

AN EXPERIMENTAL STUDY ON ABSTRACT ANAPHORA RESOLUTION IN
TURKISH WRITTEN DISCOURSE

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

RABIYE ERGİN SOMER

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF SCIENCE
IN
THE DEPARTMENT OF COGNITIVE SCIENCE

JULY 2012

**AN EXPERIMENTAL STUDY ON ABSTRACT ANAPHORA RESOLUTION
IN TURKISH WRITTEN DISCOURSE**

Submitted by **RABIYE ERGİN SOMER** in partial fulfillment of the requirements for the degree of **Master of Science in The Department of Cognitive Science, Middle East Technical University** by,

Prof. Dr. Nazife BAYKAL
Director, **Informatics Institute**

Prof. Dr. Cem BOZŞAHİN
Head of Department, **Cognitive Science**

Prof. Dr. Deniz ZEYREK BOZŞAHİN
Supervisor, **Cognitive Science, METU**

Assist. Prof. Dr. Cengiz ACARTÜRK
Co-Advisor, **Cognitive Science, METU**

Examining Committee Members

Prof. Dr. Deniz ZEYREK BOZŞAHİN
Cognitive Science, METU

Prof. Dr. Cem BOZŞAHİN
Cognitive Science, METU

Assist Prof. Dr. Cengiz ACARTÜRK
Cognitive Science, METU

Assist. Prof. Dr. Murat Perit ÇAKIR
Cognitive Science, METU

Assist. Prof. Dr. Bilal KIRKICI
Foreign Language Education, METU

Date: 05.07.2012

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 materials and results that are not original to this work.

Name, Last name : Rabiye Ergin Somer

Signature : _____

ABSTRACT

AN EXPERIMENTAL STUDY ON ABSTRACT ANAPHORA RESOLUTION IN TURKISH WRITTEN DISCOURSE

Ergin Somer, Rabiye
MS., Department of Cognitive Science
Supervisor: Prof. Dr. Deniz Zeyrek
Co-advisor: Assist Prof. Dr. Cengiz Acartürk

July 2012, 185 pages

This thesis provides an experimental approach to abstract anaphora resolution in Turkish written discourse. The core of this work consists of identifying various manifestations of abstract anaphoric expressions –*bu* vs. *bu durum*, *bu olay*, *bu iş*, *bu gerçek* (*bu* as the bare abstract object anaphor vs. *bu+label* abstract anaphors)- in Turkish discourse, and investigating whether any difference is observed in their processing. To this end, two offline experiments are conducted with human subjects, and the results indicate that *label anaphors*, compared to the bare anaphor *bu*, have a tendency to disambiguate the antecedent in some cases.

Keywords: Discourse, anaphora, anaphora resolution, referent, antecedent, ambiguity

ÖZ

TÜRKÇE YAZILI SÖYLEMDE SOYUT ARTGÖNDERİMSSEL İFADELER ÜZERİNE DENEYSEL BİR ÇALIŞMA

Ergin Somer, Rabiye

Yüksek Lisans, Bilişsel Bilimler Bölümü

1. Danışman ve Tez Yöneticisi: Prof. Dr. Deniz Zeyrek

2. Danışman: Yrd. Doç. Dr. Cengiz Acartürk

Temmuz 2012, 185 sayfa

Bu tez, yazılı Türkçe söylemde soyut artgönderimsel ifadelerin çözümlenmesine bilişsel ve deneysel bir bakış açısı sunmaktadır. Çalışmanın temelini, çeşitli soyut artgönderimsel ifadelerin *-bu* ile *bu durum*, *bu olay*, *bu iş*, *bu gerçek* (yalın bir şekilde *bu* soyut artgönderimsel ifadesiyle, yine soyut bir artgönderimsel ifade olan *bu+ad*) ifadelerinin- Türkçe yazılı söylemdeki çeşitli gösterimlerini belirlemek, ve çözümlenirken herhangi bir farklılık gözlemlenip gözlemlenmediğini incelemek oluşturmaktadır. Bu amaçla, insanlar üzerinde 2 adet deney yapılmıştır. Sonuçlar, *bu+ad* biçimindeki soyut artgönderimsel ifadelerin, yalın *bu* artgönderimsel ifadelerine kıyasla, bazı durumlarda gönderimde buldukları ifadelerdeki belirsizliği giderdiklerini göstermektedir.

Anahtar kelimeler: Söylem, artgönderimsel ifade, artgönderimsel çözümlenme, öncül, gönderge, çokanlamlılık

To my beloved one: Doğa.

ACKNOWLEDGEMENTS

I would like to start giving my special thanks to Prof. Dr. Deniz Zeyrek for her help and support from the very beginning of my attempts to establish myself in the field of Cognitive Science. It's mostly thanks to her encouragement that I found a particular future plan of action to follow, and the encouragement that I needed to achieve my career goals in the dynamic atmosphere of an interdisciplinary department. It's again mostly thanks to her that this thesis was shaped and finalized. I'm very grateful to Assist. Prof. Cengiz Acartürk, who was patient, supportive, and a real guide with the endless questions of a novice researcher in the experimental studies.

Millions of thanks to Tuna Çakar, my dear friend, for his help in the courses we took together, for his jokes, comments and ideas that made the difficulties manageable in our effort to be able to understand the fundamentals, trends, methods, techniques, approaches and various research in different fields and the combination of them in Cognitive Science. Thanks a lot to Cihan Selçuk whose friendship has been invaluable to me as well as his help in statistical and computational models during my MSc. studies at METU, and also his help for the statistical analysis of this thesis. Thanks a lot to Elif Gök for her moral support and her effort in getting this thesis signed in various cities of Turkey. I'm also very thankful to İhsan Yalçınkaya with whose help it became possible to evaluate my findings in this study through a non-parametric statistical analysis. I cannot forget my students taking part in my experiments; I thank them all.

I would like to thank METU Cognitive Science Department for the collaborative atmosphere formed by the researchers, academicians and graduate students from a variety of fields interested in all aspects of cognition and the stimulating environment to develop myself and to conduct research on linguistics and

psychology.

Special thanks to Dr. Simon Phipps for encouraging me to leave an already existing career behind at Bilkent University, apply for an MSci degree in Cognitive Science at METU.

I owe very special thanks to my dear friends, Fatma Gül Yıldırım, Hafize Eğir, Onur Özalpuk, Fatma Şen, Salih Yıldırım, Murat Yıldırım, Ali Fuat Zırhlı, Doğa Gürsoy and Onur Gülçat for taking part in my pilot studies, for their comments and feedback on my experiments, and I thank them all just because of their existence in my life.

Greatest thanks to my little ones; to Ruby for spending most of her time in front of the computer with me while I was writing this thesis, for keeping me sit down to finish it by sleeping on my lap, giving me hugs from time to time as moral support and setting the rythm for my pace with her lovely purr sounds; to Lilly for keeping me active by continuously bringing her ball for a fetch game and waking me up when I fell asleep in front of the computer; and to my miniature lion, Suzy, for guarding my study room against Lilly by not letting her in when I was fully concentrated.

Finally, I'm deeply indebted to Doğa Somer, my beloved husband, for his endless support not only in this thesis, but in every phase of my life. With love and gratitude, this thesis is dedicated to him.

TABLE OF CONTENTS

ABSTRACT.....	iv
ÖZ	v
DEDICATION	vi
ACKNOWLEDGEMENTS	vii
TABLE OF CONTENTS	ix
CHAPTER	
1.INTRODUCTION	1
1.1. Basic Notions	1
1.2. Types of Abstract Anaphora	3
1.2.1. Labelling	3
1.2.2. Situation Reference	7
1.2.3. Discourse Deixis	9
1.3. Scope of the Current Work	10
2. THEORETICAL AND METHODOLOGICAL FRAMEWORKS IN ABSTRACT OBJECT ANAPHORA STUDIES.....	12
2.1. Abstract Object Anaphora Studies	13
2.1.1. Abstract Object Anaphora Studies in English.	13
2.1.2. Abstract Object Anaphora Studies in Other Languages	16
2.1.3. Abstract Object Anaphora Studies Focusing on Contrastive Analysis of Several Languages	19
2.2 Resolution of Abstract Object Anaphors	20
2.2.1. Abstractness Hierarchy	22
2.2.2. Right Frontier Constraint.....	25
2.2.3. Segmented Discourse Representation Theory.	28

2.2.4 Inferencing within the Framework of Centering Theory	30
2.3 Ambiguous Nature of the Antecedents to Abstract Object Anaphors.....	33
2.4. Methodological Frameworks in Anaphora Annotation Studies	36
2.4.1 Krippendorff's Alpha.....	38
2.4.2 Passonneau's Proposal.....	40
3. A VIEW FROM TURKISH	42
3.1. The Data Source: the Turkish Discourse Bank (TDB)	43
3.2. Types of Abstract Object Anaphors.....	43
3.2.1. Labeling in Turkish	44
3.2.1.1. Illocutionary Nouns.....	45
3.2.1.2. Language Activity Nouns.....	47
3.2.1.3. Mental Process Nouns.....	49
3.2.1.4. Text Nouns.....	51
3.2.2. Situation Reference in Turkish	53
3.2.3. Discourse Deixis in Turkish.....	54
3.3. Anaphora Resolution in Turkish	55
3.3.1. Abstractness Hierarchy	56
3.3.2. Right Frontier Constraint in Turkish.....	58
3.4. Anaphora Studies in Turkish.....	60
4. THE EXPERIMENTS	63
4.1. The Pilot Study	64
4.2. Experiment 1	67
4.2.1 Experiment Design	67
4.2.2 Analysis of the Data	68
4.2.3 Results	71

4.2.4 Discussion	91
4.3. Experiment 2	93
4.3.1 Experiment Design	93
4.3.2 Analysis of the Data	94
4.3.3 Results	97
4.3.4 Discussion	112
4.4 General Discussion	122
5. SUMMARY, CONCLUDING REMARKS AND FUTURE WORK	125
REFERENCES.....	130
APPENDICES	136
APPENDIX 1: EXPERIMENT 1 TEST MATERIAL FOR GROUP 1.....	136
APPENDIX 2: EXPERIMENT 1 TEST MATERIAL FOR GROUP 2.....	138
APPENDIX 3: ANNOTATION GUIDELINE FOR GROUP 1.....	140
APPENDIX 4: ANNOTATION GUIDELINE FOR GROUP 2.....	143
APPENDIX 5: START-END WORD CONSISTENCY ANALYSIS FOR THE PILOT STUDY.....	146
APPENDIX 6: START WORD CONSISTENCY ANALYSIS FOR EXPERIMENT 1.....	150
APPENDIX 7: END WORD CONSISTENCY ANALYSIS FOR EXPERIMENT 1.....	153
APPENDIX 8: PASSONEAU (2005) ANALYSIS FOR EXPERIMENT 1.....	156
APPENDIX 9: PASSONEAU (2005) ANALYSIS WITHOUT ADJUNCTS FOR EXPERIMENT 1.....	160

APPENDIX 10: EXPERIMENT 2 TEST MATERIAL FOR GROUP 1.....	163
APPENDIX 11: EXPERIMENT 2 TEST MATERIAL FOR GROUP 2.....	165
APPENDIX 12: START WORD CONSISTENCY ANALYSIS FOR EXPERIMENT 2.....	167
APPENDIX 13: END WORD CONSISTENCY ANALYSIS FOR EXPERIMENT 2.....	170
APPENDIX 14: PASSONEAU (2005) ANALYSIS FOR EXPERIMENT 2.....	173
APPENDIX 15: FREQUENCY TABLE OF START&END WORDS IN EXPERIMENT 1.....	177
APPENDIX 16: FREQUENCY TABLE OF START&END WORDS IN EXPERIMENT 2.....	179
APPENDIX 17: ANALYSIS OF THE MOST FREQUENT TYPES OF THE PHRASES MARKED BY THE PARTICIPANTS IN EXPERIMENT 1.....	181
APPENDIX 18: ANALYSIS OF THE MOST FREQUENT TYPES OF THE PHRASES MARKED BY THE PARTICIPANTS IN EXPERIMENT 2.....	183

CHAPTER 1

INTRODUCTION

This thesis presents an experimental study conducted with human subjects on manifestations of abstract entity anaphors and their resolution in Turkish discourse. Just like the other linguistic devices (e.g. connectives) maintaining coherence in a text, anaphoric expressions in discourse are linguistic devices helping the text be a coherent whole rather than a random collection of words, phrases or sentences.

In this chapter, the basic notions regarding anaphoric expressions, i.e., what anaphora resolution means, types of anaphora (i.e. anaphors referring to abstract vs. concrete objects) will be explained (Section 1.1). Also, types of abstract entity anaphors will be presented (Section 1.2). To clarify these points, examples from English discourse will be provided. Finally, scope of the current work (Section 1.3) will be explained.

1.1 Basic Notions

An *anaphor* is a device defined as the dependence of an expression referring to another which has already been introduced in the discourse (Botley, 2006, p. 73). The entity to which it refers back to or for which it stands is its *antecedent* and the process of determining the antecedent of an anaphor is called *anaphora resolution* (Halliday & Hasan, 1976, p.14). While languages have anaphoric expressions

referring back to concrete individuals or objects (e.g. personal pronouns like 'he' referring to John in example (1)), they also have anaphoric expressions referring to abstract entities like situations, events, beliefs, thoughts, etc. Such kinds of anaphors are called *abstract object anaphors*¹ (Asher, 1993, p.3). To exemplify, 'that' in example (2) refers to the situation in which it is considered essential to be careful with wishes because they sometimes come true. Thus, 'that' in example (2) is an abstract object anaphor.

- (1) [John] was the manager of the company and supposed to be defending the rights of the workers, but **he** remained silent.
- (2) [Be careful with what you wish...because wishes sometimes come true.] **That's** what the Semiconductor Industry Association, which represents U.S. manufacturers, has been learning.

(Asher, 1993, p. 3)

Botley discusses the term anaphor like the ones in examples (1) and (2) (2006, p.74). In example (1), the antecedent of the anaphoric expression 'he' can be identified with little trouble by a hearer or a reader. This is because the connection between the anaphor and the antecedent is a straightforward one as the anaphor 'he' refers to a concrete individual (i.e., John). However, this is not the case for the abstract object anaphor 'that' in example (2). The antecedent of the anaphoric expression 'that' may require a more complex process of inference to be defined because the antecedent refers back to a situation, which is not a surface noun or noun phrase in its syntactic form. Dipper and Zinsmeister (2010, p.55) state that antecedents of such type of anaphora are abstract objects, like actions and events, which correspond to linguistic entities including at least a verb: partial clauses, clauses, sentences or sequences of sentences. For example, the abstract object anaphor 'that' in example (2) corresponds to the full sentence preceding itself. On the other hand, the abstract entity anaphor 'it' refers to a partial clause in example (3):

¹ The target type of anaphors, *abstract object anaphors*, have been addressed with different names in the literature such as *discourse deictics* by Levinson (1987), Webber (1991), Lakoff (1974), *anaphors to abstract objects* by Asher (1993), *situation reference* by Fraurud (1992), *indirect anaphors* by Botley (2006). Throughout the thesis, Asher's naming is followed so the target type of anaphors is addressed as *abstract object anaphors* to be consistent with the terminology.

- (3) As part of corporate streamlining programs, many companies are [extending early-retirement packages to legions of senior managers.] They see **it** as one relatively painless way to pare management ranks...

(Asher, 1993, p.3)

As pointed out by Dipper and Zinsmeister (2010, p.55), as well as corresponding to partial clauses, clauses, sentences or sequences of sentences, abstract entity anaphors can even correspond to discontinuous strings, as illustrated by the following example:

- (4) I would like to draw particular attention to the fact that [people who have made their lives here in the European Union] still do not [have the right to vote], even though the European Parliament has called for **it** on many occasions.

(Dipper & Zinsmeister, 2010, p.55)

In example (4), the antecedent of the abstract object anaphor 'it' is the fact that people have made their lives in the European Union but they do not have the right to vote. The complete antecedent is presented discontinuously in the text.

1.2. Types of Abstract Anaphora

Botley (2006, pp.75-80) distinguishes three distinct types of abstract anaphors. These are *labeling* put forward by Francis (1994), *situation reference* discussed extensively by Fraurud (1992, 1992a), and *discourse deixis* argued by Lyons (1977).

1.2.1 Labelling

Francis (1994) examines the nominal groups functioning as abstract object anaphors, which are used to connect the written discourse. He refers to them as *labeling*. He makes a distinction between *advance labels* (ALs) and *retrospective labels* (RLs). While ALs function cataphorically (forwards), RLs function anaphorically (backwards) (p.83). Examples (5) and (6) below clarify the cataphoric and anaphoric functions of the abstract label anaphors respectively:

- (5) I understand that approximately 12 per cent of the population is left-handed. Why, then, should there be such a preponderance of right-handed golfers which extends, I am informed, to club level? In reply to that question a golfing colleague of mine offered **two reasons**.

[The first was that beginners usually start with handed-down clubs, which are usually right-handed. The second was that, for technical reasons, left-handed individuals make good right-handed golfers.]

(Francis, 1994, p.84)

- (6) ...[the patients' immune system recognized the mouse antibodies and rejected them. This meant they did not remain in the system long enough to be fully effective.]

The second generation antibody now under development is an attempt to get around **this problem** by 'humanizing' the mouse antibodies, using a technique developed by...

(Francis, 1994, p.85)

In example (5), the antecedent follows the label anaphor 'two reasons' acting cataphorically. On the other hand, in example (6), the antecedent to the label anaphor 'this problem' precedes the anaphoric expression; thus, functions anaphorically.

As pointed out by Francis, RLs are far commoner than ALs, and not all the nouns listed below occur as heads of advance labels (1994, p. 89). According to the list of labels gleaned by Francis from *The Times* corpus, the labels are as follows:

Most common: approach, area, aspect, case, matter, move, problem, stuff, thing, way.

Others: accident, achievement, action, activity, advance, advantage, affair, agreement, anachronism, approach, arrangement, attempt, background, behavior, blunder, calamity, cause, challenge, change, characteristic, circumstances, combination, complication, compromise, conditions, consequence, consideration, context, contingency, contradiction, deal, deed, development, device, difficulty, dilemma, disaster, effect, element, episode, event, evidence, exercise, experience, fact, factor, fate, feature, incident,

information, issue, manner, measure, mess, method, mistake, mixture, news, objective, occasion, occurrence, operation, outcome, pattern, picture, plan, policy, possibility, practice, procedure, process, programme, project, prospect, purpose, question, reaction, reason, result, scenario, scheme, setback, sign, situation, solution, sphere, step, strategy, system, subject, tactic, task, technique, tendency, threat, topic, treatment, trend, truth.

As well as the backwards and forwards categorization of labeling, Francis (1994) also categorizes the *head nouns* in the labels as follows (p.90-93):

a. Illocutionary nouns: These are the nominalizations of the verbal processes, like the nouns *accusation, allegation, claim, criticism, explanation, indication, objection, prediction, proposal, proposition, reminder, recommendation, statement, suggestion* etc.

(7) As we left this meeting, my wife said: ['Potter has gone barmy, and they don't know what to do.'] I could not bring myself to believe she was right. I only accepted **this explanation** when my wife confined her suspicions to a friend, a psychiatrist, who exclaimed: 'That's a terrible thing to say about your child's therapist.' This level of denial convinced me that it was true.

(Francis, 1994, p.90)

In example (7), the label of the abstract entity anaphor 'this explanation', referring back to wife's utterance given in quotation marks, is a nominalization of the verb 'explain'. The derivative suffix *-tion* nominalizes the verb; thus, the head noun 'explanation' in example (7) is a clear case of an illocutionary head noun.

b. Language activity nouns: These head nouns are similar to illocutionary nouns but they do not have a cognate illocutionary verb; rather, they are the nouns corresponding to some kind of a language activity and the result of it, such as the nouns *account, ambiguity, comparison, consensus, debate, definition, distinction, dispute, example, equation, language, myth, reference, story, summary, talk, verdict, version*, etc. Consider the following example:

(8) Foster, the Fife-based organiser, said: ['So many great sporting cars are only seen as static exhibits in museums nowadays, so it is a great honour for

Scotland that it has become one of the premier venues for using these wonderful machines.']

This description is scarcely inflated. McLaren will be driving his Jaguar Lightweight E Type. John Coombes, now based in Monaco, will drive a Jaguar D Type.

(Francis, 1994, p.91)

In example (8), the label anaphor 'this description' refers back to the description provided by Foster, so it illustrates the use of 'description' as the language activity head noun in the label.

c. Mental process nouns: These are head nouns referring to cognitive states and processes and the results of these, such as *analysis, assessment, assumption, attitude, belief, interpretation, insight, knowledge, opinion, thought, theory, view, vision* etc. Consider the following example:

(9) At a press briefing in London during the inaugural meeting of the bank's board of governors, Henning Christophersen, vice-president of the European Commission, said: ['The EBRD must not be a political institution, but plainly and simply a bank.']

This view contrasted with that of Jacques Attali, the president of the European Banks, who regards the bank's role as political and economic.

(Francis, 1994, p.91)

In this example, the label of the abstract object anaphor 'this view' is both the type of language-activity label and it is the mental process type simultaneously. It is a mental process noun because the noun 'view' refers to a cognitive process of mind.

d. Text nouns: These head nouns refer to the formal textual structure of discourse without any interpretation being involved, extending the previous discourse and signaling the boundaries of a new discourse, such as *phrase, sentence, word, page, excerpt, section, paragraph, quotation, passage, term, terminology* etc. Consider the following example:

- (10) ['Projects are also introducing changes in teaching styles. Increasingly these are geared towards providing students with the opportunity to develop initiative, motivation, problem-solving skills and other personal qualities. Central to this approach is the transfer to students themselves of the responsibility for managing their own learning and applying their own knowledge.']

That quotation comes not from the Plowden report, but from the Technical and Vocational Education Initiative review of 1985. It is very different from what we found in the best primary schools?

(Francis, 1994, p.91)

In example (10), the label anaphor 'that quotation' refers back to all three sentences that precede i.e., the quoted information. The label anaphor here functions as a signal that a new discourse is coming, and it seems to have nothing with the meaning of the antecedent, rather it refers back to the textual formal structure of the discourse.

1.2.2 Situation Reference

The second type of abstract object anaphora stated by Botley (2006) is the *situation reference* discussed extensively by Fraurud (1992a, p.1). In her analysis, Fraurud (1992a) investigates the manifestations of bare anaphoric expressions 'it/this/that', named as 'sentential it/that/this' in traditional grammar and syntax theory. Fraurud (1992, 1992a) makes a distinction between object reference and situation reference in anaphora resolution: She holds that it is the *semantic class* of the antecedent as well as its *syntactic class* that distinguishes the two. As the semantic distinctions, Fraurud (1992, p. 28) differentiates between the 'objects' as referring back to entities, namely the concrete entities, and the 'situations' as *eventualities* (referring back to processes and states as well as events) and *factualities* (referring back to facts and propositions). The distinction between the object reference and the situation reference regarding the semantic class of the antecedent of the situation reference is illustrated by her own example below in (11):

- (11) Mary was fired. a. **That** happened last week
 b. **That** is true

c. **That** surprised me

(Fraurud, 1992a, p.1)

The situation referent 'that', in contrast to object referents, has a function referring back to various different higher-order entities in (11a), (11b), and (11c). Higher-order entities cover for events, processes, propositions, facts etc., rather than covering for object entities. In (11a), the anaphoric expression 'that' refers to an *event* (i.e. the firing of Mary), in (11b), the antecedent is a *proposition*, (i.e. that the firing of Mary is true), and in (11c) the antecedent is a *fact*, (i.e. the fact that Mary was fired) (Fraurud, 1992a, p.1).

The other distinction between object reference and situation reference made by Fraurud (1992, 1992a) is related to the *syntactic class* of the antecedent. The antecedents of the object reference might be limited to different kinds of NPs, whereas the syntactic class for the antecedents of the situation reference can vary. That is to say, they may range from abstract nouns and propositional pronouns to sentences and segments of the text. Consider the following examples:

- (12) Jane learnt that [John had an affair with another girl]. Upon learning **this**, she got depressed.
- (13) [An anaphor is a device defined as the dependence of an expression referring to another which has already been introduced in the discourse.] **This** is the definition for the phenomenon of 'anaphor' provided by Botley (2006, p.73).

The abstract object anaphor 'this' in example (12) refers to that 'John had an affair with another girl', which is an embedded clause in a sentence. On the other hand, the abstract object anaphor 'this' in (13) refers to the definition of anaphor in the previous sentence; that is, the antecedent is the whole sentence. As these examples show, the antecedents of the situation referents might vary in their syntactic class, which is not the case for object referents.

The focus of this thesis will be mainly on label anaphors and bare anaphors with situation reference, which have been explained so far. For the sake of completeness, another form of abstract object anaphors, *textual/discourse deixis* discussed by Lyons (1977), will also be explained below.

1.2.3 Textual/Discourse Deixis

Lyons (1977), gives the following definition for discourse deixis:

[It is] the location and identification of persons, objects, events, processes and activities being talked about, or referred to, in relation to the spatiotemporal context created and sustained by the act of utterance and the participation in it, typically, of a single speaker and at least one addressee (p.637).

Lyons (1977) makes a distinction between 'pure textual deixis' and 'impure textual deixis'. For him, the following excerpt from a conversation sets an example for pure textual deixis.

(14) A: That's a rhinoceros.

B: A what? Spell **it** for me.

(Lyons, 1977, p.667)

In this context, the abstract anaphoric expression 'it' refers to the linguistic form 'rhinoceros', not the referent of 'rhinoceros' in the real world. However, there seems to be a difference in the referent of the abstract anaphoric expression 'that' in the following context provided by Lyons (1977):

(15) A: I have never seen him!

B: **That's** a lie!

(Lyons, 1977, p.668)

In this context, the anaphor 'that', rather than the linguistic form of A's utterance, refers to the proposition expressed by the sentence uttered by A, which is defined as 'factualities' by Fraurud (1992, 1992a). As Webber (1991, p.3) states, this example is 'textual' as it has nothing to do with the linguistic form but the utterance itself, and it is 'impure' as what is being indicated is not the utterance but what is expressed by the utterance. Therefore, *impure textual deixis* (i.e., 'that' in example 15), which is defined by Lyons (1977), can be considered having a similar function to situation referent 'this' and 'that' discussed by Fraurud (1992, 1992a). However, *pure textual deixis* (i.e., 'it' in example 14) does not have an anaphoric function at all -although it might appear to have at first sight- because it is not co-referential with any antecedent expression in discourse (Lyons, 1977, pp.667-668).

1.3. The Scope of the Current Work

The current thesis is about Turkish with a focus on *abstract object anaphors* whose antecedent may be partial clauses, clauses or sequences of sentences, rather than surface nouns or noun phrases. Such kinds of anaphors refer to abstract objects, such as actions, facts, situations and events. Among the types of this kind of anaphora, the bare anaphoric expression 'bu' functioning as an abstract object anaphor in Turkish with situation reference, and label anaphors, acting anaphorically (RLs), 'bu+ durum/gerçek/olay/iş', functioning again as abstract object anaphors in Turkish, are compared to investigate whether any difference is observed in their processing. The main research questions of the thesis are:

- Is there any difference between the resolution of the bare abstract object anaphor 'bu' and label anaphors 'bu+durum/gerçek/olay/iş'?
- Do label anaphors, namely 'bu+durum/gerçek/olay/iş', have a disambiguating role in abstract anaphora resolution in Turkish written discourse?

In the current study, the Turkish head nouns used as labels in this study are *gerçek*, *durum*, *iş* and *olay*. They are just general words facts, situations, events and actions. The exact equivalents of these words do not exist in Francis' categories and examples of head nouns, which are presented in 1.2.1.

The reason why these specific labels (ie. *gerçek*, *durum*, *iş*, *olay*) were chosen is based on the inspiration stemming from Fraurud's (1992, 1992a) semantic distinction of eventualities and factualities. Recall example (10) where the anaphoric expression 'that' refers to an *event* in (10a), a *proposition* in (10b) and a *fact* in (10c). Considering that eventualities include processes and states as well as events, and factualities include facts and propositions in Fraurud's (1992, 1992a) terms, it was decided to compare and contrast the bare anaphoric expression 'bu' referring back to a *fact* with the labeled anaphor 'bu gerçek' (Eng. this fact). The one referring back to an *event* was compared and contrasted with the labeled anaphor 'bu olay' (Eng. this event). The one referring to a *state* was compared and contrasted with the labeled anaphor 'bu durum' (Eng. this situation). Lastly, the one referring back to a *process* was compared and contrasted with the labeled anaphor 'bu iş' (Eng. this process/activity) in the experimental materials.

The rest of the thesis will proceed as follows:

In Chapter 2, the general theoretical background for abstract object anaphora studies and the methods used in these studies will be reviewed. The ambiguous nature of abstract entity anaphora and the phenomenon of abstract anaphora resolution will be discussed to pave the way for the discussions that will take place in Chapters 4 and 5.

In Chapter 3, anaphora studies conducted in Turkish will be reviewed and the various manifestations of abstract anaphoric expressions in Turkish language will be discussed.

In Chapter 4, the method of the current study, the experiments conducted to have an insight into the way abstract object anaphors function in Turkish, and the results of these experiments will be presented and discussed by considering the theoretical background and ambiguous nature of anaphora explained in Chapters 2 and 3.

In Chapter 5, some conclusions will be drawn and possibilities for future work will be mentioned.

CHAPTER 2

THEORETICAL AND METHODOLOGICAL FRAMEWORKS IN ABSTRACT OBJECT ANAPHORA STUDIES

Over the last decades, with the upsurge of interest in building annotated corpora, researchers have conducted corpus-based studies on the phenomenon of abstract object anaphora in various languages -most of them are on English corpora- and, they have generated annotation schemes for marking this type of anaphors.

In this chapter, some of these pioneering corpus-based studies conducted in an attempt to pave the way in understanding the phenomenon of abstract anaphora resolution (Section 2.1) will be explained. Besides this, the accessibility relation between the antecedent and the anaphors will be discussed with respect to Consten and Knees' (2005) *abstractness hierarchy* between the anaphor and the antecedent, and also with respect to the *right frontier* of the discourse structure elaborated on by Webber (1991) (Section 2.2). Finally, two pioneering studies, i.e., Krippendorff (1980); Passonneau (2004) will be reviewed in detail (Section 2.3) as the methods in these studies are applied in the current study, which is the topic of Chapter 4.

