DISCOURSE CONNECTIVES

In most linguistic studies, the syntactic and semantic structures are investigated on the sentence-level only. The larger linguistic environment surrounding individual sentences are often ignored, or just vaguely referred to when there is a lexical or syntactic ambiguity to be resolved. However, there are other studies which look into this linguistic environment, the discourse.

3.1   Text Unity

The text, as used in linguistics is "any passage, spoken or written, of whatever length, that does form a unified whole" (Halliday & Hasan, 1976, p. 1). The native speaker can distinguish a collection of random sentences from a text without knowing exactly what makes a text a "text", but for the linguist this is an area to be explored.

Halliday and Hasan propose that the text is not a grammatical unit, built up by a string of sentences through constituency. They also defy the definition of text by its size. In fact, in some contexts like slogans or notices, a text may be shorter than a clause or a sentence; just a phrase. Instead, they regard the text as a semantic unit,

making the distinction clear with the statement "a text does not consist of sentences; it is realized by, or encoded in sentences … we shall not expect to find same kind of structural integration among the parts of a text as we find among the parts of a sentence or clause" (Halliday & Hasan, 1976, p. 2). They define the property of being a text as texture, and this texture is held together by relations between the elements in the text. A single relation is called a tie and there are five types of cohesive ties; reference, substitution, ellipsis, conjunction and lexical cohesion.

According to Halliday and Hasan, cohesion is based on the presupposition relation between the elements of a text. One element presupposes the other, and its interpretation is dependent on that element it presupposes.

Though they focus exclusively on cohesion, Halliday and Hasan do not reject the existence of a linguistic structure in text. Structure, being a unifying relation, builds a cohesive unit which has texture. But whereas structure implies texture, texture does not necessarily imply structure. Halliday and Hasan reserve the term cohesion for non-structural relations that contribute to the texture.

3.1.1   Reference

Reference items are particularly interesting for this thesis since they can be taken as arguments by discourse connectives, and the nature of the reference item sometimes determines the final analysis of the discourse relation.

Halliday and Hasan (1976) state that reference items can either refer to an entity, event or situation in the environment of the text (exophora), or to an element in the text (endophora). Endophora in turn can refer to the preceding text (anaphora), or to the following text (cataphora).

There are three types of reference according to Halliday and Hasan (1976): personal, demonstrative and comparative. Personal reference items refer to the entities in the text such as speaker, addressee, other specific persons, general persons, and animals and items. In English, *I*, *you*, *he*, *she*, *it*, *we*, *they* and *one*, and their accusative and

possessive counterparts refer to persons. *Ben*, *sen*, *o*, *biz*, *siz*, *onlar* and their inflected forms refer to persons in Turkish. In Turkish, the third person singular, *o*, does not distinguish gender or animacy.

Demonstrative reference items have a distance feature. *This*, *these*, *here* and *now* in English have closeness feature whereas *that*, *those*, *there* and *then* convey farness. Turkish counterparts of demonstrative reference items have three positions on the distance scale: *bu*, *bunlar*, *bura* refer to close referents, *şu*, *şunlar*, *şura*, refer to referents not very close but also not very far and *o*, *onlar*, *ora* refer to distant referents. Note that the distant demonstrative reference item root is *o*, same as the third person singular.

Halliday and Hasan state that the singular form of object reference in English, *it*, can also refer to a passage of text. In Turkish, *o*, can also refer to a passage of text, however, our intuition is that it is not a personal reference, but a demonstrative reference that is employed when referring to passages of texts. None of the other personal reference items refer to passages of text, whereas almost all demonstrative reference items frequently refer to passages of text.

When referring to a text passage, *o* is anaphoric, i.e., *o* refers to a passage of text in the preceding discourse. Cataphoric cases of *o* after the complementizer *ki*:"that" has been attested (Kucuk & Yondem, 2007). On the other hand, *şu* is cataphoric, i.e., *şu* refers to a passage of text in the following discourse. *Bu* is usually anaphoric, but there are cases it can be cataphoric too (see chapter 5).

Comparative reference forms cohesive links by means of identity, similarity or difference. In addition to the general comparison adjectives and their adverbial counterparts, such as *same*, *identical*, *similar*, *additional*, *other*, *different*, *else*, *identically*, *similarly*, *likewise*, *so*, *such*, *differently*, *otherwise*, particular comparison adjectives and adverbs such as *better*, *more*, and comparative forms of other adjectives form comparative reference ties, too. Turkish comparative reference items include but are not limited to: *aynı*, *benzer*, *farklı*, *başka*, *değişik*. Comparative reference items can be arguments of subordinating conjunctions as in the case of

yüzden, *sebeple* etc, to form adverbial-like discourse connectives such as *benzer sebeplerle* "because of similar reasons".

3.1.2 Substitution

During substitution a word takes the place of another word in the text. The resulting cohesive relation, according to (Halliday & Hasan, 1976), is between words. Unlike reference, which is a semantic cohesive relation, (Halliday & Hasan, 1976) take substitution, including ellipsis, to be grammatical. Therefore, reference can point to anywhere in and out of the text, but substitution is confined to the text. Even in the rare case of exophoric substitution, (Halliday & Hasan, 1976) expect to find an assumption or implication that something has been said.

Substitution has three types: nominal, verbal and clausal (Halliday & Hasan, 1976). Nominal substitution occurs when a word takes the place of the head of a nominal group. In English, *one*, *ones* and *same* can substitute nominal heads. Though Turkish can employ *biri* for nominal substitution as English employs *one*, the use of definitive morphology seems more common for this job. Where the English native speaker would use *the red one* to refer to a red dress, the Turkish native speaker would prefer *kırmızıyı* "red-DEF.ACC" or *kırmızı olanı* "red be-REL-DEF.ACC" both meaning "the red one" without substitution. The Turkish counterpart of *same* is *aynısı*. This word carries a possessive marker, morphologically indicating the cohesive relation.

Verbal substitution occurs when a word takes the place of a lexical verb, acting as the head of a verbal group. The English word for verbal substitution is *do*. Its Turkish equivalent is *yap*, and *yap* can be used as a verbal substitution item.

In the case of clausal substitution, a word does not take the place of another word or word group, but a whole clause. In English *so* and *not* are used for clausal substitution (Halliday & Hasan, 1976). In Turkish the clausal substitution can be conveyed by *öyle*. In negative situations, *öyle* is used with the appropriate negative form.

8

Substitution items can also be taken as complements by discourse connectives. They can even form discourse adverbials as *öyleyse* has done through lexicalization from an inflected form with *–se*, a subordinator-type discourse connective.

3.1.3 Ellipsis

When the discourse connective is defined by taking arguments that are abstract objects (Webber, 2004), and when the notion of abstract object depends on being a proposition, fact, description, situation, or eventuality (Asher, 1993), it becomes exceptionally important to understand the nature of ellipsis. A group of words that seem to be grouped together without an obvious predicate may constitute a proposition, fact, description, situation or eventuality, thus may be an abstract object: a valid argument for a discourse connective.

Ellipsis is not very different from substitution from a viewpoint of cohesion. In fact, (Halliday & Hasan, 1976), take ellipsis to be "substitution by zero" (p.142). Ellipsis is the case when something is not said, but is still understood.

Like substitution, ellipsis has three types: nominal ellipsis, verbal ellipsis and clausal ellipsis. Nominal ellipsis occurs within a nominal group, i.e., some part of a nominal group is missing from the utterance.

Verbal ellipsis means something in the verbal group is left unsaid. The unsaid material may be the lexical verb in the verbal group, in which case Halliday and Hasan call it a lexical ellipsis, or it may be other materials, subjects, modals, etc., in which case it is called operator ellipsis.

3.1.4 Conjunction

Conjunction is another type of cohesive link, and in some ways different from the others (Halliday & Hasan, 1976). Reference, substitution and ellipsis instruct the reader or hearer to search for an element, most of the time in the preceding or following text. Conjunction, on the other hand, instructs the addressee how to bring two parts of text together. The meaning of the conjunctive item itself is not dependent on what is presupposed.

A relation can be expressed in many ways in natural languages. Two events, A and B, in a relation can be expressed by grammatical predication, as in "A caused B", by minor predication as in "B happened because of A", by means of a subordinator as in "Because A happened, B happened", by means of an adverbial expression relating two separate sentences as in "A happened. As a result B happened." This adverbial expression is called a conjunctive adjunct or a discourse adjunct by Halliday and Hasan (1976) and a discourse adverbial by Webber (2004).

Halliday and Hasan draw a line between coordination and conjunction. They state that *and* and *or* relations in their very basic logical sense are structural and not cohesive. One of their arguments against coordination being a cohesive relation is that coordinated items form a single complex element, which behaves as simple elements behave.

Halliday and Hasan define four major types of conjunctive relations: additive, adversative, causal and temporal. These types are further specified according to too detailed criteria to mention here. For the complete "summary table of conjunctive relations" see Appendix B.

The conjunctive relations can be external or internal. Halliday and Hasan propose these terms to express functional dichotomy that might be called objective/subjective or experiential/interpersonal. The external relations exist simply between two events, or rather situations. Internal relations occur in the communication process. This dichotomy is most explicit in temporal relations. For example, in a text *after this*

might refer to after something already mentioned in the text (external, in "thesis time") or after the time the text is being realized (internal, in "thesis time").

The indication of such a division also exists in the Penn Discourse Tree Bank (PDTB) sense list in their annotation manual (Prasad, Miltsakaki, Dinesh, Lee, & Joshi, 2007). In this relatively theory independent treebank's sense hierarchy, there are four major semantic classes: temporal, comparison, contingency and expansion. These classes are further divided into types and subtypes (see Appendix B), where some senses have "pragmatic" subtypes. Pragmatic senses involve the interpretation of an argument rather than simply compositional meanings, or involve evaluation of speech acts.

One major difference between the two approaches is that Halliday and Hasan put conjunctives under certain types, for example, *thus* is put under additive, internal, apposition, exemplificatory in Halliday and Hasan's table. In PDTB annotations, on the other hand, the exact sense of a particular instance of *thus* would be clear only when the annotators put that particular *thus* into context.

3.1.5 Lexical Cohesion

Lexical cohesion occurs when semantically close words are used repetitively in a text.

Halliday and Hasan (1976) propose that lexical cohesion occurs in two ways, reiteration and collocation. Reiteration, as the name implies, is repetition of the same referent but this is not restricted to the repetition of the same word. In fact, repetition of the same word is only one of the ways reiteration can take place. Other ways are use of synonyms like *ascent-climb*, near-synonyms such as *sword- brand*, superordinates such as *Jaguar-car*, (the examples are from Halliday and Hasan, 1976, p.278) and use of general words such as *people*, *thing*, *place*, etc.

In reiteration, all the words used refer back to the same referent even though the words themselves are not the same. In collocation, on the other hand, the referents

are not the same, they even may be opposites, but the words are still cohesive. Such semantically close words often come from complementary sets as in *boy-girl*, or antonyms such as *like-hate*, members of the same ordered series, for example, *Tuesday-Thursday*, members of unordered lexical sets like *red-green*, words in a part-whole relation such as *box-lid*, or part-part relation as in *mouth-chin*, as well as words which are not easy to put under a systematic semantic class, but are related nevertheless, for instance, *comb-curl*.

Though Halliday and Hasan prefer to keep cohesion distinct from discourse structure, lexical cohesion stands close to some relations in discourse structure theories. What discourse structure theories name elaboration (Mann & Thompson, 1987, 1988) or entity relation (EntRel) (Prasad et al., 2007; Webber et al., 2005) are relations where two discourse units are related by means of providing more information about the same thing or even just being about the same thing. Unlike lexical cohesion ties, which can exist between any items in the text, both of these relations are restricted to adjacent text spans, elaboration by virtue of being an Rhetorical Structure Theory (RST) relation and EntRel by virtue of being an implicit relation which is defined at sentence boundaries. The status of elaboration as a discourse relation has been questioned. (Knott, Oberlander, O'Donnel, & Mellish, 2001)

3.2   Discourse Structure

Many researchers have observed systematic regularities in discourse that can be generalized to a discourse structure (Asher, 1993; Asher & Lascarides, 1998; Grosz & Sidner, 1986; Mann & Thompson, 1987, 1988; Webber, 2004). A few of these studies will be discussed below.

One theory of discourse structure is the Rhetorical Structure Theory (RST) (Mann & Thompson, 1987, 1988). RST sets off from the difference between a text and a random collection of sentences, which they identify to be coherence. RST assumes that coherence occurs when every part of a text is one way or an other connected to another part in the text and these connections between parts of text can be

represented by functions, i.e., plausible reasons for the presence of particular parts in the text.

RST proposes a hierarchical structure for text. Relations among clauses are analyzed independent from any lexical cue. The relations are either asymmetrical, i.e., they occur between a nucleus and a satellite, or symmetrical, meaning that they exist between parts of text that are equally prominent in the relation. In the latter case, the relation is said to be multinuclear.

A relation in RST consists of constraints on the nucleus, constraints on the satellite, constraints on the combination of the two and the effect, i.e., what the writer intended to achieve, or how this relation changes the reader's ideas. For example an EVIDENCE relation exists between a nucleus satisfying the constraint "R might not believe N to a degree satisfactory to W" and a satellite satisfying the constraint "The reader believes S or will find it credible"[1]. The constraint on the combination of these two is "R's comprehending S increases R's belief on N" and the effect of this relation is that "R's belief of N is increased" (Mann & Thompson, 1987)

Though these features seem plausible, the analyst has to guess what the writer intended in order to determine the nature of relation. Writers do not always write what they intend to. The task of analyzing low level semantic relations between parts of text is more or less mechanical, whereas the task of identifying intentions requires a deeper understanding of the text, the context and the author. What is more, one relation may be used with different intentions in different situations as in pragmatic senses of PDTB.

RST schemas define how spans of text can interact with each other. The schemas apply recursively, i.e., a text span resulting from the application of a schema can be,

---

[1] In RST, R stands for the reader, W stands for the writer, N stands for the nucleus, and S stands for the satellite. The abbreviations are restricted to this section and might be used differently elsewhere.

or rather, is expected to be the nucleus or satellite of another relation higher in the hierarchy. Below is a representation of five RST schemas.
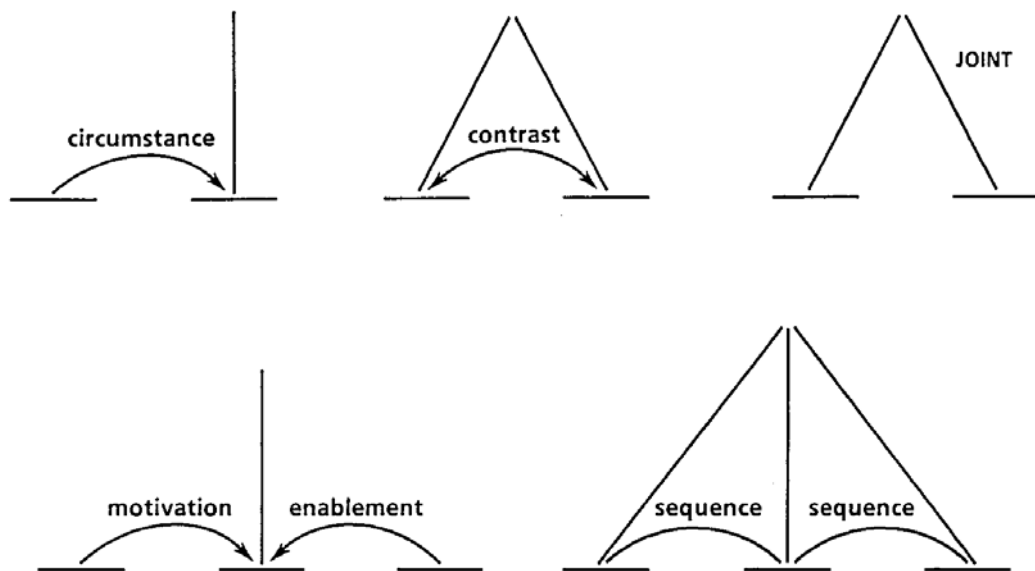
Figure 1 - RST Schemas (Mann and Thompson, 1987 p.7)

The RST schemas are applied in a way to satisfy four constraints. "Completeness" requires that the application of schemas to the entire text results in one schema application. "Connectedness" requires that all text spans in the text are either a minimal unit or take part in another schema application in the analysis. "Uniqueness" requires that schema applications are on different sets of text spans, and "Adjacency" requires that the text spans of a schema application result in another text span. (Mann & Thompson, 1987)

The schema application constraints are well defined and they are at the same time quite strict. Such strict restrictions are bound to result in consistent analyses between analysts; however, they are also likely to interfere with the analyst when determining the features of a relation. As a final note on RST, they assume that relations only

14

exist between adjacent text spans. In addition to our studies with the data from the MTC, other discourse studies show that relations may exist between non-adjacent and even non-continuous text spans (Prasad et al., 2007).

Grosz and Sidner (1986) propose a theory of tripartite discourse. They claim that discourse includes three separate components which interact with each other. The first component is the linguistic structure, which consists of a sequence of utterances. Segments of utterances are not necessarily continuous. This discourse segment structure interacts with the utterances that make up the segment. Some expressions in these utterances, i.e., "cue phrases", express information about the discourse structure, and are among the primary indicators of segment boundaries. In return, the generation and interpretation of these expressions are constrained by the discourse.

The second component is the intentional structure. It concerns the purpose of the discourse. Grosz and Sidner (1986) differentiate the purpose essential to the discourse from private purposes. The discourse purpose (DP) explains why that particular discourse is happening and why it is happening the way it does. Each discourse segment has a discourse segment purpose (DSP). DSPs make up the DP and each individual DSP indicates how the discourse segment contributes to the discourse. DSPs are structurally related by dominance and satisfaction-precedence. A DSP dominates another when the latter contributes to the satisfaction of the dominant DSP. Satisfaction-precedence relation occurs when one DSP needs to be satisfied before another DSP.

The third component is the attentional state, which concerns the focus of attention. The attentional state is represented by a focus space which defines the salient entities at that point of discourse. Naturally, the focus space is updated as the discourse progresses. A focus space, in a way, includes both (parts of) the discourse segment and the DSP, so that it represents that the conversational participants are aware of what is being discussed and why it is being discussed (Grosz & Sidner, 1986).

Other researchers such as Asher and Lascarides (1998) define a discourse structure that is above and beyond sentence level structure, as Grosz and Sidner do.

Another approach to discourse structure, Lexicalized Tree Adjoining Grammar for Discourse - D-LTAG (Webber, 2004; Webber & Joshi, 1998), looks into the structure of low level discourse, inquiring whether sentence level syntax can be extended to discourse structure. D-LTAG approach will be presented in detail in the following section.

3.3   A Lexicalized Approach to Discourse Structure

D-LTAG proposes that syntax does not come to an end at the sentence boundaries, but extend to low level discourse (Webber, 2004; Webber & Joshi, 1998; Webber, Knott, Stone, & Joshi, 1999). D-LTAG's niche is somewhere between the sentence level syntactic theories and discourse theories of higher levels. Before delving into this low level grammar formalism, it would be beneficial to go back to the sentence level and see what exactly a Tree Adjoining Grammar (TAG) is.

3.3.1   TAG and LTAG

Tree Adjoining Grammar (Joshi, 1987; Schabes & Joshi, 2002), define the language as a set of trees. Some examples of elementary trees are presented in Figure 2. Basically, there are two types of elementary trees, initial and auxiliary. Initial trees have substitution sites, allowing for simple insertion of a compatible node. Auxiliary trees have an adjunction site and identically labeled head and foot nodes. Auxiliary trees allow for recursive operations.
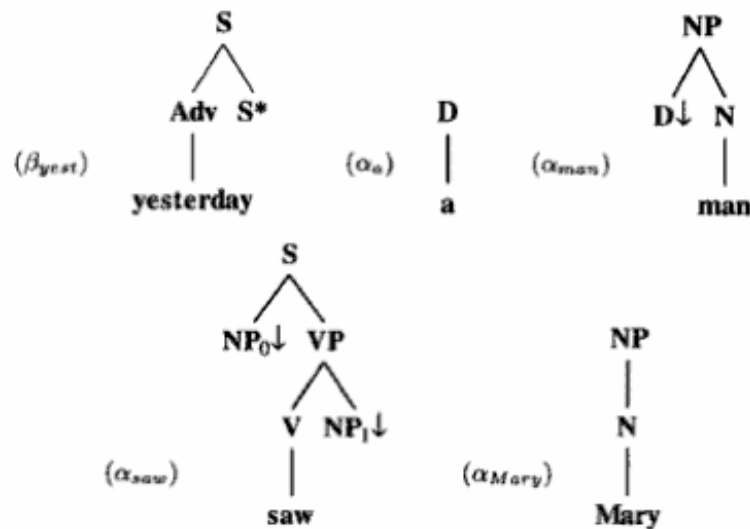
Figure 2 - Some elementary trees (Schabes & Joshi, 2002 p.75)

α trees are initial and the β tree is auxiliary

Lexicalized Tree Adjoining Grammar (LTAG) is a lexicalized version of TAG. In L-TAG, the trees are anchored by lexical entries. The set of trees of a lexical item shows which syntactic types that item can belong to, and in which syntactic configuration it can appear. An example set of a lexical item *like* is given in Figure 3. This set is far from being complete, but it is adequate to give the basic idea of what an L-TAG set for an entry looks like. Notice that the notion of lexical item is wider than the notion of lexeme. The LTAG set of *like* includes the prepositional configuration as well as the verbal configuration.