2.1 Abstract Object Anaphora Studies

Most of the studies conducted on anaphora resolution do not deal with abstract entities. The reason why many studies are limited to concrete object anaphora with NP antecedents may be stemming from the difficulty of dealing with identifying the boundaries of the antecedents to abstract object anaphors. Also, most of the work concerning this phenomenon is based on English data, and only a few of them (e.g. Dipper and Zinsmeister, 2009, 2010, 2011; Recasens, 2008; Navaretta, 2001, 2002, 2007; Lee and Song, 2010; Vieira, Salmon-Alt and Gasperin, 2005; Navaretta and Olsen, 2008) try to account for the use of abstract anaphora in other languages. These works will be reviewed in 2.1.2

2.1.1 Abstract Object Anaphora Studies in English

As it has been discussed in Chapter 1, Francis (1994) has a study on the abstract label anaphors in English gleaned from the Times Corpus (as discussed in 1.2.1). Fraurud (1992, 1992a) accounts for the English situation referents 'this/that/it' referring back to eventualities or factualities (as explained in 1.2.2). Lyons (1977) defines abstract object anaphors as *discourse/textual deixis* and makes a distinction between *pure discourse deixis* and *impure discourse deixis* in English (as discussed in 1.2.3).

Other than these early studies conducted to account for the differences between the concrete and abstract object anaphors, there are more recent and pioneering corpus-based studies which try to focus on the quantitative description of language for the use of abstract objects anaphors in English, or those which try to generate annotation schemes for marking abstract object anaphors. Some studies account for both.

Among these corpus-based studies, there is Poesio and Artstein (2008) in which a new corpus, namely ARRAU is presented. This corpus involves dialogues from the Trains-91 and Trains-93 corpus, narratives from the English Pear Stories corpus, newspaper articles from the Wall Street Journal portion of the Penn Treebank, and mixed text from the Gnome corpus. In this study, the corpus is annotated for anaphoric relations for abstract entities with an explicit representation of multiple antecedents for ambiguous anaphoric expressions. The authors conduct

annotation experiments with multiple participants (up to 20 participants annotating the same texts independently). One of the main contributions of this study is that it leads to an improved annotation scheme depending on the annotations of the participants. Also, the study presents reasonable agreement on coreference chains. The data is evaluated with Krippendorff's Alpha interval measurement (discussed in detail in Section 2.4 of the present thesis), and substantial agreement results ($\alpha=0,6-0,7$) are found. However, as a result of this study, Poesio and Arstein (2008) emphasize that it is difficult to spot ambiguity through a quantitative analysis. As they discuss, it is only implicitly possible to mention the existence of ambiguity through the disagreement results². Another implicit conclusion for ambiguity is drawn through Krippendorff's Alpha boundary measurement results ($\alpha=0,55$). Depending on these results, Poesio and Arstein (2008) conclude that it is not easy for annotators to agree on the exact boundaries of the antecedents.

The study conducted by Poesio and Artstein (2008) has been one of the main contributors to this thesis. This is because this study is one of the rare attempts to deal with the annotation of ambiguous anaphoric expressions, to establish an annotation scheme for such kind of anaphoric expressions, and to search for a reliability measurement of agreement in ambiguous texts. These concerns (discussed in detail in Chapter 4) have been central to the current study, too.

Another corpus-based study which has helped to shape the current study is the one conducted by Botley (2006) in English. This study investigates all cases of demonstratives having an indirect relationship with their antecedent, i.e., labeling, situation reference and discourse deixis. The corpora involve 100,000 word segments from the Associated Press corpus of American newswire texts, the Canadian Hansard proceedings from the Canadian House of Commons and the American Printing House for the Blind collection of literary works and motivational narrative. The target kind of abstract object anaphors were annotated using the annotation scheme

² Krippendorff's Alpha interval and boundary measurements are applied by Artstein and Poesio (2008). These measurements are also applied in this study. The results of are discussed in detail in Chapter 4.

in Botley and McEnery (2001, pp.8-10)³. The general observations made by Botley (2006) are as follows:

- a. The most frequent abstract object anaphor is the situation referents 'this/that/it', which are detailed by Fraurud (1992, 1992a).
- b. Cases of cataphora are rarer compared to cases of anaphora in all of the corpora.
- c. Proximity plays a role in identifying the antecedents in an easier way.
- d. Singular demonstratives are more than the plural ones.

Botley (2006) further analyzes each of these all three corpora for a detailed quantitative description of the occurrences of each type of abstract object anaphors (i.e., labeling, situation reference and discourse deixis). The findings will not be elaborated on here since they are not completely relevant to the current study. However, it is important to note that Botley (2006) presents comprehensive empirical insights into these three types of abstract object anaphors in English written discourse. The study also shows the difficulties of classifying the target anaphors in terms of their types because antecedents to these anaphors lack precise boundaries. As it is a challenge to specify the antecedents to these anaphors, Botley (2006) emphasizes that the inference process for retrieving an antecedent is not only complex but also unclear. This is because overlapping specifications of the antecedents make it difficult to come up with hard and fast decisions in the analysis. This is also the conclusion drawn by Artstein and Poesio (2008) through another methodology.

The study by Botley (2006) has contributed to the current study in this thesis, first, with the theoretical background it has provided regarding our decision to determine the types of abstract object anaphors to focus on. Second, the challenges in

³ The annotation scheme by Botley and McEnery (2001) categorization of the abstract anaphors depending on the following criteria: (1) 'Lexical Category of a Demonstrative' (i.e., the part of speech like a noun, a pronoun etc.), (2) 'Endophoricity' (i.e., the anaphoric vs. cataphoric functions of the referential expressions), (3) 'Exophoricity' (i.e., context-based vs. deixis), (4) 'Syntactic Category of an Antecedent' (i.e., nominals, clausal or sentential elements), (5) 'Phoric Type' (i.e., reference, the relation between meanings; substitution, the relation between the linguistic forms), and (6) 'Semantic Function of an Antecedent' (i.e., events, processes or states).

the specification of the antecedents to abstract anaphors mentioned in this study has been one of main discussions in the current study, too (as discussed in Chapter 4).

Some of the other anaphora studies conducted on English data are based on the spoken discourse, or some others have a computational perspective. These studies will not be explained in detail as they are not directly related to the current study. For example, Gundel, Hedberg and Zacharski (2005) provide an empirical analysis of pronouns without explicit NP antecedents in spontaneous conversations taken from Santa Barbara Corpus of Spoken American English. From a computational perspective, Byron (2002) introduces PHORA, the pronoun resolution algorithm, to resolve pronominal reference to either concrete and abstract entities in English discourse. This study also deals with the abstract object anaphors and provides quantitative data for the success of the algorithm. Eckert and Strube (1999) has an attempt to deal with the resolution of abstract anaphora in English spoken dialogues in one of their studies, and in another corpus-based study of theirs (i.e., Eckert & Strube, 2000), they try to generate a model of anaphora resolution in spontaneous spoken dialogues in the form of an algorithm, which is the ES-algorithm. The ES-algorithm by Eckert and Strube (1999, 2000) discriminates between the concrete and abstract object anaphors based on the predicative context in which the anaphor takes place. Although most of the algorithms deal with concrete object anaphor resolution, the ES-algorithm is an exception dealing with situation referents.

2.1.2 Abstract Object Anaphora Studies in Other Languages

Other than the studies based on English corpora, there have been studies conducted in some other languages for which abstract object anaphora has been analyzed. Dipper and Zinsmeister (2009) present a study on the annotation of the semantic type of the abstract object anaphor and the antecedent (i.e., anaphoric *das*, *die*, *was*, *es* ‘that, this, which, it’) in German texts taken from the Europarl Corpus. For the annotation of the texts, they use guidelines to identify how to locate the antecedent, how to define both the semantic type of the antecedent and the semantic type of the anaphor. They define the semantic types in terms of the following features: world-dependent, time-dependent, dynamic, telic, and modal (see e.g., Asher, 1993). What distinguishes Dipper and Zinsmeister's (2009) study from the other studies conducted in German is that it takes the semantic restrictions into

account rather than the grammatical restrictions (e.g., gender and number agreement) in anaphora resolution process. Therefore, this study mainly aims to fill this gap in anaphora studies in German. Another key point in this study is the procedure followed in the experiment. Dipper and Zinsmeister (2009) ask participants to paraphrase the antecedent (i.e., 'the paraphrase test'⁴) to determine the location of the antecedents. Inspired by this method, a similar procedure was followed both in the pilot study (Section 4.1) and in Experiment 1 (Section 4.2) of the present thesis.

For Spanish and Catalan, Recasens (2008) conducts a corpus study in written texts in an attempt to annotate abstract object anaphors, and presents an empirical analysis of 200.000 words from the AnCora corpora. Following the pioneering studies on abstract object anaphors conducted in English (i.e., Botley, 2006; Francis, 1994; Poesio & Artstein, 2008,), and the assumptions of Webber (1988a) on discourse deixis (i.e., abstract object anaphors), Recasens (2008) tries to understand whether they are true for Catalan and Spanish or not. Her assumptions are as follows:

1. Preference for demonstratives: “Subsequent reference to a sequence of clauses is most often done via deictic pronouns.”
2. Referent coercion: “Once the speaker has referred to it [discourse segment] via this/that, it must now have the status of a discourse entity since it can be referenced via the anaphoric pronoun it.
3. Required presence: “The demonstratum being something [explicitly] present in the shared context.”
4. Ambiguity: “All pointing is ambiguous . . . The listener’s choice depends on what is compatible with the meaning of the rest of the sentence.” (as cited in Recasens, 2008, p. 74)

As a result of her study, Recasens (2008) concludes that it is not very easy to specify the exact boundaries of the antecedents of the abstract object anaphors. Therefore, she draws the attention to Webber’s (1988a) ideas on non-specificity (i.e., ambiguous nature of the antecedent to abstract object anaphors and difficulty in specifying its extension) of the antecedent. She also draws the attention to Poesio,

⁴ The annotator supplements the anaphor by a paraphrase in the form of *namlich* . . . ‘namely . . .’. The part that fills the . . . corresponds to the antecedent that they are looking for (Dipper & Zinsmeister, 2009, p. 167).

Reyle and Stevenson's (2006) theory on the underspecification of anaphora i.e., The Justified Sloppiness Hypothesis.

In essence of this theory, it is asserted that readers or listeners do not always construct complete interpretations of everything they read or hear. Some meanings, especially the meanings of anaphoric expressions like pronouns, may be incomplete or underspecified. Therefore, as asserted by Poesio et al. (2006), some ambiguous anaphoric expressions may not completely be resolved, or they may not be fully identified in the context. Poesio et al. (2006) gives the following example to clarify:

(1) The engineer hooked up the engine to the boxcar and sent **it** to London.

(Poesio et al., 2006, p. 16)

In example (1) it is not clear whether the anaphoric expression 'it' refers to the engine or the boxcar. Poesio et al. (2006) asserts that such cases are truly ambiguous, and readers or listeners cannot always accurately and completely interpret everything that they read or hear.

In another study conducted on Danish, Navaretta (2001) deals with abstract object anaphors (i.e. 'situation reference' in Fraurud's (1992, 1992a) term) in Danish to identify the antecedents occurring in verbal phrases and discourse segments of the anaphors by adapting Eckert and Strube's (1999, 2000) algorithm for anaphora resolution in Danish. Navaretta (2001) modifies and extends the preference rules proposed by Eckert and Strube (1999, 2000) for identifying situation reference in Danish. To test the reliability of the algorithm, Navarretta (2001) compares the marking of the algorithm with that of humans. She finds that the same situation anaphors were identified in 83% of the cases both by the human subjects and the algorithm.

Lee and Song (2010), present preliminary work on a corpus-based study of Korean demonstratives in an attempt to identify different functions of demonstratives both in written and spoken discourse, and to examine their distributional properties with an annotation scheme similar to the one used in Botley and McEnergy (2001). This study is mainly an effort of quantitative description of the uses of Korean demonstratives in the framework of six annotation features introduced by Botley and McEnergy (2001) (as briefly mentioned in footnote 2). They find frequent uses of

adnominals compared to pronouns as the lexical category; more occurrences of anaphors compared to cataphors (in written discourse, they find no occurrence of cataphors); more deictic uses compared to context-bases uses; nominal, clausal and sentential uses of the demonstratives; and finally referential examples not substantial cases of demonstratives in their phoric type.

2.1.3 Abstract Object Anaphora Studies Focusing on Contrastive Analysis of Several Languages

Some other studies provide contrastive analysis of two or more languages. Among these studies, for example, Navarretta (2007) describes the results of a contrastive analysis of abstract anaphora in Danish, English and Italian. The results signify the similarities across these three languages i.e., the presence of a demonstrative pronoun in Danish, English and Italian discourse indicating an antecedent as an abstract entity as well as the differences. While the pronouns referring back to abstract entities are demonstratives in English, in the other two languages, personal pronouns are also frequently used to refer back to abstract entities.

Vieira, Salmon-Alt and Gasperin (2005) present another corpus study in which the use of demonstrative noun phrases in Portuguese and French written texts is investigated. They elaborate on the types of coreferential or anaphoric uses of demonstratives, the syntactic structure of the anaphors as well as the antecedents, and the semantic relations between the two. They find that the interpretation of demonstrative noun phrases in both languages is mainly context dependant. Human annotators are able to find textual chunks as antecedents for more than 80% of demonstrative NPs. For more than 40% of demonstrative NPs in the corpus Vieira et al. (2005) analyzed, it is difficult to select a precise portion of the text as an antecedent. This was the finding of Artstein and Poesio (2008a), and Botley (2006) for English, too. Similarly, Recasens (2008) draws a similar conclusion for Spanish and Catalan. Finally, Vieire et al. (2005) concludes that when the relation of a demonstrative and its antecedent is not a coreferential one, the amount of world knowledge and reasoning required for the resolution is very large.

Navarretta and Olsen (2008) attempt to extend the MATE/GNOME annotation scheme (Poesio, 2004) to account for abstract anaphora in Danish and Italian. In the core of the work, they aim to provide reliably annotated data to study abstract object anaphors in Danish and Italian. They use strict annotation guidelines to obtain better inter-annotator agreements as a result of the study. In their study, in addition to annotating the type of clausal antecedent and the semantic type of the referent (events, states, fact-like entities, etc.), they also annotate the anaphoric distance, measured in terms of clauses in between the anaphor and the antecedent. They find some important differences between Italian and Danish demonstratives used as abstract anaphors. For example, compared to Italian, they find a lot more occurrences of abstract object anaphors in Danish, i.e., out of 687 Danish texts, they find 200 occurrences of abstract object anaphors and 179 occurrences of concrete object anaphors; whereas out of 571 Italian texts, they find 59 occurrences of abstract object anaphors and 487 occurrences of concrete object anaphors.

2.2. Resolution of Abstract Object Anaphors

Anaphora resolution is the process of identifying the antecedent to an anaphor in discourse (Botley, 2006). To gain insight into the relation between the anaphor and its antecedent, the concept of accessibility has an essential role because the anaphor is an accessible linguistic device that the antecedent introduces or activates in discourse.

The morphological constraints like the gender and number agreement between the anaphor and the antecedent is the most basic strategy helping anaphora resolution in discourse by eliminating the candidates whose gender or number does not match the anaphor. Syntactic constraints are also eliminating factors as the antecedent and the anaphor usually appear in a similar syntactic position (i.e., syntactic parallelism).

Grammatical constraints (gender, number agreement) are the main guiding factors in the process of anaphora resolution in many cases where concrete object anaphors occur. Considering Lyons' (1977) categorization of anaphors and the distinction between *first-order entities* (equating approximately to objects), *second-order entities* (events), and *third-order entities* (facts/propositions), in many other

cases, grammatical constraints are insufficient to specify the correct antecedent without the help of semantic information. This difficulty is observed especially during the process of identifying the antecedents of the abstract object anaphors. Since abstract object anaphors, as the name suggests, refer back to abstract entities, grammatical restrictions may not always apply due to the non-nominal antecedents they have. As highlighted by Eckert and Strube (1999), while a concrete NP antecedent usually only refers to the individual it describes, a sentence may simultaneously denote an eventuality, a concept, a proposition and a fact (p. 38). Consider the following example, which has already been introduced in 1.1 as example (1) and repeated as example (2) here:

- (2) [John] was the manager of the company and supposed to be defending the rights of the workers, but **he** remained silent.

The concrete object anaphor 'he' is not difficult to resolve because depending on the gender, the only candidate, a male person, introduced prior to this anaphor is 'John'; thus, it seems quite straightforward to establish the relation between the anaphor and the antecedent. However, consider the following example:

- (3) It's always been presumed that when the glaciers receded, the area got very hot. The Folsom men couldn't adapt, and they died out. **That's** what is supposed to have happened. It's the textbook dogma. But it's wrong. They were human and smart. They adapted their weapons and culture, and they survived.

(Webber, 1988a, p.116)

In this example, readers may not agree on the exact antecedent. The possible antecedents might be only (a), only (b) or (a+b) as follows:

- (a) when the glaciers receded, the area got very hot.
(b) The Folsom men couldn't adapt, and they died out.
(a+b) when the glaciers receded, the area got very hot. The Folsom men couldn't adapt, and they died out.

It seems possible to resolve the anaphor 'that' in example (3) only through some inferences based on the semantic clues provided in discourse. For example, the

lexical meaning of 'happen' (i.e., the one in the clause 'That's what is supposed to have happened') can help the resolution of the abstract object anaphor 'that'. The word 'happen' directs the reader to finding some entity that can 'happen', in Fraurud's (1992, 1992a) terms, to finding an *eventuality* (i.e., a process or an event) in the preceding clauses. Both (a) and (b) involve eventualities. In (a), that the glaciers receded and the area got very hot is a process that 'happened'. Similarly, that the Folsom men couldn't adapt and they died out is another process that 'happened'. Therefore, the readers may think that either (a) or (b), or possibly the combination of both can be the antecedent to the abstract object anaphor 'that'.

As it is clear, in addition to morphological and syntactic constraints, it is essential to take 'semantic restrictions' into account and make use of the lexical semantic knowledge to be able to gain insight into abstract anaphora resolution (Dipper and Zinsmeister, 2009, p. 166).

2.2.1 Abstractness Hierarchy

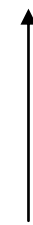
Considering the semantic constraints, the kind of difficulty in resolving the abstract object anaphor in example (3) may be said to be originating from an *abstractness hierarchy* put forward by Consten and Knees (2005). According to Consten and Knees (2005), at the bottom of the hierarchy, the 'events' as the most concrete type with spatial-temporal relations are situated, and at the top the 'propositions' as the most abstract types take place (see Table 1 below).

Since Consten and Knees (2005) is highly relevant to the current study, their ideas regarding anaphoric complexation processes will be summarized below as well as those of Consten, Knees and Schwarz-Friesel (2007).

For Consten and Knees (2005), it is only possible to resolve the abstract object anaphor when the possible antecedent is less abstract than the anaphor itself, or at least it is only as abstract as the anaphor. They call this the *anaphorical complexation*, which briefly means the process of increasing abstractness. The authors highlight the constraints during abstract anaphora resolution process in terms of ontological categories (i.e., propositions, facts, states, processes, events) with a special emphasis on bare abstract object anaphor *this* and the *label anaphors* referring back to abstract entities. In Table 1, the mapping between the degree of

abstractness (e.g., events have the lowest degree of abstractness whereas the propositions have the highest degree) and the ontological categories of events, processes, states, facts and propositions are illustrated.

Table 1 Ontological Categories of Abstractness by Consten and Knees (2005)

<i>Degree of Abstractness</i>	<i>Ontological Category</i>
high  low	proposition (pp) fact (f) [dependent on world] state (s) [-dynamic, -telic / dependent on world and time] process (p) [+dynamic, -telic / dependent on world and time] event (e) [+dynamic, +telic / dependent on world and time]

Consten and Knees (2005) propose three types of complexation process that abstract object anaphors might have:

a. The antecedent and the anaphor might be denoting the same ontological type so the ontological status of both parties remains the same:

(4) [The Americans tried to invade the building but were forced back by shots from the top floor.]_e It is said that two soldiers were injured during [**this action**]_e, one inside the house and the other one outside the house.

(Consten&Knees, 2005)

In example (4), the antecedent sentence denotes an event. This corresponds with the label anaphor 'this action', which also denotes an event. Therefore, the ontological status of both the antecedent and the anaphoric expression is the same.

b. The bare abstract object anaphor 'this' is neutral (n) in its ontological category. When 'this' is used, the ontological type of the anaphoric process is denoted by the antecedent.

- (5) [The Americans tried to invade the building but were forced back by shots from the top floor.]_e [**This**]_n happened yesterday while Mr. Rumsfeld visited Bagdad.

(Consten&Knees, 2005)

In example (5), the antecedent denotes an event so the bare anaphor 'this' is also understood as an event. The reason why 'this' is understood as an event is not only the ontological status of the antecedent, but also the lexical meaning of the verb 'happen'. As discussed in 2.2., the verb 'happen' directs the reader to searching for an eventuality in the preceding clauses to identify the antecedent.

c. Another ontological type might be denoted by the anaphoric expression because of the lexical meaning of the anaphoric expression. Thus, the anaphorical process changes the ontological type of the referent.

- (6) [The Americans tried to invade the building but were forced back by shots from the top floor.]_e [**This fact**]_f proves that the situation isn't under control yet.

(Consten&Knees, 2005)

In example (6) the antecedent referring back to an event denoted by the label anaphor 'this fact' turns into a fact. This change is initiated by the lexical meaning of the head noun 'fact' used in the label. An important point to consider is that the lexical meaning of the label anaphor has to fit the context. For example, it would not be possible to say "... ***[This fact]**_f happened yesterday." because the lexical meaning of 'fact' does not allow an event-reading while the verb 'happen' requires it (Consten, Knees & Friesel, 2007, p. 92).

Other than the types of anaphoric complexation process, Consten and Knees (2005) emphasize the possible disambiguating role of the label anaphors, which, as they state, may not be explained through purely grammatical constraints. Consider example (7):

- (7) [The Jacobs-Sisters are always in a wonderful mood and flashy.]_s [Yesterday they had a great performance in New York.]_e

(a) [**This event**]_e has surely made them even more popular.

(b) [**This quality**]_s has surely made them even more popular.

(c) [**This / that**]_n has surely made them even more popular.

In example (7), while the label anaphor 'this event' in (7a) refers back to yesterday's performance in New York, it denotes an event. The label anaphor 'this quality' in (7b) refers back to the Jacobs-Sisters' characteristics as being wonderful in mood and flashy; thus, it denotes a state. The bare abstract object anaphor 'this/that' in (7c), on the other hand, is ambiguous. It might be thought to be referring back to the state of the Jacobs-Sisters' characteristics, or the event in which the Jacobs-Sisters had a wonderful performance. It might even be thought to be referring back to both of the state and the event depending on readers' personal interpretations of the text on the text-world level.

The bare abstract object anaphors 'this/that' may bring burden on the anaphora resolution process because they may not always carry the load of denoting the precise ontological category that is referred to. Thanks to the lexical meaning of the head nouns, label anaphors can handle this burden. However, it is important to note, as discussed by Francis (1994), that labels do not necessarily refer to a clearly delimited or identifiable stretch of discourse (p.99).

2.2.2 Right Frontier Constraint (RFC)

Another pioneering theory, first discussed by Polanyi (1988) and Webber (1988b), that accounts for anaphora resolution is the right frontier operations. Webber (1991) examines the demonstrative pronouns 'this/that' in English used as *deictics* to refer to the interpretation of one or more clauses. In the essence of what Webber (1991) discusses lies the idea that only the clauses, or sequence of clauses, corresponding to discourse segments can be the referents of demonstrative pronouns, not a random sequence of them. To make it clear, Webber (1991) gives the following example:

- (8) a. For his part in their joint-project, John built a two-armed robot.
b. He had learned about robotics in CSE391.
c. For her part, Mary taught how to play the saxophone.

(Webber, 1991, p.8)

From just the previous clause (i.e., 'c'), it can be possible to derive such a subsequent utterance involving an abstract object anaphor;

d. **That** took her six months.

Or, it is also possible to derive the following subsequent utterance involving an abstract object anaphor from the previous three clauses (i.e., 'a-c').

d'. **That** earned them both As.

However, as discussed by Webber (1991, p.9), it is not possible to derive a subsequent utterance only from the previous two clauses (i.e., 'b-c') independent of the first clause (i.e., 'a'). This is because discourses are formed by smaller sequences of related clauses or sentences, which are called *discourse segments*.

Webber (1991) specifies the discourse segments as being the ones whose contribution to the discourse model is currently in focus. To clarify the focus, she presents a simple incremental tree construction algorithm showing the exact positions at which new nodes can be inserted. The reason why she presents the tree structure, which is given below in Figure 1, is to account for the discourse segments contributing to the discourse model in focus. Consider the following example:

- (9) a. John hates snakes.
b. His next-door neighbor had kept snakes,
c. and he had hated his neighbors.

(Webber, 1991, p.16)

While processing the discourse presented in example (9), the reader may consider the second clause (i.e., 'b') as the explanation of the first clause (i.e., 'a'). Therefore, in Figure 1, these two clauses constitute a segment in the first step. While processing the third clause (i.e., 'c'), the reader may recognize that the second and the third clauses actually constitute a joint segment together in step 2. Assume that another clause is added to the discourse presented in example (9), such as a fourth one (i.e., 'd') as follows;

d. Later in college, his roommate had kept snakes.

(Webber, 1991, p.17)

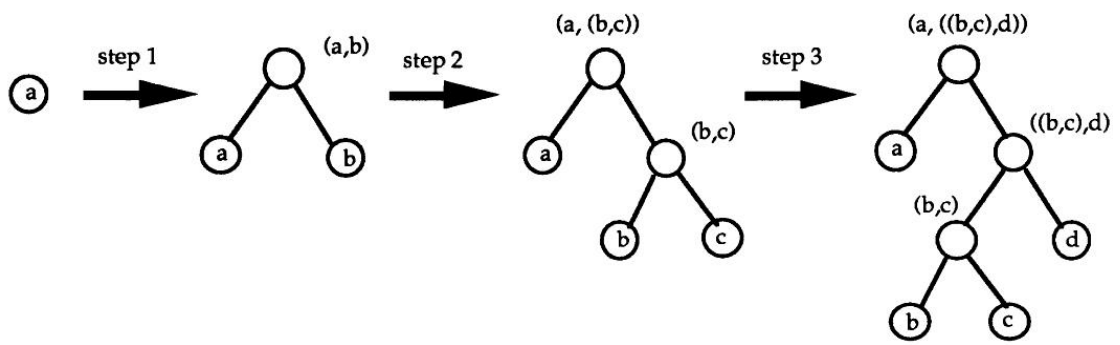


Figure 1 illustrates right frontier operations to Example (1)

As it is illustrated in Figure 1 in step 3, the exact position of a new node (i.e., 'd') is the right frontier of the tree structure. Webber (1991) claims that it is these very discourse segments at the right frontier of the tree structure which are the antecedents of the demonstrative pronouns (as illustrated in Figure 2). Discourse segments being connected to nodes not on the right frontier of the tree cannot be the antecedents of the demonstrative pronouns any more. Thus, this is called the right frontier constraint (RFC).

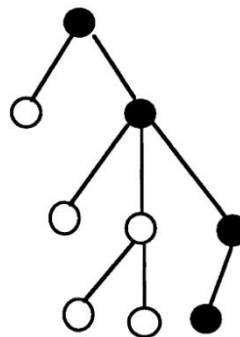


Figure 2 illustrates the right frontier of the tree structure⁵

The study by Webber (1991) is of a great deal of importance because of its contribution to the interpretation of the discourse segments involving demonstratives as abstract object anaphors in the discourse model. The rhetorical relations between the sentences may not always be linguistically marked. They may have to be inferred pragmatically from a variety of sources like the knowledge of syntax and semantics.

⁵ Webber, 1991, p. 14

This theoretical discourse model suggested by Webber (1991) describes the information flow in discourse processing and provides a computationally tractable reasoning mechanism to understand the rhetorical relations in discourse.

2.2.3. Segmented Discourse Representation Theory (SDRT)

Asher and Vieu (2005) discuss the distinction between subordinating coordinating discourse relations, which is central to Segmented Discourse Representation Theory's (SDRT) view of discourse structure. As the discussion by Asher and Vieu (2005) is highly relevant to the current study, their ideas regarding subordinating and coordinating discourse relations will be summarized below.

SDRT defines a lead sentence of a paragraph which presents that paragraph's main idea. The other sentences in the paragraph elaborate or support the main idea with the arguments they involve. Thus, some parts of a paragraph have a subordinate role compared to the parts involving the main idea. In addition, the elaborative bits of information in the paragraph are at the 'level' of detail, as named by Asher and Vieu (2005, p.592), and they coordinate together to amplify on or to support the main idea in the text.

SDRT is influenced by RFC in that the discourse constituents on the right frontier can provide attachment points for new information. Consider the following example:

- (10) a. John had a great evening last night.
b. He had a great meal.
c. He ate salmon.
d. He devoured lots of cheese.
e. He then won a dancing competition.
f. ? Then he had a great dessert.
f'. # It was a beautiful pink.
f''. John had lots of fun.

(Asher & Vieu, 2005, p.592)

(10b) elaborates on (10a) in the discourse above. (10b) and (10c) elaborate on the meal presented in (10b). (10e) also elaborates on (10a), however, it does not

elaborate on the meal in (10b). It rather forms a narrative with (10b). Therefore, (10e) cannot be a part of the discourse segment formed by (10c) and (10d). Considering this discourse structure, (10b) cannot end up on the right frontier of the tree structure. Because of RFC, (10e) cannot attach to (10b) so it goes to the right frontier of the tree structure. Thus, the new information in (10f) should attach to (10e). However, it does not attach to (10e) because it is further elaboration of (10b). Since (10b) is not on the right frontier, (10f) cannot find an attachment point. The RFC also explains the incoherence of (10a-f’). Similarly, (10f’) cannot find an attachment point on the right frontier of the tree structure. However, ‘to have fun’ in (10f’’) is anaphoric in that it may refer back to the eventuality in either the top constituent (10a) or in the lower constituent (10e) (Asher & Vieu, 2005, p.593). Figure 3 below illustrates the hierarchical structure for example (10):

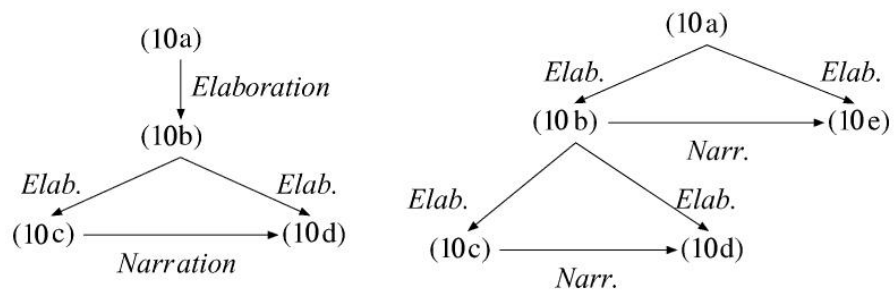


Figure 3 illustrates the Hierarchical structure of example (10).

Asher and Vieu (2005) keep the RFC in their analysis of discourse structure. The attachment of new information is governed by the right frontier in their analysis as well. The constituents may be directly linked to an open point for attachment, to the constituent involved as the last element or to the dominating constituent through some subordinating relation. Besides this, in anaphora resolution process, the accessibility of the antecedent is governed by a ‘look-left-one-step-only-or-look-up’ rule (Asher & Vieu, 2005, p.595). According to this rule, referents in the constituent where the current constituent is attached are accessible (one step, left or up) as well as those of all constituents that dominate the current constituent (up).

Asher and Vieu (2005) keep the RFC but they extend it for rhetorical relations. This is because the account of RFC for discourse structure does not distinguish between different kinds of discourse relations because it describes the

syntactic level of discourse structure. However, in a phenomenon like anaphora resolution in which the discourse structure affects the semantic and pragmatic features of the text, the pronoun assigns an antecedent not only depending on the syntactic constraints but also on the semantic constraints.

The discourse relations Asher and Vieu (2005) take into account are binary; either coordinating or subordinating. ‘Elaboration’ accounts for subordination in discourse segments and ‘narration’ accounts for coordination. The topicality is affected by subordination and coordination. The shifts in topic across the segments are coordinately connected. With coordination, before moving to the second segment in a discourse, the communicative intentions of the first segment should be satisfied. The topic of a constituent in a segment, on the other hand, is subordinate. Satisfying the communicative intentions of a subordinate constituent depends solely on satisfying the intentions of the dominating constituent.

Finally, it is important to note that subordination and coordination has an effect on the temporal order of the discourse. Coordination provides temporal progression of events, whereas subordination provides elaboration on the topic presented in the lead sentence of the discourse segment in focus.

The discussion as a result of the experiments in this thesis takes place mostly within the framework of SDRT and RFC (see Chapter 4). However, another important theory, Centering Theory, for relational structures in discourse will be explained below for the sake of completeness.

2.2.4. Inferencing within the Framework of Centering Theory (CT)

Grosz and Sidner (1986) argue the theory of discourse structure. In the discourse structure discussed by Grosz and Sidner (1986), there are three components: linguistic structure, intentional structure, and attentional state. Discourse segments are grouped by the linguistic structure. The purpose of the discourse segment is realized by the intentional structure. Finally, the relations which are salient in the discourse are established through attentional state. Grosz and Sidner (1986) define two levels of attentional state. The first one is the global level, which is concerned with the relations between discourse segments and the ways in which attention shifts between them.

The local level of attentional state, on the other hand, is concerned with changes of attention within discourse segments. The Centering Theory (CT) by Grosz, Joshi and Weinstein (1995) deals with this phenomenon. CT (Grosz et al., 1995; Walker, Joshi, & Prince, 1998) is an attempt to relate focus of attention, choice of referring expression, and perceived coherence of utterances within a discourse segment. It particularly tries to account for the local coherence within the attentional state in focus. It is also meant to be a theory of salience. This is because it attempts to predict which entities will be the most salient at any given time. Information structure of a clause bound to the discourse segment in focus instructs the reader on how to update his or her discourse model with the information in the current sentence alone. Consider examples (11) and (12):

- (11) a. Jeff helped Dick wash the car.
b. He washed the windows as Dick waxed the car.
c. He soaped a pane.
- (12) a. Jeff helped dick wash the car.
b. He washed the windows as Dick waxed the car.
c. He buffed the hood.

(Walker et al., 1998, p.1)

Inferential theories of discourse, e.g. Hobbs (1985), suggest that examples (11) and (12) are same in coherence. These theories predict that 'he' in (11c) refers back to 'Jeff' because the activity of soaping is related to the activity of washing. However, 'he' in (12c) refers back to 'Dick' because buffing is related to the activity of waxing.

According to CT, processing the discourse in (11) is easier compared to the one in (12). This is because both (11a) and (11b) present a discourse entity that is realized by Jeff. However, in (12), (12c) shifts the 'center' of discourse to an entity that is realized by Dick. Thus, unlike inferential theories of discourse that assume (11) and (12) coherent at the same level, CT considers (12) less coherent than (11).

Constituent segments like (11a), (11b) in the discourse in (11) and (12a), (12b) in (12) form the discourses in CT. For each utterance in a discourse segment, semantic entities constitute the discourse model and these semantic entities are the centers. The 'forward-looking center' $C_f(U_i, D)$ is the semantic entities that are

evoked by an utterance U_i in a discourse segment D . The 'backward-looking center' $C_b(U_i, D)$ is a special part of C_f and it is the semantic entity that U_i most centrally concerns. Thus, the set of forward-looking centers is ranked according to discourse salience (Walker et al., 1998, p.6).

There are three types of transition relations between the discourse segments in CT. The first is 'continuation' of the center from an utterance U_i in a discourse segment D not only to the next utterance in the same discourse but also to the previous utterances. The second is 'retention' of the center from one utterance to the next, and the last one is 'shifting' the center if the center is neither continued nor retained (Grosz et al., 1995, p.210). To better illustrate the transitional relations in CT, consider the examples (13) and (14):

- (13) a. John went to his favorite music store to buy a piano.
b. He had frequented the store for many years.
c. He was excited that he could finally buy a piano.
d. He arrived just as the store was closing for the day.
- (14) a. John went to his favorite music store to buy a piano.
b. It was a store John had frequented for many years.
c. He was excited that he could finally buy a piano.
d. It was closing just as John arrived.

(Grosz, Joshi & Weinstein, 1995, p. 206)

According to Grosz et al. (1995), the discourse in (13) is more coherent than the one in (14). This is because the discourse in (13) describes several actions John took; thus, it centers around a single entity, i.e., John. However, this is not the case for the discourse in (14). The center is 'John' in (14a) but it is 'it' referring back to the music store in (14b). It is again 'John' referred back by the pronoun 'he' in (14c) and it is again the music store referred back by the pronoun 'it' in (14d). The readers of the discourse in (14) would perceive a change in the entity in focus going from (14a) to (14b), (14b) to (14c) and (14c) to (14d). Briefly, the center shifts from one entity to another in (14), so (14) does not have a single clear center of attention. This lack of continuity in what the discourse is about makes this discourse less coherent compared to (13).

Discourse in (13) and (14) actually convey the same information; however, the way they convey the information differs. They are different not in terms of content but in terms of the expression. Grosz et al. (1995) suggest another important point regarding the form of expression in a discourse. They conjecture that the form of expressions substantially affects the resource demands made upon a hearer or reader in discourse processing. That is to say, different forms of referring expressions and different syntactic forms increases the inferencing load on the hearer or reader of a discourse (p.207). Thus, it is because of the syntactic form of the expressions in (14b) and (14d) that causes flipping the center back and forth in (14). Grosz et al. (1995) use the term 'inference load' for this phenomenon. It is the load that is placed upon the hearer or the reader to refer to the resources required to extract information from a discourse. This load mainly occurs as a result of the linguistic expressions used in the target discourse (p.208).

All in all, CT tries to explain the local coherence within the attentional state in focus. It is an attempt to relate focus of attention, choice of referring expression, and perceived coherence of utterances within a discourse segment. It is also called the theory of salience as it tries to predict which entities will be the most salient in a discourse.

2.3. Ambiguous Nature of the Antecedents to Abstract Object Anaphors

Some of the pioneering studies reviewed in 2.1.1 emphasize the ambiguous nature of abstract object anaphors, e.g., Botley (2006), Artstein and Poesio (2008), Recasens (2008), Vieira, Salmon-Alt and Gasperin (2005) and Poesio et al. (2006). These studies indicate that antecedents to abstract object anaphors lack precise boundaries. All these corpus based studies use coders to annotate the antecedents of anaphors, and they all report that many annotations by the coders do not exactly match, but they overlap. This is because making inferences while retrieving the antecedent to abstract object anaphor is both complex and unclear. The challenge stems from the ambiguous nature of the abstract anaphors without precise boundaries. As a result, this ambiguous nature of abstract anaphors makes it difficult to come up with hard and fast decisions in the analysis of abstract object anaphors in the studies mentioned above.

When it comes to the antecedents to the concrete object anaphors, upon eliminating the non-matching candidates in terms of structural and semantic constraints (i.e. the reference in the real world), the choice for the antecedent is made among the remaining candidates. Recall example (2) provided in 2.2: The concrete object anaphor 'he' easily resolves. This is because the structural constraint based on the gender, specifies 'John' as the only candidate to be the antecedent.

While it is easy to specify the structural constraints (e.g., gender and number agreement constraints) for concrete object anaphors, it is probable for abstract object anaphors to pass the structural constraints. Thus, they may not have a precise antecedent that coders would agree on. Consider the following discourse in which both the bare anaphor 'that' and the label anaphor 'this situation' take place:

(15) For me, the symbol for this type of woman turning into a pain in the neck for a man in any situation has been those high heels; I was going to write about those heels, about the women standing on those heels. Because, **that** was a stance. It was the stance pertaining to the claim they make in life. It was the stance pertaining to their claims everywhere and in any situation. I was not able to explain **this situation** with the point of view of the patriarchal scenarists who have displayed women as in need of protection, and men as the only heroes⁶.

In example (15), the antecedent to the bare anaphoric expression 'that' is an ambiguous one. It may be the way women stand on those high heels. It may also be the characteristics of women as being a pain in the neck for men. Or, it may be the stance of the writer as his wish to write about those women. Finally, another possibility is that the bare anaphor 'that' functions cataphorically, so it may be referring forward to the (a) the claim women make in life, (b) the claim women make everywhere and in any situation, or both (a) and (b). It is difficult to eliminate these competing candidates because they all pass both the structural and semantic

⁶ This text is translated from Turkish. The original text is as follows:

Benim için, her durumda erkeğin başına bela olan bu kadın tipinin simgesi işte o yüksek topuklar olmuştu; bir biçimde o topukları, o topukların üzerinde yükselen kadınları yazacaktım. **Bu** bir duruştu çünkü. Kadınların hayattaki iddialarına ait bir duruştu. Her yerde, her durumda, her şeye karşı gösterdikleri bir iddianın duruşuydu. Yalnızca erkeği kahraman kadını himayeye muhtaç gösteren erkek egemen senaristlerin hayat görüşleriyle açıklayamıyordum **bu durumu** (Mungan, 2004, p.13).

constraints. It is also difficult to specify the exact boundaries of these possible antecedents. As a result, the antecedent to the bare anaphor 'that' is a truly ambiguous one.

Similar confusion can arise in an effort to specify the antecedent the label anaphor 'this situation' taking place in the same discourse. The possible competing candidates might be (a) women turning into a pain in the neck for a man in any situation, (b) women standing on those heels, (c) the claim women make in life, (d) the claim women make everywhere and in any situation. While the readers of this discourse may specify the antecedent to the label anaphor 'this situation' as one of the candidates, i.e., (a), (b), (c) or (d), they may also specify a combination of several of these possible candidates. Or some readers may even specify all of these options as the antecedent. All in all, it is not easy to identify the exact boundaries of the abstract object anaphors.

As stated in 2.2., Webber (1988a) also highlights the ambiguous nature of abstract object anaphors (i.e., 'discourse deixis' in her terms), especially when it comes to specifying the boundaries of the antecedents to this type of anaphors. According to Webber (1988a), lack of inter-annotator agreement to specify the exact boundaries of the antecedent may bring different interpretations in the resolution process of the target anaphor with respect to differently specified extensions of the antecedent. In a sense, Webber (1988a) claims that it is not the anaphor itself but it is the antecedent that lacks a specific boundary. Therefore, the antecedent is the ambiguous one in the process. Recall example (3) by Webber (1988a, p.166) presented in 2.2. repeated as (16) here.

- (16) It's always been presumed that when the glaciers receded, the area got very hot. The Folsom men couldn't adapt, and they died out. **That's** what is supposed to have happened. It's the textbook dogma. But it's wrong. They were human and smart. They adapted their weapons and culture, and they survived.

(Webber, 1988a, p.116)

It has already been discussed in 2.2. that readers of this discourse may specify only (a), only (b) or (a+b) as the antecedent to the abstract object anaphor 'that'.

- (a) It's always been presumed that when the glaciers receded, the area got very hot.
- (b) The Folsom men couldn't adapt, and they died out.
- (a+b) It's always been presumed that when the glaciers receded, the area got very hot. The Folsom men couldn't adapt, and they died out.

As in example (15), readers of (16) may not agree on the precise boundaries of the antecedent.

As it is clear, the abstract object anaphors are ambiguous in nature. While it may be possible for one of the competing candidates to pass both the structural and semantic constraints, it may still not be likely to determine the exact boundaries of the antecedent, which causes abstract object anaphors to have a truly ambiguous nature.

2.4. Methodological Frameworks in Anaphora Annotation Studies

If coders can be shown to agree on the categories assigned to units to an extent determined by the purposes of the study, the data are assumed to be reliable. This is the main assumption behind the methodologies to measure intercoder agreement on the specification of antecedents in annotation studies. Provided that different coders generate consistently similar results, then it is possible to infer that they have a similar understanding of the annotation guidelines, and they can be expected to perform consistently under this understanding (Artstein & Poesio, 2008, p. 557).

One of the earliest and the simplest measure of agreement between two coders is *percentage of agreement*. According to this measurement, the agreement is the arithmetic mean of the agreement value. The agreement value agr_i for all items $i \in I$, defined as follows:

$$agr_i = \begin{cases} 1 & \text{if the two coders assign } i \text{ to the same category} \\ 0 & \text{if the two coders assign } i \text{ to different categories} \end{cases}$$

Then, the percentage of agreement over the values agr_i for all items $i \in I$ is:

$$A_o = \frac{1}{i} \sum_{i \in I} \text{agr}_i$$

(Artstein & Poesio, 2008, p.558)

However, as discussed by Artstein and Poesio (2008), percentage agreement is not a reliable measurement. This is because it allows intercoder agreements that occur by chance. For example, higher percentage agreement is expected when one category is much more common than the other.

To eliminate the agreements occurring by chance, the three best-known coefficients, S (Bennett, Alpert, and Goldstein 1954), π (Scott 1955), and κ (Cohen 1960) use the following formula:

$$S, \pi, \kappa = \frac{A_o - A_e}{1 - A_e}$$

The value $1 - A_e$ measures how much agreement over and above chance is accessible; the value $A_o - A_e$ indicates how much agreement beyond chance is actually found. The ratio between $A_o - A_e$ and $1 - A_e$ tells us which proportion of the possible agreement beyond chance is actually observed (Artstein & Poesio, 2008, p.559).

While these coefficients (S, π, κ) try to adjust the misleading agreement results due to the chance factor, they still have serious limitations. These coefficients treat the disagreements equally. However, disagreements may not always be alike due to semantic and pragmatic features of discourse. As exemplified by Artstein and Poesio (2008, p. 564), even for the relatively simple case of dialogue act tagging, a disagreement between an 'accept' and a 'reject' interpretation of an utterance is clearly more serious than a disagreement between an 'info-request' and a 'check'. As argued by Artstein and Poesio (2008, p. 564), for tasks such as anaphora resolution, the reliability is determined by measuring agreement on sets (coreference chains); thus, allowing for degrees of disagreement becomes essential. As a result, for such tasks, S, π and κ are not very useful.

Krippendorff's α , on the other hand, is based on a related formula expressed in terms of disagreement, which is more reliable for the analysis of the current study.

2.4.1 Krippendorff's Alpha

The coefficient α (Krippendorff, 1995; Krippendorff, 2004a; Krippendorff, 2004b) has similar claims to those of Fleiss Kappa measurement methods in terms of using Alpha agreement coefficients that are calculated. As stated by Artstein (2008), the calculation is based on the overall distribution of judgments so the coefficient α disregards the annotators producing those judgments. As discussed by Krippendorff (2004a) and Krippendorff (2004b), coefficient α is different from the other similar type of measurements because it is possible to obtain diversity via several weighting functions like interval and boundary. Another important note regarding Krippendorff's Alpha is, as stated by Krippendorff (1995), the way algorithm functions: it does not consider observed and expected agreements but considers observed and expected disagreements. Assuming that $\alpha = 1$ represents exact agreement and $\alpha = 0$ represents exact disagreement, the formula put forward by Krippendorff (2004a) and Krippendorff (2004b) is as follows:

$$\alpha = 1 - \frac{D_o}{D_e}$$

D_o stands for the observed disagreement and D_e stands for the expected disagreement when chance prevails. When $D_o = 0$; thus, $\alpha = 1$, and the α score indicates perfect reliability. When $D_e = D_o$, so $\alpha = 0$, α score indicates absence of reliability. α can also be a negative score. Negative values are resulted from sampling errors or systematic disagreements. Sampling errors occur for several reasons. For example, when the sample size is too small, each sample might have a large effect on α score. Systematic disagreements occur when coders agree or disagree, or when they have non-overlapping judgements of the given instructions (Krippendorff, 2004a, p.222).

As stated by Artstein and Poesio (2008) and Landis and Koch (1971), the results of kappa-like agreement measurements are mostly interpreted in six categories:

- 1- Measurement > 0.8: Perfect agreement
- 2- 0.8 > Measurement > 0.6: Substantial agreement
- 3- 0.6 > Measurement > 0.4: Moderate agreement

4- $0.4 > \text{Measurement} > 0.2$: Fair agreement

5- $0.2 > \text{Measurement} > 0.0$: Slight agreement

6- $0.0 > \text{Measurement}$: Poor agreement

It is important to denote why the Krippendorff's agreement coefficient was selected during the statistical analysis instead of others like Kappa's agreement coefficient. Kappa's agreement coefficient was introduced by Cohen (1960) and has been implemented in computational linguistic studies by Carletta (1996). It is significant to underline that Cohen's Kappa belongs to a chance-corrected agreement measurement family which aims at measuring agreement between two annotators. Moreover, by using Cohen's Kappa, it is possible to measure the degree of agreement since annotators are evaluated within a nominal scale and the chance between them are corrected (Fleiss, 1971). Cohen's Weighted Kappa (1968), a version of kappa by Cohen, could also be considered as a generalization of Cohen's Kappa. As a disadvantage and limitation for linguistic studies, both of these coefficients treat all of the disagreements equally but it is crucial to note that disagreements could not be considered alike especially for semantic and pragmatic purposes (Artstein & Poesio, 2008). Moreover, for the tasks involving coreference chains as in anaphora resolution, degrees of disagreement should be investigated where the other coefficients appear to be not useful (Artstein & Poesio, 2008). In contrary, Krippendorff's alpha coefficient -having proposed for multiple coders, different magnitudes of disagreements, and missing values (Krippendorff, 1980, 2004)- would be a more appropriate choice for this empirical study, since the main target has been the anaphora resolution. By allowing uniform reliability standards, Krippendorff's alpha could be applied to a great diversity of data, since it is applicable to a wide range of conditions including any number of values per variable, any number of observers (not just two as in traditional approaches), small and large sample sizes alike, several metrics (scale of measurements), data with missing values (within the recording units by some participants) (Krippendorff, 2004). Negative values could be result of two different kinds of errors: sampling errors and systematic disagreements. Having too small sample sizes could lead to the first type of errors, sampling errors, in which such deviations of alpha could be observed including getting below the level of minimum 0. Secondly, systematic disagreements are also

crucial phenomena that occur for the conditions in which “participants/observers agree to disagree or pursue opposing interpretations of the instructions given to them” (Krippendorff, 2004). It is clear that perfect reliability could not be reached due to all of the observed disagreements, but it is also vital to note that systematic disagreements could decrease the alpha value below the by-chance level.

However, it is important to note that achieving good agreement cannot ensure the validity of the analysis. This is because two coders of the same antecedent may well share the same prejudice and they may be objectively wrong in their annotations (Artstein & Poesio, 2008, p. 557).

More importantly, as explained in 2.1.1 and 2.1.2, all annotation studies conducted on abstract object anaphors (e.g., Botley, 2006; Francis, 1994; Poesio & Artstein, 2008; Recasens, 2008) indicate the difficulty of specifying the exact boundaries of the antecedents to such kind of anaphors. There might be some overlapping annotations-which cannot be considered as a complete disagreement. At this point, Krippendorff's α falls short. As suggested by Artstein and Poesio (2008), Passonneau's (2004) proposal on agreed sets can be a solution to this problem.

2.4.2 Passonneau's Proposal

Passonneau (2004) suggests that annotation of anaphors should be measured by using 'sets' of annotated discourse entities. These sets allows for partial agreement among anaphoric chains when two annotators agree on some part of the coreference chain but not all of it. The sets, suggested by Passonneau (2004) are as follows:

- *Main set*: This is the longest annotated antecedent
- *Subsumption*: This is a relation in which there are subsets overlapping with the main set
- *Intersection*: There is a relation in which two different annotations intersect with each other at some point.
- *Disjunct*: This is the state when two or more annotations do not overlap with one another but when they are all possible interpretations to be considered as antecedents to the target anaphor.

Passonneau (2004) calculates the agreement rate by assigning each set with a referential value. If two sets A and B are identical, they are assigned with 0; if they have subsumption relation, they are assigned with 0.33; if they have an intersection relation, they are assigned with 0.66, and if they have a disjunct relation, they are assigned with 1 (p.1505). The following formula illustrates these referential values:

$$d_p = \begin{cases} 0 & \text{if } A = B \\ 1/3 & \text{if } A \subset B \text{ or } B \subset A \\ 2/3 & \text{if } A \cap B \neq \emptyset, \text{ but } A \not\subset B \text{ and } B \not\subset A \\ 1 & \text{if } A \cap B = \emptyset \end{cases}$$

(Artstein & Poesio, 2008, p. 557)

In the current study, both Krippendorff's α coefficient and Passonneau's sets proposal were used for the analysis of data obtained from the experiments (discussed in detail in Chapter 4). Krippendorff's α coefficient was used for the agreement analysis at word boundary and interval levels. Passonneau's method was used to gain better insight into accommodating distinct interpretations of target anaphors while preserving the same units of interpretations.

CHAPTER 3

A VIEW FROM TURKISH

In Chapter 1, types of abstract object anaphors were explained. To explain these types, i.e. labeling, situation reference and discourse deixis, various examples from English discourse in which these anaphoric expressions are used were provided. In Chapter 2, the accessibility relation between the antecedent and the anaphors was discussed with respect to Consten and Knees' (2005) *abstractness hierarchy* between the anaphor and the antecedent, and also with respect to the *right frontier* of the discourse structure (Webber, 1991).

This chapter will serve as the background to the experimental analysis of the thesis and the discussions that follow. Therefore, various occurrences of abstract object anaphors in Turkish discourse will be exemplified (Section 3.2) with the examples taken from the Turkish Discourse Bank (TDB)⁷. The theory of abstractness hierarchy by Consten and Knees (2005) and right frontier operations by Webber (1991) will be adapted to Turkish discourse (Section 3.3). Finally, studies on anaphora in Turkish will be reviewed (Section 3.4).

⁷ Zeyrek, Demirşahin, Sevdik-Çallı, Ögel-Balaban, Yalçınkaya and Turan, 2010.

3.1. The Data Source: the Turkish Discourse Bank (TDB)

The Turkish Discourse Bank (TDB) is a 500,000-word-subcorpus of the two-million-word METU Turkish Corpus (MTC) (Say, Zeyrek, Oflazer & Ozge, 2002). The subcorpus comprises a wide range of genres such as fiction, interviews, memoirs, news articles etc. All the Turkish examples in the thesis are taken from the TDB, unless noted otherwise. The TDB is a resource of discourse, where discourse connectives are annotated along with their two arguments, by extending the annotation style of Penn Discourse Tree Bank (PDTB).

As in PDTB, connectives belonging to coordinating conjunctions, subordinating conjunctions and discourse adverbials are annotated in TDB. This is because discourse connectives are conveniently identified from these three syntactic classes. The major differences from the PDTB annotation scheme are as follows: (1) attribution is not annotated, (2) only overt connectives are annotated and (3) the nominal arguments of connectives are annotated where they denote an abstract object.

During the annotation cycle, the annotators went through the whole subcorpus to annotate a given connective at a time. The disagreements were discussed and solved by the project team. 60 types of discourse connectives were annotated. These are coordinating conjunctions (*ve* ‘and’, *ya da* ‘or’, *ama* ‘but’), subordinating conjunctions (*için* ‘for’, *rağmen* ‘although, despite’), converbs/simplex subordinators (*-ince* ‘when,’ *-ken* ‘while, now that’) and anaphoric connectives (*bundan başka* ‘in addition to/separate from these’, *bunun sonucunda* ‘as a result of this,’ *bunun için* ‘due to/for this reason’, *buna rağmen* ‘despite this’) and discourse adverbials (*oysa* ‘however’, *öte yandan* ‘on the other hand’, *sonradan* ‘then’).

3.2. Types of Abstract Object Anaphors

In 1.2, three types of abstract object anaphors were explained in detail. The first type presented in 1.2 was *labeling* which was put forward by Francis (1994), the second was *situation reference* by Fraurud (1992, 1992a), and the last type was *discourse deixis* argued by Lyons (1977). These types of anaphors were searched from TDB. All of them were attested in Turkish discourse, too. These will be explained in the sections that follow.

3.2.1. Labeling in Turkish

As stated in 1.2.1., as a result of the analysis of English nominal groups functioning as abstract object anaphors, Francis (1994) categorizes these nominals in two groups: *retrospective labels* (RLs) and *advance labels* (ALs). These labels occur in Turkish discourse, too.

The following discourse is an example of retrospective labels (functioning anaphorically) in Turkish:

- (1) Evlilik için hazırlıklar sürerken Ante, yeni bir görev emriyle yollara düştü. Yine aylarca süren bu sefer sonunda küçük de olsa bir başarı elde edilmişti. İstanbul'a döndüğünde Ermeni Ante hemen padişahın huzuruna çıkarak [görevinden affını istedi]. Neden olarak evliliğini gösterdi. Padişah **bu isteği** kabul etmedi, ama kendisine uzun süreli izin verdiğini söyledi.

(TDB, 1231)

While the preparations for the wedding proceeded, Ante set off with an order for a new duty. He was able to achieve success -a little one, though- as a consequence of this expedition which lasted again for months. When he came back to Istanbul, upon appearing before the Sultan, Armenian Ante asked for [his resignation of the office]. He presented this marriage as the reason. The Sultan did not accept **this request**, but he said that he gave him a long leave.

In example (1), the label anaphor 'bu istek' (Eng. 'this request') refers back to Ante's request of his resignation from his office. In Turkish, advance labels (i.e., labels acting cataphorically) differ from the ones in English. While English advance labels can be constructed as 'this+N' just like the retrospective labels, Turkish allows the 'şu+ad' (Eng. 'that+N') construction. This is because 'şu' (Eng. 'that') corresponds to the entities coming after itself in Turkish (Lewis, 1967, p.71). Consider example (2):

- (2) Charles de Gaulle Havaalanı'nda uçağa binmeden önce "Son Kadın" adlı öyküm için **şu notu** düştüm: ["Ölüm ona kadın suretinde görünmüştü." "İlk Kadın"ı ise dün bitirebildim ancak. Yüz sayfadan çok tuttu. Ağustos sıcağında Glacière Sokağı'ndaki evde yalnızdım. Eski bir yapının çatı katında

bulutlarla beraber...]

(TDB, 199170)

Before getting on the plane, at Charles de Gaulle Airport, I jotted down **the following note**⁸ for my story "The Last Woman": ["Death seemed to him in the image of a woman." I was able to finish "The First Woman" only yesterday. It has been more than a hundred pages. I was alone at home on the Glacière Street in the heat of August. In the attic of an old building together with the clouds...]

Example (2) demonstrates the occurrence of an advance label in Turkish. The antecedent to the label anaphor 'şu notu' (Eng. 'the following note') comes after the anaphor as indicated in the square brackets.

Francis (1994) categorizes the retrospective labels depending on the head nouns used in the labels in four groups: Illocutionary nouns, language-activity nouns, mental process nouns and text nouns. The occurrences of all of these categories were also observed in Turkish, as exemplified below.

3.2.1.1. Illocutionary Nouns: As stated by Francis (1994, p.90), these are the nominalizations of the verbal processes.

(3) "... Belki sadece sokak adını yanlış not ettiniz. [Belki o da sizi arıyor kentte. Bir türlü buluşamıyorsunuz.]"

"Durum böyle olsaydı, ona şü kentın en canlı yeri olan alandaki kahvelerden birinde rastlardım şimdiye kadar..."

"Belki de rastlayacaksınız ona."

"Bunca aramadan sonra..."

Ardından Jul, ellerini küçük masanın üzerinde birleştirip, okuduğu kitaptan söz etti. Ona göre, filozof da, **bu arayışa** benzer bir şeyi anlatıyordu.

(TDB, 6231)

⁸ The exact English translation of 'şu notu' is actually 'that note'. However, 'that note' disturbs the natural flow of the context; therefore, it is translated as 'the following note'.

"... Perhaps you have noted down the street name incorrectly. Perhaps she is also looking for you in the city. You somehow cannot meet."