The basic types of trees and operations are the same as TAG. (a), (b) and (c) in the Figure 3 are initial trees, whereas (d) and (e) are auxiliary trees. Tree structures can combine through *substitution* (↓) and *adjoining* (*). Figure 4 shows an example of how trees can combine. The auxiliary tree anchored by *like* (Figure 3.d) is adjoined to the NP node that dominates *an apple*. The NP *an apple* is now in the foot node of the auxiliary, and new material can be substituted to the empty sites to the right.
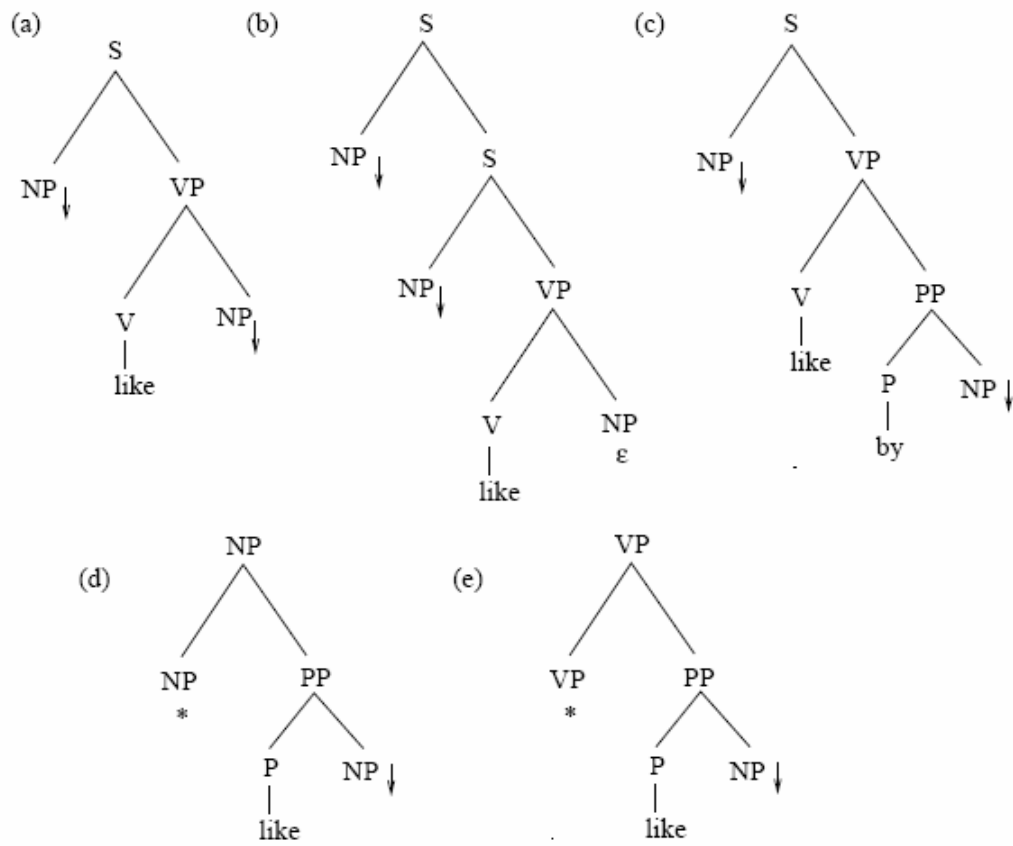
Figure 3 - Elements of the tree set of *like*. (Webber, 2004 p. 5)
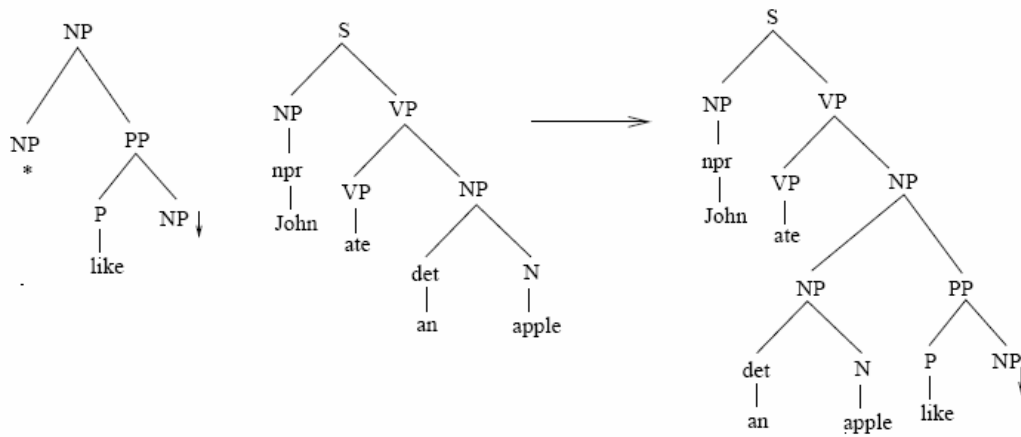
Figure 4 - An auxiliary PP tree adjoining to an initial NP tree (Webber, 2004 p.6)

3.3.2 D-LTAG

Lexicalized Tree Adjoining Grammar for Discourse (D-LTAG) builds upon L-TAG, extending the mechanisms for sentence level syntax to discourse level (Webber, 2004; Webber & Joshi, 1998). Low level discourse structure is represented by trees which are anchored by lexico-syntactic items that signify discourse relations, i.e., the discourse connectives. Discourse connectives act as predicates, similar to verbs at clausal level (Cresswell et al., 2002, 2005; Forbes-Riley, Webber, & Joshi, 2006; Forbes, Miltsakaki, Prasad, Sarkar, & Joshi, 2003; Webber, 2004, 2006; Webber & Joshi, 1998). The arguments of the discourse connectives are texts spans that can be interpreted as abstract objects, i.e., propositions, facts, descriptions, situations, or eventualities (Asher, 1993). The operations necessary to build up the discourse structure are the same as LTAG, i.e., *substitution* ($\downarrow$) and *adjoining* (*).

3.4 Discourse Connectives

D-LTAG recognizes both explicit and implicit discourse connectives. Explicit connectives are lexical items that represent discourse relations. They come from the syntactic groups of subordinating conjunctions and subordinators, lexico-syntactic anchors of parallel constructions, coordinating conjunctions, sentence modifying adverbs and some specific verb forms (Webber, 2004). Implicit connectives, on the other hand, are inserted by the annotator to represent the relations between discourse units which are not explicitly stated by a connective, but inferred by the reader (Prasad et al., 2007).

Discourse connectives can take their arguments in two ways. Some connectives like coordinators, subordinators and anchors of parallel construction, take both arguments syntactically. These connectives are referred to as structural connectives. Some other connectives such as discourse adverbials, take only one of their arguments syntactically. They retrieve the other argument anaphorically from the previous discourse. These connectives are referred to as anaphoric connectives (Cresswell et al., 2002; Miltsakaki, Cresswell, Forbes, Joshi, & Webber, 2003; Webber & Joshi,

1998; Webber, Stone, Joshi, & Knott, 2003). The discourse-syntactic trees for each type of connective and examples will be discussed in the following subsection.

Taking D-LTAG and the PDTB as starting points, the project that embarks on discourse level annotation on the MTC regards discourse connectives as anchors of discourse relations.

3.4.1   Explicit Connectives

3.4.1.1   Coordinating Conjunctions

Coordinators relate two abstract objects that are syntactically equivalent, i.e., neither is a subordinate of the other. In D-LTAG they take both arguments syntactically. Coordinator conjunctions such as the English *so* are represented by initial tress as in the figure (Webber, 2004; Webber et al., 2003).

Turkish coordinators, for example *çünkü* "because", exhibit similar structural properties; however, they are expected to have a larger syntactic tree set due to the flexible word order in Turkish. *Çünkü* will be discussed in detail later in chapter 5.

In addition to the initial tree for coordinating conjunctions, coordination is represented by auxiliary trees in case of simple coordination that conveys continuation (Forbes-Riley et al., 2006; Webber et al., 2003). In Figure 5 *and* takes both arguments structurally, one by substitution, one by adjoining.

The structural properties of the Turkish counterpart of *and*, *ve*, are yet to be explored.
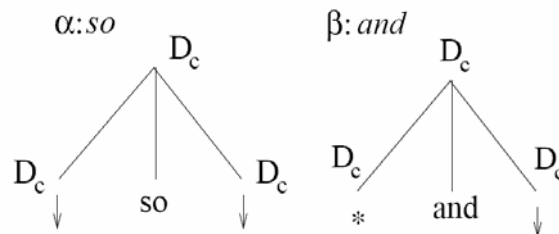
Figure 5 - Initial tree for the coordinate conjunction *so*, auxiliary tree for simple coordinator *and*

(Webber et al. 2003 p.31-32)

3.4.1.2 Subordinating Conjunctions

Subordinating conjunctions relate two arguments one of which is syntactically dependent to the other. Subordinators take both their arguments structurally (Webber, 2004; Webber, Knott, Stone et al., 1999). The trees in Figure 6 represent D-LTAG trees for English subordinating conjunctions. Since Turkish is a head-final language, another tree, α: subconj_post, should be employed to represent the post-posed subordinate clauses rather than (b) for pre-posed English subordinating conjunctions. Such trees are proposed for Turkish subordinators *için* "for", *dolayı* and *dolayısıyla* "because of" in chapter 5.
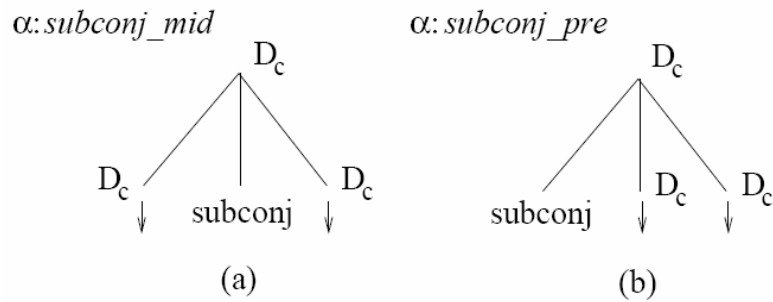
Figure 6 - Initial trees for subordinate conjunctions (Webber et al. 2003 p.29)

Subordinating conjunctions are only considered discourse connectives when both of their arguments are abstract objects, or reference items that can be interpreted as abstract objects[2]. It is quite common for a subordinator to take an anaphoric argument. Possible anaphoric arguments include, but are not restricted to, the deictic adjectives and adverbs (*bu* "this", *şu* "this, that", *o* "that, it", *böyle, şöyle, öyle* "such, so, thus") and adjectives of similarity and difference (*aynı*: "same", *benzer*: "similar", *farklı*: "different", *tersi*: "opposite", *aksi*: "opposite" lit. "reflection, mirror image, echo", *başka*: "another"). The examples of such constructions include *bunun için* "this-GEN for: for this, because of this", *aksi takdirde* "opposite-GEN case-LOC: in the

---

[2]   Turkish subordinators can take NP complements, resulting in prepositional phrases, such as *kedi için* "for the cat".

opposite case, otherwise", and *benzer sebeplerle*"similar reason-PLU-INS: with similar reasons, beacuse of similar resons".

It's possible to analyze such examples in various ways. One can take the anaphoric argument and the subordinator to be separate items in discourse structure. In this view, the anaphoric argument referring to a discourse unit is annotated as Arg2 because it is the morpho-syntactically more closely linked to the connective. The discourse unit referred to by the anaphoric element is annotated as Supp2, and the other argument, which is in fact syntactically modified by the adverbial, is annotated as Arg1.

(1) [SUPPSu alan teknem dibe vurdu sonunda.] **Onun için** *gidiyorum*.

   [SUPPMy leaking boat has finally hit the bottom.] **That is <u>why</u>** *I am leaving.*

Another way to think of these examples is to take the anaphoric element and the connective as a unit, and annotate both of them together as the connective. This discourse item behaves as a discourse adverbial, and retrieves its Arg1 anaphorically. In other words, the abstract object that resolves the anaphora of the reference item is Arg 1. Then the sentence modified by the discourse adverbial, and more if the argument is larger, is annotated as Arg2

(2) *Su alan teknem dibe vurdu sonunda*. **<u>Onun için</u> gidiyorum**.

   My leaking boat has finally hit the bottom. **<u>That is why</u> I am leaving**.

Still another way to analyze these constructions is to annotate arguments as in the adverbial approach but include the subordinator and the anaphoric expression in Arg2. Nothing is annotated as connective, and the relation type is noted as AltLex, meaning that there is an alternative lexicalization that indicates the discourse relation. For a corpus study this approach makes more sense than the adverbial approach because it limits the number of discourse connectives to a sensible number by excluding a compositional, therefore theoretically infinite (Forbes, 2003; Knott, 1996), set of connectives from the explicit connectives list.

(3) *Su alan teknem dibe vurdu sonunda.* (IMPLICIT: AltLex) **Onun için gidiyorum**.

My leaking boat has finally hit the bottom. (IMPLICIT: AltLex) **That is why I am leaving.**

In this thesis, the emphasis is on the discourse connectives rather than the discourse relations. Therefore the AltLex approach which does not single out a connective is not practical. Among the remaining alternatives, the supplement-free adverbial approach will be used for the sake of simplicity. This decision was made for two reasons: the first one is to keep the already complex annotations which represent both discourse components and information structure components from getting further complicated, and the second is to free the reader from the cumbersome task of searching for the anaphora resolving supplement for every such example. Personal experience shows that it is easier to interpret an example at a glance when the annotations are kept simpler.

3.4.1.3 Parallel Constructions

The syntactic items that anchor discourse relations are not always continuous. For example, there are parallel constructions in English such as on *one hand ... on the other hand, either ... or, neither ... nor*, and *not only ... but also*. Such parallel constructions are represented with initial trees (Webber, 2004; Webber et al., 2003). They take both their arguments through substitution.
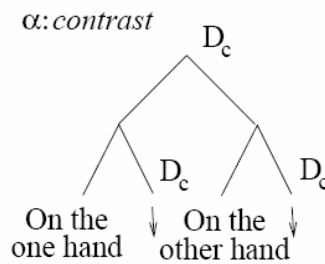
Figure 7 - Initial tree for a parallel construction (Webber, 2003 p. 30)

The Turkish equivalent of English parallel constructions is an area to be researched.

3.4.1.4 Discourse Adverbials

A group of adverbials relate two abstract objects and contribute to the discourse structure. This group of discourse connectives has been called conjunctive adjuncts, discourse adjuncts, and discourse adverbials throughout the literature (Forbes, 2003; Halliday & Hasan, 1976; Knott, 1996; Webber & Joshi, 1998). Where structural versus non-structural distinctions were made, these connectives have been called presuppositional or anaphoric (Cresswell et al., 2002; Forbes, 2003; Forbes & Webber, 2002; Miltsakaki et al., 2003; Webber & Joshi, 1998; Webber et al., 2003). Throughout this thesis, they will be referred to as discourse adverbials. This term is intended to differentiate them from sentential adverbials, which take only one argument which is the clause they syntactically belong to (Forbes, 2003).

Discourse adverbials differ from other connectives significantly. While other connectives take both of their argument structurally, discourse adverbials take only one of their arguments structurally. They retrieve the other argument anaphorically from the previous (or following, if cataphora is involved) discourse. The anaphorically retrieved argument is not represented on the D-LTAG tree, because the process is taken to be resolved by means that are extraneous to syntax.
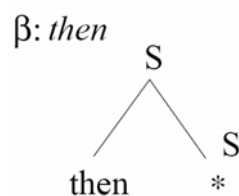
Figure 8 - Auxiliary tree for the discourse adverbial *then* (Webber, 2003 p. 32)

Discourse adverbials such as *yine de* "then again" and *tersine* "conversely, the opposite" has been attested in Turkish discourse, too (Zeyrek et al., 2008). *Yine de* presupposes an antecedent by means of its lexical meaning and *tersine* presupposes an antecedent by virtue of its morphology. The possessive case on *tersine* "opposite-POSS-DAT" looks for a genitive cased partner to complete the noun-noun complement as in *karısının tersine* "wife-POSS-GEN opposite-POSS-DAT: contrary to his wife".

Zeyrek et al. (2008), proposes that the anaphoric properties of discourse adverbials hold for Turkish, as well.

3.4.2 Implicit Connectives

D-LTAG allows for implicit connectives which do not appear in the written discourse material. The relations between adjacent sentences that are not connected by a discourse connective are assumed to be anchored by punctuation. Implicit relations are represented by auxiliary trees. They take one of their arguments by substitution and the other by adjoining (Cresswell et al., 2005; Webber et al., 2003).

β: *punct1*

D_c

D_c * · D_c ↓
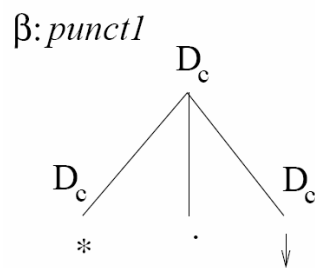
Figure 9 - Auxiliary tree for implicit relation anchored by punctuation (Webber, 2003 p.32)

When annotating an implicit discourse relation, the PDTB group tries to insert an explicit connective that could express the relation between adjacent sentences. When they can, the relation is labeled with the inserted connective, for example an implicit *because*. When they cannot, they look for alternative ways to express the discourse relation. If the insertion of an explicit connective is blocked by the existence of some expression that is not a discourse connective, or a connective that was not practical to annotate as discussed for Turkish subordinators, the implicit relation is labeled AltLex, which stands for alternative lexicalization. If the relation between the sentences is not expressed by any lexical expression, but exists because the two sentences talk about the same thing, the relation is labeled EntRel, which means there is an entity relation between the sentences. If none of these options apply, if there seems to be no relation between two sentences, the implicit relation is labeled NoRel, i.e., no relation.

CHAPTER 4

INFORMATION STRUCTURE AND TURKISH WORD ORDER

4.1   Information Structure

Information structure refers to the information exchange throughout a discourse. For an utterance that belongs to an existing discourse, some material has already been established by the previous context and there is some new material being conveyed about it. The old, contextually given information in the utterance is called the theme (topic), and the newly introduced information about the theme is called the rheme (focus) (Halliday & Matthiessen, 2004). See the following constructed example for a very simple demonstration of topic and focus:

(4) a. Which movie did you see yesterday?

    b. [$_T$ Yesterday I saw] [$_R$ **the Illusionist.**]

In (1b), the information already established by (1a), *Yesterday I saw* is the theme, and the new information conveyed, *the Illusionist* is the rheme.

There is more to information units than just what is old and what is new. The question in (1a) does not just establish a theme for the participant and time, but also creates a set of answers, consisting of movies, among which the given answer was selected. Similarly in the next constructed example, the question mentions two participants.

(5) a. So, you saw the Illusionist yesterday. Which movie did your sister see?

b. [$_T$ [$_{T-K}$ My sister] saw] [$_R$ **the Phantom of the Opera.**]

The question in (2a) builds a set of possible themes, {*I, my sister*} and in (2b) *my sister* is selected among a set of possible themes for (2b). This smaller partition of the information unit, which was distinguished among its alternatives is the kontrast of the unit it belongs to (Steedman, 2000a). In the above example, *my sister* is the theme kontrast. Both theme and rheme information units can have kontrast. The complementary of kontrast is background.

The main hypotheses of information structure account by Komagata (1999) are as follows:

a. "The theme is necessarily contextually-linked.
b. The rheme is not necessarily contextually-linked.
c. The theme is not necessarily contrastive.
d. The rheme is necessarily contrastive.
e. A proposition is a semantic composition of a Theme and a Rheme."[3]

(Komagata, 1999, p. 55)

Languages have been attributed different means for how they express information structure throughout the literature, such as word order, prosody and other grammatical means such as particles (as in Japanese).

---

[3]   Note that the existence of a Theme is not obligatory.

Though this thesis will be constructed over written data, it is informative to mention the prosodic manifestation of information structure, and how syntax and prosody interact. The main reason to include prosody-related information structure account is the strong link between intonation and word order in Turkish. Word order has effect on both the position of the discourse connective in the argument, and the order of the arguments when both arguments are in the same sentence, for example in the case of subordinators.

When there is no prosodic information, in order to show the information units of an utterance, the information units will be bracketed, as in the above examples with labels indicating the information structure (IS) status of the bracketed part.

Theme: T	Rheme: R
Theme background: T-B	Rheme background: R-B
Theme kontrast: T-K	Rheme kontrast: R-K

In some cases, there may be more than one information structure in one example. In such cases, the IS partitions are indexed. For example, T1 is the theme of the first information structure and T1-K is the kontrast in the theme of the first information structure.

The prosodic information and the information structure of a sentence can be integrated into its syntax. Steedman (1991, 2000a, 2000b) provides a derivation scheme which takes lexical items and their prosodic information and derives syntax, semantics and information structure simultaneously. Before presenting his syntax and information structure interface, a short introduction to Combinatory Categorial Grammar (CCG) will be presented.

4.2 Combinatory Categorial Grammar

Categorial grammars, Combinatory Categorial Grammar (CCG) in particular, provide a semantic interpretation for syntax (Steedman, 1995). The lexicon consists of categories, which include syntactic and semantic information. Each form is

associated with a syntactic type which sums up the place of the item in a predicate argument structure, and a lambda term that provides the semantics in logical form.
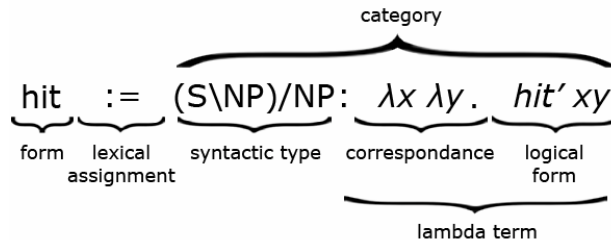
Figure 10 - A category in the lexicon.

The slashes indicate directionality.

The categories are combined to form constituents through universal combinatories. These combinatories include:

1. Forward Application

    X/Y   Y $\rightarrow$ X

2. Backward Application

    Y   X\Y $\rightarrow$ X

3. Composition (B)

    X/Y   Y/Z $\rightarrow$ X/Z

4. Type Raising (T)

    X $\rightarrow$ T/(X\T) (Forward Type Raising)

    X $\rightarrow$ T\(X/T) (Backward Type Raising)

A sample derivation with CCG is as follows:

(6) John   hit   Mary

    NP: *john'*   (S\NP)/NP: *λxλy. hit' xy*   NP: *mary'*
    ——→
                        S\NP: *λy. hit' mary' y*
    ←——
    S: *hit' mary' john'*
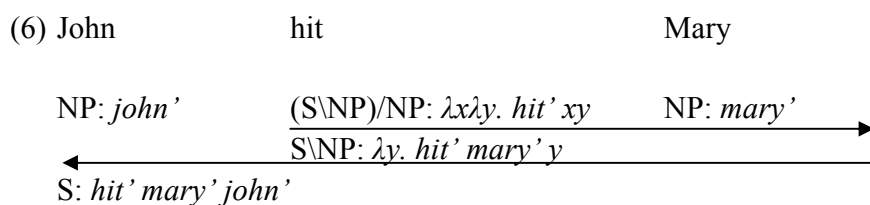
CCG allows type-raising, i.e., assigning higher order functions to categories. By means of type-raised categories, CCG is able to build non-traditional constituents.

Therefore the explanatory power of CCG extends beyond English to cover a broader range of languages as compared to other grammars (Komagata, 1997). In addition, assigning higher order functions to categories allows for an incremental derivation of the same sentence:
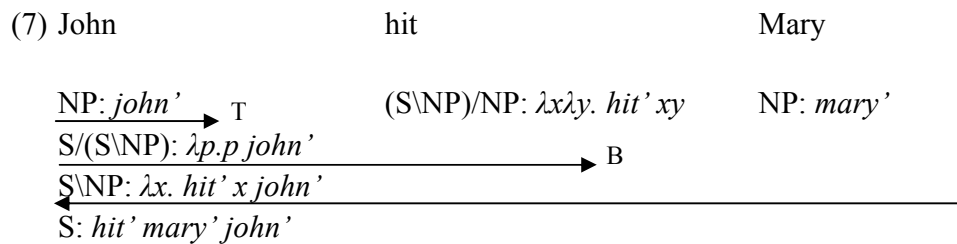
(7) John   hit   Mary

$\underrightarrow{\text{NP: } john'}$ T
$\overline{\text{S/(S\backslash NP): } \lambda p.p\ john'}$
$\underline{\text{S\backslash NP: } \lambda x.\ hit'\ x\ john'}$
$\text{S: } hit'\ mary'\ john'$

(S\NP)/NP: $\lambda x \lambda y.\ hit'\ xy$ NP: *mary'*

B

By means of type-raised categories, CCG is able to build non-traditional constituents, such as [John hit], providing ways to coordinate unconventional constituents, such as [John hit] and [Mark comforted] as in (5).