"If this had been the situation, I would have run into her till now in one of the cafes in that square which is the liveliest place of the city..."

"Maybe you will run into her."

"After so much search..."

Later on Jul gathered his hands on the little table, mentioned the book he was reading. For him, the philosopher was also telling something similar to **this quest**.

In example (3), the label anaphor 'bu arayış' (Eng. 'this quest'), refers back to the man's search for the woman. The head noun 'arayış' (Eng. 'quest') is the nominalization of the verb 'aramak' (Eng. 'to search for'). The derivative suffix *-Iş* nominalizes the verb. Thus, the head of this label anaphor demonstrates the occurrence of illocutionary head nouns in Turkish discourse. Other examples attested in TDB are as follows:

- (4) AB Komisyonu sözcülerinden Jean-Christophe Filori, dün basına yaptığı açıklamada, Türkiye'nin AB'ye katılım şansını arttıran reformların TBMM'den geçmesini memnuniyetle karşıladıklarını söyledi. Filori, [reformların hayata geçirilmesinin AB Komisyonu tarafından "yakından izleneceğini"] anlattı. [AB Komisyonu'nun, onaylanan reform paketini analiz edeceğini ve Ankara'dan bazı hususları aydınlatıcı bilgiler isteyeceğini] belirten Filori, ekim ayında aday ülkelere ilişkin olarak yayımlanacak yıllık değerlendirme raporunda **bu gelişmelerin** dikkate alınacağını duyurdu.

(TDB, 10560000)

One of the spokesmen from EU Council Jean-Christophe Filori, in the statement he gave to the press yesterday, said they were pleased that the reforms increasing Turkey's chance to get into the EU were approved by TBMM⁹. Filori explained that [the reforms being put into practice were going

⁹ The Grand National Assembly of Turkey

to be "watched closely" by the EU Council]. Pointing out that [EU Council will analyze the approved reforms and will request information enlightening some points], Filori announced that **these developments** were going to be taken into consideration in his annual evaluation report which was going to be published in October regarding the candidate countries.

- (5) [MİT'in istihbari rapor göndermesinin önüne geçmek] isteyen hükümet, diğer yandan da [Cumhurbaşkanı'nın üçlü atamalardaki yetkisine son vermeye hazırlanıyor.] Hükümet, **bu çalışmalarıyla** "irticai soruşturmalara uğramış AKP kadrolarını" devletin üst düzeyine taşıyabilmek için mevcut engelleri ortadan kaldırmayı hedefliyor.

(TDB, 10630000)

The government trying [to preclude MİT¹⁰'s sending intelligence report], on the other hand, [is getting prepared for calling the President's authority in triple designations off.] The government, with **these practices**, is aiming to terminate the potential obstacles to move the JDP¹¹ staff that experienced reactionary questioning" to the top of the government.

In example (4), the head noun 'gelişme' (Eng. 'development') is the nominalization of the verb 'gelişmek' (Eng. 'to develop'). In example (5), the head noun 'çalışma' (Eng. 'practice') is the nominalization of the verb 'çalışmak' (Eng. 'to practise'). The derivative suffix *-mA* nominalizes the verbs in these examples.

3.2.1.2. Language Activity Nouns: As stated by Francis (1994, p. 91), these head nouns refer to some kind of a language activity and the result of it. They are very much similar to illocutionary nouns; therefore, the distinction between illocutionary nouns and language activity nouns is blurred in terms of their functions in discourse (Francis, 1994, p. 90). However, language activity nouns are not the nominalizations of the verbal processes. Consider example (6):

¹⁰ The National Intelligence Agency

¹¹ The Justice and Development Party

- (6) DSİ su ürünleri mühendisleri ise [1 litre suda 2 miligram amonyak bulunmasının]tatlısu canlıları öldürmeyeceğini ancak Ergene Nehri'nde **bu oranın** 27 miligram olduğunu belirlediklerine dikkat çektiler.

(TDB, 10570000)

DSİ¹² aquaculture engineers, on the other hand, drew the attention to the fact that [the presence of 2 milligram of ammonia in 1 liter of water] would not kill freshwater creatures but **this proportion** was 27 milligram in the River Ergene.

In example (6), the use of 'oran' (Eng. 'proportion') as the language activity head noun in the label is illustrated. It refers back to the proportion of ammonia to 1 liter of water. The head noun 'oran' demonstrates the occurrence of language activity head nouns in Turkish discourse. Some other examples attested in TDB are as follows:

- (7) Bu yönetimin, Sayın Derviş aracılığı ile uyguladığı politika, [kredi adı altında, borç al ve bununla borçları öde politikasıdır]... Ekonomik geleceğimiz IMF'ye ve Sayın Derviş'e teslim edilmiştir. İki yılı aşkın bir süredir uygulanan **bu politikalarla** hiçbir yere gitmediğimiz ve gidemeyeceğimiz ortadadır.

(TDB, 10590000)

The policy this government carries out through Mr. Derviş is [the policy, in the name of credit, to borrow and pay the debts with this.] ... Our economic future is handed in to IMF and Mr. Derviş. It is obvious that we have not proceeded and will not be able to proceed with **these policies** that have been in practice for more than two years.

- (8) Bir Cinayet Romanı'nda da budur durum. Üstelik [gerçeklikle kurmaca arasındaki ilişki okuru, romanı anlamlandırırken bocalatacak kadar karmaşıktır] da. Her romanın, oyunun, öykünün gerçek dünya ile uzak ya da yakın ilişkisi vardır, ama Bir Cinayet Romanı'nın yapısı ve dile getirdiği roman anlayışı **bu soruna** özel bir yer veriyor kanımca.

(TDB, 26231)

¹² The State Water Supply Administration

This is the situation in *The Novel of a Murder*, too. Moreover, it is also [so complicated to confuse the reader while making sense of the connection between the reality and the fiction.] Every novel, play, story has more or less a connection to the real life, but the plot of *The Novel of a Murder* and the literary approach it has brought into the language give special attention to **this problem** in my opinion.

The head noun 'politika' (Eng. 'policy') in example (7) and the head noun 'sorun' (Eng. 'problem') in example (8) illustrate the occurrences of language activity nouns in the labels in Turkish discourse.

3.2.1.3. Mental Process Nouns: As stated by Francis (1994, p. 91), these head nouns refer to cognitive states and processes and their results. Consider the following example:

(9) ["Evet, memnunum hayatımdan. Ama şu anda onunla bir bağlantı kuramıyorum. Sanki ustaca koparıldım ondan. Şimdi yeni bir kimlikle dolaşıyorum dünyada. Düşünmeye başladım; acaba benim hayatımı, düşlerimi, özgürlüğümü, anılarımı ve şu anda ulaşamadığım, bu içinde olduğum boyutta varlıklarına rastlayamadığım sevdiklerimi kim almak istiyor elimden?"]

...

"Ustaca işlenmiş bir cinayet. Ortada hiçbir ipucu yok. Çünkü öldürülen yok. Ama bir insanın rayı değiştiriliyor; başka bir yaşamın içine sokuluyor. Dediklerinizi uzun zaman unutamam. Ama, acaba **bu düşünceniz** gerçek mi?" diye sordu Hacı Murat. "Acaba abartıyor musunuz gerçekleri biraz diye düşünüyorum. Niçin böyle bir şey yapılsın? Niçin durup dururken bir insanın kimliği, yaşamı, şu hayattaki konumu değiştirilsin? Tuhaf bir şey bu."

(TDB, 2113)

["Yes, I am pleased with my life. But now, I cannot come into contact with it. It is as if I was skillfully removed from it. Now I am wandering in the world with a new identity. I have started to think: Who an earth wants to deprive me of my life, my dreams, my freedom, my memories and the beloved ones

whom I cannot reach now, whose presence I cannot run across in this dimension I am inside now?"]

...

"It is a skillfully committed murder. There is no clue at all. Because there is nobody killed. But a person's track is changed; he is led into another life. I will not forget what you have said for a long time. But, is **this thought** of yours real?" asked Hacı Murat. "I am thinking whether you are exaggerating the truth a little? Why could such a thing be done? Why should a person's identity, life, position in this life be changed for no reason at all? This is such a weird thing."

In example (9), the label anaphor 'bu düşünce' (Eng. 'this thought'), refers back to the thoughts of the person presented in the first paragraph. The head noun 'düşünce' (Eng. 'thought'), referring to a cognitive process of mind, demonstrates the occurrence of mental process head nouns in Turkish discourse. Other examples from TDB are as follows:

- (10) Kırklareli Çevre İl Müdür Vekili Mehmet Güler de, kesin olmamakla birlikte [balık ölümlerinin metan gazından kaynaklanabileceğini söyledi.] Ölümlerin yaşandığı yerin göl olması nedeniyle **bu kaniya** vardıklarını, ancak yine de kesin tahlil sonuçları sonrası net açıklama yapacaklarını belirten Güler, "Bu hafta içinde tahlillerden sonuç alınmasını bekliyoruz" dedi.

(TDB, 10570000)

Kırıkkale Provincial Deputy Director of Environment Mehmet Güler also said, despite not being certain, that [death of fish might have originated from marsh gas.] Stating that they arrived at **this conviction** because of the reason that the place the deaths took place is a lake, but still they will make a statement upon precise results of analysis, Güler said "We expect to obtain the results of the analysis this week".

- (11) Atatürk, Türkiye Cumhuriyeti'nde İslam dinine inanan bireyler, dünya işleri dışında olup bitenleri de anlasın istemişti. İlk adım olarak da [ezan

Türkçeleştirilmişti]. Türkçe ezanı Süleymaniye Camisi başıamamı, tenor sesli Hafız Kemal 'den dinlemiş olan Atatürk, coşkuyla vermişti **bu kararı**.

(TDB, 10660000)

Atatürk desired that individuals in Turkish Republic believing in Islam understood what was going on outside other than worldly deeds. As the first step, [the call to the payer was translated into Turkish.] Having listened to the Turkish call to the prayer from the chief Imam of Süleymaniye Mosque Hafız Kemal with a tenor voice, Atatürk made **this decision** with enthusiasm.

In example (10), the label anaphor 'bu kararı' (Eng. 'this conviction') refers to the conviction related to the reason for the death of fish. In example (11), the head noun 'karar' (Eng. 'decision') refers to Atatürk's having the call for the prayer translated into Turkish. Both of these head nouns refer to the cognitive processes of mind.

3.2.1.4. Text Nouns: As stated by Francis (1994, p. 91), these head nouns refer to the formal textual structure of discourse without any interpretation being involved extending the previous discourse and signaling the boundaries of new discourse. Consider the following example:

(12) [" [...] Türk Tarihî ile Türk Milletinin bütünlüğü, şahsi rekabetlerin ve her şeyin üzerinde tutulmalıdır. Biz Atatürk'ün emrettiği Türk Tarihî araştırmalarına hayat vermiş ve yemin etmiş kimseleriz. Babamıza bile taviz veremeyiz. Tabii olarak gerçeği ve kaynakları en iyi bilerek."]

Dikkat ve ibretle okunması gereken **bu satırlar**, bize bazı Tarihçilerimizin düşünce yapılarını ve düzeylerini göstermesi bakımından çok ilgi çekicidir.

(TDB, 24220)

["[...] The unity of Turkish History and Turkish Nation should be considered superior to personal competitions and any other thing. We are the ones vitalizing and swearing in the research of Turkish History that Atatürk ordered. We cannot appease even our fathers. By naturally knowing the truth and the sources."]

These lines, which should be read carefully and as a lesson, are very appealing with respect to showing us the way and the level of thinking of some of our historians.

In example (12), the label anaphor 'bu satırlar' (Eng. 'these lines') refers back to the preceding words presented in the quotation marks. The label anaphor here functions as a signal that a new discourse is coming, and it seems to have nothing to do with the meaning of the antecedent, rather it refers back to the textual formal structure of the discourse. Some other examples from TDB are as follows:

- (13) Gazetecilere de şu demeci verdi: ["Bursa'ya geldim, olay hakkında ilgililerden bilgi aldım. Gerçekte olayın fazla bir önemi yoktur. Cahil gericiler herhalde Cumhuriyet adliyesinin pençesinden kurtulamayacaklardır. Olaya dikkatimizi özellikle çevirmemizin nedeni, dini siyasete ve herhangi bir kışkırtmaya vesile yapacaklara asla hoşgörü göstermeyeceğimizin bir daha anlaşılmasıdır. Sorunun özü din değil, dil'dir. Kesinlikle bilinmelidir ki Türk ulusunun ulusal dili ve ulusal benliği tüm yaşantısında temel kalacak ve hizmetinde olacaktır."] Resmi kayıtlara da **bu demeç** geçti.

(TDB, 10530000)

He gave the following statement to the journalists: ["I came to Bursa, received information about the event from the concerning parties. The event is not very important in reality. The illiterate conservatives will certainly not be able to escape from the clutch of the Republic's judiciary. The reason why we pay special attention to the event is once again to make it clear that we will not show tolerance to those who will use religion as a means of provocation against politics or anything else. The essence of the problem is not the religion, but the language. It must be understood that the national language and national identity of Turkish nation will remain as the basis throughout its life and will serve it.] **This statement** was noted down in the official records, too.

- (14) [Niçin vakıfların tümünü kapsayan yeni bir düzenleme yerine, böyle bir "konu sınırlaması" tercih edilmiş?] **Bu soruya** yanıt olarak, bu düzenlemenin

konusunu oluşturan vakıf türü açısından ivedi bir çözüm gereksinmesinden söz edilebilecektir.

(TDB, 10550000)

[Why is such a "scope restriction" preferred to a new regulation involving all of the charities?] As an answer to **this question**, in terms of the type of charities that constitute the scope of this regulation, it will be possible to mention the necessity of an urgent solution.

In example (13), the head noun 'demeç' (Eng. 'statement') is a text noun. It does not deal with the content of the statement that it refers to. Likewise, the head noun 'soru' (Eng. 'question') in example (14) does not refer to the meaning of the question which is the antecedent to this label anaphor.

3.2.2. Situation Reference in Turkish

As stated in 1.2.2., the second type of abstract object anaphor is the situation reference, discussed extensively by Fraurud (1992, 1992a). The situation referent 'this' ('bu' in Turkish) is used in Turkish discourse with the same purpose; that is to say, to refer back to eventualities and factualities. Consider the following examples:

- (15) Masallar, efsaneler bedel ödemekten söz ederler. [Geçmişte yaptığınız kötülükler] için, kötülük yaptığınız bir insanın ruhu sizi yer yüzünde, bir sonraki hayatınızda bulur ve siz **bunun** bedelini ödersiniz.

(TDB, 7121)

Tales and legends mention paying a price. The spirit of a person that you harmed finds you on the Earth, in your next life for [the evils you did in the past] and you pay the price for **this**.

- (16) Kardeşimin dediği gibi bu ülkede [mahkûm sayısı çok fazlaymış]. Ekonomik bir yükümü **bu**. Bir kısmının ölmesi gerekiyormuş sessiz sedasız.

(TDB, 7221)

As my brother says, [the number of the prisoners in the country is a lot]. **This** is said to be an economic burden. Allegedly, some have to die quietly.

(17) Bu teorik analizde; modern sosyalist kuramların ulaştığı, ["sosyalizm aşamasının uzun olduğu, bu aşamada sınıf mücadelesinin devam ettiği, sosyalizmin kendinden kaynaklanan bazı eşitsizlikleri barındırdığı, dolayısıyla devrimi sürdürmenin zorunlu olduğu"] gibi sonuçların çekirdeğini görebiliyoruz. **Bu** da Marx'ın "deha"sı olsa gerek.

(TDB, 12112)

In this theoretical analysis; we can see the core of the outcomes that modern socialist theorems have reached, like ["socialism is a long process, during which the fight of classes continue, socialism accommodates some inequalities stemming from itself; thus, it is difficult to maintain the revolution"]. **This** must be the "genius" of Marx.

In the examples above, the anaphoric expression "this" in (15) refers back to an *event* (i.e. the evils done in the past), the one in (16) refers back to a *state* (i.e. the number of the prisoners being a lot in a country), and the one in (17) refers back to a *proposition* (i.e. the theoretical analysis put forward by Marx).

As cited by Lewis (1967), Turkish has three main types of demonstrative pronouns; bu 'this', şu 'this/that', and o 'that' (p.71). According to a recent corpus-based study conducted by Sevdik-Çallı (2012) on TDB, the most frequently used demonstrative in Turkish has been found to be 'bu' as well as being the most preferred demonstrative for abstract object reference, whereas 'şu' and 'o' are rarely used to refer to abstract objects (Sevdik-Çallı, 2012, p.4).

3.2.3. Textual/Discourse Deixis in Turkish

As explained in 1.2.3, Lyons (1977) makes a distinction between *pure* and *impure textual deixis*. The following example illustrates the occurrence of pure textual deixis in Turkish:

(18) Avlunun dışındaki küçük parkın ağaçları altında iki köylü davul zurna çalıyorlar: ["Ankara'dan top atıldı, atlı yayana katıldı, Mustafa Kemal kıyam etti, Yunan aklını şaşırdı..."] Zaferi, Kurtuluş'u hâlâ kanıksamış değiller: ["Biz Yunanı tepeledik..."] Düşündükçe hâlâ övünç ve inanmazlıkla çalıp oynuyorlar: ["Şu Allah'ın işine bak!"]

Tüm bayram günlerinin fon müziğidir **bu**, neşeli bir yürek gibi vuran davulla kıvanç çığlıkları atan zurnanın sesi...

(TDB, 8113)

Two peasants are playing the drum and shrill pipe under the trees of the park outside the garden: ["The gun was fired in Ankara, the cavalry joined the infantry, Mustafa Kemal rebelled, the Greek got shocked..."] They are not aware of the victory, the Salvation: ["We smashed the Greek..."] They play and dance in pride and disbelief as they think of it: ["It's the act of God!"]

This is the background music of all celebration days, the drum beating like a joyous heart together with the sound of the shrill pipe shrieking with pride.

In this context, the abstract anaphoric expression 'bu' (Eng. 'this') refers to the lyrics of the song which is discontinuously presented in the text. In this example, 'bu' refers to the linguistic form of the lyrics, not the referent of it in the real world. In the following example, 'bu' refers not the linguistic form of the antecedent, but to the factuality (being jealous of other guys or not) that it refers to. Thus, it exemplifies impure textual deixis.

(19) ["Parkta onu bekleyen diğer erkekleri kıskanmıyor musun?"]

Bir an durdu.

"Hiç düşünmedim **bunu**," dedi doğallıkla.

(TDB, 2213)

["Aren't you jealous of the other guys waiting for her in the park?"]

He hesitated for a moment.

"I have never thought of **this**," he said in a natural manner.

3.3. Anaphora Resolution in Turkish

As explained in 2.2.1., Consten and Knees (2005) discuss a hierarchy of abstractness. They use the term 'anaphorical complexation process' to refer to the process of increasing abstractness in anaphors and in their antecedents. This theory is highly relevant to the scope of the current study. Thus, in the first part of this section,

some examples of Turkish discourse are analyzed within the framework of this complexation process argued by Consten and Knees (2005) as well as Consten, Knees and Schwarz-Friesel (2007). In the second part of this section, the Right Frontier operation of Webber (1991) is adapted to the processing of Turkish demonstrative 'bu' (Eng. 'this').

3.3.1. Abstractness Hierarchy

Consten and Knees (2005) propose three types of complexation processes that abstract object anaphors might have. The occurrences of these types of complexation can be observed in Turkish discourse, too.

a. The antecedent and the anaphor might be denoting the same ontological type so the ontological status of both parties remains the same:

(20) [Morgun önünde korsan gösteri yapan bir grup polisle çatışmış, bir kişi de orada ölmüştü.]_e [**Bu olay**]_e morgda buz gibi yatan Mesut'un içini sızlatmıştır kuşkusuz.

(TDB, 3221)

[A group raising an illegal hue and crying in front of the morgue fought against the police, and a person died there.]_e [**This event**]_e must have deeply hurt Mesut lying in the morgue ice cold.

In example (20), the antecedent sentence denotes an event. As the name suggests, this corresponds with the label anaphor 'bu olay' (Eng. 'this event'), which also denotes an event. Therefore, the ontological status of both the antecedent and the anaphoric expression is the same.

b. The bare abstract object anaphor 'bu' (Eng. 'this') is neutral (n) in its ontological category. In the cases 'bu' is used, the ontological type of the anaphoric process is denoted by the antecedent.

(21) [Morgun önünde korsan gösteri yapan bir grup polisle çatışmış, bir kişi de orada ölmüştü.]_e [**Bu**]_n morgda buz gibi yatan Mesut'un içini sızlatmıştır kuşkusuz.

[A group raising an illegal hue and crying in front of the morgue fought against the police, and a person died there.]_e [**This**]_n must have deeply hurt Mesut lying in the morgue ice cold.

In example (21), the antecedent denotes an 'event' so the bare anaphor 'bu' is also understood as an event. This is because of the ontological status of the antecedent.

c. Another ontological type might be denoted by the anaphoric expression because of the lexical meaning of the anaphoric expression. Thus, the anaphorical process changes the ontological type of the referent. Consider example (22), which exemplifies an events turning into a state:

(22) [Morgun önünde korsan gösteri yapan bir grup polisle çatışmış, bir kişi de orada ölmüştü.]_e [**Bu durum**]_s 21. yüzyılda hala demokrasiyle alakalı ciddi sorunlarımız olduğunun göstergesidir.

[A group raising an illegal hue and crying in front of the morgue fought against the police, and a person died there.]_e [**This situation**]_s is an indicator of still having serious problems concerning democracy in the 21st century.

In example (22) the label anaphor 'bu durum' (Eng. 'this situation') refers back to the event in the first sentence. However, the antecedent referring back to an event denoted by the label anaphor 'bu durum' becomes subject to an abstraction process and turns into a situation. This change is initiated by the lexical meaning of the head noun 'durum' used in the label.

In addition to the types of anaphoric complexation process, Consten and Knees (2005) emphasize the possible disambiguating role of the label anaphors, which was explained in 2.2.1. The same role of label anaphors can also be observed in Turkish. Consider the following example, with the labels in (a), (b) and (c) for illustration purposes:

(23) [İstanbul'da cenaze töreni için izin verilmemişti.]_s [Morgun önünde korsan gösteri yapan bir grup polisle çatışmış, bir kişi de orada ölmüştü.]_e

(TDB, 3221)

- a. [**Bu olay**]_e morgda buz gibi yatan Mesut'un içini sızlatmıştır kuşkusuz
- b. [**Bu kısıtlama**]_s morgda buz gibi yatan Mesut'un içini sızlatmıştır kuşkusuz
- c. [**Bu**]_n morgda buz gibi yatan Mesut'un içini sızlatmıştır kuşkusuz.

[A funeral in Istanbul was disallowed.]_s [A group raising an illegal hue and crying in front of the morgue fought against the police, and a person died there.]_e

- a. [**This event**]_e must have deeply hurt Mesut lying in the morgue ice cold.
- b. [**This restriction**]_s must have deeply hurt Mesut lying in the morgue ice cold.
- c. [**This**]_n must have deeply hurt Mesut lying in the morgue ice cold.

In example (23), the label anaphor 'bu olay' (Eng. 'this event') in (23a) refers back to the fight between the group and the police; therefore, it denotes an event. The label anaphor 'bu kısıtlama' (Eng. 'this restriction') in (23b) refers back to the funeral being disallowed in Istanbul; thus, denotes a state. The bare abstract object anaphor in (23c), on the other hand, seems to be quite ambiguous. It might be thought to be referring back to the state in which the funeral was disallowed, or the event in which the group and the police fought against each other, or it might even be thought to be referring back to both of the state and the event depending on the personal interpretations of the text on the text-world level. Therefore, it can be concluded that label anaphors eliminates the competing candidates as antecedents thanks to the lexical meaning of the head noun in the label.

3.3.2 Right Frontier Constraint in Turkish

As explained in 2.2.2., Webber (1991) argues that only clauses, or sequence of clauses, corresponding to discourse segments can be the referents of demonstrative pronouns, not a random sequence of them. She specifies the discourse segments as being the ones whose contribution to the discourse model is currently in focus. She claims that it is these very discourse segments which are the antecedents of the demonstrative pronouns. Thus, discourse segments being connected to the

nodes on the right frontier can be the antecedents to the demonstrative pronouns, not the ones on the left. This anaphora resolution model can also be applied to Turkish discourse as in example (24):

- (24) a. Ahmet, bir gün gazeteyi bir açmıştı, Mesut yok!
b. Pusuya düşürülmüş oracıkta can vermişti.
c. İki kişiymişler.
d. Öbürünün adı Halim.
e. Tanımıyordu.
f. Hafif yaralı kurtulmuş o.
g. **Bu**, derinden sarsmıştı Ahmet'i.

(TDB, 3221)

- a. Ahmet, one day opened the newspaper, Mesut disappeared!
b. He was trapped and got killed there.
c. They were two people.
d. The other's name is Halim.
e. He didn't know.
f. He got away slightly injured.
g. **This** afflicted Ahmet deeply.

While processing the discourse presented above, the reader may consider the second clause (b) as the explanation of the first clause (a). Therefore, these two clauses constitute a segment. However, the other clauses (c-f) are the elaborative elements of the event, which can be inferred from the segment constituted by both (a) and (b) i.e., the event in which Mesut was captured and got killed. These elaborative elements can only attach to the already existing discourse segment constituted by (a) and (b) simply because of they give the details of the event presented in (a) and (b). Therefore, they cannot constitute a new discourse segment. As a result, the tree structure for the discourse in (24), looks like the one as follows:

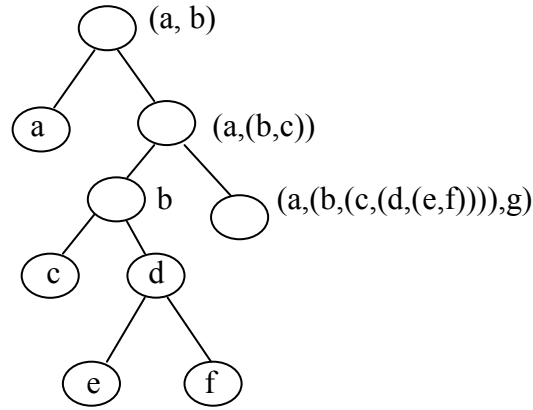


Figure 3 illustrates the right frontier operations to example (24).

The discourse segment constituted by (a) and (b) are on the right frontier of the tree structure. As discussed by Webber (1991), demonstratives with situation reference should attach to the discourse segments on the right frontier. Therefore, the Turkish demonstrative 'bu' (Eng. 'this') in (g) should attach to either (a) or (b) or the combination of them (a+b), not any single one of the elaborative elements or the combination of them in which (a) or (b) does not take place.

3.4. Anaphora Studies in Turkish

In this section, the main aim is to briefly review anaphora studies in Turkish briefly, not to elaborate on them. This is because anaphora resolution studies conducted so far in Turkish are not directly related to the scope of the current study. However, for the sake of completeness, the general framework will be explained.

Most of the studies in Turkish conducted on anaphoric relations and anaphora resolution have a computational perspective. BABY-SIT, by Tın, Erkan and Akman (1998), is one of the earliest computational studies, in which the resolution of zero anaphors is investigated. Yüksel and Bozşahin (2002), on the other hand, describe a system contextually appropriate for anaphor and pronoun resolution in Turkish by using Binding and Centering Theories. They show that the generation of contextually appropriate pronouns and anaphors can be modeled. Besides these, Yıldırım, Kılıçaslan and Aykaç (2004), present a computational model for anaphora resolution in Turkish that is based on Centering Theory. Considering that reflexives and reciprocals are too complicated to capture in their binding relations, Yıldırım, Kılıçaslan and Aykaç (2004) leave them aside and deal solely with pronominal

anaphors. This study yields an effective implementation of the centering approach for Turkish but also identifies some defective sides of this theory.

Another study, conducted by Tüfekçi and Kılıçaslan (2005), provides a computational model for pronominal anaphora resolution in Turkish based on the attempt to implement a system that is in the framework of Hobbs' Naïve Algorithm for pronominal anaphora. This algorithm processes only the surface syntax of sentences in a given text. The system does not account for the processing of higher-level information. Furthermore, Yıldırım, Kılıçaslan and Yıldız (2007) present a corpus-based machine learning model to resolve anaphoric relations in Turkish in an effort to identify a unique antecedent for each pronoun. The model resolves pronouns in six stages: annotation, generating examples, filtering generating tests and training sets, building a classifier, and evaluation.

Tüfekçi, Küçük, Turhan-Yöndem and Kılıçaslan (2007), in another study, present a comparison of two pronoun resolution systems developed for identifying the antecedents of third-person personal pronouns; Küçük and Turhan-Yöndem (2007) provide a knowledge-poor pronoun resolution system for Turkish which resolves third person personal and reflexive pronouns referring to proper person names; Yıldırım, Kılıçaslan and Yıldız (2009) work on a machine learning model, involving a decision tree component and a rule-based algorithms learning from an annotated corpus, to estimate the antecedents of anaphorically used pronouns in Turkish discourse; and Kılıçaslan, Güner and Yıldırım (2009) attempt to investigate several machine learning models developed for pronoun resolution in Turkish by offering an evaluation of the classification performances of the learning models.

All of the studies mentioned so far have a focus on personal pronouns or other pronominal anaphora referring to concrete objects. There seems to be no attempt to deal with the semantic or pragmatic phenomenon of abstract anaphora resolution. One of the few studies of that type can be considered to be the one having been conducted by Turan (1997), who analyzed a total of 56 demonstratives, 'bu' (Eng. 'this') and 'şu' (Eng. 'this/that'), on the Bilkent University e-database consisting of newspaper articles and novel texts. Another study is by Çokal-Karadaş (2010), which is an effort of investigating the pronominal uses of 'bu' and 'this' and 'şu' and 'that' in a contrastive manner, from the written academic discourse based on

Rhetorical Structure Theory (Mann and Thompson, 1988). As an outcome of the study, Çokal-Karadaş (2010) indicates the presence of rhetorical relations in which 'bu-şu' and 'this-that' take place. While 'bu' and 'this' function similarly with their role as the provider of interpretation, explanation and reason relations in both languages, they have distinct functions in textual organization and hypothetical relations. 'Şu', on the other hand, functions in a different way compared to 'that' and 'this' with respect to rhetorical relations. While 'şu' is used in the subtypes of elaboration relation (i.e. elaboration-set-member, elaboration part-whole), 'that' does not seem to have such relations but used in antithesis, list and contrast relations, where 'şu' is not used.