(8) John   hit   and   Mark   comforted   Mary

$\underrightarrow{\text{NP}}$ T   (S\NP)/NP   (X*X)/*X $\underrightarrow{\text{NP}}$ T   (S\NP)/NP   NP
$\overline{\text{S/(S\backslash NP)}}$ $\overline{\text{S/(S\backslash NP)}}$
$\overline{\text{S\backslash NP}}$ $\overline{\text{S\backslash NP}}$

                    (S\NP)*(S\NP)

$\overline{\text{S\backslash NP}}$
S
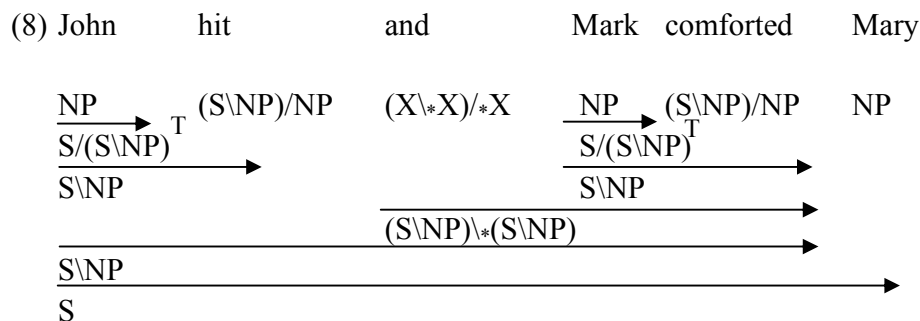
Though traditionally CCG takes words as lexemes, it is possible to apply the same approach at the morpheme level, syntactically accounting for the productivity of inflectional morphemes. A combinatory morphemic lexicon that allows CCG to process at a morphemic level can model the processing of English plural, Turkish case marking, subordination, control, relativization, possessives and syntactic compounds, without inefficiency or overgeneralization (Bozşahin, 2002). Though this account will provide an invaluable tool for Turkish subordinators which are composed of suffixes or which take arguments with distinctive morphological

properties, in this thesis, Turkish subordinators will be analyzed at word level for simplicity.

Bozşahin (2002) provides another useful asset for CCG: a lexical rule of contraposition.

(9) a. NP   →   S/(S/NP $_{+topic}$)   (< T $_x$)

    b. NP   →   S\(S\NP $_{-topic}$)   (> T $_x$)

The lexical rule in (a) allows left displacement of an item whereas the rule in (b) allows right displacement. A revised rule of contraposition with information structure components will be presented in the following

4.3 Prosody and Information Structure

The English language associates certain intonation contours with information units (Pierrehumbert, 1980; Steedman, 2000a). Theme tunes in English are θ-markers: L+H*, L*+H, and rheme tunes are ρ-markers: H*, L*, H*+L, H+L*.[4] Intermediate boundaries are marked with -, and the utterance final boundaries are marked with %. The intermediate boundaries indicate intermediate phrases. Intermediate phrases build intonational phrases (Pierrehumbert & Hirschberg, 1990).

CCG provides a phonology-syntax interface that combines prosody with syntax (Steedman, 1991, 2000a, 2000b). The tune is marked in all parts of the category.

(10)   hit:= $(S_\rho \backslash NP_\rho) / NP_\rho : \lambda x \lambda y. *hit' \, xy$

     H*

Unaccented forms are marked by a null tune, η.

(11)   hit:= $(S_\eta \backslash NP_\eta) / NP_\eta : \lambda x \lambda y. \, hit' \, xy$

---

[4]   θ and ρ denote theme and rheme, respectively. H denotes a relatively high pitch and L denotes a relatively low pitch.

Throughout the derivation, the accent projects over the result[5].

(12)   John   hit   Mary

                               L+H*   LH%

$S/(S\backslash NP)$: $\lambda p.p\ john'$ $(S_\theta\backslash NP_\theta)/NP_\theta$: $\lambda x\lambda y.\ ^*hit'\ xy$ NP: *mary'*

$\overline{S_\theta\backslash NP_\theta}$: $\lambda x.\ ^*hit'\ x\ john'$ B

The projection of tunes is restricted by phrasal tones (9), and intonational phrase by boundary tones (10).

(13)   L, H := $S\$_\iota \backslash S\$_\eta$: $\lambda f.\ \eta' f$

(14)   L% := $(S\$_\Phi \backslash S\$_\eta) \backslash (S\$_\iota \backslash S\$_\eta)$: $\lambda f\lambda g.[S]\ (fg)$

       H% := $(S\$_\Phi \backslash S\$_\eta) \backslash (S\$_\iota \backslash S\$_\eta)$: $\lambda f\lambda g.[H]\ (fg)$

Variable η stands for theme/rheme marking θ', ρ'. ι will no longer unify with η, θ or ρ and Φ unifies only with itself and ι, thus combines only with intermediate or complete intonational phrases.

As a result, the prosodically realized information units are projected onto syntax, and therefore semantics. In accord with the spirit of the CCG, information structure is derived hand in hand with syntax and semantics. A single derivation accounts for all information an utterance provides.

4.4   Turkish Word Order and Information Structure

Traditionally, Turkish information structure has been regarded as word order based. Certain pragmatic functions have been associated with certain positions relative to the main predicate of the sentence; topic has been associated with sentence initial position, focus has been associated with pre-predicate elements, and background and afterthought has been associated with post-predicate position (Erguvanlı, 1979).

---

[5]   Variable η is taken to be the default value, and is not displayed in the derivation for reasons of simplicity.

Erguvanlı uses the term background to indicate material that can be predicted or recovered from the previous discourse, and material that does not need to be contrasted. Afterthought is different from backgrounding in that it does not necessarily share the predictability or recoverability. The sole reason it is in the post predicate position is because the speaker remembered it only after the sentence was uttered.

Erguvanlı (1979) sees stress as a different tool for expressing information structure. She proposes that the immediately pre-predicate position receives a neutral stress and any pre-predicate material can receive empathic stress. Since in her view word order and stress are distinct strategies to mark information structure, they can contradict, which results in ungrammatical constructions. For example she predicts that it should be ungrammatical to stress the verb in a marked, i.e., non-canonical/non-SOV order, sentence. Özge (2003) disagrees, claiming that her ungrammatical examples become grammatical in appropriate context.

Other information structure approaches, such as topic-comment where comment has been further partitioned into focus and ground (Vallduvi, 1990), have been mapped onto Turkish word order, too (Hoffman, 1995). Hoffman associates topic with only sentence-initial position. Sentence initial material is not necessarily topic all the time, but if there is a topic, it needs to be sentence initial. The focus is again immediately pre-predicate and the post-predicate items are ground material.

Some researchers (İşsever, 2003; Kılıçaslan, 1994, 2004; Vallduvi & Engdahl, 1996) have argued that the information structure of Turkish makes use of both syntactic and prosodic strategies. According to İşsever, the two strategies are used for two different types of foci, presentational (p-focus) and contrastive (c-focus). C-focus is a position for material from a set defined by the context. C-focus is not restricted to a sentence position and is marked prosodically. P-focus, on the other hand, needs to be in the immediate pre-predicate position.

Kılıçaslan (1994) argues that both syntactic and prosodic strategies for focusing are available in Turkish, but the syntactic approach is preferred. Later (Kılıçaslan, 2004)

he drops his syntactic strategy argument, but he argues that elements "may" undergo syntactic operations to fulfill the "informational requirements" of the sentence.

Özge (2003) proposes a tune based approach to Turkish information structure, and demonstrates how prosodic restrictions can motivate the word order of an utterance. Özge states that the rheme on Turkish is marked by H* LL% contour, regardless of the position it occupies in an utterance. The thematic kontrast in Turkish is marked by L* H-, H*+L H-, or L+H* L- contour. Thematic background is deaccented. Özge proposes the restrictions laid upon the word order by prosodic features as:

- "Thematic-contrast must come before rheme as rheme contour causes flooring, rendering the announcement of thematic-kontrast impossible to its right.

- Rheme should come before the main functor, as announcing a rheme contour is impossible after the main functor again due to flooring."

(Özge, 2003, p. 79)

In (12), a, b, and c convey different information structures, though the word order is the same.

(15) a. Maymun   elma-yı   ye-di.

       L+H* L-   H*   LL%

      b. Maymun   elma-yı   ye-di.

       H*+L H-   H*   LL%

      c. Maymun   elma-yı   ye-di.

       H* L-L%   <   -F-   >

Bozşahin and Özge propose a radically lexicalist approach for Turkish intonation, information structure and word order. Similar to Steedman (2000), they propose that

the syntax and intonation-based information structure go hand in hand, and the word order is a result of constraint interaction between multiple facets of the derivation of an utterance. In this approach, surface positions are not predicative over information structure.

Bozşahin and Özge characterize two phrasal tunes;

(16) "Phrasal Tunes:

 a. A tune: A nonflat tune ending with the L-L% boundary.

 b. B tune: A nonflat tune not ending with an L%."

 (Özge & Bozşahin, 2008)

In addition to the A and B tunes, they define an F domain, which is for the unaccented domains after the rheme or the main functor.

(17) "F domain: A low and flat prosodic domain."

 (Özge & Bozşahin, 2008)

The Turkish intonation patterns are given as below:

BA: topic-comment sentence with declarative mood

BB, AB: Incomplete utterances that indicate more material is coming

AF: The tune is destructured (Büring, 1997) after the focus.

AA: two predicates, ungrammatical unless interpreted as two clauses

Bozşahin and Özge state that with verb-final constructions, Turkish has a flexible intonation, but with right-displaced constructions, the intonation is not as plastic. They propose that this might be because of predication, and the closely related rheme which is not allowed to the right of the verbal complex, rather than any phonological notion, such as stress. Following (Göksel & Özsoy, 2000), Bozşahin and Özge advocate that rheme is exclusively prosodically determined by an H*L- contour.

(18)   "Turkish rheme realization (material between vertical lines represents rheme):

   a.... |XV|...

   b.... |X|V...

   c....|X|... V...

   d.... |V|...

   e. *... V... |X|...."

(Özge & Bozşahin, 2008)

Theme is expressed by L*H-, with H*+L H- and L+H* L- marking thematic kontrast as in (Özge, 2003). The former gives partial information and the latter is corrective.

(19)   "Information structure, surface order, tunes and boundary tones in Turkish grammar:

a.   Rheme is signaled by the H*L- contour.

b.   Thematic-kontrast is signaled by B-tuned phrases. L*H- is the theme contour.

c.   The final tune in a complete declarative sentence is A, followed only by an F.

d.   Lexical categories of boundary tones: $L- := S\$_\iota \backslash S\$_\rho : \lambda f.\rho'f$

$H- := S\$_\iota \backslash S\$_\theta : \lambda f.\theta'f$

$L\% := S\$_{\Phi,\beta} \backslash S\$_\iota : \lambda f.\Phi'f$

e.   Categories of pitch accents:

   H* decorates the item in the string with ρ (rheme) feature.

   L* decorates the item in the string with θ (theme) feature.

f.   The revised lexical rule of rightward NP contraposition: $NP_\delta \quad S_\beta \backslash (S_\beta \backslash NP_\beta) (> T_\times)$"

(Özge & Bozşahin, 2008)

Notice the β indices on the right displaced items. With a combination of syntactic types of boundary tunes, contraposition rule and the directionality, the restrictions on

the accent to the right of the verb is accounted for without constraining information structure on the surface positions.

The information structure sensitivity of right displacement is attested in other frameworks, too (Temürcü, 2001, 2005). Temürcü states in his Minimalist Program[6] based research that rightward displacement of a subject through extraposition (EP) is a discourse-sensitive operation[7]. He also suggests that for reasons of economy, the rightward displacement of object should be handled the same way rather than proposing a new rule.

The rightward displacement is a syntactic phenomenon that results in a distinctive information structure configuration. The right displaced items in Turkish will be backgrounded because they will end up in the F domain, the flat, unaccented domain that cannot receive stress because of the flooring effect of the main predicate and the rheme. The contraposition rule is defined over noun phrases for the time being, but the information structure status of other right displaced items will uncover whether they utilize a similar rule or not.

---

[6]   (Chomsky, 1995 and following work)

[7]   This operation takes into consideration the stress status of constituents. The stressed constituents do not go through EP.

37

CHAPTER 5

CONNECTIVE AND ARGUMENT POSITIONING

Whether the word order has a pragmatic function and determines the information structure, or the word order is determined by the intended information structure by means of prosodic constraints, these two are undeniably related. In other words, the order of constituents in Turkish utterances reflects some part of the information structure in writing. Unfortunately, the prosodic information, which would have provided a measurable property to rely on when determining information structure, is lost when an utterance is written. When working with the written text, one is left with one prominent clue to Turkish information structure and that is the word order. Likewise, when one is looking into the interaction between information structure and of discourse in written text, the most easily accessed clue is the word order and the relative positioning of the discourse connective and its arguments.

Having the previous discourse at hand, it is possible to analyze how information structure partitions are distributed over an utterance. In this study such analysis are done following the theme-rheme distinction based on the novelty of the information (Halliday & Matthiessen, 2004) and kontrast-background distinction based on how "interesting" the part of information unit is (Steedman, 2000). As explained in

Chapter 4, the information unit that conveys old, already known information is marked as the theme and the information unit that conveys new information is marked as the rheme. Kontrast of each unit is again decided based in the discourse, and when there is ambiguity, the intuition of the native speaker was held to be the most possible information structure.

Disregarding the lexical properties of any one particular discourse connective, discourse connectives in general can precede, intervene in or follow the second argument (Conn-Arg2: "Nevertheless, I went early", Arg2-Conn-Arg2: "I nevertheless went early", Arg2-Conn: "I went early nevertheless"). Because of the syntactic definition of the second argument, i.e., because it is the argument that syntactically hosts the connective, the connective should not be able to intervene in the first argument. The second argument, in turn, can either precede or follow the first argument, or in the case of subordinators and parenthetical constructions, intervene in the first argument (Arg2-Arg1: "Because I wanted a front seat, I went early", Arg1-Arg2: "I went early because I wanted a front seat", Arg1-Arg2-Arg1: "I, because I wanted a front seat, went early"). Up to this date, there has not been a case where the first argument is attested as intervening in the second argument, but these orders will be kept in the inventory as possible orders.

As a result, there are thirteen possible arrangements for the connective and its arguments.[8] We will refer to this orderings as Connective Argument Orders (CAO). The list of possible CAOs is given in Table 1. The leftmost coloumn shows the order of arguments for a group of CAOs. Next coloumn assigns each CAO a number that will be used to refer that CAO throughout this thesis. Next, the configuration of the arguments and the connective is given. The dashes indicate the syntactic association of the connective. When a connective or an argument intervenes in another argument, the interrupted argument is displayed twice, once before the intervening element(s),

---

[8]   This list of possible arrangements excludes the possibilities for parallel connectives such as "Not only … but also" and "either … or" in English and "Ya … ya" (Either … or) and "Hem … hem" (both … and) in Turkish. Such constructions, interesting as they are, are out of the scope of this thesis.

and once more after the intervening element(s). Finally, the rightmost column indicates the position of the connective relative to Arg2 (i: initial, m: medial, f: final)

Table 1 - Possible Connective Argument Orders

Arg1-Arg2	CAO-1	Arg1		Conn-		Arg2	i
	CAO-2	Arg1		Arg2	-Conn-	Arg2	m
	CAO-3	Arg1		Arg2		-Conn	f
Arg2-Arg1	CAO-4	Conn-		Arg2		Arg1	i
	CAO-5	Arg2	-Conn-	Arg2		Arg1	m
	CAO-6	Arg2		-Conn		Arg1	f
Arg1-Arg2-Arg1	CAO-7	Arg1	Conn-	Arg2		Arg1	i
	CAO-8	Arg1	Arg2	-Conn-	Arg2	Arg1	m
	CAO-9	Arg1		Arg2	-Conn	Arg1	f
Arg2-Arg1-Arg2	CAO-10	Conn-	Arg2	Arg1		Arg2	i
	CAO-11	Arg2	-Conn-	Arg1		Arg2	m
	CAO-12	Arg2		Arg1	-Conn-	Arg2	m
	CAO-13	Arg2	Arg1		Arg2	-Conn	f

This chapter attempts to shed some light on which of these thirteen possibilities are employed by Turkish discourse connectives, and what the connective-argument positioning reveals about the information structure of these constructions.

5.1 *Çünkü*

Çünkü, a structural discourse connective of Persian origin (Lewis, 1967), is one of the most common causal connectives in Turkish. This connective displays a very rigid argument order. In almost all examples in the MTC, Arg1 precedes Arg2. The only exceptions to this ordering are parenthetical expressions, which will be discused below.

Çünkü expresses that its Arg2 is the cause of its Arg1.

(20)   [T1*Ortada*] [R1*hiçbir ipucu yok.*] **[T2Cünkü]** [R2**öldürülen yok.**]

  Around   any   clue absent Because   kill-PASS-REL absent

  [R1*There aren't any clues*] [T1*around.*] [T2**Because]** [R2**nobody was killed.**]

The Turkish information structure is mapped directly on the English translation for the reader to be able to keep up with the Turkish analyses. The realization of a similar utterance with the same information structure might be quite different in naturally generated English.

In example (20), *çünkü* connects two finite sentences. Arg1 is composed of a theme and a rheme. Arg2 introduces previously unknown information. The information status of the connective seems to be ambiguous: it can be a part of the rheme, or it can carry a theme contour on its own. In such ambiguous cases, information units are marked based on native speaker intuition.

Several researchers (Halliday & Matthiessen, 2004; Özge & Bozşahin, 2008; Steedman, 2000; Vallduvi, 1990) defined the theme or focus as the part that links the current utterance to the previous discourse. This definition seems closely related to the definition of a discourse connective. As a result, without intonation information one is tempted to include the discourse connectives in the theme. In this study instead of giving into this temptation, the information provided by the discourse and the native speaker intuitions about the possible intonations were relied on.

Lewis, (1967) observes the behaviour of *çünkü* (and *zira*) as follows: They "almost always begin a sentence but, like the English 'for', always explain the preceding statement."

Çünkü, indeed, always explains a preceding statement; however, it would be more accurate to say that the most common position for *çünkü* is clause-initial, rather than sentence initial. Examples like (21), where *çünkü* connects two finite clauses in the same sentence, occur quite frequently in the MTC.

(21)   [R1*Belki de o anda Tanrı'ya yakarıp yardım istiyordu*], [T2**çünkü**] [R2**çok dindar bir adamdı**.]

Maybe that moment-LOC God-DAT pray-GER help want-PROG-PAST because very religious a he.be-PAST

[R1*Maybe at that moment, he was praying to God and asking for help*,] [T2**because**] [R2**he was a very religious man.**]

Judging from the previous text, both pieces of information is new; as a result, neither clause can be labeled as theme. There are a few possibilities to consider while considering the information structure of such constructions. The first one is that these clauses belong to the same information structure partition, which is rheme. This seems unlikely because the comma between the clauses compel the reader to insert a phrase boundary. The second reason is that when *çünkü* connects two finite clauses in the same sentence, the information structure is similar to the cases where *çünkü* connects two finite sentences; the clauses have their independent information structures and *çünkü* tends to link the clauses as a theme or theme-kontrast unit.

In the following example (22), the arguments of *çünkü* are two nominalized clauses which are the complements of an attribution verb. The themes and the rhemes are labeled according to the preceding text.

(22)   [T1253 - 1255 yılları arasında Moğol hakanını, elçi olarak ziyaret eden Wilhelm von Rubruk,] [T1-K*Moğolların elbiselerini*] [R1*asla yıkamadıklarını,*] [T2**çünkü bunu yaparlarsa**] [R2**Tanrı'nın hiddetlenerek yıldırım ve gök gürültüsü göndereceğine inandıklarını,** belirtir.]

[T1Wilhelm von Rubruk, who visited the Mongol khan as an ambassador between the years 1253 – 1255, asserts] [T1-K*that the Mongols*] [R1*never wash their clothes,*] [T2**because**] [T2**they believe**] [R2**that the God will get furious and send lightning and thunder**] [T2**if they do that**.]

There seems to be two sets of themes and rhemes in this example. However, in both examples, the comma implies not only a phrase boundary, but an intermediate phrase boundary, i.e., the reader might interpret this example as having one main predicate, not two. However, the connective, together with a substitution expression (*bunu yaparlarsa*: "if they do that"), seems to belong to a theme unit, intervening in two rheme units, eliminating infroations structures with one continuous rheme unit. In addition the reader might read the sentence as to main predicates, because this is a long sentence with a heavy subject modified by a relative clause and two finite clauses.

Tshe third possibility is that a sentence can have only one information structure set, independent of the number of finite verbs. In this case, clauses with intermediate boundaries would pose no problem; they would join with the sentence as any theme partition. However the presence of new information in these clauses cannot be explained, unless the rheme is discontinuous. With the discontinuous rheme analysis, both the rheme-theme partitioning and the reader's phrasing intuitions are satisfied, but this is a possibility which would require an empirical study focused on complex sentences.

When the arguments, especially the second argument of *çünkü* consists of multiple clauses (23) or include other discourse connectives (24), the information structure is more complex than these examples, but can still accommodate the same analysis.

(23) [T1-K*Korkuyormuş*,] [T2-K**çünkü]** [T2**Müslüman ordularının başındaki komutan kimsenin aşamadığı surları aşamadığında**] [R2**vazgeçmiyor**] [R3**hep yeniden yeniden surları aşma isteği duyuyormuş.**]

[T1-K *(He was) afraid*,] [T2-K**because**] [T2**the commander leading the Muslim armies**] [R2**wouldn't give up**] [T2**when he cannot pass through the walls that no one can pass through,**] [R3**(but) he would yearn to pass through the walls again and again.**]

(24) Giyildiği ve iyice yıpranmadığı sürece elbiseleri yıkamak yasaktır. [R1*Burada bizce bir ifade bozukluğu veya çeviri yanlışı bahis konusu olabilir*,] [T2**çünkü elbiseler sanki**] [T2-K**giyildiği sürece ve yıpranmamışken**] [R2**yıkanamaz,**] [T3**fakat**] [T3-K**daha sonra**] [R3**yıkanabilirmiş gibi bir anlam taşımaktadır.**]

It is forbidden to wash the clothes as long as they are being used and they are not worn out. [R1*In our opinion, there may be a faulty expression or a mistranslation*,] [T2**because it bears the meaning as if the clothes**] [R2**cannot be washed**] [T2-K**as long as they are being used and they are not worn out,**] [T3**but**] [R3**they can be washed**] [T3-K**afterwards.**]

In (23), the connective is not only a part of the theme, but it is marked as theme kontrast. It may be the case that in this example, the relation itself (i.e., reason) is as as prominent as the arguments of the relation. In contrast, in (24), the connective is a part of the theme with other elements, whereas the theme-kontrast is the part which is compared to the theme-kontrast of the next clause. Hoewever, it seems possible to read the connective in (23) as a theme-kontrast, and the connective in (24) as a theme. Changing the intonation does not seem to change the information conveyed by either sentence. There is a slight change in how the information is conveyed, but the difference does not seem vital to the overall information structure of the sentences.