Sevdik-Çallı (2012) investigates Turkish demonstrative anaphora comprising bare demonstrative uses and demonstrative NP uses on a 20K subpart of the TDB. In 10 different texts, the antecedents of the target anaphors with both concrete and abstract reference have been analyzed. It has been found that the *overall* occurrences of reference to concrete objects by Turkish demonstratives 'bu' (Eng. 'this'), 'şu' (Eng. 'this/that'), and 'o' (Eng. 'that') outweigh the ones to abstract objects. While the total occurrences of these three demonstratives mostly used to denote concrete objects, the total occurrences of 'bu' (Eng. 'this') on its own has a different pattern. The results indicate that occurrences of the reference to abstract and concrete objects by the demonstrative 'bu' (Eng. 'this') is almost the same in number. This shows that it is mostly the demonstrative 'bu' (Eng. 'this'), which is used as situation referent in Turkish discourse.

All in all, to the best of our knowledge, there has been no empirical attempt to study the resolution of abstract object anaphora in Turkish, which makes this study a significant one to fill this gap in the literature regarding anaphora studies in Turkish.

CHAPTER 4

THE EXPERIMENTS

The main aim of the study is to understand the resolution of 'bu' (Eng. 'this') and the label anaphors 'bu durum', 'bu olay', 'bu iş' and 'bu gerçek' (Eng. 'this situation', 'this event', 'this activity', 'this fact') in Turkish written discourse. The experimental material consists of anaphors having abstract objects as referents. The test material used in the experiments was selected from METU Turkish Discourse Bank (TDB), and the novel *Yüksek Topuklar* (High Heels) written by Murathan Mungan. To gain an insight into the distinction between the resolution of these two types of abstract object anaphors, first, a pilot study was carried out. Then, two offline experiments were conducted. The first experiment was done to address the main research questions presented in 1.3, and the second experiment was conducted to verify the generalizability of the findings obtained from experiment 1.

In this chapter, the procedures and results regarding the pilot study (Section 3.1), experiment 1 (Section 3.2) and experiment 2 (Section 3.3) will be presented. The results of these experiments will be presented and discussed.

4.1. The Pilot Study

The main aim of the study was to investigate whether there is any difference in determining the antecedents of the bare anaphor 'bu' and four label anaphors in Turkish discourse. More specifically speaking, the focus was the following questions:

- Is there any significant difference between the resolution of the bare abstract object anaphor 'bu' and the label anaphora 'bu iş/durum/gerçek/olay'?
- Considering that anaphors have the tendency to cause ambiguity in discourse, is there a disambiguating role of the label anaphor in the resolution process?

The pilot experimental study consisted of four texts selected from the TDB and the novel *Yüksek Topuklar* (High Heels) written by Murathan Mungan. The lengths of each text were equalized to 60-64 words in total. Each text had one anaphor referring back to an abstract object to be determined by the participants. The texts were original in that they occurred with the bare anaphor 'bu'. The participants were presented with the text with the bare anaphor as well as the manipulated texts containing the label anaphors. The antecedents to these anaphors were all continuous strings to avoid unreliable inter-coder analysis. The participants were given all four texts in which the target abstract object anaphor was highlighted, and they were asked to find the antecedent to the target anaphor by underlining the span of the antecedent.

The pilot study was conducted with 10 participants between the ages 27 and 32. All of the participants were native speakers of Turkish and university graduates. They can all be considered linguistically naive with no previous experience of marking referential chains in a text. There were two groups in the study. The first group consisted of five participants. They received all four texts with the bare abstract object anaphor 'bu' (see Appendix 1). The second group also had five participants and, received the same contexts in which the bare abstract object anaphors were replaced with label anaphors (see Appendix 2).

Prior to the experiment, all of the participants were provided with a very brief training session about the conventions on how to find the antecedents of the target anaphors in the texts (see Appendix 3 and Appendix 4). During the training, participants were introduced with the following concepts:

- A coherent discourse
- An anaphor
- An antecedent
- The length of the referent (The participants were informed that it could be a phrase, a sentence or more than one sentence)
- The distance between the referent and the anaphor (i.e., The participants were informed the referent and the anaphor may not be adjacent to each other and that there may be some other clauses or sentences that intervene)
- The minimality principle, which is the principle applied for the Penn Discourse TreeBank. In the framework of this principle, it is important to mark the minimally required and sufficient information for the interpretation of the relation between the arguments (Prasad, Miltsakaki, Dinesh, Lee, Joshi & Webber 2006, p.13).

During this brief training session, the participants were given two examples of an anaphor in texts and asked what their possible antecedents could be. Their answers were elicited and they were given feedback on their answers.

A comprehensive statistical analysis was not done on the results of this pilot study. Due to limited number of participants, a qualitative evaluation was done, rather than any quantitative analysis. The main goal of this qualitative evaluation was to check the tendency for any observable difference between the groups and to construct the main test material of the main experiment.

The preliminary evaluation of the results indicated that there were differences in the way the antecedent of the third paragraph was marked. In Group 1, most of the participants (four out of five) receiving the text with the bare anaphor 'bu' (Eng. 'this') specified the antecedent as 'kadın bedeni çizme tutkusu' (Eng. 'the passion for

drawing woman body'). In this marking, the segment starts with the 9th word and ends with the 12th word. In Group 2, most of the participants (four out of five) receiving the same text with the label anaphor 'bu iş' (Eng. 'this activity'), specified the antecedent of the label anaphor as 'kadın bedeni çizme' (Eng. 'drawing woman body') by leaving out 'tutkusu' (Eng. 'the passion'). The segment starts with the 9th word, but end with the 11th one. In brief, it has been observed that participants had a tendency to associate the label anaphor 'bu iş' with the activity of drawing, whereas they had a tendency to associate the bare anaphor 'bu' with the passion of drawing¹³. Figure 3 below illustrates the differences in end words between Group 1 and Group 2 (see Appendix 5 for the analysis of start and end words of the data obtained from paragraphs 1,2 and 4).

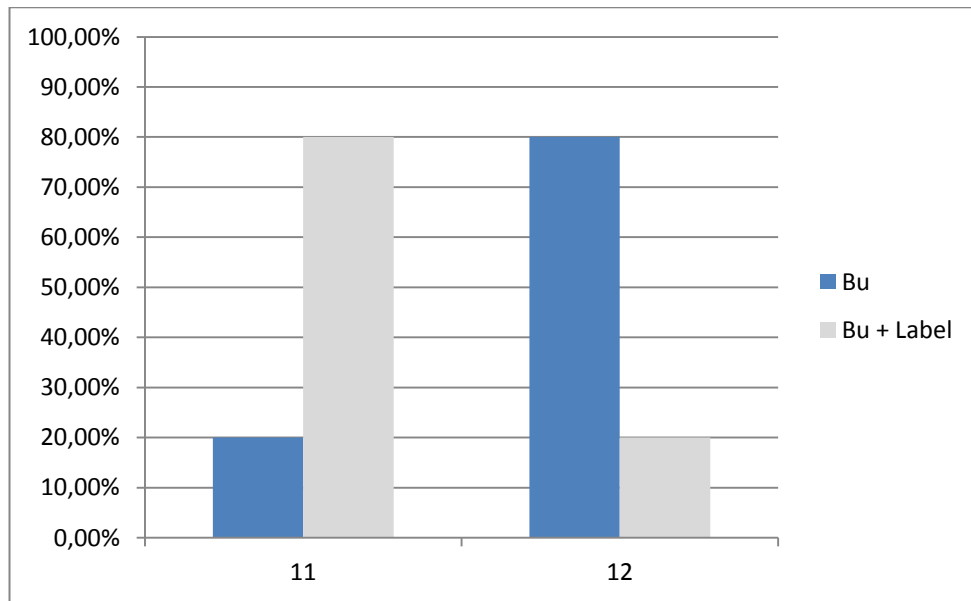


Figure 3. The difference between Group 1 and Group 2's end words in paragraph 3

The underlying reason behind this difference between the marking of the antecedents to the bare abstract object anaphor 'bu' (Eng. 'this') and the label anaphor 'bu iş' (Eng. 'this activity'), might have stemmed from the lexical meaning of the head noun in the label. Speaking intuitively, participants might have looked for an activity

¹³ Considering our criteria for the antecedents to the abstract object anaphors, noun phrases cannot be the antecedents to the target anaphors. Participants were trained where to search for the antecedents. However, as they were told that a phrase could also be an antecedent, the antecedents specified as noun phrases like 'kadın bedeni çizme tutkusu' (Eng. 'the passion for drawing woman body') were not excluded from the data analysis. Moreover, the noun 'passion' already refers to an abstract entity; hence, may qualify as and abstract object, though not in the sense of Asher (1993).

in the text and specified the phrase 'kadın bedeni çizme' as referring to an activity, but not 'kadın bedeni çizme tutkusu'.

To wrap up, the qualitative evaluation of the pilot study indicates that there is a difference between the resolution of label anaphors and the bare anaphors, at least in some conditions. The result of the pilot study suggests that more participants converge on a specific antecedent when the anaphor has a label than when it is bare. This is particularly true when the head noun of the label anaphor (e.g., this activity) confirms that the antecedent is an activity.

Since it was observed that there were some differences between the resolution of bare abstract object anaphor 'bu' and the label anaphors 'bu + head noun', it was decided to use the material prepared for the pilot study for Experiment 1, as explained in more detail in Section 4.2. below.

4.2 Experiment 1

4.2.1 Experiment Design:

Participants: 96 participants took part in the experiment from both genders (31 females and 65 males). 49 of them were in Group 1, and 47 were in Group 2. A demographic questionnaire was given to the participants prior to the experiment. Participants were asked to write their ages, departments, and the cities they grew up to learn which language/languages they were exposed to. All of the participants were native speakers of Turkish. The ages of the participants, all of whom are undergraduate freshmen from various departments at METU, ranged between 18 and 21 (Mean age: 19.2).

Materials: Experiment 1 was an offline paper-and-pencil test. Participants were provided with four texts and asked to underline the antecedent of the highlighted abstract object anaphor. As in the pilot study, participants were provided with the same brief training session prior to the experiment (see Appendix 3 and Appendix 4). They were informed of the basic concepts like discourse, anaphor, antecedent, and Minimality Principle.

Stimuli: The stimuli were exactly the same four texts taken from TDB and Yüksek Topuklar (High Heels written by Murathan Mungan) that were used in the pilot study (see Appendix 1 and Appendix 2).

Procedure: The experiment was conducted by the researcher. It was completed within nearly 20 minutes, including the training session. All of the participants returned the test material with their annotations.

4.2.2 Analysis of the Data

First of all, cataphoric (forward) annotations, not the focus of the study, were excluded from the analysis. This is because such outlier cases may distort the results to a great extent; as a consequence, the results may not be reliable. The exclusion of the cataphoric cases was done upon discussion with an expert linguist who has expertise in Turkish linguistics and cognitive studies.

Besides this, some markings in the data were identified as structurally or semantically non-matching to be the antecedents to the target anaphors. Those markings were also excluded from the analysis upon consulting two expert annotators from the TDB group. Table 2 shows the non-matching markings in paragraphs 1, 3, and 4. The markings for paragraph 2 do not involve any non-matching marking. In Table 2, the reasons for the exclusion of non-matching markings from the data are also provided Table 3 presents the percentage of the excluded data in each group and each paragraph.

Table 2 Non-Matching Markings in Experiment 1

The Paragraph	Non-Matching Marking	Reasons for Non-matching
1. Kadınlar her şeyi çok daha kolay kabullenir, değişimlere daha çabuk rıza gösterirler. Nazlı da, komünist günlerini geride bırakıp hiçbir şeyin artık eskisi gibi olmadığını anladığı anda, yeni hayatlarına daha çabuk uyum göstermiş, kocasını da kendi zamanı içinde bunu/bu gerçeği kabulleneceği noktaya gelmesini sessizce beklemeye başlamıştı. Arada bir hapisanedeki eski dostlara, tutuklu ailelerine, derneklere para yardımı yaparak sol vicdanlarını yatıştırmak kalmıştı onlara.	<ul style="list-style-type: none"> • komünist günlerini geride bırakıp hiçbir şeyin artık eskisi gibi olmadığını anladığı anda • komünist günlerini • Nazlı da, komünist günlerini geride bırakıp hiçbir şeyin artık eskisi gibi olmadığını anladığı anda, yeni hayatlarına daha çabuk 	<ul style="list-style-type: none"> • None of these markings can occur with the verb 'kabullenmek' (Eng. 'to accept') • They also do not match the head noun 'gerçek' meaning wise. They do not explain a 'fact' that the anaphors denote in the text.

<p>3. İlla ki kadın bedeninin gizemi olmalıydı Halil'in gözlerinin önünde. Kadın bedeni çizme tutkusu yaşama amacı haline gelmişti. Çareler düşünüyor, bulamıyordu, çünkü içinde bulunduğu ortamda bunu/bu işi başarabilmesi neredeyse olanaksızdı. Artık bir şey çizemez olmuştu. Altı ay önce bitirdiği bir resmi uzun süre dayanması ve renklerini koruması için verniklediği bir gece ansızın bir tekme savurarak üst kata çıktı. Hiçbir tablosu değerli değildi onun için</p>	<ul style="list-style-type: none"> • çareler düşünüyor • çizemez olmuştu 	<ul style="list-style-type: none"> • This marking cannot occur with the verb 'başarmak' (Eng. 'to accomplish'), i.e., 'çareler düşünmek' (Eng. 'thinking of solutions') is not an entity to accomplish. It is a process. • Cataphoric reference
<p>4. Kemal'i eliyle yana itip yukarı koştı Ahmet Reşat. Kemal kapıda beti benzi atmış Hakkı Efendi'yi görünce paldır küldür inmeye başladı merdivenleri. Hakkı Efendi, Kemal'e titreyerek az önce yaşadıklarını naklediyordu ki, Ahmet Reşat giyinmiş olarak geri geldi. İşgalciler, Ziya Paşaların konağına el koymuşlardı. Hakkı Efendi'yle Ahmet Reşat alelacele çıkıp Paşa'nın konağına gittiler. Bu/bu olay, bir yıkımdı Ziya Paşa için. Büyük bir hakaretti devlet nazırına bu yapılanlar</p>	<ul style="list-style-type: none"> • Ahmet Reşat alelacele çıkıp Paşa'nın konağına gittiler • Kemal'i eliyle yana itip yukarı koştı Ahmet Reşat. Kemal kapıda beti benzi atmış Hakkı Efendi'yi görünce paldır küldür inmeye başladı merdivenleri. Hakkı Efendi, Kemal'e titreyerek az önce yaşadıklarını naklediyordu ki, Ahmet Reşat giyinmiş olarak geri geldi. İşgalciler, Ziya Paşaların konağına el koymuşlardı. Hakkı Efendi'yle Ahmet Reşat alelacele çıkıp Paşa'nın konağına gittiler • İşgalciler, Ziya Paşaların konağına el koymuşlardı. Hakkı Efendi'yle Ahmet Reşat alelacele çıkıp Paşa'nın konağına gittiler 	<ul style="list-style-type: none"> • The meaning of the first marking does not match the verb phrase 'bir yıkımdı' (Eng. 'was a destruction')i.e., "Going to Pasha's house immediately was a destruction" is not the meaning the anaphors denote in this text. • None of these markings properly account for the event that the anaphors denote.

Table 3 Percentage of Excluded Data in Experiment 1

Experiment	Group	Paragraph	Number of Participants	Number of Excluded Data	Percentage of Excluded Data
1	Bare Anaphor	1	49	4	8,16%
1	Bare Anaphor	2	49	0	0,00%
1	Bare Anaphor	3	49	3	6,12%
1	Bare Anaphor	4	49	2	4,08%
<hr/>					
1	Label Anaphor	1	47	3	6,38%
1	Label Anaphor	2	47	0	0,00%
1	Label Anaphor	3	47	0	0,00%
1	Label Anaphor	4	47	4	8,51%

Method: Four different methods were used to investigate the collected data.

Step 1: The first investigation was to analyze the histograms of the start and end words that each participant underlined in each group (see 4.2.3.1.).

Step 2: Upon the preliminary results from the analysis in Step 1, Krippendorff's Alpha, which was explained in 2.3.1, was used as a non-parametric inter-coder analysis method (see 4.2.3.2).

Step 3: To have a better understanding of the data and the results obtained from Krippendorff's Alpha, bar graphs were drawn so that the data could be visually represented (see 4.2.3.3).

Step 4: As discussed in 2.3.1, Krippendorff's Alpha is not the best solution to account for computing the credibility of referential chains because it fails to detect some equivalent classes which are more alike than others (Passonneau, 2004, p.1505). Because of this, in the last step, a new analysis was undertaken using the method explained by Passonneau (2004). Passonneau analyzes the antecedents of

referential chains by classifying them as being subsumed by the antecedent, intersecting with it, and being disjunct from it (see Section 2.3.2 for more detail of her work). In the present thesis, the underlined antecedent spans were analyzed as being subsumed and intersecting with the largest span of the antecedent for the bare or label anaphor, as well as being disjunct from it (see 4.2.3.4).

4.2.3 Results

4.2.3.1. Consistency Analysis on the Start and End Words

The first step in the data analysis was to identify the *start word* and the *end word* of each marking by each participant in both groups. A consistency analysis was conducted. The histograms for each paragraph were drawn for both the words the participants started and ended underlining. A 15%-threshold was created to specify the most frequent start and end words (see Appendix 15, Tables 18 and 19).

Consistency in Start Words¹⁴: For paragraph 1 (see Appendix 6, Figure 7 and Figure 8), the results indicate that the most frequent start words in Group 1 (who marked the bare abstract anaphor 'bu') are the 14th, 18th and 26th whereas the most frequent ones are the 14th and 18th words in Group 2, who marked the antecedent to the label anaphor 'bu gercek' (Eng. 'this fact').

For paragraph 2 (see Appendix 6, Figure 10 and Figure 11), the most frequent word for both groups are the 1st and 35th words. However, the start words vary in Group 1, compared to the ones in Group 2. There is also a higher frequency of starting with the 35th word in Group 2 compared to the one in Group 1. This may imply the disambiguating role of label anaphor 'bu durum' (Eng. 'this situation') in this item. This is because participants more consistently start with the 35th word, which shows a better agreement of participants on the start boundary of the antecedent to the label anaphor 'bu durum'. On the other hand, the distribution is more scattered in Group 1, which shows a lower level of agreement on the start boundary of the antecedent to the bare anaphor 'bu' (Eng. 'this').

¹⁴ A 15%-threshold was created to specify the most frequent start and end words.

To better illustrate this analysis, the histograms in Figure 4 and Figure 5 below will clarify the words the participants started with in paragraph 3:

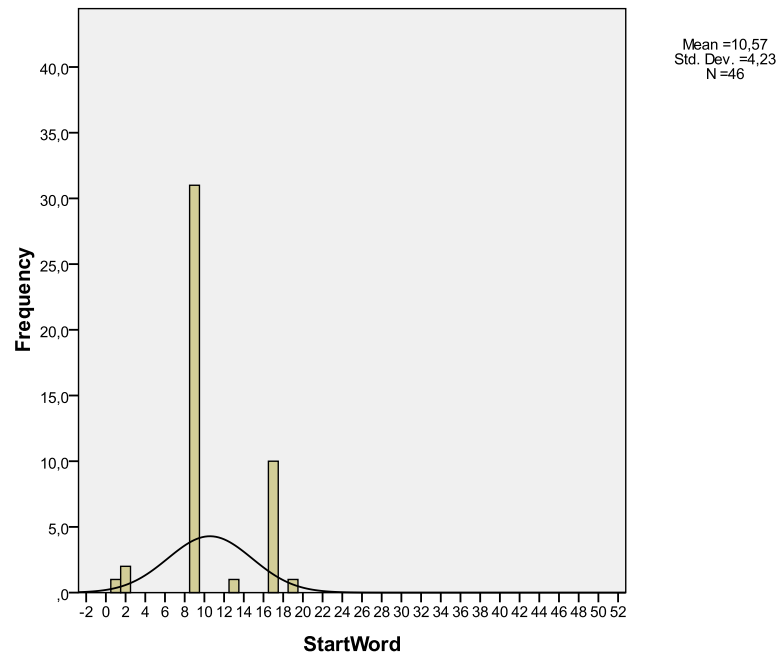


Figure 4¹⁵. The frequency of start words of the participants in Group 1 for paragraph 3 with the most common start words being the 9th and 17th ones

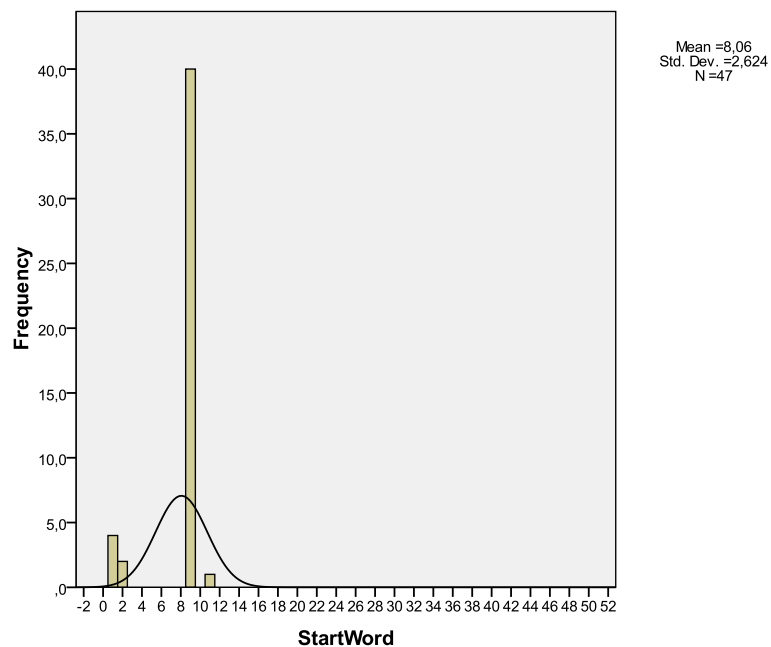


Figure 5. The frequency of start words of the participants in Group 2 for paragraph 3 with the most common start word being the 9th one

¹⁵ X axis shows the words in the paragraphs. Y axis shows the frequency to start with nth word.

As it is clear from Figures 4 and 5, for both groups, the most frequent start word is the 9th in paragraph 3 but with a higher frequency in Group 2. Also, the start words vary in Group 1. In the marking of this paragraph, there is also a high number of participants starting with the 17th word in Group 1, while there is nobody starting with the 17th word in Group 2. This may again indicate the disambiguating role of the label anaphor i.e., 'bu iş' (Eng. 'this activity') in this item. In other words, while the bare anaphor 'bu' allows for multiple interpretations, which is visible in the variety of the start words, i.e., the 9th or the 17th word), the label anaphor 'bu iş' reduces the alternative interpretations with convergence on one of the alternatives. For this case, participants converge on the 9th word.

For the last paragraph, most of the participants from both groups converge on the 36th and 37th words (see Appendix 6, Figure 12 and 13), again with higher consistency on the 36th word in Group 2, who marked the antecedent to the label anaphor 'bu olay' (Eng. 'this event'). The high consistency in the start words for this paragraph in both groups may be due to the lexical meaning of the head noun in the label. In this paragraph, the head noun 'olay' was used in the label. Being the most concrete entities among the factualities and eventualities, events lie at the bottom of the abstractness hierarchy (Consten & Knees, 2005). Therefore, the inter-coder agreement is high for marking of the antecedents to anaphors denoting events. This is discussed in detail in 4.2.4.

Consistency in End Words: For paragraph 1, the histograms are almost alike and the most frequent end word is the 23rd one for both groups (see Appendix 7, Figures 14 and 15). However, there is also some frequency to end with the 31st word in Group 1. For paragraph 2, the most frequent end words are the 42nd and 43rd. In Group 1, differently from Group 2, there is also an observable frequency of the 50th as the end word (see Appendix 7, Figure 16 and 17). Once again, this shows a more scattered distribution in marking the boundaries of the antecedents to the bare anaphor 'bu' in this paragraph.

For paragraph 3, while the most frequent end words are the 11th and the 12th in both groups, there is also a frequency of the 19th word in Group 1. Alternative markings are fewer in Group 2, i.e., ending with the 19th word, is not observed in Group 2. The histograms for paragraph 3 are given in Figures 6 and 7 below.

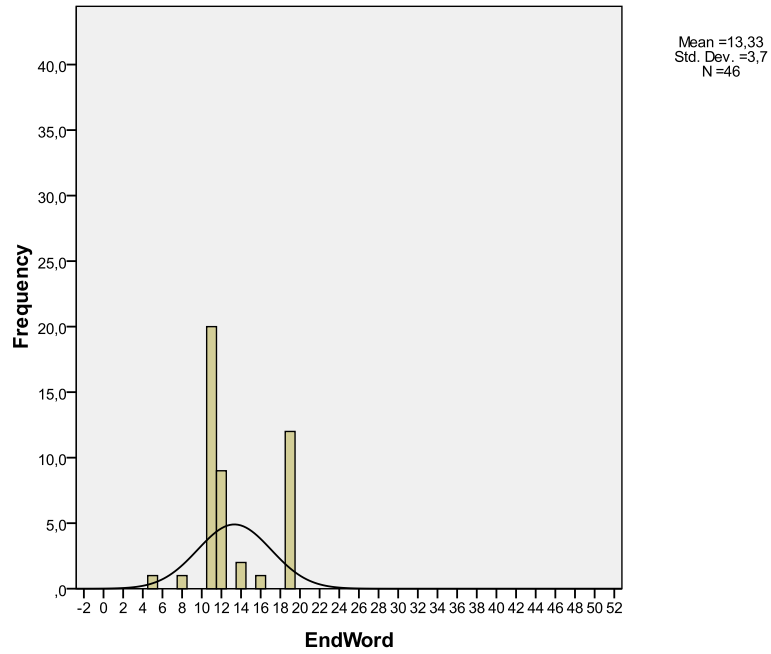


Figure 6. The frequency of end words of the participants in Group 1 for paragraph 3 with the most common end words being the 11th, 12th and 19th words

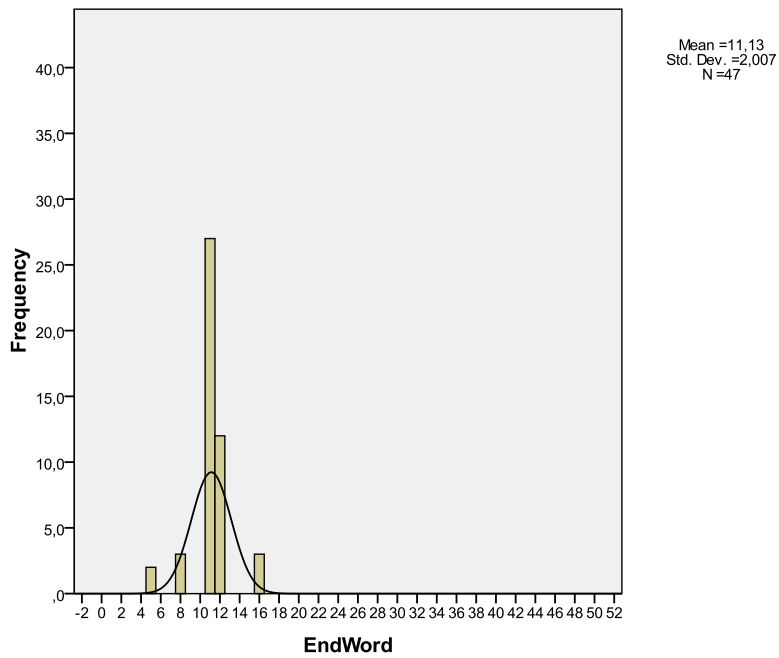


Figure 7. The frequency of end words of the participants in Group 2 for paragraph 3 with the most common end words being the 11th and 12th words

For paragraph 4 (see Appendix 7, Figure 18 and 19), all of the participants from both groups converge on the 41st word to end with. This may be a factor demonstrating the absence of ambiguity in this paragraph.

Interim Summary:

To investigate the data, the start and end words that participants in Groups 1 and 2 underlined as the antecedent to the target anaphor were analyzed. The aim was to gain an insight into whether there was any observable difference in the way the participants specified the boundaries of the antecedents to bare anaphors and label anaphors. The analysis indicates higher consistency in the start words in paragraph 3 and consistency in the end words in paragraph 2 while marking the antecedent to the label anaphors. There is also an observable difference in paragraph 3 both in start and end words, which indicates that label anaphors have a tendency to be more precise in specifying the boundaries of the antecedent. Table 4 below summarizes the start words from both groups and Table 5 summarizes the end words.

Table 4 Frequency of Start Words in Experiment 1¹⁶

Paragraph #	Group 1 (Bare Anaphor)		Group 2 (Label Anaphor)	
	Start Word	Percentage of Frequency	Start Word	Percentage of Frequency
1	14 th	37,78%	14 th	45,45%
	18 th	44,44%	18 th	38,64%
	26 th	15,56%		
2	1 st	22,45%	1 st	17,02%
	35 th	34,69%	35 th	57,45%
3	9 th	67,39%	9 th	85,11%
	17 th	21,74%		
4	36 th	70,21%	36 th	86,05%
	37 th	27,66%	37 th	13,95%

¹⁶ see Appendix 15, Table 18 for the frequency of other markings on start words.

Table 5 Frequency of End Words in Experiment 1¹⁷

Paragraph #	Group 1 (Bare Anaphor)		Group 2 (Label Anaphor)	
	End Word	Percentage of Frequency	End Word	Percentage of Frequency
1	23 th	66,67%	23 rd	68,18%
	31 st	17,78%		
2	42 th	48,98%	42 nd	48,94%
	43 th	24,49%	43 rd	29,79%
	50 th	18,37%		
3	11 th	43,48%	11 th	57,45%
	12 th	19,57%	12 th	25,53%
	19 th	26,09%		
4	41 st	100,00%	41 st	100,00%

4.2.3.2. Inter-coder Agreement Analysis through α Value

The analysis in Step 1 gives some clues on the differences between the resolution of the bare abstract object anaphor 'bu' and the label anaphors: A cautious conclusion is that label anaphors have a disambiguating role as indicated by the higher consistency rates on both the start words and the end words. To be able to reach a reliable conclusion, it was decided to analyze the data with Krippendorff's Alpha (α value) as a non-parametric measurement of inter-annotator agreement (as explained in 2.3.1). This is because analyzing only the end words and the start words fall short to account for the inter-coder agreement rate on completely or partially overlapping strings that are underlined by the participants.

¹⁷ see Appendix 15, Table 19 for the frequency of other markings on end words.

Krippendorff's Alpha measurements were taken for the word boundaries and word intervals¹⁸ of the annotated texts. The scores obtained through α values are as follows:

As shown in Table 6, Krippendorff's Alpha results for word boundaries confirm the results of the previous analysis of consistency in start words and end words. The α values indicate moderate agreement in both groups for paragraph 1 ($0.6 > \alpha > 0.4$), substantial agreement in both groups for paragraph 2 ($0.8 > \alpha > 0.6$), and a perfect agreement in both groups for paragraph 4 ($\alpha > 0.8$). These results reinforce the initial intuition arising as a result of the pilot study as well as the consistency analysis on the start and end words. The results suggest that the bare anaphor and the label anaphor were resolved in similar ways. On the other hand, there is a difference in the inter-coder agreement scores between the groups in paragraph 3. While there is almost only slight agreement in Group 1 ($0.2 > \alpha > 0.0$), there is moderate agreement in Group 2 ($0.6 > \alpha > 0.4$). In other words, there is a quantitative increase in Group 2's agreement measures. This shows that the strings specified as antecedents to the label anaphor by Group 2 overlaps more compared to the ones in Group 1. This is a contributing finding to the initial hypothesis that label anaphor 'bu iş' (Eng. 'this activity') help readers specify the boundaries of the antecedents in a more precise manner.

Table 6 Analysis of Data with Krippendorff's Alpha - Word Boundary Results for Experiment 1

Experiment 1 - Word Boundary			
Paragraph #	Group 1 (Bare Anaphor) α value	Group 2 (Label Anaphor) α value	Label
1	0.555	0.566	gerçek
2	0.671	0.724	durum
3	0.223	0.442	iş
4	0.925	0.962	olay

¹⁸ Taking the starting and ending of the annotated text as context unit is called the boundaries, and taking all the coding units that span between the starting and ending of the annotations are called the interval (Yalçınkaya, 2010, p. 65).

The results for word interval data are shown in Table 7. As a difference from the boundaries, the word interval measurements consider the partial disagreements in the coding units. The results are mostly compatible with the ones obtained through word boundary data, i.e., the α values for intervals indicate almost the same results for paragraphs 1, 2 and 4. However, the agreement rates decreases even more for paragraph 3. The decrease especially in Group 2 is an unexpected result. As the α value for intervals measure partial disagreements, the small number of disagreements in paragraph 3 Group 2 data might have brought this unexpected result. Also, we conjecture that the poor disagreements ($0.0 > \alpha$) in this paragraph may be stemming from the ambiguity that lexical meaning of the head noun 'iş' (Eng. 'activity') causes. This is further discussed in 4.2.4.

Table 7 Analysis of Data with Krippendorff's Alpha - Word Interval Results

Experiment 1 - Word Interval			
Paragraph #	Bare Anaphor (Group 1)	Label Anaphor (Group 2)	Label
	α value	α value	
1	0.555	0.500	gerçek
2	0.579	0.590	durum
3	-0.125	-0.410	iş
4	0.989	0.996	olay

Krippendorff's Alpha Measurements without Adjuncts:

Prior to the experiment, participants were provided only with little training on how to annotate the data. They were not asked to leave the adjuncts out. As Kroeger (2005) defines, adjuncts are closely related to the meaning of the predicate and helps understanding the flow of the story in a text by giving information about time and place of an event as well as the way an action is done or an event takes place (p. 58). However, as Kroeger (2005) informs us, adjuncts are always possible to be omitted because they are not directly linked to the predicate; thus, they are never obligatory.

On a stricter annotation guideline, the participants could have been instructed to exclude the adjuncts in their markings of the antecedents. However, with the concern to cause bias in the decisions of the participants, using a strict annotation guideline was not preferred (as also not preferred by Müller (2007) because of the same reason). As a result, the adjuncts were left to participants' discretion.

As an expected result of this preference, it was observed that not all of the participants obeyed the minimality principle, and they also underlined the adjuncts. It was thought that this might have blurred the results on agreement rates in paragraph 2 and 3. The distinction between the groups, especially in paragraph 3, might have been caused by the involvement of adjuncts in the markings. Considering what could have happened if participants had been instructed with a stricter guideline, the following two manner adjuncts were taken out of the data, and Krippendorff's Alpha was computed again:

Paragraph 2: *Şimdi Kanada'larda sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, [boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu] öğrenmek doğrusu sevindirmişti beni. Elbet bir gün Kanada'dan dönecekti! **Bunu/bu durumu** açıklarken onun o kırmızıya dönen yüzünü seyretmek hoş olacaktı doğrusu.*

Learning that *the buck-toothed ugly know-it-all girl wandering idly now in Canada in a social scientist manner, acting as if love and men were not important, staring at the emotional jolts others go through with objective and blank eyes even by looking down on them as if she was analyzing an insect* was [secretly reading photo novels in her free time] delighted me indeed. She was certainly going to come back from Canada! It was going to be nice to watch her face turning into red while she was explaining **this/this situation**.

Paragraph 3: *İlla ki kadın bedeninin gizemi olmalıydı Halil'in gözlerinin önünde. Kadın bedeni çizme tutkusu yaşama amacı haline gelmişti. Çareler düşünüyor, bulamıyordu, çünkü içinde bulunduğu ortamda **bunu/bu işi** başarabilmesi neredeyse olanaksızdı. Artık bir şey çizemez olmuştu. Altı ay önce bitirdiği bir resmi uzun süre dayanması ve renklerini koruması için*

vernikleđiđi bir gece ansızın bir tekme savurarak üst kata çıktı. Hiçbir tablosu değerli değildi onun için.

The mystery of woman body must have lied before Halil's eyes *necessarily*. The passion for drawing woman body had turned into his purpose in life. He was looking for solutions, not able to find one because it was almost impossible to achieve **this/this activity** in the environment he lived in. He was no longer able to draw anything. At a night when he varnished a painting that he finished six months ago so that it could be durable and the colors could be protected, he kicked it suddenly and went upstairs. None of his paintings were valuable for him.

The only adjuncts from both experiments were identified as these two. The first adjunct written in italics is from Experiment 1, Paragraph 2, and the second one is from Experiment 1, Paragraph 3 indicated in italics. As these are the only adjuncts identified in the markings of the participants, Krippendorff's Alpha was computed only for paragraph 2 and 3.

As shown in Table 8, the results for word boundary data do not change upon taking the adjuncts out. In the previous case, the α value indicated substantial agreement in both groups for paragraph 2. As stated in Table 7, the α value again indicates substantial agreement in both groups after taking the adjunct in paragraph 2. In paragraph 3, the previous analysis with adjuncts indicated slight agreement in Group 1, and moderate agreement in Group 2. This remains the same after taking the adjunct the adjunct in paragraph 3, too.

Table 8 Analysis of Data without Adjuncts in Experiment 1 - Word Boundary Results

Experiment 1 - Word Boundary			
Paragraph #	Bare Anaphor (Group 1) α value	Label Anaphor (Group 2) α value	Label
2	0.688	0.734	durum
3	0.224	0.446	iş

As shown in Table 9, the results for word interval data differ for paragraph 2. In the first analysis in which the adjuncts were included, moderate agreement ($0.6 > \alpha > 0.4$) was found in both groups. However, the results indicate perfect agreement ($1 > \alpha > 0.8$) after taking the adjunct (1) out. For paragraph 3, the results for word interval data remains almost the same indicating again poor disagreement ($0.0 > \alpha$).

Table 9 Analysis of Data without Adjuncts in Experiment 1 - Word Interval Results

Experiment 1 - Word Interval			
Paragraph #	Bare Anaphor	Label Anaphor	Label
	(Group 1) α value	(Group 2) α value	
2	0.811	0.917	durum
3	-0.120	-0.276	iş

As indicated by the results in the Tables 6 and 7, while there seems to be a slight increase in inter-annotator agreements in word boundary results for both of the paragraph 2 and 3, there is an observable increase in both of the paragraphs in word interval results. As is clear, the omission of the adjuncts helps obtaining better agreement results.

Interim Summary:

In the second step of the analysis, Krippendorff's Alpha was computed to analyze inter-coder agreement through α value. The findings indicate a difference between the agreement rates of the two groups in paragraph 3. This result is compatible with the findings in the start/end word consistency analysis, as well as the qualitative evaluation in the pilot study.

4.2.3.3. Visual Illustration of the Whole Data

To gain better insight into the data generating the agreement and disagreement rates indicated by α values, it was thought that it could have been more meaningful to have a visual representation of the markings. This is because although the scores obtained through α values give an idea regarding the inter-coder

agreement in the paragraphs, they do not fully account for where participants completely or partially agree with their markings. Moreover, it is not possible to see the alternative interpretations of the antecedent stemming from ambiguities through the agreement rates obtained through α values. To this end, for each paragraph in both groups, the following figures have been prepared. The bars stand for the boundaries of the annotation of each participant. These bars have been ordered as beginning with the longest marking and ending with the shortest one. In other words, the first bar in the figures stands for the longest marking, and the last one stands for the shortest marking.

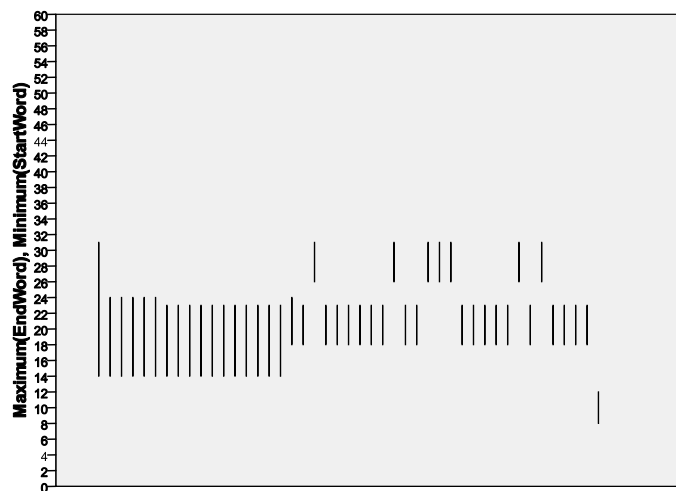


Figure 8. The markings of the antecedents by each participant for Paragraph 1 from Group 1¹⁹

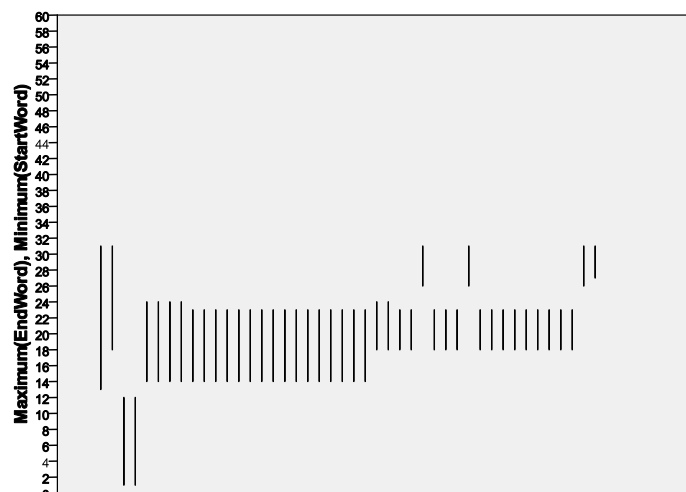


Figure 9. The markings of the antecedents by each participant data for Paragraph 1 from Group 2

¹⁹ Y axis shows the n^{th} word each participant started marking the antecedent and the n^{th} word s/he ended marking

The Figure 8 and Figure 9 visually represent the markings for the antecedents in Paragraph 1. The bars indicate that the dominant marking of the antecedent in both groups is between the 14th and 24th words. The alternative marking in both groups is between the 26th and 32nd words. There seems to be no difference on the data obtained from both groups in this paragraph.

Figures 10 and 11 illustrate the markings for paragraph 2 in which the bare anaphor 'bu' (Eng. 'this') and label anaphor 'bu durum' (Eng. 'this situation') are compared:

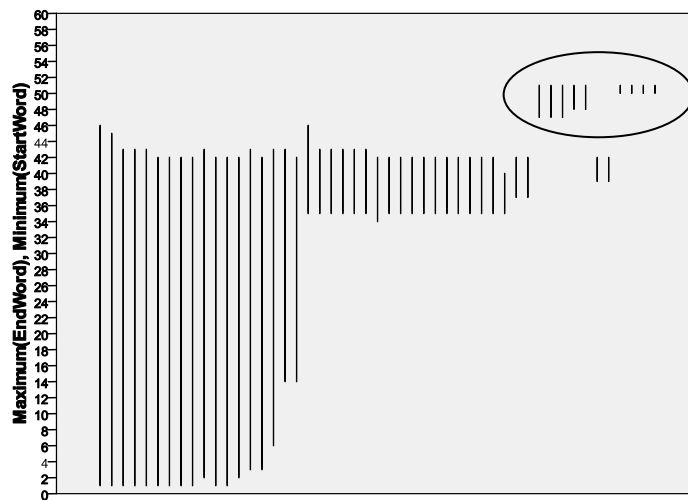


Figure 10. The markings of the antecedents by each participant for Paragraph 2 from Group 1

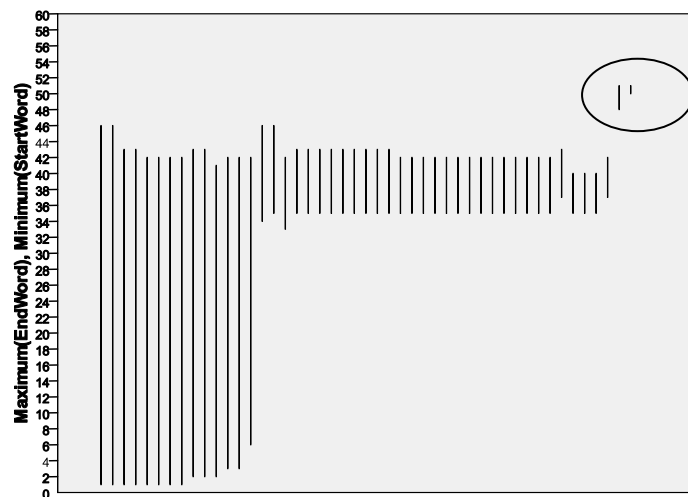


Figure 11. The markings of the antecedents by each participant for Paragraph 2 from Group 2

According to the scores obtained from α values, substantial agreement was found in both groups for paragraph 2. The α value considers non-overlapping markings (circled above) as a disagreement with the most common marking. Therefore, the agreement rates naturally go down. However, those non-overlapping markings can actually be possible alternative interpretations of the target abstract anaphors in the texts. As it was explained in 2.3., abstract anaphors are ambiguous in nature, and they allow alternative interpretations. The α value fails to capture this semantic possibility. Figures 10 and 11, on the other hand, clearly show the possible alternative interpretations as well as the overlapping ones.

This alternative interpretation in paragraph 2 is between the 48th and 51st words circled in the right upper corner. While approximately 18% of the participants in Group 1 selected this alternative interpretation, only 4% of the participants selected this interpretation in Group 2 (marking the label anaphor 'bu durum' (Eng. 'this situation')), which may be indicating the disambiguating role of label anaphor 'bu durum' for this paragraph. This is because the label anaphor minimizes an alternative interpretation possibly due to the lexical meaning of the head noun 'durum' (Eng. 'situation'). This is further discussed in 4.2.4.

Figures 12 and 13 illustrate the markings for paragraph 3 in which the bare anaphor 'bu' (Eng. 'this') and label anaphor 'bu iş' (Eng. 'this activity') are compared:

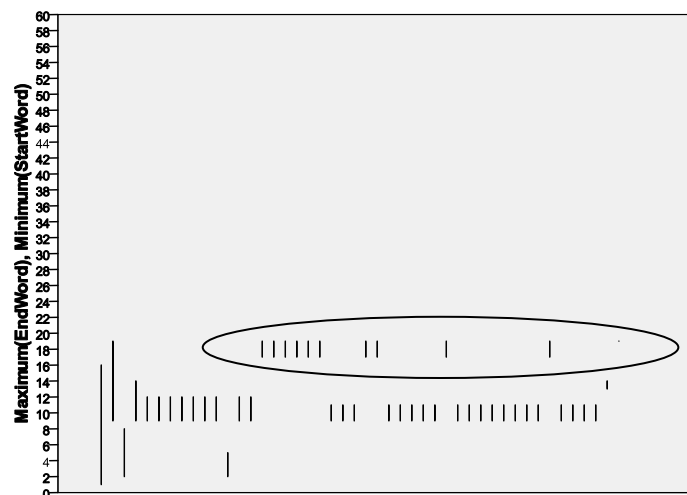


Figure 12. The markings of the antecedents by each participant for Paragraph 3 from Group 1

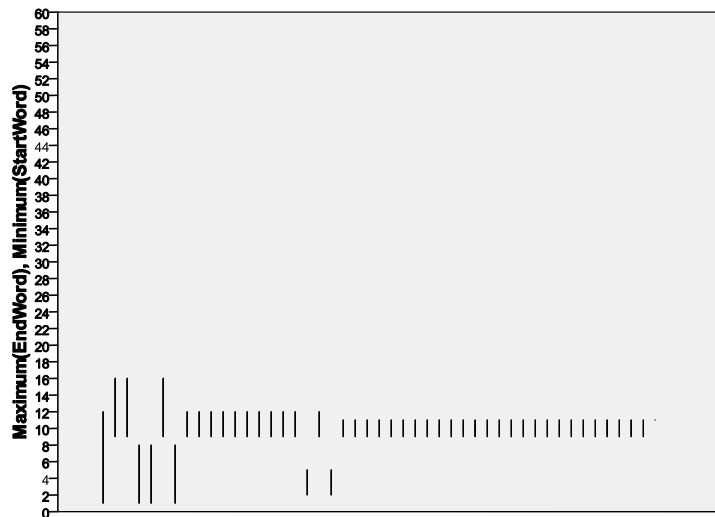


Figure 13. The markings of the antecedents by each participant for Paragraph 3 from Group 2

There is a notable distinction between the markings of Group 1 and Group 2 worth mentioning: It is the marking of 17th and 19th words, which is underlined by almost 22% of the participants in Group 1 while it does not occur in Group 2 at all. Depending on this difference between the marking of the two groups, it becomes possible to draw the conclusion that the text is ambiguous when it is presented with the bare anaphor 'bu' (Eng. 'this'). However, when it is presented with the label anaphor 'bu iş' (Eng. 'this activity'), the label anaphor seems to be eliminating one of the competing candidates in the interpretation of the antecedent.

Finally, for paragraph 4 in which the bare anaphor 'bu' and label anaphor 'bu olay' (Eng. 'this event') are compared, Figures 14 and 15 illustrate the annotations:

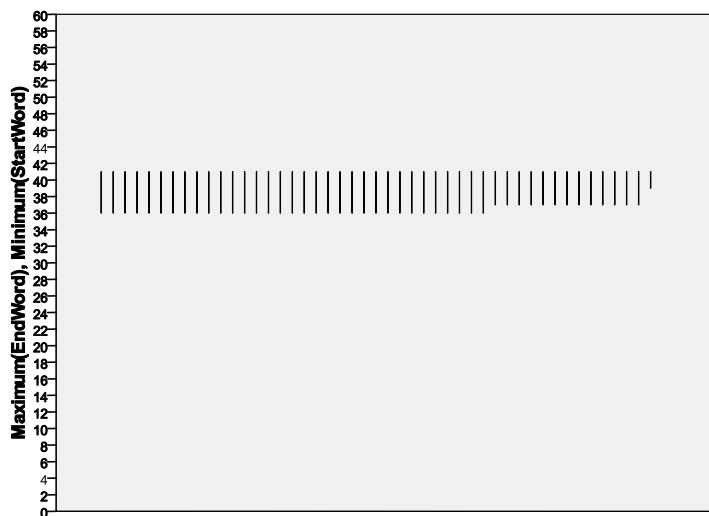


Figure 14. The markings of the antecedents by each participant for Paragraph 4 from Group 1

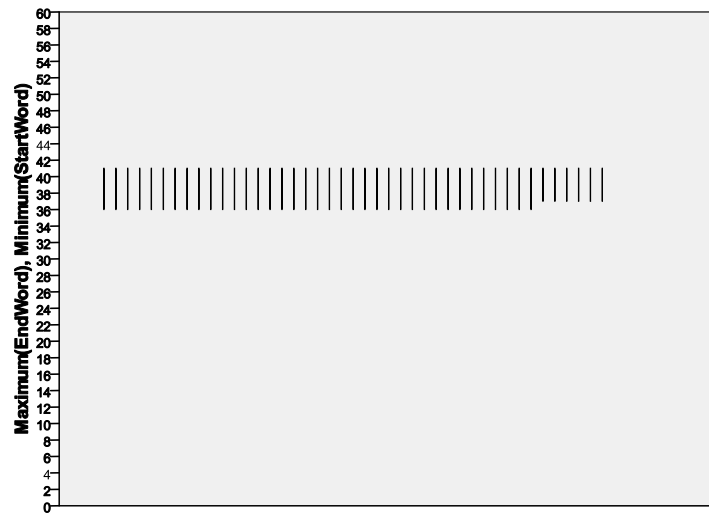


Figure 15. The markings of the antecedents by each participant for Paragraph 4 from Group 2

Perfect agreement was obtained through α values in paragraph 4. Almost all of the markings in both groups converge on the same marking. This is obvious from the figures, too. These results confirm the absence of ambiguity in this paragraph.

Interim Summary:

A visual representation of the markings through bar graphs have been prepared to gain better insight into the data generating the agreement rates indicated by α values. The main aim was to see 'where' participants completely or partially agreed or disagreed with their markings, and also to see 'the alternative interpretations' of the antecedent caused by ambiguous nature of the abstract object anaphors. The analysis showed that bare abstract anaphor 'bu' (Eng. 'this') have a tendency to allow alternative interpretations. On the other hand, label anaphors have a tendency to eliminate the competing candidates to be antecedents; therefore, minimize the occurrences of the alternative interpretations.

4.2.3.4. Passonneau's Proposal

From the visually presented data, seeing that there were both overlapping and non-overlapping antecedents marked by the participants, to have a better understanding into possible *semantic constraints* the bare anaphor 'bu' (Eng. 'this') and the label anaphors 'bu+iş/durum/gerçek/olay' (Eng. 'this + activity/situation/fact/event) might have gone through in the resolution process, the last method followed in the analysis of data was based on the approach put forward

by Passonneau (2004) as explained in 3.2.2. The data was classified in terms of the following relations (see Appendix 8):

- **Main set:** The longest annotated antecedent, indicated as **A**
- **Subsumption:** The subsets overlapping with the main set, indicated as **A1, A2, A3...**
- **Intersection:** Two different annotations intersecting with each other at some point indicated with the symbol ' \cap '.
- **Disjunct:** Annotations not overlapping with one another but all are possible interpretations to be considered as antecedents to the target anaphor, i.e., **A** and **B** has a disjunct relation between each other.

For the reliability of the analysis, the mismatching²⁰ annotations were excluded upon consulting two expert annotators from the TDB group. Besides this, adjunct in paragraph 2²¹, as stated in 4.2.3.b. was excluded from the analysis because when it is included in the analysis, the number of the subsets to the main set reaches up to a non-manageable level (see Appendix 9).

Passonneau's method was used to gain better insight into the distinct interpretations of target anaphors while preserving the same units of interpretations²². To illustrate this analysis, the classifications for paragraph 3, in which previous analysis proved the presence of an observable difference between the groups, is as follows (for the rest of the analysis, see Appendix 8):

Paragraph 3:

A: kadın bedeninin gizemi olmalıydı Halil'in gözlerinin önünde. Kadın bedeni çizme tutkusu yaşama amacı haline gelmişti.

A1: kadın bedeninin gizemi olmalıydı Halil'in gözlerinin önünde.
Kadın bedeni çizme tutkusu

²⁰ The ones which cannot be the antecedents to the target anaphors.

²¹ (1) Şimdi Kanada'larda sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan

²² However, the referential values²² assigned by Passoneau (2004) to each type of set was found impractical due to the number of participants in the current study; thus, it was not applied.

A2: kadın bedeninin gizemi olmalıydı Halil'in gözlerinin önünde

A3: kadın bedeni çizme tutkusu yaşama amacı haline gelmişti

A4: kadın bedeni çizme tutkusu yaşama amacı

A5: kadın bedeninin gizemi olmalıydı

A6: kadın bedeni çizme tutkusu

A7: kadın bedeni çizme

A8: yaşama amacı

A9: çizme

B: çareler düşünüyor, bulamıyordu

B1: bulamıyordu

A4 \cap **B:** kadın bedeni çizme tutkusu yaşama amacı haline gelmişti. Çareler düşünüyor, bulamıyordu

No Match: çareler düşünüyor

Paragraph 3:

A: the mystery of woman body must have lied before Halil's eyes. The passion for drawing woman body had turned into his purpose in life.

A1: the mystery of woman body must have lied before Halil's eyes. The passion for drawing woman body

A2: the mystery of woman body should must lied before Halil's eyes

A3: the passion for drawing woman body had turned into his purpose in life.

A4: the passion for drawing woman body ... his purpose in life.

A5: the mystery of woman body must have lied

A6: the passion for drawing woman body

A7: drawing woman body

A8: his purpose in life.

A9: drawing

B: He was looking for solutions, not able to find one

B1: not able to find one

A4 ∩ B: The passion for drawing woman body had turned into his purpose in life. He was looking for solutions, not able to find one.

No Match: He was looking for solutions

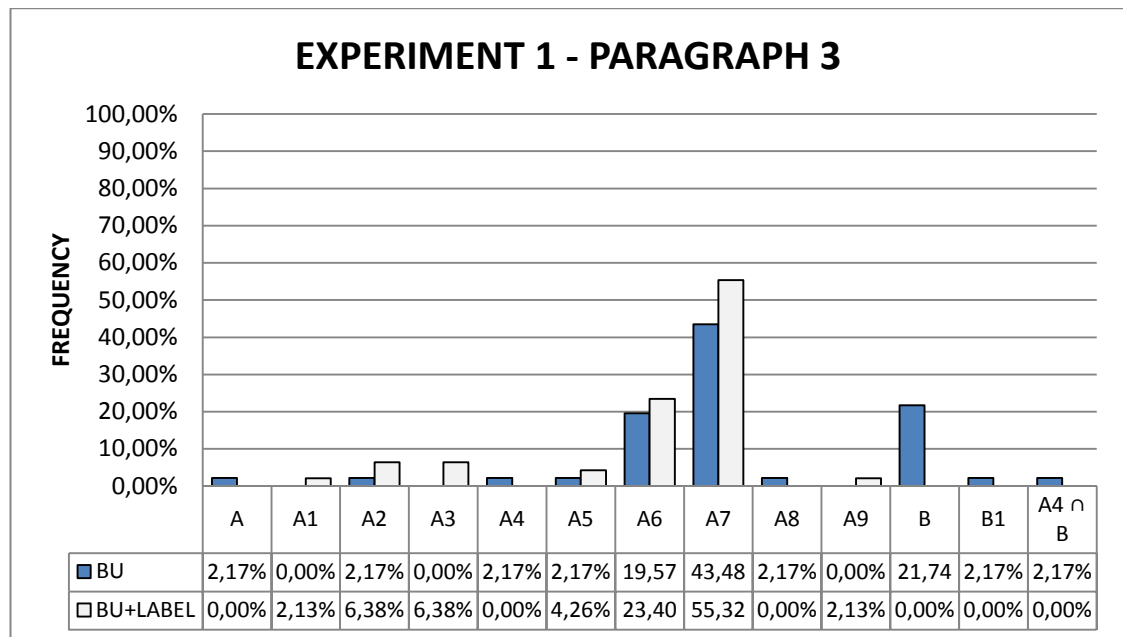


Figure 16²³. Pasonneau's classification of main set, subsumption, intersection and disjunct relations for Paragraph 3 from both groups

Figure 16 shows that the sets B, B1 and A4 ∩ B are observed in Group 1 who received the text with the bare anaphor 'bu', but not in Group 2 who received the same text with the label anaphor 'bu iş'. In the visual illustrations of the data (provided in 4.2.3.3.), Figure 12 and Figure 13 already revealed this difference. However, it was only possible to see that the strings involving the 17th and 19th words were marked by almost 22% of the participants in Group 1 while it did not occur in Group 2 at all. This analysis, on the other hand, helps seeing 'what' those strings actually are, and speculate on the possible reasons on these non-overlapping strings between the groups.

²³ Y axis shows the frequency of the classifications. Below the graph, the percentages of each set's occurrence in Group 1 (bu) and Group 2 (Bu + label) are provided

While the antecedents to the label anaphor 'bu iş' were identified as the subsets of A by the participants, those to the bare anaphor 'bu' involved the alternative interpretations as B, B1 and $A4 \cap B$. The lexical meaning of the head noun 'iş' in the label seems to have limited the decisions of the participants. It is obvious that they associated the head noun 'iş'²⁴ with 'the mystery of the woman's body lying before Halil's eyes', or with 'the passion for drawing woman body', or 'that passion's turning into his purpose in life'. However, they did not associate 'iş' with looking for solutions, i.e., there is no occurrences of B, B1 and $A4 \cap B$ in Group 2. This is possibly because 'bu iş' in this context was perceived by the participants to be referring back to a proposition or an process, i.e., 'the proposition that the mystery of woman body lies before Halil's eye's', 'the proposition of passion for drawing woman body', 'the process of that passion's turning into his purpose in life'. However, the sets B, B1 and $A4 \cap B$ involve an atelic action, 'bulamıyordu' (Eng. 'he was not able to find one'), which is an incompleted action because of the continuous aspect marker *-yor*. It is obvious that the participants do not associate the head noun 'iş' with this atelic action.