In addition to occupying the clause-initial, and by extension argument-initial position, the connective can also be located argument-finally.

(25)   Kadın ona bir şarkı yazmış. [_{T1}*Adı:*] [_{R1}*Deli Mavi*]. [_{R2}**Adamın gözleri maviymiş**] [_B**<u>çünkü</u>**.]

Woman he-DAT a song write-PAST name crazy blue man-gen eye-POSS blue.BE-PAST-PERF because

The woman wrote him a song. [_{T1}*Its name:*] [_{R1}*Deli Mavi (Crazy Blue)*]. [_B**<u>Because</u>**] [_{R2}**the man's eyes were blue.**]

When the discourse connective is in the post-predicate position, like any other element in the clause, it is backgrounded.

Although argument medial constructions sound acceptable to the native speaker with proper context or intonation, such *çünkü* examples are not freely attested in MTC (cf. *zira*). However, when *çünkü* is in the post-predicate position together with other elements of the sentence, it is possible to get argument-medial *çünkü*. Since the ordering of the backgrounded elements should not affect the information structure of an utterance, such cases are not expected to have different information structures. The connective, as in the argument-final constructions, is backgrounded.

(26)   [_{T1}*Divitin yanında biriken talaşları çekmecesinden çıkardığı lal rengi kadife bir keseye,*] [_{T1-K}*keseyi rahleye bitişik tutarak,*] [_{R1}*bir damlası bile yere düşmeyecek şekilde dolduruyor.*] [_{R2}**Bir gün cenazesini yıkamak için ısıtılacak suyun ateşinde yakılacak**] [_B**<u>çünkü</u> o talaşlar...**]

[_{T1}*He puts the sawdust that piles up near the inkwell into a garnet colored velvet pouch he took out of his drawer,*] [_{T1-K}*holding the pouch next to the book-rest,*] [_R*(he puts them) in such a way that not a single drop will fall on the floor.*] [_B**<u>Because</u> that sawdust**] [_{R2}**will be burned in the fire that will heat the water which will be warmed up to wash his corpse.**]

There are few exceptions to the rigid Arg1-Arg2 order of *çünkü*'s arguments. All exceptions occur where a parenthetical expression intrudes in the middle of a clause, to which it is linked by means of *çünkü*.

(27) [T1*Biz yasalar karşısında evli sayılacak,*] [T1-K*ama gerçekte evli iki insan gibi değil de*] ([R2**evlilikler sıradanlaşıyordu**] [B**çünkü**,] [R3tekdüze ve sıkıcıydı;] [R4biz farklı olacaktık]), [R1*aynı evi paylaşan iki öğrenci gibi yaşayacaktık.*]

[T1*We would be married under the law,*] [T1-K*but in reality*] [R*we would live like two students sharing the same house*] [T1-K*rather than two married people*] ([B**because**] [R2**marriages were getting ordinary,**] [R3(they were) monotonous and boring;] [R4we would be different]).

In this example (27), the connective is backgrounded as usual for the argument final position. The second argument may be extended up to the semicolon, to cover "they were monotonous and boring". This reading would be even more prominent if *tekdüze ve sıkıcıydı* is backgrounded. The resulting structure would be similar to the preceding example: an argument-medial *çünkü* in the post-predicative region of the clause which is backgrounded due to post-rheme flooring. It is also possible to have this example with argument-initial connective. Then the connective would constitute a theme or theme-kontrast partition.

It should be noted that not all parenthetical expressions linked to the main clause by *çünkü* are intervening arguments. Most of the time, the parenthetical explains why the immediately previous part of the utterance, not necessarily an abstract object, was verbalized as it is. Even in the above example, in addition to the annotation given, one may judge that the parenthetical gives the reason of *evli iki insan gibi değil* "rather than two married people – lit. not like two married people". Though this example is ambiguous, the context often clarifies when the parenthetical is pragmatically motivated.

Çünkü does not necessarily connect two finite arguments. As in the example below, its arguments may be relative caluses, too. In this case, the arguments are not expected to have independent information structures. Both *çünkü* and its argument(s) may belong to the same, larger information structure partitions.

(28)   [T1sürekli aynı doğrultuda yapılan,] [T1*yapılacağı baştan belli olan,*] [T1-K**çünkü huy haline getirilmiş** yanlışları] [R1istesem de unutamam]

[R1I cannot forget] [T1the mistakes that are persistently made in the same direction, *that it is obvious they will be made*, **because they have become habits**,] [R1even if I wanted to.]

In (28), *çünkü* belongs to the same information unit with its second argument and the syntactic head of the relative clause.

One should note the fact that all *çünkü* examples examined so far were in theme partitions does not necessarily mean that *çünkü* cannot be a part of the rheme. These examples are from written text, which is often constructed with an attempt at dramatizing the content. The sentences are often longer than ordinary daily speech and commas are used abundantly, resulting in intermediate boundaries with rising tones, hence an excess of theme partitions.

Since *çünkü* has a rigid argument order, there seems to be two D-LTAG trees for *çünkü*:
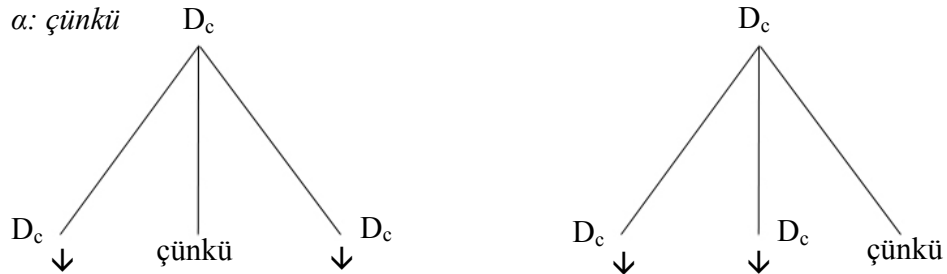
Figure 11 - D-LTAG trees for *çünkü*

Just like it is not possible to display an intervening connective with a simple D-LTAG tree, it is not possible to display an intervening argument with a simple tree structure, either. Thus, there will not be separate D-LTAG trees for intervening connective or argument configurations. The intervening argument is assumed to be the latter argument in tree representations. It should be noted that these trees are not adequate at representing the argument order and connective positions but they are

quite useful and practical for representing anchors, the type of operations necessary to take arguments, and the resulting hierarchy. For this purpose one tree for each connective (or each use of a connective, see *dolayısıyla*) would be sufficient. Nevertheless this study follows the D-LTAG literature (Webber, 2004; Webber & Joshi, 1998; Webber, Stone, Joshi, & Knott, 2003, and others) in representing different argument order and connective positions with different trees when possible for the time being. The argument orders and the position of the connective in the argument are represented in CAO profile tables separately. The CAO profiles will also display the information structural tendencies of the connectives for theme (T), rheme (R) and kontrast (K). It is highly probable that an information unit marked as theme can be included in theme-kontrast and vice versa and theme and rheme partitions may extend shrink in cases of narrow versus broad rheme readings. The subscript B, on the other hand, indicates that the information unit is always backgrounded due to phonological constraints on Turkish intonation system.

The CAO profile of *çünkü*, representing the possible connective-argument combinations is below:

Table 2 - The CAO profile for *çünkü*

Group	CAO						Res
Arg1-Arg2	CAO-1a	Arg1	$\textit{çünkü}_T$ -			Arg2	i
	CAO-1b	Arg1	$\textit{çünkü}_{T-K}$ -			Arg2	i
	CAO-2	Arg1		Arg2	- $\textit{çünkü}_B$ -	$Arg2_B$	m
	CAO-3	Arg1		Arg2		- $\textit{çünkü}_B$	f
Arg1-Arg2-Arg1	CAO-7a	Arg1	$\textit{çünkü}_T$ -	Arg2		Arg1	i
	CAO-7b	Arg1	$\textit{çünkü}_{T-K}$ -	Arg2		Arg1	i
	CAO-9	Arg1		Arg2	- $\textit{çünkü}_B$	Arg1	f

There are two instances of CAO-1 and CAO-7 because the connective has displayed more than one information structure feature in the analyses.

5.2 *Zira*

Like *çünkü, zira* is of Persian origin (Lewis, 1967) and has a rigid Arg1-Arg2 order. With the exception of parenthetical constructions where the second argument intervenes in the first, in all cases of *zira* in the MTC the first argument precedes the second.

(29) [T1*Bu kumpanya ziyaretinin*] [R1*bana pek faydası dokundu*]. [T2**Zira**] [R2**sıralarda aklıma zehir gibi bir fikir geldi.**]

This   company   visit-POSS-GEN me-DAT   very   be-beneficial-PAST

because   row-LOC   mind-DAT   poison like   an   idea come-PAST

[T1*This company visit*] [R1*has been very beneficial for me.*] [T2**Because**] [R2**an ingenious idea occurred to me among the rows.**]

In (29), *zira* connects two finite sentences. As expected, both sentences have their own independent information structures and *zira* links the second sentence to the previous one as the theme (or part of the theme depending on how broad the rheme is).

Like *çünkü*, it is quite common for *zira* to connect two finite clauses in the same sentence (30). The analysis for the cases where *çünkü* connects finite clauses is also true for the following example where two finite clauses in the sentence are connected by *zira*.

(30) [R1*Fırtınanın yaklaşmakta olduğunu sezinlediğim an kaçardım,*] [T2**zira bana göre dayak yemek**] [R2**çok alçaltıcı bir muamele sayılırdı**.]

[R1*I used to run away the moment I sensed the storm coming*], [T2**because I thought getting a beating**] [R2**to be a very humiliating treatment**.]

Zira can occupy post-predicate position (31), and can be a part of the parenthetical construction (32).

(31) [T1-K*Ne kadar geç büyümesi mümkünse,*] [R1*o kadar geç büyüsün.*] [T2**Büyümek**] [R2**zor ve acıtıcı**] [B<u>**zira**</u>.]

[R1*Let her grow up*] [T1-K*as late as possible.*] [B<u>**For**</u>] [T2**growing up**] [R2**is hard and painful**.]

(32) [T1-KŞimdi *ilkokul birden başlayarak*] - [T2-K<u>**zira** daha önce</u>] [R2**belli başlı bir sosyalleşme hayatım olmamıştı**] − [T1-K*düzenli olarak kavga etmiş* bir anne,] [R1bu durumda, nasıl tavsiyede bulunur, ne der ne eder ; söyleyebilir misiniz]

[T1-KNow,] [R1can you tell me what advice does] [T1-Ka mother *who has been fighting regularly since the first year of primary school*] − [T2-K<u>**because**</u>] [R2**I didn't have a major social life**] [T2-K**before that**] − [R1give, what does she say, what does she do?]

In the post predicate position, *zira* is backgrounded like *çünkü*, as well as any other element that is deaccented due to post rheme flooring. In the parenthetical intrusion, on the other hand, the connective is interpreted as theme-kontrast. The connective seems to be more prominent when it introduces a parenthetical expression than when it connects linearly sequenced clauses.

In this analysis, the mother's not having a major social life is taken to be the reason why she started fighting regularly at the first year of primary school. There is another possible analysis: The parenthetical might explain why the author mentions the first year of the primary school. In this latter - so called pragmatic sense – analysis, the Agr2 will not intervene in the Arg1.

The following example (33) derived from (26), by replacing *çünkü* with *zira*, is both grammatical and acceptable. As a result, the background connective and argument positioning is available to *zira*.

(33) [_{T1}*Divitin yanında biriken talaşları çekmecesinden çıkardığı lal rengi kadife bir keseye,*] [_{T1-K}*keseyi rahleye bitişik tutarak,*] [_{R1}*bir damlası bile yere düşmeyecek şekilde dolduruyor.*] [_{R2}**Bir gün cenazesini yıkamak için ısıtılacak suyun ateşinde yakılacak**] [_B<u>**zira o talaşlar...**</u>]

[_{T1}*He puts the sawdust that piles up near the inkwell into a garnet colored velvet pouch he took out of his drawer,*] [_{T1-K}*holding the pouch next to the book-rest,*] [_R*(he puts them) in such a way that not a single drop will fall on the floor.*] [_B<u>**Because**</u> **that sawdust**] [_{R2}**will be burned in the fire that will heat the water which will be warmed up to wash his corpse.**]

The information structure for (33) is, as expected, the same as that of (26). The connective is backgrounded together with *o talaşlar* and cannot be prominent in the information structure.

Lewis (1967) mentions that *zira* could replace *çünkü* in the examples he provided. So far, all the examples given in this thesis confirm this observation. However, the MTC examples present a point of diversion. Argument-medial *zira* is attested in a non-background position, whereas *çünkü* has no such examples.

(34) [_{R1}*Bunların burnunun dibinde sallayabileceğiniz sarmısak demeti de, mizah anlayışıdır.*] [_{T2}**Mizah**] [_{T2-K}<u>**zira**</u>] [_{R2}**siyaseten yanlışçılıktan ibarettir,**] [_B**böyle bir kelime yazı turalamak durumunda kalırsak.**]

[_{R1}*The garlic bunch that you can swing under their noses is sense of humor.*] [_{T2-K}**Because**] [_{T2}**humor**] [_{R2}**consists of political wrongdoerness;**] [_B**if we have to toss-up such a word.**][9]

As mentioned before, in spoken language argument-medial *çünkü* sounds acceptable to the native speaker. Yet it is a noticeable fact that eventhough *çünkü* occurs a lot

---

9   This word-play is a little unintelligible in Turkish, too.

more frequent than *zira* in the MTC, pre-verbal argument-medial *çünkü* is not attested, whereas such a *zira* example is present, albeit only once.

In (34), the connective is in a theme-kontrast unit. For this is the only example at hand, it will be recorded as kontrast in CAO profile, but with the caveat that because of the lack of argument medial *çünkü* and *zira* examples in the MTC, it cannot be judged whether the kontrast feature is a tendency or not.

Although there seems to be a slight difference, *çünkü* and *zira* are quite similar as Lewis (1967) mentioned. This slight difference in the possible argument-medial occurences cannot be represented in the current tree scheme. As a result, the D-LTAG trees for the *zira* are similar to those for *çünkü*.
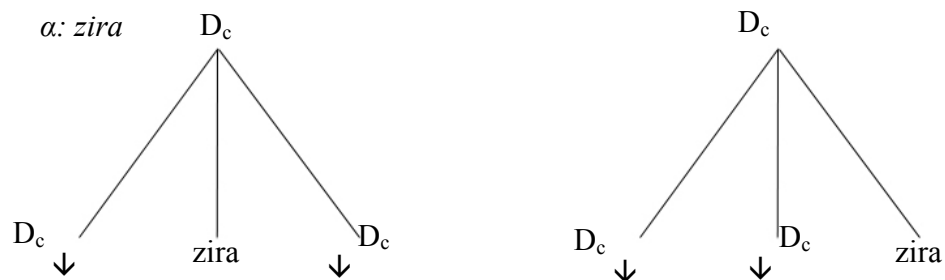
Figure 12 - D-LTAG trees for *zira*

The CAO profile of *zira* is below. 2.a shows the ordering not attested for *çünkü*.

Table 3 - The CAO profile for *zira*

Arg1-Arg2	CAO-1	Arg1		$zira_T$ -	Arg2	i
	CAO-2a	Arg1	Arg2	- $zira_{T\text{-}K}$ -	Arg2	m
	CAO-2b	Arg1	Arg2	- $zira_B$ -	$Arg2_B$	m
	CAO-3	Arg1	Arg2		- $zira_B$	f
Arg1-Arg2-Arg1	CAO-7	Arg1 $zira_{T\text{-}K}$ -	Arg2		Arg1	i
	CAO-9	Arg1	Arg2	- $zira_B$	Arg1	f

5.3 *İçin*

İçin is a causal particle that requires an NP or a nominalized VP to its immediate left. As a discourse subordinator, the sense of *için* depends on the morphological properties of its second argument. The syntactic head of the subordinate clause can be an infinitival clause (-*mek*), or a subordinate clause with a nominalized head. This nominalized head can be constructed with one of the non-factive (–*me*-AGR) or factive (–*dik*-AGR) [10] nominalizers. *İçin*'s complement can also be a deictic expression. Then the resolution of its meaning will depend on the resolution of the anaphora, which at times can be ambiguous. Individual cases will be analyzed in detail in the following subsections.

Lewis (1967) states that *için* is also used with third person imperative. The resulting form is -*sin için*, similar to –*sin diye* "so that" (as in *Amerika'nın derdi Saddam... O gitsin diye Irak Kürtlerini kullanmaya bakar.* "America's problem is with Saddam… (America) will use the Iraqi Kurds <u>so that</u> he will leave."). He mentions that this construction is uncommon, and presents one example from a piece of Turkish text, where Lewis judges the use of this unusual form to avoid double use of *diye*.

(35)   Dün Köprüden geçerken Fatih camiinin minaresinde bayrak çekildiğini gördüm. Yarın öbürgün bu âdet de yerleşirse Demokrat Parti zamanında yerleşmiştir diye tarih kitapları yazsın için, ben de buraya yazıyorum. (B. Felek)

'Yesterday while crossing the Bridge I saw that flags had been hoisted on the minaret of the Fatih Mosque. I am writing <this> here so that if this custom too takes root, tomorrow or the next day, the history-books may write that it took root in the Democrat Party era.'

(Lewis, 1967: p.288)

---

[10]   Following Kornfilt (1997), non-factive nominalizers will be glossed as action nominals (ANOM), and factive nominalizers will be glossed as factive nominals (FNOM).

Lewis does not give complete references for his examples but given the context, I judge the time of this unusual utterance to be 1950's. This construction may have become obsolete in time, or more probably, might be a slip of tongue – or pen in this case. It seems very probable that Felek intended to write either *–sin diye* or *–mesi için*, both of which are common ways of expressing purpose. Even if it was written on purpose, since this construction sounds strongly ungrammatical to the native speaker of the present day, and since it is not attested in the MTC, it will not be examined in detail.

5.3.1 Purposive *İçin*

When the syntactic head of the second argument of *için* is an infinite or non-factive nominalized VP, the second argument expresses the purpose of the first argument, i.e., Arg1 is desired to be realized so that Arg2 may come to be. In this thesis, in order to differentiate it from other uses of *için*, this particular use will be mentioned as purposive *için* from hereafter.

Purposive *için* takes an infinite complement (VP-*mek*, Inf) when the subject of the subordinate clause is co-referential with the subject of the matrix clause. Otherwise the head of the complement is non-factive, nominalized (VP-*me*, ANom) and agrees with the subject of the subordinate clause (Kornfilt, 1997).

(36)   [T-K**Yine bir yanlışlık yapmamak <u>için</u>**] [T-K*içeriye girip*] [R*hususi banyoların tam yerini öğrenmeliydi.*]

  [T-K**<u>In order not to</u> make a mistake again,**] [T-K*he had to go inside*] [R*and learn the exact location of the private baths.*]

(37)   [R1Sağlıksız bir topluluk içinde sağlıklı olmak rastlantıya bağlıdır.] [T2-KO halde] [R2*toplumu düzeltmek gereklidir,*] [B**kişileri iyi koruyabilmek <u>için</u>**.]

  [R1In an unhealthy community, being healthy is coincidental.] [T2-KThen] [R2*it is necessary to set the society right*] [B**<u>in order to</u> protect the individuals**.]

(38)   [T-K**Onu kolayca tanıyabilmen <u>için</u>**] [R*sana iki sır vereceğiz.*]

    [R*We will give you two secrets*] [T-K**<u>for</u> you to recognize it easily.**]

(39)   [R*Sağlam bir iskele yapmıştı*] [B**teknelerin rahatça yanaşabilmesi <u>için.</u>**]

    [R*He had built a sturdy pier*] [B**<u>for</u> the boats to dock easily**].

As (36-39) show above, both Arg1-Arg2 and Arg2-Arg1 orderings are possible with both an infinitival and a nominalized complement. When a purposive *için* clause occurs in the background position, all elements of the subordinate clause occur post-predicatively, and thus the whole clause is backgrounded (39). When the subordinate clause is in the canonical position, the examples indicate towards a tendency for theme-kontrast intonation contour. This theme tendency, which is based on native speaker intuition, is also consistent with the proposition that subordinating conjunctions in English present the knowledge as if it was given (Quirk, Greenbaum, Leech, & Svartvik, 1985). Unlike the coordinating conjunctions analyzed above, which can be an IS unit on their own in addition to belonging to a larger IS unit, the subordinators are always expected to belong to a larger IS unit rather constituting an IS unit on their own. More specifically the subordinator is expected to be in the same IS unit with its verbal complement because of the morpho-syntactic association.

Since *için* must follow its complement, it cannot be argument initial. But it is possible to have argument medial *için* by postposing a constituent of the second argument other than the syntactic head in addition to the connective. The following examples are derived from examples (38) and (39) respectively, and are acceptable.

(40)   [T-K**Kolayca tanıyabilmen <u>için</u> onu**] [R*sana iki sır vereceğiz.*]

    [R*We will give you two secrets*] [T-K**<u>for</u> you to recognize it easily.**]

(41)   [R*Sağlam bir iskele yapmıştı*] [B**rahatça yanaşabilmesi <u>için</u> teknelerin.**]

    [R*He had built a sturdy pier*] [B**<u>for</u> the boats to dock easily**].

As for the argument-final position of the connective, the argument medial connectives belong to a theme-kontrast unit when in the canonical position and backgrounded together with its clause when in the post-predicate position.

The subordinate clause can also intervene in the matrix clause, resulting in Arg1-Arg2-Arg1 ordering. This ordering is available for both the infinitival and the nominalized subordinate *için* clauses.

(42)   [T-K*Korku Çağında*] [T*hepimiz* **iktidarı ele geçirmek <u>için</u>**] [R*büyük bir Çeteye bağlı minik çetelerin üyesi olmak istemiyor muyduk?*]

   [T-K*In the Age of Fear*] [R*didn't we all want to be a member of the mini gangs connected to a big Gang*] [T<u>**in order to**</u> **seize the power?**]

(43)   [T*Kardeşim,*] [T-K**hayatımı ülkemdeki bir hapishanede tamamlamam <u>için</u>**] [R*elinden geleni yapıyor.*]

   [T*My brother*] [R*is doing his best*] [T-K<u>**for**</u> **me to live the rest of my life in a prison in my country**.]

There seems to be a slight difference between the information structures of two sentences, since in (42) *için* seems to belong to a theme unit, whereas in (43) it belongs to a theme-kontrast unit. Though these sentences are morpho-syntactically different, the difference of information structures is probably due to the construction of the sentence and their surrounding discourse rather than the morphologic features of the verbal complement. If the reason was morphology related, such differences would be observed for other argument orders, too. It is also possible to right-displace items from subordinate clauses such as *iktidarı* or *hayatımı* to get argument-medial connective in an intervening argument, with same information structure.