When the text is presented with the bare anaphor 'bu' (Eng. 'this'), on the other hand, there is not such a limiting factor on the participants' decisions. As discussed in 2.2.1, the bare anaphor 'bu' does not always carry the load of denoting the precise ontological category that is referred to. Consten and Knees (2005) argue that the bare abstract object anaphor 'this' is neutral in its ontological category. In the cases 'this' is used, the ontological type of the anaphoric process is denoted by the antecedent. Therefore, the participants might have associated the bare anaphor 'bu' with 'looking for solutions', 'the mystery of woman body lying before Halil's eyes', as well as 'the passion for drawing woman body', or 'that passion's turning into his purpose in life', i.e., the number of subsumption and disjunct relations increase with 'bu'; thus, the number of competing elements as a potential antecedent.

²⁴ As the direct translation of the head noun 'iş', the English translations are given as 'activity'. However, as is clear from the subsets, the label anaphor 'bu iş' actually does not refer to a concrete entity, rather it refers to a process or a proposition.

Interim Summary:

Passonneau's (2004) classification method was used to gain better insight into distinct interpretations of target anaphors while preserving the same units of interpretations. This analysis helps understanding 'what' distinct interpretation strings actually are. As it was also discussed in 4.2.3.3, when the text is presented with the label anaphor 'bu iş', the label anaphor seems to be eliminating one of the competing candidates in the interpretation of the antecedent due to the lexical meaning of the head noun.

4.2.4 Discussion


All in all, to investigate the collected data, four different methods were used. The first was the consistency analysis on the start and end words that each participant underlined in each group. This analysis indicated the presence of higher consistency in the start and end words while marking the antecedent to the label anaphors in some cases, specifically in paragraph 3. Secondly, Krippendorff's Alpha was used to obtain inter-coder agreement through α values. The findings indicated a notable difference between the agreement rates of the two groups in paragraph 3. Thirdly, visual representations of the markings through bar graphs were prepared to see 'where' participants completely or partially agreed or disagreed with their markings. The analysis showed that bare abstract anaphor 'bu' (Eng. 'this') had a tendency to allow alternative interpretations. On the other hand, label anaphors had a tendency to eliminate the competing candidates as potential antecedents. Finally, Passonneau's (2004) classification of subsumption, intersection and disjunct relations among the markings of the participants was adapted to the current data. This analysis clarified 'what' exactly the non-overlapping strings were.

All of these methods proved that participants from both groups partially agreed on their markings in paragraph 1 in which the bare anaphor 'bu' and the label anaphor 'bu gerçek' (Eng. 'this fact') were compared, and also in paragraph 2 in which the bare anaphor 'bu' and label anaphor 'bu durum' were compared. Their markings showed significant difference with a bias for the label anaphor for paragraph 3 in which the bare anaphor 'bu' and label anaphor 'bu iş' (Eng. 'this activity') were compared. Finally, their markings almost fully overlapped for

paragraph 4 in which the bare anaphor 'bu' and label anaphor 'bu olay' (Eng. 'this event') were compared.

These findings confirm the complexation of abstract anaphora resolution in the abstractness hierarchy discussed by Consten and Knees (2005), which was summarized in Section 2.3. According to this hierarchy, the abstract anaphors denoting events are the most concrete, thus, less complex compared to processes, states, facts and propositions. Therefore, they lie at the bottom of the hierarchy. Recall that the hierarchy follows this order; *propositions* at the very top, followed by *facts*, *states*, *processes* and *events* at the very bottom.

Table 10 Degree of abstractness, ontological categories and Krippendorff's Alpha Scores

Degree of Abstractness	Ontological Categories	Word Boundary α Values in Group 1	Word Boundary α Values in Group 2
high  low	bu (pp) vs. bu iş (pp) ²⁵	0.223	0.442
	bu(f) vs. bu gerçek (f) ²⁶	0.555	0.565
	bu (s) vs. bu durum (s) ²⁷	0.671	0.723
	bu (e) vs. bu olay (e) ²⁸	0.925	0.962

The α values indicate perfect agreement scores for paragraph 4 in which the bare abstract object anaphor 'bu' and the label abstract anaphor 'bu olay' are compared. In this paragraph, the common antecedent specified by the participants from both groups is the event in which Ziya Paşa's house was seized by the invaders. According to Consten and Knees' (2005), the events lies at the bottom of the abstractness hierarchy, and with the highest rate of agreement, they hold the lowest level of ambiguity. Besides this, α values indicated substantial agreement scores for paragraph 2 in which the bare abstract object anaphor 'bu' and the label abstract anaphor 'bu durum' were compared. Also, α values indicated moderate agreement

²⁵ Eng. 'this activity'

²⁶ Eng. 'this fact'

²⁷ Eng. 'this situation'

²⁸ Eng. 'this event'

scores in which the bare abstract object anaphor 'bu' and the label abstract anaphor 'bu gerçek' were compared. Compatible with the suggestion of Consten and Knees (2005), 'bu durum' denoting a situation in the text and 'bu gerçek' denoting a fact in the text lie in the middle of the hierarchy.

However, 'bu' and 'bu iş', demonstrate the lowest agreement scores being the most complex ones to be processed among these 4 paragraphs. The lowest agreement scores indicate the presence of ambiguity in paragraph 3 no matter whether it is presented with the bare anaphor or the label anaphor. However, the level of ambiguity increases when it is presented with the bare anaphor. This is because, as discussed in 4.2.3.4., the lexical meaning of the head noun is a limiting factor on the number of the competing potential antecedents.

4.3. Experiment 2

To confirm the reliability of the findings in Experiment 1, the second experiment was conducted with the following aim:

- Is the data set obtained from Experiment 1 generalizable enough to allow disambiguating role of label anaphors?

4.3.1. Experiment Design:

Sample: 66 participants, all native speakers of Turkish, took part in the experiment from both genders. 36 of them were in Group 1, and 30 were in Group 2. As in Experiment 1, participants were given a demographic form to learn about their ages, departments etc. The ages of the participants differed from 18 to 21. All of them were undergraduate freshmen from various departments at METU.

Materials: It was an off-line study like the one conducted for the pilot study in Experiment 1. Paper and pencil were used. Participants were provided with 4 texts (see Appendix 10 and Appendix 11). These texts were different from the ones used in Experiment 1.

As in Experiment 1, participants were asked to underline the antecedent of the highlighted abstract object anaphor. They were provided with the same brief training session prior to the experiments (see Appendix 3 and Appendix 4).

Stimuli: 4 texts different from each other taken from the TDB, the novel *Yüksek Topuklar* (High Heels written by Murathan Mungan) and *Veda* (Farewell written by Ayşe Kulin) were used to be given to each group.

Text Manipulation: As in the pilot study and Experiment 1, the length of the each text was equalized to 60-64 words in total. Each text had one anaphor referring back to an abstract object to be determined by the participants. The original texts occurred with the bare anaphor 'bu'. In the manipulated version, the bare anaphors were replaced with the label anaphors with the same head nouns used in Experiment 1. The participants were presented with the text with the bare anaphor as well as the manipulated texts containing the label anaphors.

Procedure: The experiment was conducted by the researcher. It was completed within nearly 20 minutes with the training session included as in Experiment 1. Group 1 received the texts with the bare abstract object anaphor 'bu' (Eng. 'this) (see Appendix 10) and Group 2 received the same texts in which the bare anaphors were replaced with label anaphors, 'bu+iş/gerçek/olay/durum' (Eng. 'this+activity/fact/event /situation) (see Appendix 11). All of the participants returned the test material with their annotations.

4.3.2 Analysis of the Data

As in Experiment 1, some markings by the participants were identified as structurally or semantically non-matching to be the antecedents to the target anaphors. Those markings were also excluded from the analysis upon discussion with two expert annotator from TDB group. Cataphoric cases were also excluded upon discussion with an expert linguist who has expertise in Turkish linguistics and cognitive studies. Table 11 shows the excluded data and reasons for excluding them from the analysis. Table 12 presents the percentage of the excluded data from each group and from each paragraph.

Table 11 Non-Matching Markings in Experiment 2

The Paragraph	Non-Matching Marking	Reasons for Non-matching
<p>1. Taş kaynaklarına baktığımızda, doğal olarak çok taş malzemesi var; bazalt, kireçtaşı, tüfler, volkanik taşlar çok fazla var. Aşıklı'ya yerleşenler ilk geldiklerinde kerpiç mimariyi biliyor. Yapılarını önceleri kerpiçten yapıyorlar, ama sonra kerpiçle taşı birlikte kullanmayı öğreniyorlar. Mimarlık açısından çok önemli bu/bu iş, çünkü bir yapı malzemesini başka bir malzemeyle beraber kullanmayı, ilk defa burada görüyoruz. Evler kerpiçten, ama tapınağın temelinde taş var; çevre duvarı neredeyse tamamen taştan.</p>	<ul style="list-style-type: none"> • birlikte kullanmayı 	<ul style="list-style-type: none"> • This marking does not fully account for the activity 'bu/bu iş' refers to. 'Kerpiçle taşı' (Eng. 'mud and stone') should have been involved in this marking.
<p>2. Ahmet, bir gün gazeteyi bir açmıştı, Mesut yok! Pusuya düşürülmüş oracıkta can vermişti. İki kişiymişler. Öbürünün adı Halim. Tanımıyordu. Hafif yaralı kurtulmuş o. Bu/Bu olay derinden sarsmıştı Ahmet'i. Evden ayrılmıştı. Okula gitmiyordu artık. Kimseyle görüşmüyordu. Köyüne gömmüşlerdi Mesut'u. 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.</p>	<ul style="list-style-type: none"> • Pusuya düşürülmüş oracıkta can 	<ul style="list-style-type: none"> • The meaning of this marking does not match the verb phrase 'bir yıkımdı' (Eng. 'was a destruction'). This marking does not explain an event.
<p>3. Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı hiçbir zaman içine sindirememiş olan ve bu konuda duyduğu ezikliği saklayamayan Turgay'a, bu konuda sürekli takılırlar. (Aslında şimdi ben de onu yapmış oldum.) Çünkü bu/bu durum, onun daima kanayan açık yarasıdır, camiada herkesçe bilinir ve kabul edersiniz ki, insan ilişkilerinde en çabuk öğrendiğimiz şeylerden biri, başkalarının yarasını kullanmaktır.</p>	<ul style="list-style-type: none"> • Yıllarca hapis yatmış militan bir komünist • Turgay'a, bu konuda sürekli takılırlar. 	<ul style="list-style-type: none"> • The first marking is incomplete because it does not explain a situation as the anaphors in this text denote. • This markings does not explain a situation as the anaphors in this text denote.
<p>4. Dört kişiydik, törenden sonra deniz kenarında bir kahvede oturmuştuk. Kıştı, kar yerden kalkmamıştı. Sen ağlamıştın. Mutluluktan,</p>	<ul style="list-style-type: none"> • gizlemeye 	<ul style="list-style-type: none"> • This marking does not account for a 'fact' that the bare anaphor and the label

yalnızca mutluluktandı göz yaşların. Sulu gözlüydün Nesli, ama itiraf etmeye utanıyordun değil mi? Ne kadar gizlemeye çalışsan da bunu/bu gerçeği , başaramıyordun. Biz yasalar karşısında evli sayılacak, ama gerçekte evli iki insan gibi değil de (evlilikler sıradanlaşıyordu çünkü, tekdüze ve sıkıcıydı; biz farklı olacaktık), aynı evi paylaşan iki öğrenci gibi yaşayacaktık.	<ul style="list-style-type: none"> • evli iki insan gibi değil de (evlilikler sıradanlaşıyordu çünkü, tekdüze ve sıkıcıydı; biz farklı olacaktık), aynı evi paylaşan iki öğrenci gibi yaşayacaktık 	<p>anaphor 'bu gerçek' (Eng. 'this fact') denote.</p> <ul style="list-style-type: none"> • Cataphoric reference
---	---	--

Table 12 Percentage of Excluded Data in Experiment 2

Experiment	Group	Paragraph	Number of Participants	Number of Excluded Data	Percentage of Excluded Data
2	Bare Anaphor	1	30	1	3,33%
2	Bare Anaphor	2	30	0	0,00%
2	Bare Anaphor	3	30	1	3,33%
2	Bare Anaphor	4	30	1	3,33%
2	Label Anaphor	1	36	0	0,00%
2	Label Anaphor	2	36	1	2,78%
2	Label Anaphor	3	36	2	5,56%
2	Label Anaphor	4	36	1	2,78%

Method: The same steps as in Experiment 1 were followed in the data analysis:

Step 1: Consistency analysis was done by means of the histograms of the start and end words that each participant underlined in each group (see 4.3.3.1).

Step 2: Krippendorff's Alpha was computed to obtain inter-coder agreement scores (see 4.3.3.2).

Step 3: Bar graphs were drawn so that the data could be visually represented (see 4.3.3.3).

Step 4: Passonneau's (2004) categorization of the main sets, subsumptions, intersection and disjunct analysis of the annotations was applied (see 4.3.3.4).

4.3.3 Results

4.3.3.1. Consistency Analysis on the Start and End Words

The first step in the data analysis was to conduct a consistency analysis to identify the start word and end word of each annotation specified by each participant in both groups. A 15%-threshold was created to specify the most frequent start and end words (see Appendix 16, Tables 20 and 21).

Consistency in Start Words: For paragraph 1, the results indicate that the most frequent start words (see Appendix 12, Figure 23) for Group 1 marking the antecedent to the bare anaphor 'bu' (Eng. 'this') is the 31st one. This is the same for Group 2 (see Appendix 12, Figure 24) marking the antecedent to the label anaphor 'bu iş' (Eng. 'this activity'). However, there is also a high frequency to start with the 25th word in Group 2 receiving the test material with the label anaphor. This is not compatible with the results in Experiment 1. This is because, in Experiment 1, the distribution of the start words in Group 2 was less scattered compared to Group 1. Besides this, a notable difference between the start words was found for the paragraph in which 'bu' and 'bu iş' were compared in Experiment 1 with a bias for 'bu iş'. However, there is not a remarkable difference between the start and end words of the antecedents to the bare anaphor and the label anaphor 'bu iş' in Experiment 2.

For paragraph 2 in which the bare anaphor and the label anaphor 'bu olay' (Eng. 'this event') were compared, the most frequent words for both groups are the 6th and 8th words (see Appendix 12, Figure 25 and Figure 26), which indicates no difference between the groups.

For paragraph 3 in which the bare anaphor and the label anaphor 'bu durum' (Eng. 'this situation') were compared, while the participants converge on the 1st word

in Group 2, the convergence of participants in Group 1 is distributed between the 1st and the 8th words. Figures 17 and 18 below illustrate this difference:

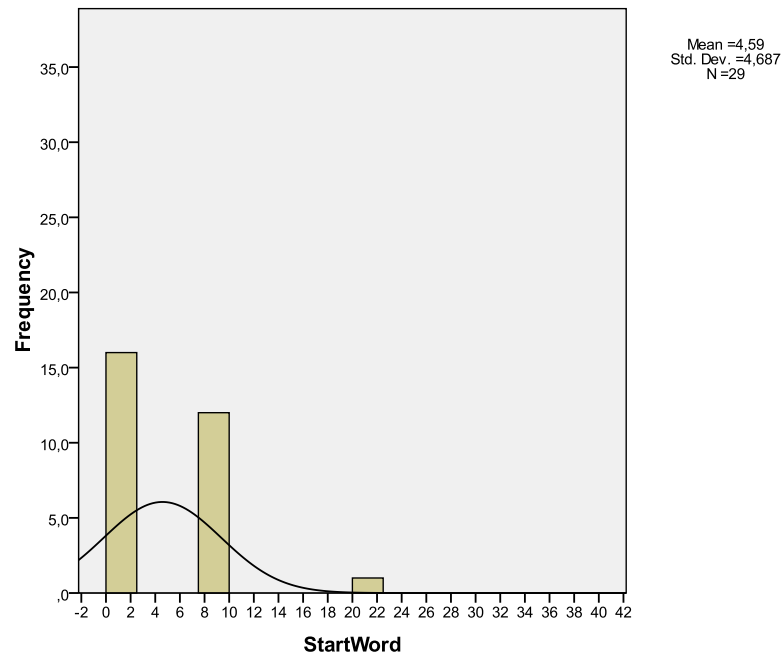


Figure 17. The frequency of start words of the participants in Group 1 for paragraph 3 with the most common start words being the 1st and the 8th words.

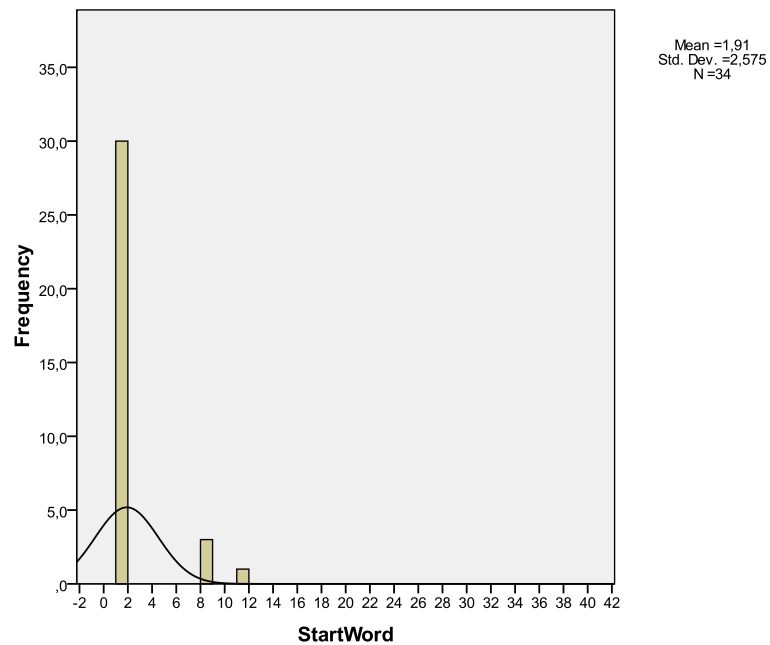


Figure 18. The frequency of start words of the participants in Group 2 for paragraph 3 with the most common start word being the 1st one.

For the last paragraph in which the bare anaphor and the label anaphor 'bu gerçek' (Eng. 'this fact') were compared, the most frequent start words are the 20th and 13th for both groups (see Appendix 12, Figures 27 and 28). This indicates no difference between the groups.

Consistency in End Words: For paragraph 1 in which the bare anaphor 'bu' and the label anaphor 'iş' were compared, the histograms are almost alike. The most frequent end words are the 34th and 35th for both groups (see Appendix 13, Figures 29 and 30).

For paragraph 2 in which the bare anaphor and the label anaphor 'bu olay' were compared, there is high consistency in both of the groups to end with the 12th word (see Appendix 13, Figures 31 and 32).

For paragraph 3 in which the bare anaphor and the label anaphor 'bu durum' were compared, despite some variations, there is a high tendency to end with the 14th word in both groups. This shows the lack of a notable difference (see Appendix 13, Figures 33 and 34).

For paragraph 4 in which the bare anaphor and the label anaphor 'bu gerçek' were compared, most of the participants from both groups ended their markings with the 21st word.

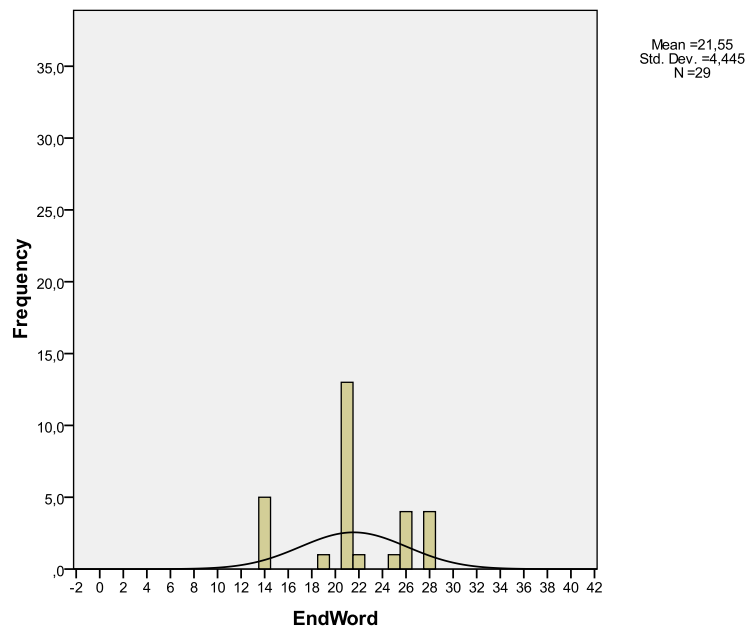


Figure 19. The frequency of end words of the participants in Group 1 for paragraph 4 with the most common end word being the 14th one.

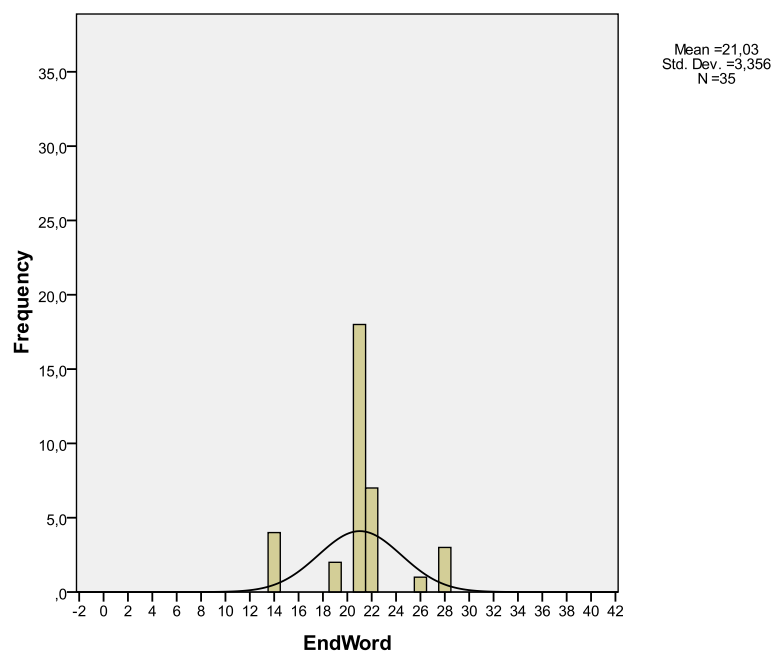


Figure 20. The frequency of end words of the participants in Group 2 for paragraph 4 with the most common end word being the 14th one.

Interim Summary:

As in Experiment 1, to investigate the consistency in the start and end words, histograms were prepared for the markings of both Groups 1 and 2. The aim was to gain an insight into whether there was any observable difference in the way the participants specified the boundaries of the antecedents to bare anaphors and label anaphors. The analysis indicates higher consistency in the start words of Group 2 in paragraph 3. The markings in the rest of the paragraphs do not indicate a notable difference between the boundaries of the bare anaphor and the label anaphor. However, the higher standard deviations in the end words of Group 2 makes it difficult to generalize the role of label anaphors in specifying the boundaries of the antecedents in a precise way. Table 13 below summarizes the start words from both groups and Table 14 summarizes the end words:

Table 13 Consistency in Start Words in Experiment 2²⁹

Paragraph #	Group 1 (Bare Anaphor)		Group 2 (Label Anaphor)	
	Start Word	Percentage of Frequency	Start Word	Percentage of Frequency
1			25 th	25,00%
	31 st	82,76%	31 st	69,44%
2	6 th	40,00%	6 th	37,14%
	8 th	40,00%	8 th	45,71%
3	1 st	55,17%	1 st	88,24%
	8 th	41,38%		
4	13 th	24,14%	13 th	20,00%
	20 th	62,07%	20 th	74,29%

Table 14 Consistency in End Words in Experiment 2³⁰

Paragraph #	Group 1 (Bare Anaphor)		Group 2 (Label Anaphor)	
	End Word	Percentage of Frequency	End Word	Percentage of Frequency
1	34 th	31,03%	34 th	52,78%
	35 th	68,97%	35 th	44,44%
2	12 th	86,67%	12 th	82,86%
3	14 th	65,52%	14 th	52,94%
4	14 th	17,24%	21 st	51,43%
	21 st	44,83%	22 nd	20,00%

²⁹ see Appendix 16, Table 20 for the frequency of other markings on start words.

³⁰ see Appendix 16, Table 21 for the frequency of other markings on end words.

4.3.3.2. Inter-coder Agreement Analysis through α Values

The second step in the analysis of the data, was to take Krippendorff's Alpha measurements for the word boundaries and word intervals of the data in Experiment 2. Table 15 shows the α values for word boundaries:

Table 15 Analysis of Data with Krippendorff's Alpha - Word Boundary Results for Experiment 2

Experiment 2 - Word Boundary			
Paragraph #	Group 1	Group 2	Label
	(Bare Anaphor) α value	(Label Anaphor) α value	
1	0.536	0.624	iş
2	0.430	0.569	olay
3	0.135	0.438	durum
4	0.805	0.725	gerçek

The α values indicate moderate agreement in both groups for paragraph 1 and paragraph 2 ($0.6 > \alpha > 0.4$), and substantial agreement in paragraph 4 ($0.8 > \alpha > 0.6$). The results suggest that the bare anaphor and the label anaphor are resolved in similar ways in these paragraphs. However, there is a notable difference in the scores between the groups in paragraph 3: There is only slight agreement in Group 1 ($0.2 > \alpha > 0.0$), whereas there is moderate agreement in Group 2 ($0.6 > \alpha > 0.4$). As also observed in Experiment 1, the difference between the agreement scores produced by the bare anaphor 'bu' and the label anaphor 'bu durum' might have caused by label anaphors' specifying the boundaries of the antecedent with the help of the lexical meaning of the head noun in the label. The results obtained through Krippendorff's Alpha word boundary scores are compatible with the preliminary analysis conducted through histograms. The histograms on the start words for paragraph 3, as shown in Figures 16 and 17, also indicated a notable difference between the groups in paragraph 3.

The results for word interval data are shown in Table 16 below. As it has already been mentioned in 4.2.3.2., the word interval measurements consider the partial disagreements in the coding units. The word interval results in paragraph 1 and 4 are compatible with those in word boundary results although there is only a slight increase in the scores of paragraph 1, and only slight decrease in the scores of paragraph 4 (cf. Table 15 & Table 16). These are not significant differences between the word boundary and word interval measurements for paragraph 1 and 4.

Table 16 Analysis of Data with Krippendorff's Alpha - Word Interval Results for Experiment 2

Experiment 2 - Word Interval			
Paragraph #	Group 1	Group 2	
	(Bare Anaphor) α value	(Label Anaphor) α value	Label
1	0.647	0.760	iş
2	0.172	0.219	olay
3	-0.241	0.463	durum
4	0.708	0.717	gerçek

However, the agreement rates decrease in paragraph 2, in which both the bare anaphor 'bu' and the label anaphor 'bu olay' denoted an event. This result is not compatible with the results for the paragraph in Experiment 1 in which the bare anaphor 'bu' and the label anaphor 'bu olay' were compared. Perfect agreement was found for that paragraph in Experiment 1 for both groups, whereas the same label and the bare anaphor denoting an event display almost poor disagreement scores in Experiment 2.

For paragraph 3, the word interval results are compatible with the word boundary results (cf. Table 13 & Table 14). The agreement rate decreases even more in Group 1 ($\alpha > 0.0$). This score is compatible with the word boundary results as well as the consistency analysis on the start and end words conducted for Experiment 2.

Interim Summary:

The only notable difference obtained through α values between the groups in Experiment 2 is in paragraph 3 in which the bare anaphor 'bu' and the label anaphor 'bu durum' (Eng. 'this situation') were compared. In other paragraphs, there is no significant difference found between the groups.

Both the word interval and word boundary measurements in Experiment 2 is compatible with Experiment 1 in that label anaphors have a tendency to produce better agreement scores compared to the bare anaphor 'bu'. However, the individual scores are non-compatible with the ones from Experiment 1 (cf. Tables 5 and 6 & Tables 13 and 14). Perfect agreement was found in the text where label anaphor 'bu olay' (Eng. 'this event) and the bare anaphor 'bu' referring back to an even were used in Experiment 1. In Experiment 2, on the other hand, the text in which these anaphors were used (i.e., paragraph 4) demonstrates only moderate agreement both in word interval and word boundary scores. Moreover, the lowest agreement scores in Experiment 1 were obtained from the word boundary and word interval results of the text in which the label anaphor 'bu iş' (Eng. 'this activity') and the bare anaphor 'bu' were compared. In Experiment 2, on the other hand, the text presented with 'bu iş' vs. 'bu' presents moderate agreement in word boundaries and substantial agreement in word intervals. Besides this, in Experiment 1, there was not a significant difference between the scores of groups for the text in which the label anaphor 'bu durum' (Eng. 'this situation') and the bare anaphor 'bu' was compared. However, in Experiment 2, there is an observable difference.

4.3.3.3. Visual Illustration of the Whole Data

As in Experiment 1, bar graphs were used in the visually illustration of the data so as to gain better insight into the markings generating the agreement and disagreement rates indicated by α values. As in Experiment 1, the bars stand for the boundaries of the annotation of each participant. These bars have been ordered as beginning with the longest marking and ending with the shortest one.