There are also examples of purposive *için* in rheme units. For example in *Bir şey anlatmak için gelmiştin buraya*" You came here to tell something", and *Evet, buraya bir şey anlatmak için gelmiştim* "Yes, I came here to tell something" the subordinate

clause is in the rheme unit in Arg2-Conn-Arg1 and Arg1-Arg2-Conn-Arg1 configurations respectively.

With the morphosyntactic restrictions on the order of the connective and the second argument, the following D-LTAG trees will only reflect the arrangements of the arguments:

α: p-için

D_c — $D_c \downarrow$ — için — $D_c \downarrow$

D_c — $D_c \downarrow$ — $D_c \downarrow$ — için
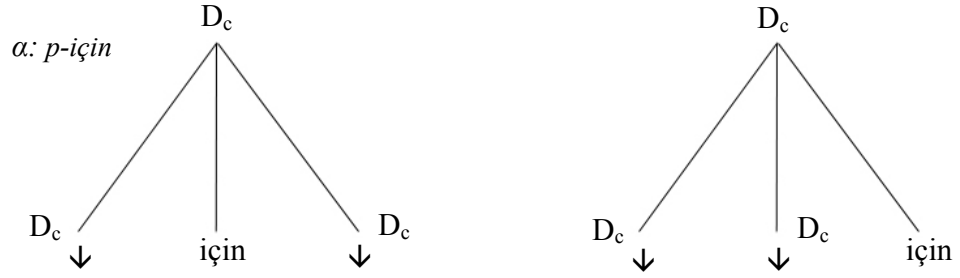
Figure 13 - D-LTAG trees for purposive subordinator *için*

The CAO profile of purposive *için* is given in table.

Table 4 - The CAO profile for purposive *için*

Arg1-Arg2	CAO-2	Arg1		$Arg2_B$	- $için_B$ -	$Arg2_B$	m
	CAO-3	Arg1		$Arg2_B$		- $için_B$	f
Arg2-Arg1	CAO-5	Arg2	- $için_{T-K}$ -	Arg2		Arg1	f
	CAO-6a	Arg2		- $için_{T-K}$		Arg1	f
	CAO-6b	Arg2		- $için_R$		Arg1	f
Arg1-Arg2-Arg1	CAO-8a	Arg1	Arg2	- $için_T$ -	Arg2	Arg1	m
	CAO-8b	Arg1	Arg2	- $için_{T-K}$ -	Arg2	Arg1	m
	CAO-9a	Arg1		Arg2	- $için_T$	Arg1	f
	CAO-9b	Arg1		Arg2	- $için_{T-K}$	Arg1	f
	CAO-9c	Arg1		Arg2	- $için_R$	Arg1	f

5.3.2 Reason-bearing *İçin*

When the syntactic head of the second argument of için is a factive, nominalized VP, the second argument expresses the reason of the first argument. This use of the connective will be referred to as reason-bearing *için*.

(44) [_T**Tüm gücünü kullandığı <u>için</u>**] [_R*ter içinde kalmıştı.*]

    [_T**Because he used all his power**] [_R*he was soaking with sweat.*]

(45) [_TBöyle durumlarda almak istediği yanıtları bilir,] [_Rona göre konuşurdum.] [_R*Biraz da sıkılırdım*] [_B**onu böyle kandırdığım <u>için.</u>**]

    [_TIn situations like these I knew the answer he wanted to get] [_Rand I talked accordingly.] [_R*I used to feel a little embarrassed*] [_T**for deceiving him like that.**]

(44) and (45) demonstrate that both Arg1-Arg2 and Arg2-Arg1 orders are available to reason-bearing *için*. The information structure is similar to that of purposive *için*. The subordinate clause in canonical position is interpreted as theme, theme-kontrast or rheme (it seems equally possible between these possibilities in this particular example (44)), and the right-displaced subordinate clause is backgrounded in (45). The connectives belong to the same information unit with their complement.

The postposed subordinate clause may appear as an independent sentence fragment. In such cases, the subordinate clause doesn't constitute a full, finite sentence. It is not traditional to fragment a sentence and when it is done, it is easy to disregard this as a stylistic use. However, this use is expected to be used commonly in spoken language too, and this fragmentated use of subordinate clause might have information structural reasons.

In example (46), both the author's being in jail and the reason for his being in jail are introduced for the first time in the discourse. They both should be read as rhemes.

(46)   [R1Seksen beş yaşında filan değilim.] [R2*Sadece bir hapishanedeyim.*] [R3**Bir adam öldürdüğüm için**.]

[R1It's not like I'm eighty five years old.] [R2*I'm just in a prison.*] [R3**Because I killed a man.**]

This example can be reconstructed in several ways:

(47)   [R1Seksen beş yaşında filan değilim.] [R2*Sadece bir hapishanedeyim,*] [B**bir adam öldürdüğüm için**.]

In this reconstruction in (47), the subordinate clause is in the background position. Thus the reason clause is backgrounded and cannot be prominent in the information structure. It is not possible for a right-displaced subordinate clause to carry a rheme tune.

(48)   [R1Seksen beş yaşında filan değilim.] [R**Bir adam öldürdüğüm için** *sadece bir hapishanedeyim*.]

When the subordinate clause is between the previous sentence and the adverbial *sadece* "just" which modifies the clause *bir hapishanedeyim* "I am in a prison" as in (48) the utterance loses the strong comparison between being eighty five years old and being in a prison. It is possible to get the meaning "I killed a man and my only sentence is being in a prison". This unintended meaning is valid for the discourse, so the context wouldn't have disambiguated the expression. The information structure of this sentence would be ambiguous. With a broad-rheme reading, the subordinate clause can be included in the rheme as marked on the example. It is also possible to read the subordinate clause as a theme or theme-kontrast. The subordinate clause loses its "new information" effect when it is embedded in the matrix clause. This can be related to the proposition that subordinate clauses tend to present the information they carry as "old information" (Quirk, Greenbaum, Leech, & Svartvik, 1985).

(49)   [R1Seksen beş yaşında filan değilim.] [T-K*Sadece*] [R**bir adam öldürdüğüm**
       için *bir hapishanedeyim*.]

In (49) the subordinate clause occurs after the adverbial. The marked IS structure
displays what seems to convey the intended meaning (based on personal intuition).
Like (48), the subordinate clause can carry a variety of information structure features.
It is possible to misread this version as "I'm in a prison just because I killed a man"
and "I'm in a prison because I killed just a man" with different intonations. A comma
after *sadece* may disambiguate the latter but not the first in written text. In both cases
(48, 49) the clarity of the narration is lost.

(50)   [R1Seksen beş yaşında filan değilim.] [R2*Sadece bir hapishanedeyim*.]
       [T**Çünkü**] [R3**bir adam öldürdüm**.]

(51)   [R1Seksen beş yaşında filan değilim.] [T-K*Sadece bir hapishanedeyim*,] [T-
       K**çünkü**] [R**bir adam öldürdüm**.]

İçin can be replaced with *çünkü* with the same meaning but different morpho-syntax.
When the clauses are independent sentences (50), the meaning and information
structure remains intact, with the exception of the theme tendency of the connective.
When the clauses belong to the same sentence (51), it is possible to read *sadece bir
hapishanedeyim* as either theme or theme-kontrast, or rheme since it is a finite
clause.[11]

As a result, the author's choice to fragment a sentence and use a hanging subordinate
clause in the text might be due to a syntactic, semantic, information structural or
stylistic reason, or a combination of those. More fragmented *için* clauses exist in the
MTC. It is highly probable that a purposive *için* clause can be constructed in the
same way, but such an example was not encountered during this study.

---

[11]   These occurrences of *çünkü* can be contraposed, too. The IS analysis will differ as previously
discussed for *çünkü*.

Like purposive *için* clauses, it is possible for a subordinate clause headed by reason bearing *için* to intervene in the matrix clause (52):

(52)   [T*Bunu*] [R**annemi kıskandığım <u>için</u>**] [B*yapıyorum.*]

    [B*I am doing*] [T*this*] [R<u>**because**</u> **I am jealous of my mother.**]

The subordinate clause in the immediately pre-predicate position is highly prominent in the information structure (Göksel & Özsoy, 2000). The connective belongs to the rheme with its complement, and the main predicate is backgrounded due to post-rheme flooring. An intervening reason clause is not necessarily a rheme all the time, nor does it necessarily occur at the immediate left of the main predicate. (Consider *Bunu annesini kıskandığı için yaptığını tahmin ediyorum ama emin değilim*: "I guess he is doing this because he is jealous of his mother, but I am not sure". The contrast between guessing and being sure seems to be more prominent. The subordinate clause is not expected to be the rheme in this utterance.)

Reason-bearing *için* can occur in argument-medial positions, but it's not common. Usually the postposed element of the subordinate clause is an adverbial. The information structure in such cases is not radically different from argument-final connective examples. The subordinate clause and the connective tend to have theme features.

(53)   [T-K**Cocteau'nun sözünü bilmediği <u>için</u> henüz**,] [R*parmaklarıyla söyler*] [B*şarkısını.*]

    [T-K<u>**Since**</u> **he doesn't know Cocteau's quote yet,**] [R*with his fingers, he sings*] [B*his song*].

The subordinate clause in the above example can be located in the post-predicate position and will still be acceptable. In this case, the connective will be backgrounded together witht the whole subordinate clause.

(54)   [R*Parmaklarıyla söyler*] [B*şarkısını* **Cocteau'nun sözünü bilmediği <u>için</u> henüz.**]

However, the acceptability decreases considerably when Arg2 intervenes in Arg1 unless the intervening argument is a parenthetical, or the information units are kontrasted noticeably. One such information structure is suggested in (55). Distinct theme-kontrast units make the subordinate clause stand out, increasing acceptability.

(55) [$_{T-K}$*Şarkısını,*] [$_{T-K}$**Cocteau'nun sözünü bilmediği <u>için</u> henüz,**]
[$_{R}$*parmaklarıyla söyler.*]

The D-LTAG trees and the connective argument orders of reason bearing *için* are not different from the purposive *için*:

α: r-için

D_c ↓   için D_c ↓

D_c D_c ↓   için
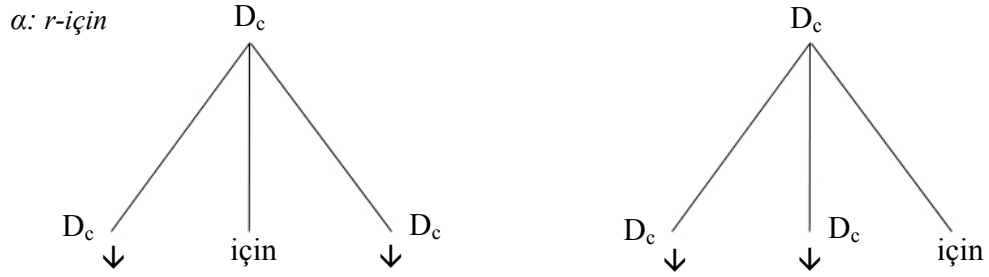
Figure 14- D-LTAG trees for reason bearing subordinator *için*

The CAO profile of reason-bearing *için* is given in Table 5.

Table 5 - The CAO profile for reason-bearing *için*

Group	CAO						
Arg1-Arg2	CAO-2	Arg1		Arg2$_B$	- *için*$_B$ -	Arg2$_B$	m
	CAO-3a	Arg1		Arg2$_B$		- *için*$_B$	f
	CAO-3b	Arg1		Arg2		- *için*$_R$	f
Arg2-Arg1	CAO-5	Arg2	- *için*$_{T-K}$ -	Arg2		Arg1	m
	CAO-6a	Arg2		- *için*$_T$		Arg1	f
	CAO-6b	Arg2		- *için*$_R$		Arg1	f
Arg1-Arg2-Arg1	CAO-8	Arg1	Arg2	- *için*$_{T-K}$ -	Arg2	Arg1	m
	CAO-9a	Arg1		Arg2	- *için*$_{T-K}$	Arg1	f
	CAO-9b	Arg1		Arg2	- *için*$_R$	Arg1	f

5.3.3 *İçin* with an Endophoric Complement

When *için* takes an endophoric complement, the resolution of the endophora determines whether *için* is a discourse connective or not, as well as the sense of that particular *için*.

(56)   [T1*Her kent*] [R1*kendi depremini kendi gözlemlemelidir.*] [T2-K**Bunun için**;] [R2**kamusal ve özel girişim ve kurumlar desteklenmelidir.**]

[T1*Each city*] [R1*has to monitor its own earthquake.*] [T2-K**For this,**] [R2**public and private enterprises and institutions must be supported.**]

(57)   Kardeşim için ben bu düzenin dışında bir şeyim. Beni bu hapishaneden kurtarıp kendi ülkemin hapishanelerine tıktırmak... Bu **onun için** kahramanlık olabilir.

For my brother I am something out of this order. To save me from this prison and to thrust me into the prisons of my own country… This might be a heroic act **for him.**

Here *için* is certainly a discourse connective in (56). The only potential referent that makes sense is the previous clause. In (57), on the other hand, *için* is not a discourse connective. *Bu* "this" refers to the previous infinitival fragment but the complement of *için, o* "this"*,* refers to the brother.

In (56) the discourse adverbial is a theme-kontrast. It is expected for adverbial *için* to display similar features to the subordinator *için*, which constructs adverbial clauses.

(58)   *Köpek kahverengiymiş, bazı yerleri beyaza daha yakınmış*, **Kahve onun için daha uygun bir isim.**

   a.   The dog is brown, some places closer to white, Coffee is a more suitable name for him.

   b.   *The dog is brown, some places closer to white,* **that's why, Coffee is a more suitable name.**

In this example, the prominent reading is (58a), where *için* is not a discourse connective. However it is also possible to get the discourse connective reading in (55b). This reading will be quite acceptable with correct intonation in speaking and by inserting a comma after *onun için* in writing. Example (58) serves to demonstrate that the flexibility of anaphora resolution can result in ambiguous reading when determining whether a connective is a part of discourse structure or not.

(59)   a. [_T*Burada*] [_{T-K}*1980'ler Türkiye'sinde*] [_R*İslâmcı hareketin kadına ne tür bir mesaj verdiğine bakmak istiyorum.*] [_{T-K}**Bunun için önce**] [_T**başka bir çalışmada (Acar, 1989) ayrıntılı biçimde incelenmiş olan**] [_R**üç İslâmcı kadın dergisinin içerik analizi sonuçlarını özetle sunacağım.**]

[_T*Here*] [_R*I would like to have a look at what sort of message the Islamic movement gives to women*] [_{T-K}*in Turkey of 1980's.*] [_{T-K}**For that**, **first**] [_R**I will briefly present the results of content analysis of three Islamist woman's magazines**] [_T**that were examined in detail in another study (Acar, 1989).**]

     b. [_T*Burada*] [_{T-K}*1980'ler Türkiye'sinde*] [_R*İslâmcı hareketin kadına ne tür bir mesaj verdiğine bakmak*] [_B*istiyorum.*] [_{T-K}**Bunun için once**] [_T**başka bir çalışmada (Acar, 1989) ayrıntılı biçimde incelenmiş olan**] [_R**üç İslâmcı kadın dergisinin içerik analizi sonuçlarını özetle sunacağım.**]

[_T*Here*] [_BI would like] [_R*to have a look at what sort of message the Islamic movement gives to women*] [_{T-K}*in Turkey of 1980's.*] [_{T-K}**For that**, **first**] [_R**I will briefly present the results of content analysis of three Islamist woman's magazines**] [_T**that were examined in detail in another study (Acar, 1989).**]

Example (59) shows a case where *için* is a discourse connective, but whether it is a purposive *için* or a reason-bearing *için* depends on the resolution of the anaphora. The ambiguity is due to the fact that Arg1 includes two salient discourse referents for the deictic *bu*. The reading in (59a) can be rephrased as … *bakmak istediğim için...* "… since I want to have a look at …" whereas (59b) can be rephrased as … *bakmak için...* "… in order to have a look at …" Both readings are valid, and in this context,

the choice of the referent doesn't make a major change in the meaning. Note that in (59b) the connective takes a lower clause, the complement of *istiyorum:*"I want to" as its argument. Access to syntactically unavailable arguments is taken to be a hint of anaphoric properties of discourse connectives in both English and Turkish (Webber et al., 2003; Zeyrek et al., 2008).

The IS markings on (59) are suggestions for intended discourse structures. Notive that the main predicate is not included in the rheme in (59b) and is backgrounded due to post rheme flooring. It is not necessary to exclude the main predicate to get the same reading but with this reading is more likely to induce the intended reading. In both cases, the discourse adverbial is theme-kontrast.

The argument order of *için* depends on the nature of the complement, too. For the anaphoric complements *bu* (60) and *o* (58b, 63) to be resolved, Arg1 needs to precede the complement. If the complement is cataphoric *şu* (62) and sometimes *bu* (61) then it is necessary for Arg1 to follow the complement.

(60)   [R1*Şimdi de arabamı mı bozmak* istiyorsun?] [T2**Bunun için**] [R2**gelecek haftayı beklemen gerekecek.**]

[R1*And now you want to break my car down?*] [T2**For that,**] [R2**you will have to wait until next week.**]

(61)   [T-KO gencecik üslûbun tezahürü topu topu birkaç cümleydi ya,] [R**medyanın bazı ayakları da <u>bunun için</u> vardı**] [B**zaten.**] [T*Üç kesik cümleden destan yaratırken*] [R*bin cümlelik bir konuşmayı duyulmaz hâle getirmek için.*]

[T-KThe manifestation of that fresh style was all in all a few sentences yet] [R**some branches of the media existed <u>for this reason</u>**] [B**anyway.**] [R*In order to render a speech of a thousand sentences inaudible*] [T*while creating a saga out of a speech of three interrupted sentences.*]

(62)   [_R**Bu açıklama <u>şunun için</u> gerekti:**] [_{T-K}*Kemalist Devrim, din özgürlüğünü,*] [_R*dinin özgürlüğü olarak tanımlamadı,*] [_R*o zaman lâiklik diye bir şey kalmazdı.*]

[_R**This announcement was necessary <u>for this:</u>**] [_{T-K}*the Kemalist Revolution*] [_R*did not define the freedom of religion as the freedom of the religion,*] [_R*(if so) then there would be no such thing as secularity.*]

In (60), the discourse adverbial is a theme. In (61) and (62), the adverbial is a part of the rheme. As in the case of the subordinator *için*, when the connective occurs to the left of the main predicate, there is a tendency for the connective and its complement to be included in the rheme. But it is possible to have argument-medial connective in example (60) just by adding the pro-dropped subject back, as in *Senin bunun için gelecek haftayı beklemen gerekecek* "For that, you will have to wait until next week." The subject will be a part of the theme with the now argument-medial connective and the rest of the information structure will remain the same.

The rheme status of the adverbial in (61) and (62) may also be induced by the cataphoric property. The second argument in example (62) can be reconstructed, as *Şunun için bu açıklama gerekti* or *Şunun için gerekti bu açıklama* in order to get argument-initial cataphoric adverbial, and the adverbial still tends to be included in the rheme in both constructions. The theme reading is possible if tried, but a theme-kontrast reading does not seem possible. In addition, there are no backgrounded argument-final examples of cataphoric *için* in the MTC and such examples cannot be constructed from the examples at hand. It is possible to get argument-final cataphoric *için* from (62), for example *bu açıklamanın gerekmesi şunun için:* "The necessity of this announcement is because of that". In this construction the adverbial is the main predicate and is not backgrounded.

The following examples demonstrate that *için* and its anaphoric complement can occupy argument-initial (63), argument-medial (65) and argument-final (64) positions. The argument-final *için* is backgrounded as expected. Argument-initial adverbial *için* is marked as theme-kontrast, but it is ambiguous between theme-

kontrast and rheme. Regarding this ambiguity, rheme will be added to the CAO for this position. Argument-medial *için* is included in the rheme.

(63)   [R*Bu bizim bünyemize uygun, değildi.*] [T-K**Onun için**] [R**değiştirdik.**]

[R*This was not appropriate for our constitution.*] [T-K**As a result**] [R**we changed it.**]

(64)   [R*Dinleyeceksin biliyorum,*] [R**sana anlatıyorum**] [B**onun için**.]

[R*I know that you will listen,*] [B**that's why**] [R**I'm telling you.**]

(65)   [T1-KKorkulan olmuş,] [R1köpekbalıkları takımın onuncu kilometreden sonrasını parçalayıp sürüklemişti.] [R2*Saatlerce koca denizde takımın yarısını aradık.*] [R3**Siyah bayraklı şamandıralar bunun için takılıyordu**] [B**takıma,**] [T4ancak] [R4onlardan eser yoktu.]

[T1-KWhat was feared had happened] [R1and sharks had torn and dragged the net after the tenth kilometer.] [R2*We searched for the half of the net in the great sea for hours.*] [R3**The buoys with black flags were attached to the net for this,**] [T4but] [R4they were nowhere to be seen.]

The D-LTAG tree for *için* with an endophoric complement:

β: a-için   D_c

A/C için   *
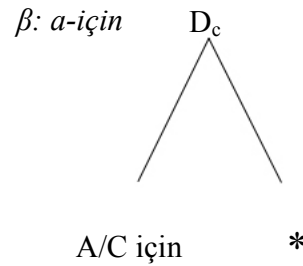
Figure 15 - D-LTAG tree for *için* with deictic complement

The CAO profiles of *için* with anaphoric and cataphoric complements are given separately below:

Table 6 – The CAO profile for *için* with an anaphoric complement

	CAO-1a	Arg1	*A için*$_T$ -		Arg2	i
	CAO-1b	Arg1	*A için*$_{T-K}$ -		Arg2	i
	CAO-1c	Arg1	*A için*$_R$ -		Arg2	i
Arg1-Arg2	CAO-2a	Arg1	Arg2	- *A için*$_T$ -	Arg2	m
	CAO-2b	Arg1	Arg2	- *A için*$_R$ -	Arg2	m
	CAO-3	Arg1	Arg2		- *A için*$_B$	f

Table 7 - The CAO profile for *için* with a cataphoric complement

	CAO-4	*C için*$_R$ -		Arg2	Arg1	i
Arg2-Arg1	CAO-5	Arg2	- *C için*$_R$ -	Arg2	Arg1	m
	CAO-6	Arg2		- *C için*$_R$	Arg1	f

5.4 *Dolayı(sıyla)*

5.4.1 Subordinators

5.4.1.1 *Dolayı* with VP complements

Dolayı takes a variety of VP complements. The morphological inflections on these possible complement types include non-factive nominal (66), factive nominal (67), infinitive (68), and gerund (69), as well as the ablative case *dolayı* assigns to all its complements.

(66)   [$_T$*Burada evler*] [$_R$**sık sık iklim değişmesinden <u>dolayı</u>** *tamir görmüş.*]

   [$_T$*Here the houses*] [$_R$*had been repaired* **<u>because</u> of the frequent climate changes.**]

(67) [T-K**Kitlesel imha silahları çok korkunç olduğundan ve kullananı da aynı anda yok ettiğinden <u>dolayı</u>,**] [R*şimdi de duygu ve düşünceyi esir eden, denetim altına alan elektromanyetik silahlar üretiyor*] [B*adına gelişmiş denilen ülkeler.*]

[T-K<u>**Because**</u> **the mass destruction weapons are too terrifying and they destroy the user at the same time,**] [B*now the so called developed countries*] [R*are producing electromagnetic weapons which enslave and control emotions and thoughts.*]

(68) [T19 Aralık tarihinden itibaren halkın ziyaretine açılacak olan Osmanlı Bankası Müzesi'nin basına tanıtım toplantısında konuşan Garanti Bankası Genel Müdürü Ergun Özen,] [R**bankanın tarihi mirasına sahip çıkmaktan <u>dolayı</u>** *gururlu olduklarını*] [B*söyledi.*]

[TErgun Özen, General Manager of Garanti Bank who spoke at the press conference of the Osmanli Museum which will be open to public visitors as of December 19th,] [Bsaid] [R*that they were proud* <u>**because**</u> **they claimed the historical heritage of the bank.**]

(69) [T-K**Devletin borç stokunun yüksek oluşundan <u>dolayı</u>**] [T-K*aktif bir tahvil bono pazarının bulunması*] [R*yatırım fonları portföylerini de etkiliyor.*]

[T-KThe fact *that there is an active bond market*] [T-K<u>**because**</u> **the debt stock of the state is high**] [Raffects the investment fund portfolios.]

In (66) and (68) the subordinator is a part of the rheme. In (67) and (69) it is part of a theme-kontrast. Like subordinator *için*, *dolayı* belongs to the same information unit with its complement.

This use of *dolayı* is akin to the subordinator use of *için* in several ways. Both subordinators must immediately follow the head of the subordinate clause. Both construct subordinate clauses that can precede, intervene or follow the matrix clause, as shown in (67), (66), and (70) respectively. The subordinate clause can be an independent sentence fragment as in (70). The fragment is the rheme like in (46).

(70)   [_{R1}*Aslında Ayyıldız Apartmanı, Ayyaş Yazar'a edebiyat konusunda hiç güvenmiyordu.*] [_{T2-K}**Edebiyattan anladığından değil,**] [_{R2}**Ayyaş Yazar'ın başarı derecesini gördüğünden <u>dolayı.</u>**]

[_{R1}*In fact, Ayyıldız Block did not trust the Drunkard Author with literature at all.*] [_{T2-K}**<u>Not because</u> it knew about literature**,] [_{R2}**<u>but because</u> it had seen Drunkard Author's degree of success.**]

However, *dolayı* and *için* have a major difference. The type of VP complement determines the sense of the discourse relations anchored by *için*, whereas all discourse relations anchored by *dolayı* have the same "reason" sense. In other words, the second argument of *dolayı* is always the reason of its first argument.

This fixation of sense might be due to the ablative case on the head of the subordinate clause. Ablative case on (preferably but not necessarily factive) nominalized or infinitival VPs can convey causative discourse relation by itself, as well as together with subordinators like *dolayı* and *ötürü*. *Dolayı* and *ötürü* are almost always replaceable, but they are not always omissible. One cannot always rely on the ablative VP to handle the discourse relation. For example, in (68) *dolayı* cannot be omitted. *Bankanın tarihi mirasına sahip çıkmaktan gururlu olduklarını söyledi* is not as acceptable, because without *dolayı*, the ablative infinitival VP is expected to be the complement of being proud, and it calls for another construction: *Bankanın tarihi mirasına sahip çıkmaktan gurur duyduklarını söyledi*[12]. In addition, the bare ablative sometimes loses some acceptability when the nominal is the main predicate of the sentence[13]. The examples (71) and (72) below are derived from acceptable corpus examples (66, 67).

---

[12]   Personal intuition. All native speakers may not agree.

[13]   Such a construction is not attested for *dolayı* or *dolayısıyla* in the MTC, but it is attested for *sebebiyle*, *nedeniyle* and *yüzünden*. These subordinators are also causal discourse subordinators and their complements share the same morphological properties with those of *dolayısıyla*.

(71)   [_T-K_*Burada evlerin tamir görmesi,*] [_R_**sık sık iklim değişmesin<u>den</u>dir.**]

(72)   *?* [_T-K_*Adına gelişmiş denilen ülkelerin şimdi de duygu ve düşünceyi esir eden,*] [_T-K_*denetim altına alan elektromanyetik silahlar üretmesi,*] [_R_**kitlesel imha silahları çok korkunç olduğundan ve kullananı da aynı anda yok ettiğin<u>den</u>dir.**]

The first difference to be noticed between these sentences would be the different morphology of the VP. The first example includes a non-factive nominalizer where the second has a factive nominalizer. The second example becomes acceptable when the head is non-factive.

(73)   [_T-K_*Adına gelişmiş denilen ülkelerin şimdi de duygu ve düşünceyi esir eden,*] [_T-K_*denetim altına alan elektromanyetik silahlar üretmesi,*] [_R_**kitlesel imha silahlarının çok korkunç olmasından ve kullananı da aynı anda yok etmesin<u>den</u>dir.**]

However, it would be impetuous to attribute the difference in acceptability to this morphological difference, and declare that the non-factive nominals are acceptable as main predicates whereas factive nominals are not; because the first example is still somewhat acceptable with a factive VP (74), and even more acceptable without the optional copula –*dir* (75).

(74)   *? Burada evlerin tamir görmesi,* **sık sık iklim değiştiğin<u>den</u>dir.**

(75)   *Burada evlerin tamir görmesi,* **sık sık iklim değiştiğin<u>den</u>.**

Argument medial *dolayı* is not attested in the MTC, and it is not easy to derive a completely acceptable argument-medial *dolayı* construction from the *dolayı* samples in MTC, so argument-medial ordering will not be included in the list.

D-LTAG trees for subordinator *dolayı*:

α: s-dolayı   D_c

D_c   dolayı   D_c

D_c

D_c   D_c   dolayı
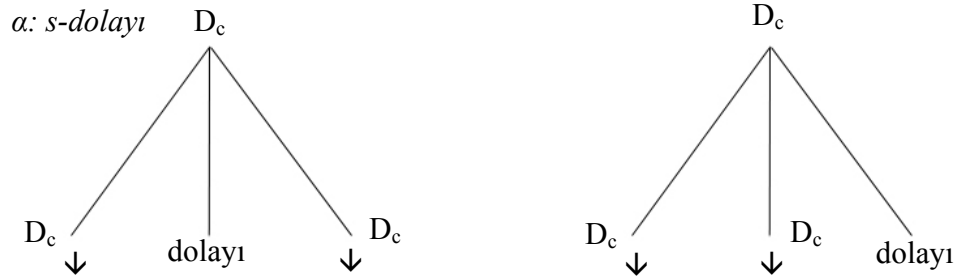
Figure 16 - D-LTAG trees for subordinator *dolayı*

Table 8 below shows the connective argument orders available to *dolayı*:

Table 8 - The CAO profiles for subordinator *dolayı*

Arg1-Arg2	CAO-3a	Arg1	Arg2$_B$	- *dolayı* $_B$	f
	CAO-3b	Arg1	Arg2	- *dolayı*$_R$	f
Arg2-Arg1	CAO-6a	Arg2	- *dolayı*$_{T-K}$	Arg1	f
	CAO-6b	Arg2	- *dolayı*$_R$	Arg1	f
Arg1-Arg2-Arg1	CAO-9	Arg1	Arg2	- *dolayı*$_R$	Arg1	f

5.4.1.2   Dolayısıyla as subordinator

This subordinator is the inflected form of *dolayı* with possessive and instrumental case. It is not as common as the subordinator *dolayı* or the adverbial *dolayısıyla*. The attested complements of subordinator *dolayısıyla* are either non-factive nominals or gerunds.

(76)   [$_{T-K}$**3 saat 52 dakikalık bir film olması <u>dolayısıyla da</u>**] [$_R$*o günlerin en uzun filmi özeliğini taşımaktaydı.*]

[$_{T-K}$**Because it was a movie of 3 hours and 52 minutes,**] [$_R$*it was the longest movie of those days.*]

72

(77)   [T-KPir, program soru - cevap biçiminde yapıldığı için] [R**Özyılmaz'ın ayrılışı**

dolayısıyla *programı iptal ettiklerini söyledi.*]

[RPir said *that they cancelled the program* **because of** the departure of

Özyılmaz] [T-Kbecause they did the program in question-answer format.]

Like *dolayı*, *dolayısıyla* tends to be in in the theme-kontrast or rheme units. Tough

not common, the subordinator clause with *dolayısıyla* can be right-displaced, and

information structurally backgrounded, too. The following example (78) is derived

from (76).

(78)   [R*O günlerin en uzun filmi özeliğini taşımaktaydı*] [B**3 saat 52 dakikalık bir**

film olması dolayısıyla (?da).]

Without the context the support the modifier particle *de* at the end, the sentence is

odd. But without the modifier, the postposed subordinate clause is acceptable. The

intervening subordinate clause is acceptable without the particle *de* as well, as

demonstrated by (79) again derived from (76). In (79) the intervening argument,

which was the theme-kontrast originally, is now included in the rheme.

As with *dolayı,* argument medial *dolayısıyla* was not attested and could not be

constructed from existing examples.

(79)   [T-K*O günlerin en uzun filmi özeliğini*] [R**3 saat 52 dakikalık bir film olması**

dolayısıyla *taşımaktaydı*.]

D-LTAG trees for subordinator *dolayısıyla* are given in Figure 17.
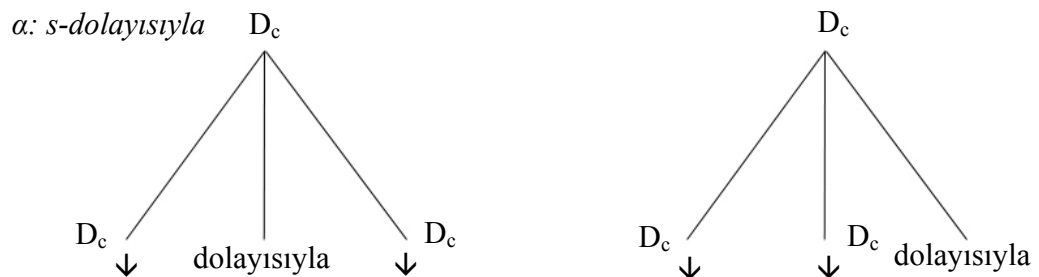
Figure 17 - D-LTAG trees for subordinator *dolayısıyla*

73

Table 9 below shows the CAO profile of subordinator *dolayısıyla*:

Table 9 - The CAO profiles for subordinator *dolayısıyla*

Arg1-Arg2	CAO-3	Arg1	Arg2$_B$		- *dolayısıyla* $_B$	f
Arg2-Arg1	CAO-6a	Arg2	- *dolayısıyla*$_{T-K}$		Arg1	f
	CAO-6b	Arg2	- *dolayısıyla*$_R$		Arg1	f
Arg1-Arg2-Arg1	CAO-9	Arg1	Arg2	- *dolayısıyla*$_R$	Arg1	f

5.4.2   Adverbials

5.4.2.1   Dolayı with a discourse deictic complement

Dolayı can take a discourse deictic complement like *için*. The deictic expression will be assigned ablative case like the VP complements of *dolayı* and must be adjacent to the connective. When the deictic is anaphoric, the first argument needs to precede the second argument so that the anaphora will be resolved as in (80). Though no such examples are attested, it is expected the first argument will follow the second if the deictic expression is cataphoric.

(80)   [$_T$*Temizlediği enginarlar*] [$_R$*ellerini kahveyle kızıl arası bir renge boyamış.*]
[$_R$**Bundan dolayı sinirli.**]

[$_T$*The artichokes he peeled*] [$_R$*painted his hands brownish red color.]* [$_R$**Because of that he is irritated.**]

The connective can intervene in the second argument, and can occupy pre-predicate position.

(81)   [$_{T-K}$*Bu arada*] [$_R$*polise 'sen' diye hitap ettim,*] [$_{T-K}$**ancak**] [$_R$**hemen bundan dolayı özür diledim.**]

[$_{T-K}$*In the meanwhile*] [$_R$*I called the police 'you' (impolite),*] [$_{T-K}$**but**] [$_R$**I apologized for that immediately.**]

In both examples (80, 81) adverbial *dolayı* is in the rheme. The reason for this rheme tendency is not clear. It might be a lexical tendency of this use of *dolayı* or it might be because in both examples *dolayı* is located at the immediate left of the main predicate. In another position, as in *Bundan dolayı, 18. yüzyıla Aydınlanma Çağı denir.* "As a result of this, the 18th century is called the Age of Enlightment", *dolayı* and its complement constitute the theme-kontrast. This sentence can be reconstructed with argument-medial connective: *18. yüzyıla bundan dolayı Aydınlanma Çağı denir.* In this position, the adverbial is ambiguous between theme-kontrast and a broad rheme reading.

There are two interesting points that results from the analysis of *dolayı* in the MTC. First, the postposed adverbial *dolayı* is not attested. In the structural vs. anaphoric discussion so far, it is established that anaphoric connectives occupy a wider range of positions in languages like English. In Turkish, however, the case seems to be different. The subordinators offer a variety of CAOs, whereas the adverbials make use of a more restricted order. In addition to the restrictions imposed by anaphora resolution, some acceptable connective-argument orders are not employed in the MTC at all.

The other peculiarity is that the only deictic complement used with *dolayı* was *bu* "this". The cataphoric counterpart *şu* is usually less frequent as a discourse deictic complement and this might be the reason for its lack of use together with *dolayı*. The distant demonstrative *o* "that", in contrast, is as frequent as *bu*, if not more. There has to be another reason for its absence.

This deictic seems to be lexicalized to express causal relation with ablative case.

(82)   *öğlende tarlaya babama yemek götürdüm*, **ondan geç kaldım.**

   At noon, I carried food to the fields for my father, **that's why I was late.**

It may be redundant to use the same structure, i.e., the ablative *o*, as the complement of another causal connective because it already conveys causality. Though this sounds like a logical explanation at first glance, it is challenged by the fact that

factive nominals with ablative case, which express causal relation by themselves as well, appear as the complements of causal connectives in the MTC.

The preference of *bu* over *o* may also be due to the difference in meaning. *O* denotes an object farther than the object *bu* denotes. Since very few texts in the corpus include dialogue and the referent of the deictic is always close, usually in the immediately preceding sentence, the distance expressed by *o* may not be necessary.

In the MTC samples, *dolayı* is not used as frequently as *dolayısıyla* as an adverbial. This might be another reason for the lack of variety, but as mentioned in the later subsection, adverbial *dolayısıyla* has no argument-final occurrences either. For the sake of completeness, the acceptable orders will be explored and included in the possible connective-argument orders. The following examples are derived from (80) and (81), and are acceptable.

(83) [R1**Şundan dolayı sinirli:**] [T2-K*Temizlediği enginarlar ellerini*] [R2*kahveyle kızıl arası bir renge boyamış.*]

(84) [T1-K**Sinirli olması**] [R1**şundan dolayı:**] [T2-K*Temizlediği enginarlar ellerini*] [R2*kahveyle kızıl arası bir renge boyamış.*]

(85) [T1-K**Hemen**] [R1**şundan dolayı özür diledim:**] [R2*Bu arada polise 'sen' diye hitap etmiştim.*]

(86) [R*Bu arada polise 'sen' diye hitap ettim,*] [T-K**ancak**] [R**hemen özür diledim**] [B**bundan dolayı.**]

The right-displaced adverbial in (86) is backgrounded as expected. In examples (83) – (85) the cataphoric adverbial *dolayı* displays a strong tendency to be in the rheme unit. In (84) the adverbial is the main predicate and in (83) and (85) it is to the immediate left of the main predicate. However, when the second argument of (85) is rearranged so as to put some distance between the cataphoric adverbial and the main predicate, for example as *Şundan dolayı hemen özür diledim*, the whole argument becomes a rheme unit, including the previously excluded time adverbial. When one

tries to read *şundan dolayı* as a theme kontrast as it is possible for *bundan dolayı*, though the utterance remains meaningful, the link between the cataphor and its referent seem to be severed. Following native speaker intuition, only rheme will be included in the CAO of the .cataphoric adverbial

The D-LTAG tree for *dolayı* with discourse deictic complement is in Figure 18:

β: a- dolayı D_c
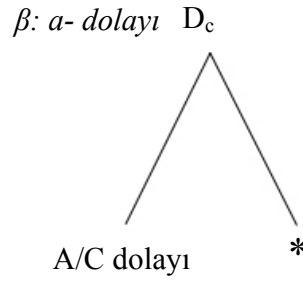
A/C dolayı   *

Figure 18 - D-LTAG tree for *dolayı* with deictic complement

The CAO profiles of *dolayı* with anaphoric and cataphoric complements:

Table 10 - The CAO profile for *dolayı* with an anaphoric complement

	CAO-1a	Arg1	*A dolayı*$_{T-K}$ -		Arg2	i
	CAO-1b	Arg1	*A dolayı*$_R$ -		Arg2	i
Arg1-Arg2	CAO-2a	Arg1	Arg2	- *A dolayı*$_{T-K}$ -	Arg2	m
	CAO-2b	Arg1	Arg2	- *A dolayı*$_R$ -	Arg2	m
	CAO-3	Arg1	Arg2		- *A dolayı*$_B$	f

Table 11 - The CAO profile for *dolayı* with a cataphoric complement

	CAO-4	*C dolayı*$_R$ -		Arg2	Arg1	i
Arg2-Arg1	CAO-5	Arg2	- *C dolayı*$_R$ -	Arg2	Arg1	m
	CAO-6	Arg2		- *C dolayı*$_R$	Arg1	f

5.4.2.2 Dolayısıyla as a discourse adverbial

Unlike *için* and *dolayı*, *dolayısıyla* does not take an anaphoric complement when it is a discourse adverbial. Its anaphoric properties are due to morphological features: the possessive case marking it carries (Zeyrek et al., 2008). The possessive marked *dolayısıyla* presupposes a genitive marked counterpart. In order to resolve this anaphora, *dolayısıyla* needs a salient referent in the previous discourse, resulting in strict Arg1-Arg2 order. In this respect, *dolayısıyla* is still very similar to other connectives that take a deictic complement and act as discourse adverbials, such as *için* and *dolayı*.

(87) [R1*Uzun aralıklarla yemek yendiğinde vücut yağı artar,*] [R2**dolayısıyla şişmanlık oluşur**.]

[R1*When the meals are eaten after long intervals the body fat increases,* [R2**as a result fatness occurs.**]

The adverbial in (87) is included in the rheme. *Dolayısıyla* can also be theme-kontrast in Arg1-Conn-Arg2 order. One such example is (92) below, but the presence of another connective may raise doubts about that example. There are other examples in the MTC where *dolayısıyla* is not a part of the rheme. For example in *Dolayısıyla, bilimsel teorilerin bazı kavramları oluşturup kullanabilmeleri, toplumun ideolojik düzeyinin o kavramların ortaya çıkmasına elveren bir olgunluğa ulaşmış olmasını gerektirir* "Consequently, for the scientific theories to create and use some concepts it is necessary that the ideological level of the society have reached a maturity that can enable the creation of those concepts", the adverbial is a theme-kontrast unit.

Dolayısıyla can occur argument-medially. Interestingly, this position, which is expected to be the distinguishing position for discourse adverbials, is not preferred much frequently in the MTC. Below is the only example attested for argument-medial adverbial *dolayısıyla*. In this example the adverbial is theme-kontrast.

(88)   [T1-KGelgelelim,] [T1-Kçemberin üzerinde olsun duramamış,] [R1çemberden
hep içeri doğru davranmıştır.] [T2-KSon cüretli edimi de,] [T2-Kdolayısıyla,]
[R2onun çemberi kırıp dışarı çıkma isteğinin bir belirtisi olarak
değerlendirilemez,] [Bdiye düşünüyorum.]

[T1-KHowever,] [T1-Khe did not even stay on the circle;] [R1he has always acted
towards inside the circle.] [T2-KHis last daring act,] [T-Kconsequently,] [Rcannot
be judged to be a hint for that he wants to break free of the circle] [BI think.]

It is possible for argument-medial dolayısıyla to be a part of the rheme, for example
in a sentence inspired by (87) such as *Uzun aralıklarla yemek yendiğinde vücut yağı
artar, kişi de dolayısıyla şişmanlar* "When the meals are eaten after long intervals the
body fat increases; as a result, one gets fat." In this sentence, the particle *de* comes
between the NP *kişi* and *dolayısıyla* and prevents the reading where the NP is the
complement of *dolayısıyla*. Probably the possibility of such ambiguity, i.e., the
possibility of mistaking the elements before *dolayısıyla* for its argument is the reason
for the scarcity of argument-medial adverbial *dolayısıyla*.

Like the adverbial use of *dolayı*, *dolayısıyla* has no argument final instances in the
MTC. Again, a position that is acceptable to native speakers was not employed by the
authors. Below is an example of acceptable argument final *dolayısıyla*, derived from
(87). The right-displaced adverbial is backgrounded.

(89)   [R1*Uzun aralıklarla yemek yendiğinde vücut yağı artar,*] [R2şişmanlık oluşur]
[Bdolayısıyla.]

Dolayısıyla connects VPs and subordinate and relative clauses as well as finite
clauses.

(90)   [T-KOna göre bu], [T-K*arz-talep dengeleriyle oynamak ve üretim fazlası
yaratmak,*] [Rdolayısıyla kaynak israf etmek anlamını taşıyordu.]

[T-KAccording to him,] [Rthis meant] [T-K*playing with the supply-demand balance
and creating a production surplus,*] [Rconsequently wasting resource.]

In such constructions, *dolayısıyla* is not expected to occur argument-finally frequently, because the syntactic heads of most subordinate clauses and other VPs have infinitive or nominal inflections which can also be taken as a complement of *dolayısıyla*. Argument-final connective would raise an ambiguity that is hard to disambiguate both in writing and in speech.

(91)   # Ona göre bu, *arz-talep dengeleriyle oynamak ve üretim fazlası yaratmak*, **kaynak israf etmek <u>dolayısıyla</u>** anlamını taşıyordu.

The arguments of *dolayısıyla* may also be coordinated by another connective. The use of one connective together with another connective in the same relation is taken to be a hint that at least one of the connectives involved in this relation takes one of its arguments anaphorically (Webber, Knott, & Joshi, 1999). In this example two connectives constitute a theme-kontrast unit.

(92)   [T-K12 Mart'ta sol açısından belirleyici olan,] [T-K*birtakım insanların bu düzeni aşmak istemeleri*] [T-K**<u>ve dolayısıyla</u>**] [R**yasalara ters düşmeleriydi**.]

[T-KWhat was significant about March 12th for the Left was] [T-K*that some people wanted to overcome this regime,*] [T-K**and consequently**] [R**contradicted the laws**.]

The D-LTAG tree for *dolayısıyla* as discourse adverbial:

β: a- dolayısıyla D_c
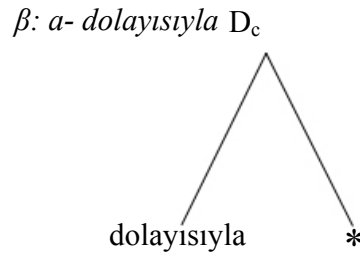
dolayısıyla   *

Figure 19 - D-LTAG tree for *dolayısıyla* as discourse adverbial

The CAO profile of *dolayısıyla* is in table (12):

Table 12 - The CAO profile for *dolayısıyla* as discourse adverbial

	CAO-1a	Arg1	*A dolayısıyla*T-K -		Arg2	i
	CAO-1b	Arg1	*A dolayısıyla*R -		Arg2	i
Arg1-Arg2	CAO-2a	Arg1	Arg2	- *A dolayısıyla*T-K -	Arg2	m
	CAO-2b	Arg1	Arg2	- *A dolayısıyla*R -	Arg2	m
	CAO-3	Arg1	Arg2		- *A dolayısıyla*B	f

5.4.3 Summary of CAO Profiles

The complete CAO profiles of the discourse connectives we have analyzed so far is summarized in Table 13. Looking at this table, one cannot but notice certain patterns. The reasons for these patterns and the implications that come with them will be discussed in the following chapter.

Table 13 - The CAO profiles of all investigated discourse connectives

	Arg1-Arg2			Arg2-Arg1			Arg1-Arg2-Arg1		
	CAO-1	CAO-2	CAO-3	CAO-4	CAO-5	CAO-6	CAO-7	CAO-8	CAO-9
çünkü	T, T-K	B	B				T, T-K		B
zira	T	T-K,B	B				T-K		B
p-için		B	B		T-K	T-K,R		T, T-K	T, T-K, R
r-için		B	B, R		T-K	T,R		T-K	T-K, R
a-için	T, T-K, R	T, R	B						
c-için				R	R	R			
s-dolayı			B, R			T-K, R			R
s-dolayısıyla			B			T-K, R			R
a-dolayı	T-K, R	T-K, R	B						
c-dolayı				R	R	R			
a-dolayısıyla	T-K, R	T-K,R	B						
	Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f

CHAPTER 6

DISCUSSION AND CONCLUSION

In previous chapter, the connective position, argument order and information structure of the discourse connectives in written Turkish text were analyzed using examples from Metu Turkish Corpus. In this chapter, the results of the analyses will be interpreted in an attempt to identify the tendencies of these discourse connectives.

6.1 The Connective-Argument Order Profiles of Connective Types

Now that we have an idea about how certain Turkish discourse connectives tend to behave, we can compare and contrast them to work out what lies beneath their similar and different behaviors. First, we would like to look at what we have in hand in full. Deserving its reputation as a free word order language, Turkish indeed employs nine expected CAO possibilities, but not through one type of discourse connective.

Table 13 suggests that discourse connectives align in groups, mainly depending on their syntactic type. First, we look at the CAO profiles of the two coordinating conjunctives, *çünkü* and *zira*.

The coordinating conjunctives are restricted to Arg1-Arg2 order in our analysis. If we set aside the parenthetical constructions for now, we see that coordinating conjunctives can occupy all of the argument-initial, argument-medial and argument-final positions. In the case of parenthetical expressions, the coordinating conjunctions do occupy argument-initial and argument-final positions.

Table 14 - The CAO Profiles for Coordinating Conjunctives[14]

	Arg1-Arg2			Arg2-Arg1			Arg1-Arg2-Arg1		
	CAO-1	CAO-2	CAO-3	CAO-4	CAO-5	CAO-6	CAO-7	CAO-8	CAO-9
çünkü	T, T-K	B	B				T, T-K		B
zira	T	T-K,B	B				T-K		B
	Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f

The lack of argument-medial connective in these constructions (CAO9) is interesting because there seems to be no syntactic obstacle for them. One possibility is that argument-medial coordinating conjunctions are usually constructed by contraposing some additional material with the connective, and parenthetical expressions are usually short clauses with few constituents. Since items that are important for information structure cannot be contraposed, even fewer of these constituents can be contraposed in addition to the connective. As a result, it would be hard to create an environment that would support an argument-medial coordinator in a parenthetical, unless the speaker/writer is particularly trying to do so, as we do in (90).

(93)   *Cumartesi bütün gün dinlenmelerine rağmen*, (**Cuma akşamı çok uzun bir yoldan gelmişti <u>çünkü</u> misafirlerimiz,**) *Pazar günü hala yorgunlardı.*

Although *they rested all Saturday* (**<u>because</u> our guests had arrived from a very long journey on Friday**) *they were still tired on Sunday.*

---

[14]   T, T-K and B in a box indicates the CAOs that were attested in the MTC examples and their information structural features. The bottom row shows the position of the connective relative to the second argument in that CAO.

Keeping in mind that parenthetical constructions are not very common in our examples, and heavy parentheticals are even less so, we will conclude that coordinating conjunctions have Arg1-Arg2 order and they can occupy all three of the argument-initial, argument-medial and argument-final positions.

Next we consider subordinating conjunctions. Table 15 demonstrates CAO profiles available for the clause subordinating uses of purposive *için*, reason-bearing *için*, *dolayı* and *dolayısıyla*.

Table 15 - The CAO profiles for clause subordinating conjunctives

	Arg1-Arg2			Arg2-Arg1			Arg1-Arg2-Arg1		
	CAO-1	CAO-2	CAO-3	CAO-4	CAO-5	CAO-6	CAO-7	CAO-8	CAO-9
p-için		B	B		T-K	T-K,R		T, T-K	T, T-K,R
r-için		B	B, R		T-K	T,R		T-K	T-K, R
s-dolayı[15]			B, R			T-K, R			R
s-dolayısıyla			B			T-K, R			R
	Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f

As morpho-syntactic constraints dictate, the argument-final orders dominate the CAO profile of subordinators, and argument-initial orders are completely banned.

There is an intriguing difference between the subordinators. Both purposive and reason-bearing *için*s occupy argument-medial position, whereas neither *dolayı* nor *dolayısıyla* does so. The major difference between these subordinator types is the morphological differences of the complements. *İçin* does not assign a case to its complements like *dolayı* and *dolayısıyla* does. This might indicate that there is a difference between subordinators that assign case and those do not. Another difference is that the sense of *için* depends on the morphological properties of its complement, indicating that there should be two different lexical entries for discourse connective *için*. One *için* takes an infinitival or nonfactive nominal complemet and has purposive sense, whereas the other *için* takes a factive nominal complement and

---

[15] The prefix s- stands for subordinator use.

has reason sense. This distinction might be the sign of another tendency. However, before jumping to conclusions, one should always keep in mind that the frequency of *için* in MTC is considerably higher than that of *dolayı* or *dolayısıyla*. Unfortunately, we leave this subject aside inconclusively, for anything we propose cannot go beyond mere speculation without further research.

Table 15 shows that the discourse subordinators can accommodate all three possible argument orders. Arg2 can precede, follow or intervene in Arg1. Remember that when Arg2 follows Arg1, either the whole subordinate clause follows the matrix clause, and therefore is backgrounded, or the subordinate clause stands as an independent fragment. We have argued in the previous chapter that this can be due to several reasons, including but not restricted to stylistic and information structural considerations.

The subordinators we call discourse connectives come from a syntactic group that makes their clause act similar to adverbials (Kornfilt, 1997). It is not surprising to see that adverbial clauses exploit all positions available to adjuncts. It is also expected that when the clause is replaced with an endophoric item that refers to an abstract object, the resulting constructions act as adverbials do. Now we have a look at those discourse adverbials.

Table 16 - The CAO profiles of anaphoric discourse adverbials

	Arg1-Arg2			Arg2-Arg1			Arg1-Arg2-Arg1		
	CAO-1	CAO-2	CAO-3	CAO-4	CAO-5	CAO-6	CAO-7	CAO-8	CAO-9
a-için	T, T-K,R	T, R	B						
a-dolayı	T-K, R	T-K, R	B						
a-dolayısıyla	T-K, R	T-K	B						
	Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f

Table 17 - The CAO profiles of cataphoric discourse adverbials

	Arg1-Arg2			Arg2-Arg1			Arg1-Arg2-Arg1		
	CAO-1	CAO-2	CAO-3	CAO-4	CAO-5	CAO-6	CAO-7	CAO-8	CAO-9
c-için				R	R	R			
c-dolayı				R	R	R			
	Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f

As we have seen in Chapter 5, the CAO profiles of adverbials are heavily dependent on whether the complement is anaphoric or cataphoric. Tables 16 and 17 show how the CAO of the same head with one complement can be so definitely different from the CAO of the same head with another complement. An anaphoric complement, or morphology with anaphoric consequences, needs its antecedent in the previous discourse, imposing strict Arg1-Arg2 orders on the adverbial's arguments. A cataphoric argument, on the other hand, imposes Arg2-Arg1 orders on the adverbial's arguments, because it needs its referent in the following discourse.

As stated above, one would expect adverbial clauses to occupy a variety of positions. This expectation is even stronger because of the fact that the adverbials we have analyzed share the same syntactic head with adverbial clauses which can occupy all positions available to adjuncts in a sentence. However, this expectation is not satisfied by the use of these adverbials in MTC. Notice that argument-final adverbials constructed with *dolayı* and *dolayısıyla* were not attested, even though this position is syntactically available. We believe that a quantitative study over the discourse adverbials in MTC, after the annotations are complete over a substantive part of the corpus, will reveal the word order tendencies more clearly, and guide a more focused research as to how and why some discourse adverbials are constrained as such.

Table 18 is a rearrangement of Table 13 according to connective types, showing the clear distinctions between the CAO profiles of different types of discourse connectives. The coordinating conjunctions stand as one group as expected. Subordinating conjunctions stand as one major group, but with a question mark

hanging. Whether there is a distinction between subordinators that assign case to their complements and subordinates that do not remains to be clarified by future subordinator-focused research.

The next type, discourse adverbials, have clear-cut distinctions between anaphoric and cataphoric groups. Though these groups do not have a single common CAO, their distinction is motivated by exactly the same reason, and their CAO profiles differ predictably.

Table 18 - The CAO profiles sorted according to connective type

			Arg1-Arg2			Arg2-Arg1			Arg1-Arg2-Arg1		
			CAO-1	CAO-2	CAO-3	CAO-4	CAO-5	CAO-6	CAO-7	CAO-8	CAO-9
S T R U C T U R A L	Coordinating Conjunctions	*çünkü*	T, T-K	B	B				T, T-K		B
		zira	T	T-K,B	B				T-K		B
	Subordinating Conjunctions	*p-için*		B	B		T-K	T-K, R		T, T-K	T, T-K,R
		r-için		B	B, R		T-K	T, R		T-K	T-K, R
		s-dolayı			B, R			T-K, R			R
		s-dolayısıyla			B			T-K, R			R
E N D O P H O R I C	Anaphoric	*a-için*	T, T-K, R	T, R	B						
		a-dolayı	T-K, R	T-K, R	B						
		a-dolayısıyla	T-K, R	T-K, R	B						
	Cataphoric	*c-için*				R	R	R			
		c-dolayı				R	R	R			
			Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f	Arg-i	Arg-m	Arg-f

6.2 Information Structural Properties of Discourse Connectives

One should note that the information structure markings on the examples are only suggestions based on two factors: information provided by the surrounding discourse and the native speaker intuition for intonation. It is usually possible to read the same sentence in more than one way, and it is sometimes possible to force an intonation that is not intuitive, or compatible with the previous discourse. In our examples, we tried to present information structures that are both intuitive and compatible with the previous discourse. When an example forced only one of the intuitively available intonations, we tried to provide alternatives to cover a wide range of Information structures. Nevertheless, the information structural features of the connectives are only tendencies. We do not claim that the features we suggest for the connectives are exhaustive.

That said, Table 13 and Table 18 show that connectives of the same syntactic type tend to have similar information structure features when they occupy similar positions.

The coordinating conjunctions *çünkü* and *zira* tend to be included in a theme or theme-kontrast unit when they are in the so-called canonical position, which is clause-initial, and for the purposes of this study, argument-initial. Both connectives can be right-displaced, and therefore backgrounded, to occupy argument-medial and argument-final positions. It should be noted that once the connective is contraposed, there should be no difference between the argument-medial and argument-final positions from an information structure perspective. Because all the backgrounded, the connective and the material that comes before or after that, are deaccented, i.e., they carry no weight in the information flow.

Zira seems to be able to occupy argument-medial position without contraposition, and in the one and only example attested, it is included in a theme-kontrast unit. Without further examples we cannot come to a conclusion about this difference, especially because constructed examples with argument-medial *çünkü* sounds well constructed to the native speaker, too. We believe that a quantitative study of the

intonation patterns of naturally generated speech data will be necessary for a final decision on this matter.

The subordinating conjunctions, *için*, *dolayı*, and *dolayısıyla* tend to have similar information structure features, too. These subordinators and their complements can be a part of the theme or rheme when the subordinator connective is in its canonical argument-final position. When the connective is argument-medial, which occurs when some material from the subordinate clause is right-displaced from the subordinate clause, the subordinate clause tends to be included in the theme-kontrast unit, unless the caluse as a whole is right-displaced to the right of the main predicate. In these contraposed cases, the connective and its complement are backgrounded. A seemingly right-displaced subordinate clause may appear as an independent fragment, an incomplete sentence. These fragmented subordinate clauses can have a complete information structure, and the subordinator is a part of the rheme of this fragment. Purposive use of *için* and subordinator use of *dolayısıyla* were not attested in fragments, but the construction seems available to them, as well.

The information structure of endophoric connectives seem to be divided into two distinct groups for anaphoric and cataphoric connectives. Just like their argument order, their information structure tendencies differ radically but predictably. Anaphoric connectives can be included in any theme, theme-kontrast or rheme unit when in argument-initial or argument-medial positions. In argument-final positions, anaphoric connectives tend to be backgrounded. Whereas it is possible to make the connective the main predicate of a sentence and give it a rheme reading, such examples were not attested, and the constructed examples did not have total acceptability[16].

The cataphoric connectives, on the other hand, have displayed an interesting information structure profile. In all examples attested and derived from the MTC, the cataphoric connectives tend to be included in the rheme. This difference between

---

[16] Examples with modifier particles such as *de* were acceptable, but we cannot speculate on the effect of such particles on the information structure for now.

information structures of anaphoric and cataphoric connectives is sensible. Anaphora needs to be resolved from the previous discourse; as a result, anaphoric adverbials connect the utterance to the previous discourse like coordinating and subordinating conjunctions. Cataphora, on the other hand, resolves from the following discourse. The cataphoric connective point forward in the discourse, and introduces information to come. It is expected that it will display a different information structure tendency.

The distribution of information structure tendencies supports the distinction between groups of connectives such as coordinating conjunctions, subordinating conjunctions, endophoric connectives and the anaphoric vs. cataphoric distinction within endophoric connectives. Further study including spoken data and a detailed analysis of the syntactic, semantic and pragmatic factors should will reveal more about how and why the information structure of two lexical items, which come from different categories but do the same job, differ so.

6.3 Argument Order and Information Structure

With such abundance in connective-argument order and IS tendencies of connectives, does the information structure of the utterance depend on the order of its elements? Consider example (17), repeated as (94) below.

(94)   *Ortada hiçbir ipucu yok.* **<u>Çünkü</u> öldürülen yok.**

    There is no clue around. **<u>Because </u>there are no victims.**

The utterance is a licit response to the question "Why aren't there any clues around?" The connective introduces the rheme "Nobody was killed".

(95)   *Ortada hiçbir ipucu yok*. **Öldürülen yok <u>çünkü</u>**.

The variation where the connective is contraposed is also a valid answer to the same question, (maybe with the slight change in meaning where the lack of a victim is promoted to the sole reason for the absence of clues among a set of possible reasons.)

(96)   *Ortada hiçbir ipucu yok*. **<u>Zira</u> öldürülen yok**.

(97)   *Ortada hiçbir ipucu yok.* **Öldürülen yok <u>zira</u>**.

It is possible to replace çünkü with zira in both orders with the same meaning and information structure.

(98)   **Öldürülen olmadığı <u>için</u>** *ortada hiç bir ipucu yok*.

(99)   *Ortada hiç bir ipucu yok*, **öldürülen olmadığı <u>için</u>**.

(100) *Ortada hiçbir ipucu yok*, **öldürülen olmaması <u>dolayısıyla</u> / olmamasından <u>dolayı</u>**.

(101) **Öldürülen olmaması <u>dolayısıyla</u> / olmamasından <u>dolayı</u>** *ortada hiç bir ipucu yok*.

While the utterance is still an illicit response to the question when the subordinate clause precedes the theme, it is not acceptable when the subordinate clause is in the post-verbal position. Since the backgrounded elements carry no weight in the information structure, the backgrounded elements cannot be the answer to a preceding question.

Clearly argument order does not have as much significance as one would guess on the information structure. When replacing the coordinating conjunct in (94) with subordinating conjuncts (98-101), the reversed order is more acceptable. What makes a reversed argument order more acceptable is the fact that the answer to a question should be prominent but backgrounded elements cannot be prominent in the information structure, a restriction we take to be originating from prosodic constraints as mentioned in Ch.4 (Özge, 2003; Özge & Bozşahin, 2008).

Clearly, these connectives are not necessarily synonymous and one might argue that the different argument orders result from different relations or meanings. This is, in fact, the point we would like to conclude. The given, old information is not necessarily sentence-initial and the new information is not necessarily immediately pre-predicative. The position of a discourse argument with certain information

structure depends on a combination of factors, including the lexical differences between individual connectives.

Our analysis concurs with Özge and Bozşahin's statement that information structure should not be linked directly to surface position, and we propose to add, neither on sentence level nor on discourse level.

6.4 Conclusions

In this thesis, five Turkish discourse connectives were analyzed in their eleven uses. The connective argument orders and the distribution of information units on the arguments were inspected.

Turkish exploits nine of possible connective argument orders in one way or the other. Subordinating conjunctives that construct adverbial clauses have the most flexible positioning, with one restriction that dictates subordinators cannot be argument initial due to mopho-syntactic restrictions. The discourse adverbial counterparts of discourse subordinators are not as flexible because of argument orders imposed by endophora resolution.

The information structure tendencies of the discourse connectives reveal that the syntactic type of the connective has some defining role on its information structure as well as the position of the connective and the order of the argments, but it is not the only factor as the clear distincltion between the anaphoric and cataphoric adverbials demonstrates.

The information structure of an utterance does not necessarily correlate with the order of the arguments. An expression can be conveyed with exactly opposite argument order with a different discourse connective but same information structure and meaning.

REFERENCES

Asher, N. (1993). *Abstract Objects in Discourse.* Dordrecht: Kluwer.

Asher, N., & Lascarides, A. (1998). The semantics and pragmatics of presupposition. *Journal of Semantics, 15*(3), 239-300.

Bozşahin, C. (2002). The combinatory morphemic lexicon. *Computational Linguistics, 28*(2), 145–176.

Büring, D. (1997). *The meaning of topic and focus : the 59th Street Bridge accent.* London ; New York: Routledge.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Cresswell, C., Forbes, K., Miltsakaki, E., Prasad, R., Joshi, A., & Webber, B. (2002). *The Discourse Anaphoric Properties of Connectives* Paper presented at the 4th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC).

Cresswell, C., Forbes, K., Miltsakaki, E., Prasad, R., Joshi, A., & Webber, B. (2005). The predicate argument structure of discourse connectives: a corpus based study. In A. Branco, T. McEnery & R. Mitkov (Eds.), *Anaphora Processing, Linguistic, Cognitive and Computational Modelling*: John Benjamins Publishing.

Erguvanlı, E. E. (1979). *The Function of Word Order in Turkish Grammar.* Unpublished Ph.D, University of California, Los Angeles.

Forbes-Riley, K., Webber, B., & Joshi, A. (2006). Computing Discourse Semantics: The Predicate-Argument Semantics of Discourse Connectives in D-LTAG. *Journal of Semantics, 23*, 55-106.

Forbes, K. (2003). *Discourse Semantics of S-Modifying Adverbials.* Unpublished PhD, University of Pennsylvania.

Forbes, K., Miltsakaki, E., Prasad, R., Sarkar, A., & Joshi, A. (2003). D-LTAG System: Discourse Parsing with a Lexicalized Tree-Adjoining Grammar. . *Journal of Logic, Language and Information, 12*(3), 261–279.

Forbes, K., & Webber, B. (2002). *A Semantic Account of Adverbials as Discourse Connectives* Paper presented at the 3rd SIGDial Workshop on Discourse and Dialogue, Philadelphia PA.

Göksel, A., & Özsoy, S. A. (2000). *Is there a focus position in Turkish?* Paper presented at the Studies on Turkish and Turkic Languages; Proceedings of the Ninth International Conference on Turkish Linguistics, Harrassowitz, Wiesbaden.

Grosz, B. J., & Sidner, C. L. (1986). Attention, intention and the structure of discourse. *Computational Linguistics, 12*(3), 175-204.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Halliday, M. A. K., & Matthiessen, C. (2004). *An Introduction to Functional Grammar* (Third ed.). New York: Oxford University Press.

Hoffman, B. (1995). *The Computational Analysis of the Syntax and Interpretation of "Free" Word Order in Turkish.* Unpublished Ph.D, University of Pennsylvania.

İşsever, S. (2003). Information structure in Turkish: the word order–prosody interface *Lingua, 113*(11), 1025-1053.

Joshi, A. (1987). An Introduction to Tree Adjoining Grammar. In A. Manaster-Ramer (Ed.), *Mathematics of Language*. Amsterdam: John Benjamins.

Kılıçaslan, Y. (1994). *Information Packaging in Turkish.* Unpublished Ph.D., University of Edinburgh.

Kılıçaslan, Y. (2004). Syntax of information structure in Turkish. *Linguistics, 42*(4), 717.

Knott, A. (1996). *A Data Driven Methodology for Motivating a Set of Coherence Relations.* Unpublished PhD, University of Edinburgh.

Knott, A., Oberlander, J., O'Donnel, M., & Mellish, C. (2001). Beyond elaboration: The interaction of relations and focus in coherent text. In T. Sanders, J. Schilperoord & W. Spooren (Eds.), *Text Representation:Linguistic and psycholinguistic aspects* (pp. 181-196): John Benjamins Publishing.

Komagata, N. (1997). *Efficient parsing for CCGs with generalized type-raised categories.* Paper presented at the Proceedings of the International Workshop on Parsing Technologies (IWPT), Cambridge, MA.

Komagata, N. (1999). *A computational analysis of information structure using parallel expository texts in English and Japanese.* Unpublished Ph.D, University of Pennsylvania.

Kornfilt, J. (1997). *Turkish*. New York: Routledge.

Kucuk, D., & Yondem, M. T. (2007). *Automatic identification of pronominal Anaphora in Turkish texts.* Paper presented at the ISCIS 2007. 22nd International International Symposium on Computer and Information Sciences.

Mann, W. C., & Thompson, S. A. (1987). *Rhetorical Structure Theory: A Theory of Text Organization* (No. ISI/RS-87-190). Los Angeles, CA: Information Sciences Institute. (M. d. Rey o. Document Number)

Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text, 8*(3), 243-281.

Miltsakaki, E., Cresswell, C., Forbes, K., Joshi, A., & Webber, B. (2003). *Anaphoric arguments of discourse connectives: Semantic properties of antecedents versus non-antecedents.* Paper presented at the EACL Workshop on Computational Treatment of Anaphora, Budapest, Hungary.

Özge, U. (2003). *A Tune-Based Account of Turkish Information Structure.* Unpublished M.S. Thesis, Middle East Technical University, Ankara.

Özge, U., & Bozşahin, C. (2008). Intonation in the Grammar of Turkish. *In submission*.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation.* Unpublished Ph.D. Thesis, Massachusetts Institute of Technology.

Pierrehumbert, J. B., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan & M. Pollack (Eds.), *Intentions in Communications* (pp. 271-312). Cambridge, MA: MIT Press.

Prasad, R., Miltsakaki, E., Dinesh, N., Lee, A., & Joshi, A. (2007). *The Penn Discourse Tree Bank Annotation Manual*.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

Say, B., Zeyrek, D., Oflazer, K., & Özge, U. (2002). *Development of a Corpus and a Treebank for Present-day Written Turkish.* Paper presented at the 11th International Conference on Turkish Linguistics.

Schabes, Y., & Joshi, A. (2002). Tree-adjoining grammars In G. Rozenberg & A. Salomaa (Eds.), *Handbook of Formal Languages* (Vol. 3, pp. 69-124). Berlin: Springer-Verlag.

Steedman, M. (1991). Structure and Intonation. *Language, 67*(2), 260-296.

Steedman, M. (1995). Computational aspects of the theory of grammar. In D. N. Osherson, L. R. Gleitman & M. Liberman (Eds.), *An invitation to cognitive science, Vol. 1: Language.* Cambridge, MA: MIT Press.

Steedman, M. (2000a). Information Structure and the Syntax-Phonology Interface. *Linguistic Inquiry, 31*(4), 649-689.

Steedman, M. (2000b). *The Syntactic Process*. Cambridge, MA: MIT Press.

Temürcü, C. (2001). *Word Order Variations in Turkish: Evidence from Binding and Scope.* Unpublished M.S., Middle East Technical University, Ankara.

Temürcü, C. (2005). The Interaction of Syntax and Discourse in Word Order Variability: Data from Turkish. In G. König, I. Özyıldırım, D. Aydın & A. Altan (Eds.), *Dilbilim ve Uygulamaları* (pp. 123-159). İstanbul: Multilingual.

Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam/Philadelphia: John Benjamins.

Vallduvi, E. (1990). *The Informational Component.* Unpublished Ph.D., University of Pennsylvania.

Vallduvi, E., & Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics, 34*, 459-519.

Webber, B. (2004). D-LTAG: Extending Lexicalized TAG to Discourse *Cognitive Science, 28*(5), 751-779.

Webber, B. (2006). *A Lexical Perspective on Discourse Structure and Semantics.* Paper presented at the SPRIK conference, Oslo.

Webber, B., & Joshi, A. (1998). *Anchoring a lexicalized tree-adjoing grammar for discourse.* Paper presented at the Coling/ACL Workshop on Discourse Relations and Discourse Markers, Montreal, Canada.

Webber, B., Joshi, A., Miltsakaki, E., Prasad, R., Dinesh, N., Lee, A., et al. (2005). *A Short Introduction to the Penn Discourse TreeBank* Paper presented at the Copenhagen Working Papers in Language and Speech Processing.

Webber, B., Knott, A., & Joshi, A. (1999). *Multiple Discourse Connectives in a Lexicalized Grammar for Discourse* Paper presented at the Third International Workshop on Computational Semantics, Tilburg, The Netherlands.

Webber, B., Knott, A., Stone, M., & Joshi, A. (1999). *Discourse Relations: A Structural and Presuppositional Account using Lexicalised TAG* Paper presented at the Meeting of the Association for Computational Linguistics, College Park MD.

Webber, B., Stone, M., Joshi, A., & Knott, A. (2003). Anaphora and Discourse Structure. *Computational Linguistics, 29*, 545-587.

Zeyrek, D., Turan, Ü. D., & Demirşahin, I. (2008). *Structural and Presuppositional Connectives in Turkish.* Paper presented at the Constraint in Discourse III, Potsdam, Germany.

Zeyrek, D., & Webber, B. (2008). *A Discourse Resource for Turkish: Annotating Discourse Connectives in the METU Corpus.* Paper presented at the The 6th workshop on Asian Language Resources, The 3rd International Joint Conference on Natural Language Processing (IJNLP), Hyderabad, India.

APPENDICES

APPENDIX A – ANNOTATION SAMPLE

Agreed annotation of çünkü by 5 annotators

1. Koca bir duvar taşıyordun yüreğinde kimsenin aşamayacağı, aşmaya cesaret bile edemeyeceği. Dışa karşı güçlüydü, ama içe, kendi yüreğine yıkılmak üzereydi. Anılarla örülmüş, acılarla harçlanmış bu duvara tırmanmak, onu aşabilmenin ilk şartıydı. Vazgeçmek kolaydı, ertelemek de. *Ama tırmanmaya başlandı mı bitirilmeli!* <u>Çünkü</u> **her seferinde acımasız bir geriye dönüş vardı.** Bıraktığın her sefer bir başlangıca gebeydi. Bir aşsaydım bu duvarı benim olacaktın, kucaklayacaktın beni. Kırgınlıkların, korkuların eriyip gidecekti, hepsi benim olacak, bana geçecekti. Ben kıvranacaktım, ben acı çekip işkencelere gönüllü katlanacaktım senin yerine, sen bilmeden.
2. *Artık gündüzleri terasta kalmak zorundaydı,* <u>çünkü</u> **hamam gündüzleri açıktı.** Dikkat çekmemek için geceden gidip terasa gizleniyordu. Hamamcı kadın geldiğinde iyice kıvrılarak köşede soluğunu tutuyordu. Zaten kadın akşamdan astığı çamaşırları aceleyle topluyor ve çevresine hiç bakmadan hemen gidiyordu. Halil ' i görmesi olanaksızdı.
3. Ustaca işlenmiş bir cinayet. (IMP) *Ortada hiçbir ipucu yok.* <u>Çünkü</u> **öldürülen yok**. Ama bir insanın rayı değiştiriliyor ; başka bir yaşamın içine sokuluyor.
4. Kimi gün Neslihan oluyor yanında. Koluna giriyor Vedat ' ın. Gözlükleri yine burnunun ucunda, saçları kısacık. Keyfi yerindeyse çok konuşuyor Vedat. *Nesli ' nin, koluna asılıp gözlüklerinin üstünden şehla şehla bakmasını ve gülmesini çok*

seviyor, <u>çünkü</u> **gülmek çok yakışıyor ona.** Benim amcam kaptandı, diyor durup dururken ve Nesli ' nin araştıran şımarık gözlerine bakıyor. İnan ki, diyor. Ben ip cambazı olduğunu sanıyordum, diyor Nesli de. Evet, küçük amcamdı o, nur içinde yatsın, yetmişlik bir rakıyı devirip ipi sek sek geçmeye kalkmış ; kaptan olan amcam ise kocaman bir gemiyi sulara gömdü. Aylardan kasımdı, ben çocuktum, çok iyi anımsıyorum, fırtınalı bir gecede, Karadeniz ' in batısında batmışlardı. Kaptandı, ama yüzme bilmezdi amcam. Bir namaz tahtasına sarılmış olarak kıyıya vurduğunda kollarını zor açmışlar, yarı yarıya donmuş. *Belki de o anda Tanrı ' ya yakarıp yardım istiyordu,* <u>çünkü</u> **çok dindar bir adamdı.** Ama artık değil ; küp gibi içip meyhanelerde keman çalıyor. Sonra da Nesli ' nin ilgiyle çatılmış alnına bakıp gülüyor: Çok istavritsin! Ne demek o, diyor. Yani oltaya çabuk geliyorsun. Amaan, diyor Nesli koluna yeniden asılarak, İyi atıyorsun, soprano teyzelerinin İtalya serüvenlerine ne zaman geçeceksin. Keyifsiz olduğu günler de konuşuyor Vedat, ama daha az. Sanayi çarşısına gidelim, diyor örneğin, sana orada türlerine yabancılaşmış bahtsız köpekler göstereceğim, bahtsız ve kadın adıyla çağrılan ve kadın adıyla çağrılınca koşarak gelen ve gelirken kuyruğunu kaldırıp sallayan. *Nesli gülüyor hiçbir şey anlamadan.* Oyun gibi geliyor anlattıkları. <u>Çünkü</u> **gülmek çok yakışıyor ona.** Ayı, kendini dine vermiş, ne radyo ne televizyon. Sakalı her gün biraz daha uzuyormuş. Bir gün herkes, Ayı ' nın, cüppesini dalgalandırarak ilçe semalarında uçtuğunu görecek. Ayı uçup gidecek, bir daha da geri dönmeyecek. Bir Sıntır ' ı tanıtacağım sana, diyor ; Nesli de, Kimi kimi? deyip asılıyor koluna. Sıntır ' ı, Sıntır ' ı, diyor, duymadın mı hiç? İnsan kendine acımamalı Nesli, ama bize yazık oldu biliyor musun? Bak İlhami ' ye, soru sormadan yaşamasını biliyor. Yazgı, diyor, iş bitiyor.

5. Sulu gözlüydün Nesli, bunu ne kadar gizlemeye çalışsan da (utanıyordun değil mi, bir devrimciye yakışmayan güçsüzlüklerdi bunlar) başaramıyordun. *Biz yasalar karşısında evli sayılacak, ama gerçekte evli iki insan gibi değil de* (**evlilikler sıradanlaşıyordu** <u>çünkü</u>, tekdüze ve sıkıcıydı ; biz farklı olacaktık), *aynı evi paylaşan iki öğrenci gibi yaşayacaktık.* Sürekli yenilenen bir birliktelikti amacımız. Annelerimiz, amcalarımız, ağabeylerimiz gibi olmamalıydık. *Oysa evlilik üzerine hiçbir şey bilmiyorduk.* <u>Çünkü</u> **hiç evlenmemiştik,** yaşanmadan nasıl bilinirdi ki...

6. Şam tatlısı satan, suratında kocaman beni olan bir adam vardı. İspirto içerdi. Bir kış günü sinemanın önünde ölüsü bulundu. Her iki salondaki sandalyelerden kıçının izi silinmemiştir Vedat ' ın. Hatta bir bilet alıp o salona girse, film o anda gösteriliyor olsa (Kızılderilileri acımasızca katleden John Wayne ya da dişleri arasına sıkıştırdığı bir kamayla timsahların kaynadığı suya dalan Tarzan) ve el yordamıyla oturacak bir yer arasa kendine... Yanı başında Vedat ' ın oturduğunu ve kucağında bir tomar resimli romanı sıkı sıkı tuttuğunu neden sonra fark etse...

Birinci film bitip de (**dört film birden gösteriliyor** <u>çünkü</u>, günlerden pazar) ışıklar yanınca onun ardından kalksa, gazoz ve sigara içilen o bölümde kitap alıp sattığını ya da değiştirdiğini görse... Göz göze gelirlerdi, inanıyordu. Ama küçük Vedat tanımazdı onu ve acı dolu yanlış yollardan yürümemesini haykıran sesini de duymazdı... Neden giremiyordu şimdi o salona? Engel olan, onu tutan neydi? İçeri girdiğinde görmek istemedikleri mi? Anılar mı? Yıldız Sineması. Sandalyeler beşer onar yan yana getirilip arkalarından çıtalarla birbirine çakılmış. O sandalyeler bile değişmemiştir, bunca şey değişir, çözülüp giderken. Bir pazar matinesinde, yıllar önce, öndeki sırada kafaları sıfır numara tıraşlı askerler otuzbir çekerlerken, beşi birden sıra ile birlikte nasıl da gerisin geri devrilmişlerdi... Herkes gülerken, kimi de yuh çekerken - askerdir, hoş görülür aslında - onlar da önlerini kapamaya çalışırken neye uğradıklarını şaşırmışlardı.

7. *Dilindeyse kırık dökük bir ezgi, bozuk bir plak gibi aynı şeyi yineliyor : Seni yakacaklar... Ardından da bir sazın telleri gibi titreşiyor dili. Onu bırakıyor bir başkasını yakalıyor ; Belki çıkmam sabaha...* <u>Çünkü</u> **Yakup ' un kötü teybinde hep bunlar çalıyor.** Yakup ve Yakup ' un müşterileri bunları seviyor. Sigara gibi, alışılınca bir türlü bırakılamayan, iyice saran ve yarını belirsiz bir yaşamın ezgileri olan kara sesler.

8. Evden ayrılmıştı. Okula da gelmiyordu. Görüşemiyorlardı. Köyüne gömmüşlerdi. İsteseydi, bir çırpıda o dağ köyüne gider, son görevini yapardı. Görev! Ama o, Mesut ' u bir köy alanında gülümserken bırakmıştı. Mesut hep gülümsesin diye omuzlamak istememişti kanlı bir tabutu. Kimse bilmiyor bunu. Kınıyorlar ; yoktun, diyorlar. *Yoktum,* <u>çünkü</u> **Mesut da yoktu**, diyemiyor. Senden bunu beklemezdim. Neyi beklemezdin Nesli? Her yer karla kaplıydı, bembeyaz. Avlanırız, demişti Mesut, ateş ederiz! Şimdi orada yatıyor. O köyde. Yanından geçtiğimiz ve duvarına yaslanıp sigara yaktığımız o mezarlıkta

9. Bir gün korkunç bir temizlik yapacağım. Kitaplarımı yaktım, biliyor musun? Bir gece ellerim titreyerek yaktım onları. Utanç içindeydim. Sanki bir yararı olacaktı, parçalayıp parçalayıp sobaya atıyordum. *Buna şaşmayacaksınız,* **yeni bir şey değil** <u>çünkü</u>. Bizden önceki kuşaklar da yakmıştı. Olanları kitaplardan okurken, artık böyle şeyler yaşanmaz, diye düşünmüştük. Dizelerini kabaralı çivilerle beynime çaktığım o şiirleri de attım sobaya. Alevler daha sarmamıştı arka kapaktaki yüzünü, bir başkaydı gülüşü.

10. Çayı bitiyor. Çay dediği ne ki, hepsi üç yudum. Yakup ' a bakıyor, bugün ilkyaz, diyor, kırlara doğru süreceğim bisikletimi ; *ama sözcükleri ulaşamıyor ona.* <u>Çünkü</u> **Yakup kendi havasında** ve Şaşı Ramazan ' a para üstü veriyor kirli önlüğünü bir yayık gibi sallayarak. Göz göze gelseler, Tazeleyim mi abi, diyecek, biliyor.

11. Hani aydın kimliğiniz nerede? Yobazlar yarın hepimizi kesecekler! Tarafsız kalanları bile! Onlardan olmayan herkesi yakacaklar anlıyor musunuz? Off, gene

testerelerle paranoyak ve modası geçmiş laisizm edebiyatına kırdın direksiyonu güzelim, diyor Metin. *Bence sen bu gözlemcilik işini bırakıp bir tavuk çiftliği kurmalısın*, <u>çünkü</u> **aklının ermediği kaygan zeminlerde gezinmektesin.** Her an bir kazaya uğrayabilirsin. Sen konuşma, hedefim sen değilsin, sen ananı bile satarsın, diyor Tuhfe. Kesin artık, diye bağırıyor Erkböke, tamam, öncelikli tavukların ve domateslerin kurtarılması gerekiyor. Haydi, iç rüzgar dalga geçiyoruz, yarasın! Bardağını kaldırıyor.

12. Yarın sabah mı gideceksin? diye soruyorum.

Akşam. Gündüz yolculuğunu hiç sevmem. Ne zaman döneceğin belli değil tabii.

Değil. Dükkânla dayım ilgilenecek. *Uzun sürmez umarım.*

Evet, <u>çünkü</u> **ben de bekleyeceğim seni**. Sen olmadan yazmayı sürdüremem.

.

Appendix B – SUMMARY TABLE OF CONJUNCTIVE RELATIONS

Table 19 - Summary Table of Conjunctive Relations (Halliday and Hasan, 1976 p. 242-243)

	External/internal	Internal (unless otherwise specified)		
Additive	Additive, simple: Additive — *and, and also* Negative — *nor, and . . . not* Alternative — *or, or else*	Complex, emphatic: Additive — *furthermore, in addition, besides* Alternative — *alternatively* Complex, de-emphatic: Afterthought — *incidentally, by the way*	Apposition: Expository — *that is, I mean, in other words* Exemplificatory — *for instance, thus*	Comparison: Similar — *likewise, similarly, in the same way* Dissimilar — *on the other hand, by contrast*
Adversative	Adversative 'proper': Simple — *yet, though, only* Containing 'and' — *but* Emphatic — *however, nevertheless, despite this*	Contrastive: Avowal — *in fact, actually, as a mater of fact* Contrastive (external): Simple — *but, and* Emphatic — *however, on the other hand, at the same time*	Correction: Of meaning — *instead, rather, on the contrary* Of wording — *at least, rather, I mean*	Dismissal: Closed — *in any case, in either case, whichever way it is* Open-ended — *in any case, anyhow, at any rate, however it is*
Causal	Causal, general: Simple — *so, then, hence, therefore* Emphatic — *consequently, because of this* Causal, specific: Reason — *for this reason, on account of this* Result — *as a result, in consequence* Purpose — *for this purpose, with this in mind*	Reversed causal: Simple — *for, because* Causal, specific: Reason — *it follows, on this basis* Result — *arising out of this* Purpose — *to this end*	Conditional (also external): Simple — *then* Emphatic — *in that case, in such an event, that being so* Generalized — *under the circumstances* Reversed polarity — *otherwise, under other circumstances*	Respective: Direct — *in this respect, in this regard, with reference to this* Reversed polarity — *otherwise, in other respects, aside from this*
Temporal	Temporal, simple (external only): Sequential — *then, next, after that* Simultaneous — *just then, at the same time* Preceding — *previously, before that* Conclusive: Simple — *finally, at last* Correlative forms: Sequential — *first . . . then* Conclusive — *at first . . . in the end*	Complex (external only): Immediate — *at once, thereupon* Interrupted — *soon, after a time* Repetitive — *next time, on another occasion* Specific — *next day, an hour later* Durative — *meanwhile* Terminal — *until then* Punctiliar — *at this moment*	Internal temporal: Sequential — *then, next, secondly* Conclusive — *finally, in conclusion* Correlative forms: Sequential — *first . . . next* Conclusive — *. . . finally*	'Here and now': Past — *up to now, hitherto* Present — *at this point, here* Future — *from now on, henceforward* Summary: Summarizing — *to sum up, in short, briefly* Resumptive — *to resume, to return to the point*

Appendix C - PDTB SENSE HIERARCHY

TEMPORAL
- Asynchronous
- Synchronous
  - precedence
  - succession

CONTINGENCY
- Cause
  - reason
  - result
- *Pragmatic Cause*
  - *justification*
- Condition
  - hypothetical
  - general
  - unreal present
  - unreal past
  - factual present
  - factual past
- *Pragmatic Condition*
  - *relevance*
  - *implicit assertion*

COMPARISON
- Contrast
  - juxtaposition
  - opposition
- *Pragmatic Contrast*
- Concession
  - expectation
  - contra-expectation
- *Pragmatic Concession*

EXPANSION
- Conjunction
- Instantiation
- Restatement
  - specification
  - equivalence
  - generalization
- Alternative
  - conjunctive
  - disjunctive
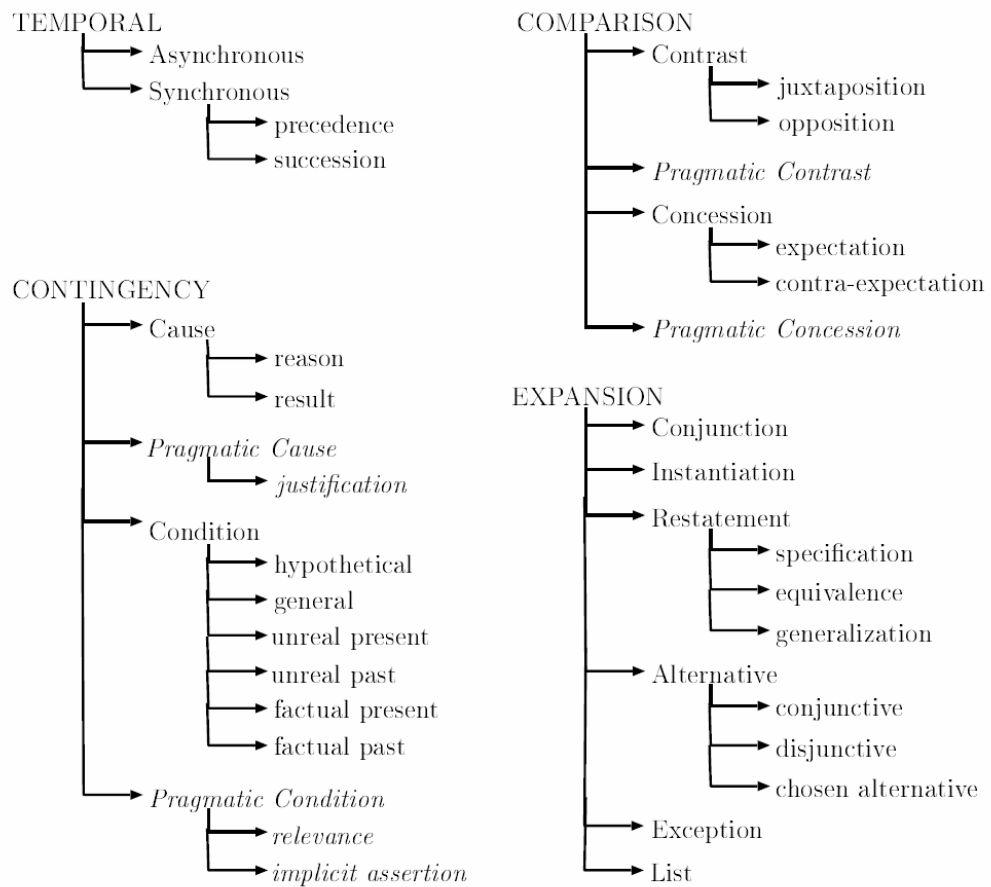 - chosen alternative
- Exception
- List

Figure 20 - The hierarchy of sense tags in PDTB

(Prasad et al., 2007 p. 32)