The Figures 21 and 22 below illustrate the markings of the antecedents in paragraph 1:

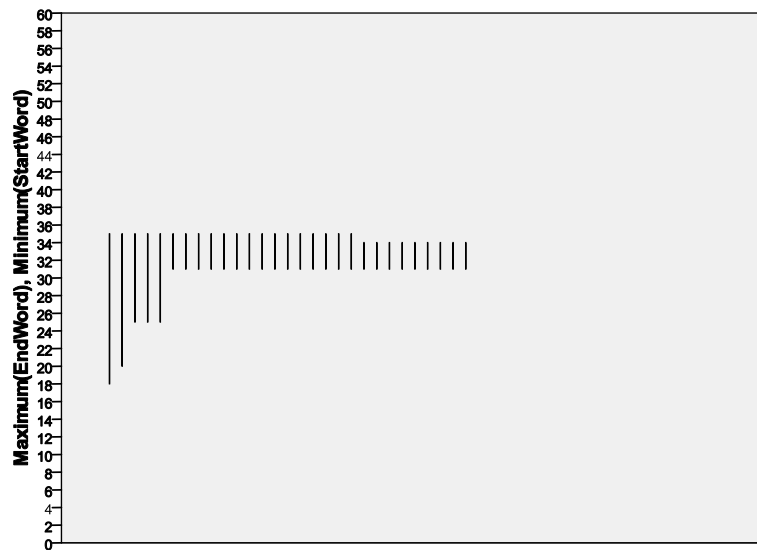


Figure 21. The markings of the antecedents by each participant for Paragraph 1 from Group 1

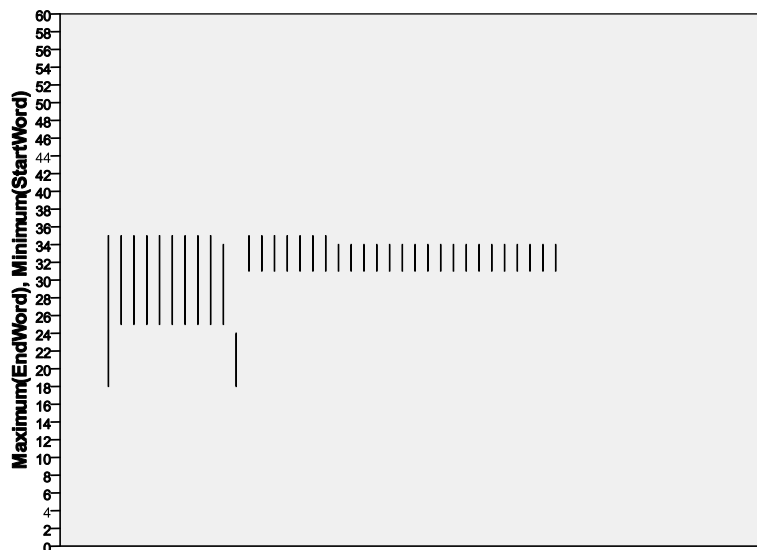


Figure 22. The markings of the antecedents by each participant for Paragraph 1 from Group 2

The bars in Figures 21 and 22 show that the dominant marking of the antecedent is between the 31st-34th and 31st-35th words in both groups. There seems to be no difference on the data obtained from both groups in this paragraph.

Figures 23 and 24 below represent the markings in paragraph 2:

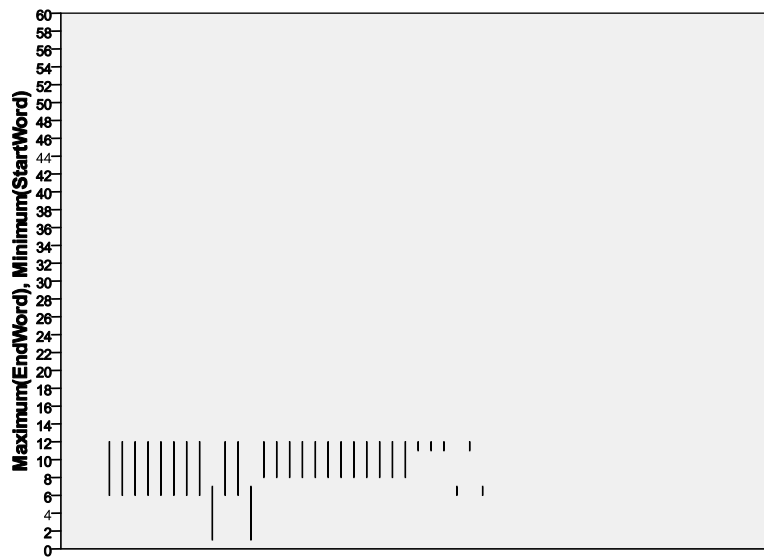


Figure 23 The markings of the antecedents by each participant for Paragraph 2 from Group 1

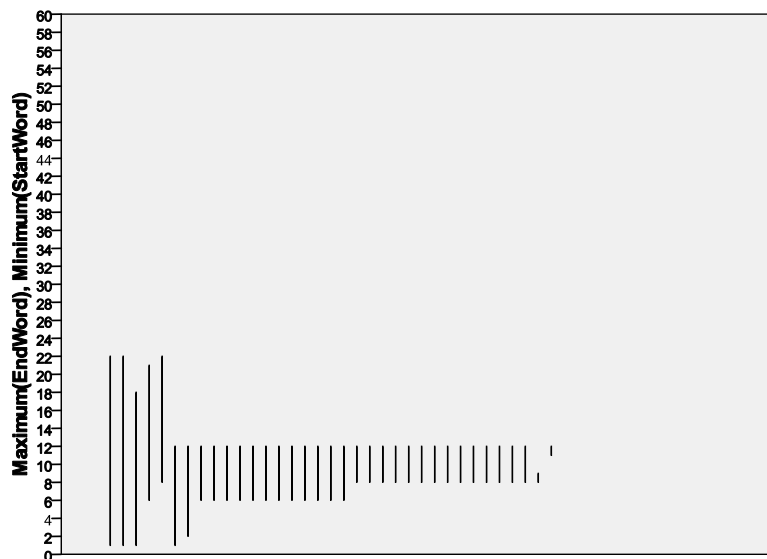


Figure 24. The markings of the antecedents by each participant for Paragraph 2 from Group 2

The bars in Figures 23 and 24 indicate that the dominant marking in both groups is between the 6th-12th and 8th-12th words. However, in Figure 24, it is observed that the boundaries of the antecedents to the label anaphor 'bu olay' specified by the participants vary more compared with the ones to the bare anaphor. This is not compatible with the findings in Experiment 1. This is because most of the findings in Experiment 1 indicated that label anaphors were more precise in

specifying the boundaries of the antecedents and in producing consistency in the markings.

Figures 24 and 25 below represent the markings in paragraph 3:

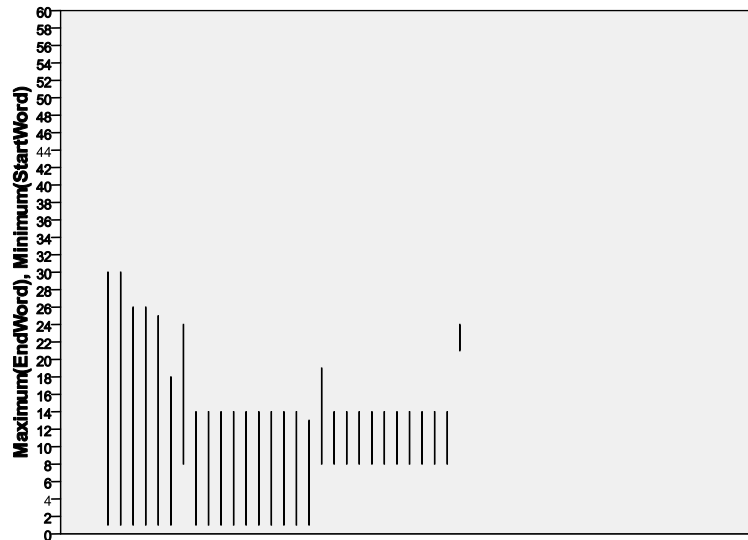


Figure 25. The markings of the antecedents by each participant for Paragraph 3 from Group 1

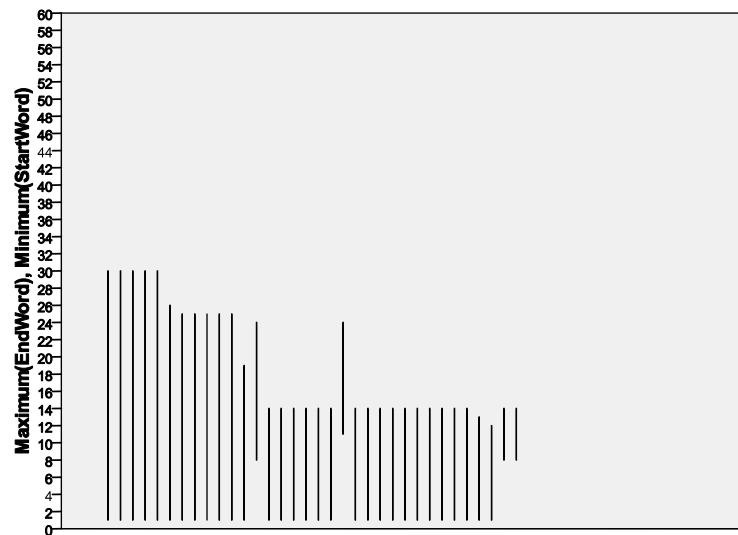


Figure 26. The markings of the antecedents by each participant for Paragraph 3 from Group 2

The quantitative analysis conducted through histograms in Step 1 and the Krippendorff's Alpha scores indicated a difference between the groups in paragraph 3. The bars in Figures 25 and 26, on the other hand, do not present a visible certain pattern to be interpreted as a difference.

The Figures 27 and 28 below illustrate the markings of the antecedents in paragraph 4:

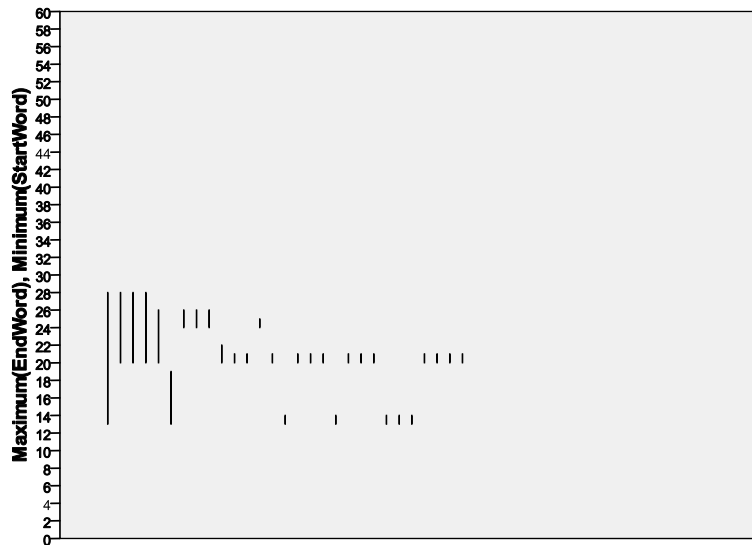


Figure 27 The markings of the antecedents by each participant for Paragraph 4 from Group 1

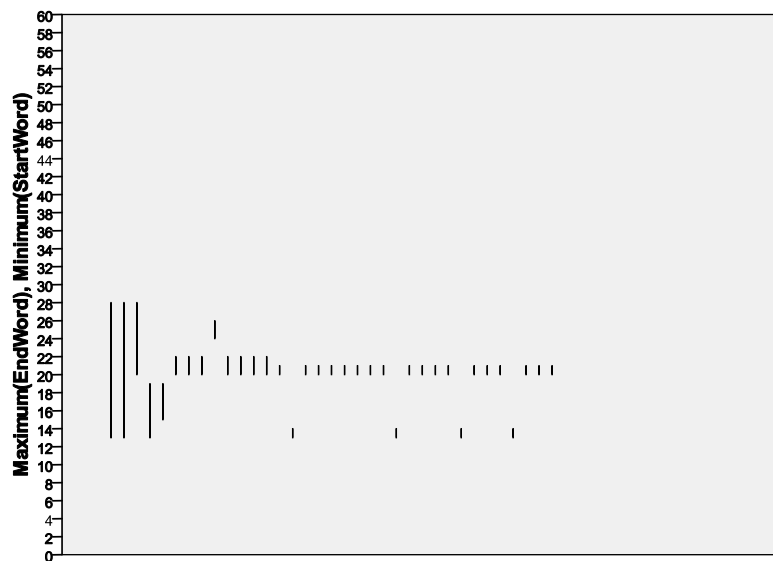


Figure 28 The markings of the antecedents by each participant for Paragraph 4 from Group 2

The bars in Figures 27 and 28 show that the dominant marking of the antecedent is between the 20th and 21st words in both groups. There seems to be no difference on the data obtained from both groups in this paragraph.

Interim Summary:

In the bar graphs provided, the annotated antecedents in all of the paragraphs in both groups seem to be at least partially overlapping with each other. In Experiment 1, it was observed that the bare abstract anaphor 'bu' had a tendency to allow alternative interpretations in some cases. On the other hand, label anaphors had a tendency to eliminate the competing candidates to be antecedents; therefore, minimize the occurrences of the alternative interpretations. The same analysis for the data obtained in Experiment 2 does not reveal such a difference between the markings of the groups.

4.3.3.4. Passonneau's Proposal

The categorization of subsumption, disjunct and intersection by Passonneau (2004) was also applied for the data set obtained from Experiment 2 (see Appendix 14). While all of the three relations i.e., subsumption, disjunct and intersection were observed in Experiment 1, there is only the subsumption relation observed in Experiment 2.

To illustrate this analysis, the analysis of antecedents for paragraph 1 is as follows:

Paragraph 1:

A: Aşıklı'ya yerleşenler ilk geldiklerinde kerpiç mimariyi biliyor. Yapılarını önceleri kerpiçten yapıyorlar, ama sonra kerpiçle taşı birlikte kullanmayı öğreniyorlar

A1: ilk geldiklerinde kerpiç mimariyi biliyor. Yapılarını önceleri kerpiçten yapıyorlar, ama sonra kerpiçle taşı birlikte kullanmayı öğreniyorlar.

A2: Yapılarını önceleri kerpiçten yapıyorlar, ama sonra kerpiçle taşı birlikte kullanmayı öğreniyorlar

A3: Yapılarını önceleri kerpiçten yapıyorlar, ama sonra kerpiçle taşı birlikte kullanmayı

A4: Aşıklı'ya yerleşenler ilk geldiklerinde kerpiç mimariyi biliyor

A5: kerpiçle taşı birlikte kullanmayı öğreniyorlar

A6: kerpiçle taşı birlikte kullanmayı

No Match:

- birlikte kullanmayı

Paragraph 1:

A: Those settling in Aşıklı knows mud architecture when they first came. They first build their structures with mud, but later they learn how to use the mud together with stone

A1: knows mud architecture when they first came. They first build their structures with mud, but later they learn how to use the mud together with stone

A2: They first build their structures with mud, but later they learn how to use the mud together with stone

A3: They first build their structures with mud, but later ... how to use the mud together with stone

A4: Those settling in Aşıklı knows mud architecture when they first came

A5: they learn how to use the mud with stone

A6: to use the mud together with stone

No Match:

- how to use together

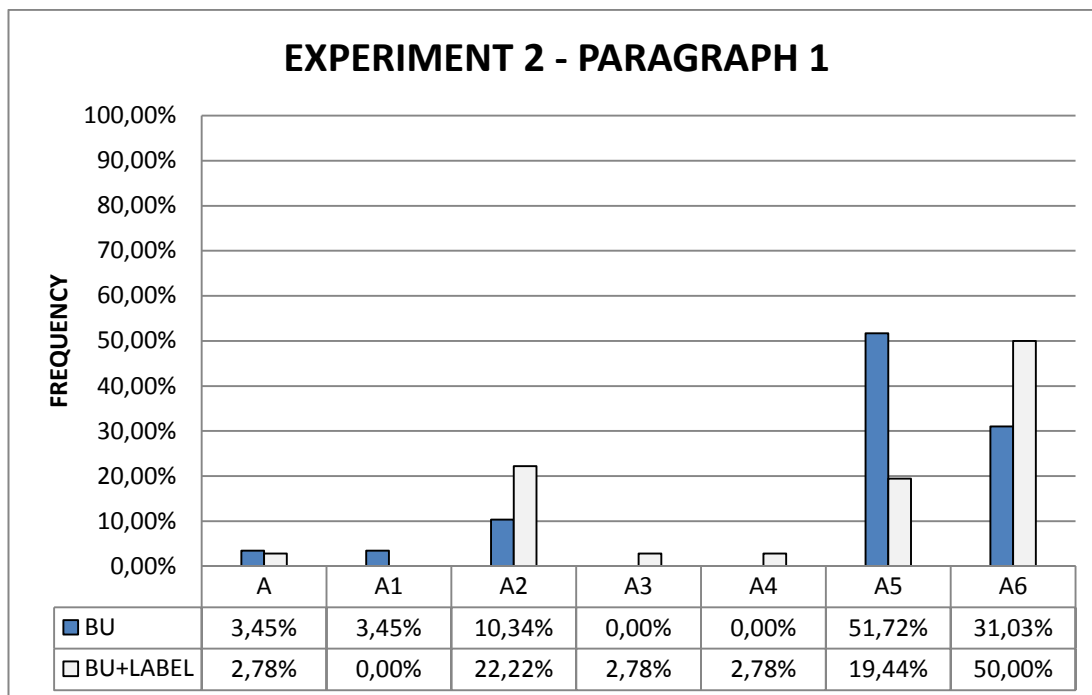


Figure 29³¹. Passonneau's classification of main set and subsumption relations for Paragraph 1 from both groups

Figure 29 reveals an interesting case that was not possible to see in the previous three steps. There is an asymmetry between the subsets A5 and A6. There is a higher number of participants in Group 1 receiving the text with the bare anaphor 'bu' (Eng. 'this') who specify the antecedent as 'kerpiçle taşı birlikte kullanmayı öğreniyorlar' (Eng. 'they learn how to use the mud with stone'). On the other hand, there is a higher number of participants in Group 1 receiving the same text with the label anaphor 'bu iş' (Eng. 'this activity') who specify the antecedent as 'kerpiçle taşı birlikte kullanmayı' (Eng. 'to use the mud together with stone') by leaving out 'öğreniyorlar' (Eng. 'they learn'). This shows that participants have a tendency to associate the label anaphor 'bu iş' with the activity of *using* the mud together with stone, whereas they have a tendency to associate the bare anaphor 'bu' with *learning* how to use the mud together with stone. This finding is compatible with the result obtained as a result of the qualitative evaluation of the pilot study. Recall: Participants had a tendency to associate the bare anaphor 'bu' with 'kadın bedeni

³¹ Y axis shows the frequency of the classifications. Below the graph, the percentages of each set's occurrence in Group 1 (bu) and Group 2 (Bu + label) are provided.

çizme tutkusu' (Eng. 'the passion for drawing woman body'), and the label anaphor 'bu iş' with the activity 'kadın bedeni çizme' (Eng. 'drawing woman body').

Interim Summary:

As in Experiment 1, Passonneau's (2004) classification method was used to gain better insight into the overlapping interpretations of target anaphors. This analysis helped to see the differences in the resolution of the bare anaphor 'bu' and 'bu iş'. As it was also discussed as a finding in the pilot study, participants have a tendency to associate 'bu' and 'bu iş' with different entities. While they associate 'bu iş' with an activity (e.g., *drawing* woman body, *using* mud together with stone), they associate 'bu' with more abstract entities (e.g., *passion* for drawing woman body, *learning* how to use the mud together with stone).


4.3.4 Discussion

The main aim to conduct Experiment 2 was to understand whether the data set obtained from Experiment 1 was generalisable enough to allow the disambiguating role of label anaphors. To analyze the collected data, four different steps were followed. The first was to investigate the consistency in the start and end words each participant specified while finding the antecedents to the target anaphors. This analysis did not indicate a notable difference -except for the higher consistency in the start words of Group 2 in paragraph 3- between the boundaries of the bare anaphor and the label anaphor specified by the participants. The second was to compute Krippendorff's Alpha to obtain agreement scores to investigate the differences in agreement rates between the groups. It was found that both the word boundary and word interval scores indicated a slightly better agreement in the antecedents to the label anaphors. A remarkable difference between the agreement rates of the groups was found only in paragraph 3 in which the bare anaphor 'bu' and the label anaphor 'bu durum' (Eng. 'this situation') were compared. 'Bu durum' produced better agreement scores. The third step was to visually represent the data with bar graphs. This analysis did not provide any visible difference between the markings of the two groups. The final step was to apply Passonneau's (2004) method. This analysis revealed the differences in the resolution of the bare anaphor 'bu' and

the label anaphor 'bu iş'. Participants have a tendency to associate 'bu' and 'bu iş' with different entities.

It was discussed in 4.2.4. that the findings in Experiment 1 confirmed the complexation of abstract anaphora resolution in the abstractness hierarchy discussed by Consten and Knees (2005). According to this hierarchy, *events* should be at the bottom being the most concrete entities, and they are followed by *states*, *facts* and *propositions*. However, the findings in Experiment 2 do not completely confirm this hierarchy. Table 17 below shows the hierarchy based on the word boundary α values. In the table, the anaphors denoting a fact lie at the bottom, followed by those denoting a proposition, an event and a state:

Table 17 Degree of abstractness, ontological categories and Krippendorff's Alpha Scores

Degree of Abstractness	Ontological Categories	Word Boundary α Values in Group 1	Word Boundary α Values in Group 2
high  low	bu (s) vs. bu durum (s) ³²	0.135	0.438
	bu (e) vs. bu olay (e) ³³	0.430	0.569
	bu (pp) vs. bu iş (pp) ³⁴	0.536	0.624
	bu(f) vs. bu gerçek (f) ³⁵	0.805	0.725

The findings indicate that the scores obtained through α value cannot be accounted for with the abstractness hierarchy suggested by Consten and Knees (2005).

Upon a closer look at the structures in the texts, syntactic factors seem to be having a role in the non-compatible α scores with the abstractness hierarchy. First, it was thought whether a distinction in the types of the phrases participants underlined

³² Eng. 'this situation'

³³ Eng. 'this event'

³⁴ Eng. 'this activity'

³⁵ Eng. 'this fact'

had an effect on the scores. Whether participants had a tendency to specify, for example, consistently a verb phrase or a noun phrase as the antecedent depending on the type of the anaphor (i.e., bare anaphor or label anaphor) was analyzed. The analysis of categorizing the subsumption, disjunction and intersection relations by Passonneau (2005) applied in 4.3.3.4 and 4.2.3.4 was expanded for this. A frequency analysis of the type of the phrases specified by the participants was conducted. To do this, the types of the phrases in each category (e.g. A, A1, A2, B, etc.) specified as the antecedents to the target anaphors by the participants were determined as NP, VP AdjP, etc., (see Appendix 17 for Experiment 1 and Appendix 18 for Experiment 2). The percentage of each marking distributed in the whole data was calculated. As a result of this analysis, a specific pattern of consistently choosing a VP or NP as the antecedent to either the bare anaphor or label anaphor was not observed. The most frequent markings, decided on a %15-threshold, varied as VP as well as NP regardless of the type of the target anaphor.

Other factors like level of embedding and the number of elaborative clauses giving details of the discourse segment in focus (as discussed in 3.3.2 as the right frontier operations) can be thought to be the reasons for the α scores and the order in the hierarchy given in Table 15.

Bu vs. Bu gerek:

The most common antecedent (A7 as the subset)³⁶ specified by the participants (see Appendix 14) in both groups to the bare anaphor 'bu' and the label anaphor 'bu gerek' is not in an embedded clause. This might have eased the resolution of the abstract anaphors and resulted in the highest agreement scores obtained through α values.

Drt kiŐiydik, trenden sonra deniz kenarında bir kahvede oturmuŐtuk. KıŐtı, kar yerden kalkmamıŐtı. Sen aĐlamıŐtın. Mutluluktan, yalnızca mutluluktandı gz yaŐların. [Sulu gzlydn]_{A7} Nesli, ama itiraf etmeye utanıyordun deĐil mi? Ne kadar gizlemeye alıŐsan da **bunu/bu gereĐi**, baŐaramıyordun. Biz yasalar karŐısında evli sayılacak, ama gerekte evli iki insan gibi deĐil de

³⁶ Please note that these are not syntactic level of bracketing. Rather, they indicate the antecedents specified by the participants.

(evlilikler sıradanlaşıyordu çünkü, tekdüze ve sıkıcıydı; biz farklı olacaktık), aynı evi paylaşan iki öğrenci gibi yaşayacaktık.

We were four; we sat down at a cafe on the seaside after the ceremony. It was winter; the snow was still on the ground. You cried. Because of happiness, your tears were only because of happiness. [You were a crybaby]^{A7} Nesli, but you were ashamed to confess, weren't you? No matter how much you tried to hide **this/this fact**, you were not able to succeed. We were going to be considered married under the law, but in reality we were going to live not like a married couple (because marriages would become ordinary, they were monotonous and boring; we were going to be different) but like two students sharing the same apartment.

Bu vs. Bu iş:

As stated in 4.3.3.4, the most common subsets specified by the participants in both groups as the antecedents to the bare anaphor and the label anaphor in this paragraph are A5 and A6 (indicated in square brackets below). A6 is in an embedded clause and A5 involves an embedded clause. Compared with the previous paragraph, the grammatical structure of the potential antecedents is more complex in this paragraph. Thus, we believe, the scores obtained through α value are lower.

Taş kaynaklarına baktığınızda, doğal olarak çok taş malzemesi var; bazalt, kireçtaşı, tüfler, volkanik taşlar çok fazla var. Aşıklı'ya yerleşenler ilk geldiklerinde kerpiç mimariyi biliyor. Yapılarını önceleri kerpiçten yapıyorlar, ama sonra [[kerpiçle taşı birlikte kullanmayı]^{A6} öğreniyorlar]^{A5}. Mimarlık açısından çok önemli **bu/bu iş**, çünkü bir yapı malzemesini başka bir malzemeyle beraber kullanmayı, ilk defa burada görüyoruz. Evler kerpiçten, ama tapınağın temelinde taş var; çevre duvarı neredeyse tamamen taştan.

Upon analyzing the stone sources, there are naturally a lot of stone materials; there are a great number of basalt, limestone, tuff, volcanic rocks. Those settling in Aşıklı knows mud architecture when they first come. They first build their structures with mud, but later they [learn how to [use the mud together with stone]^{A6}]^{A5}. **This/This activity** is very important in terms of

architecture, because we see it here for the first time that a structure material is used with another material. The houses are made of mud, but there is stone in the foundation of a temple; the surrounding wall is almost totally made of stone.

Bu vs. Bu olay:

The most common subsets specified by the participants in both groups as the antecedents to the bare anaphor and the label anaphor in this paragraph are A7 and A8 (see Appendix 14). It is not possible to mention a complexity of structure due to level of embedding in this paragraph. However, one structural difficulty might stem from the fact that one of the main arguments, i.e., Mesut, of the verb 'can vermek' (Eng. 'to get killed') is not located within the same sentence with the verb. Therefore, the participants might have had difficulty on agreeing on the span of the antecedent in this paragraph. In other words, participants might not have preferred the clauses where the subject was dropped, selecting instead a larger span including the argument in the antecedent. As a result, this might have lowered the agreement scores for this paragraph compared to the previous two.

The antecedents underlined by the participants are indicated in brackets below:

Ahmet, bir gün gazeteyi bir açmıştı, [Mesut yok! [Pusuya düşürülmüş oracıkta can vermişti.]]_{A8}]_{A7} İki kişiymişler. Öbürünün adı Halim. Tanımıyordu. Hafif yaralı kurtulmuş o. **Bu/Bu olay** derinden sarsmıştı Ahmet'i. Evden ayrılmıştı. Okula gitmiyordu artık. Kimseyle görüşmüyordu. Köyüne gömmüşlerdi Mesut'u. 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.

Ahmet, one day opened the newspaper, [Mesut disappeared! [He was trapped and got killed there.]]_{A8}]_{A7} They were two people. The other's name is Halim. He didn't know. He got away slightly injured. **This/This event** afflicted Ahmet deeply. He left home. He wouldn't go to school any more. He wouldn't see anyone. They buried Mesut in his village. If he had

wanted, he would have gone to that village in the mountains immediately; he would have fulfilled his last duty. Duty! But he left Mesut in a village square while smiling. He did not want to shoulder a bloody coffin so that Mesut would smile all the time.

Another reason to decrease the agreement rates in the markings in this paragraph, compared to the previous two, might be the elaborative elements (as discussed in 3.3.2.) of the event denoted by the target anaphors. The subsets for this paragraph (see Appendix 14) show that the antecedents specified by the participants vary.

The variations in the span of the markings can be accounted for with SDRT (Asher & Vieu, 2005), as explained in 2.2.3. To have a better illustration of the constituents in the discourse structure of this paragraph, the sentences are given in a linear order below:

- a. Ahmet, bir gün gazeteyi bir açmıştı, Mesut yok!
- b. Pusuya düşürülmüş oracıkta can vermişti.
- c. İki kişiymişler.
- d. Öbürünün adı Halim.
- e. Tanımıyordu.
- f. Hafif yaralı kurtulmuş o.
- g. **Bu/Bu olay**, derinden sarsmıştı Ahmet'i.

(a) and (b) in this paragraph constitute a discourse segment, which introduces the event in which Mesut got killed and the way he got killed. The constituents (c-f) are elaborative elements giving the details of the event introduced in the discourse segment in focus. As discussed by Webber (1991), and as explained in detail in 2.2.2 and 3.3.2 as the right frontier constraint (RFC), the target anaphors attach to the discourse segment in focus. Following RFC, the constituent (g) involving the abstract anaphors 'bu' and 'bu olay' should attach to the discourse segment formed by (a) and (b). None of the participants violate RFC. All of the markings (see Appendix 14) specified by the participants in both groups indicate that either (a), or (b) or a combination of both is underlined by the participants.

- h. Evden ayrılmıştı.

- i. Okula gitmiyordu artık.
- j. Kimseyle görüşmüyordu.

The second discourse segment, starting with the clause (h), introduces a situation which is further elaborated by (i) and (j). This segment is coordinated with the previous segment in a cause and effect chain. The situation presented in this segment can be considered to have resulted from the event in the first discourse segment formed by (a) and (b).

- k. Köyüne gömmüşlerdi Mesut'u.
- l. İsteseydi, bir çırpıda o dağ köyüne gider, son görevini yapardı.
- m. Görev!
- n. Ama o, Mesut'u bir köy alanında gülümserken bırakmıştı.
- o. Mesut hep gülümsesin diye omuzlamak istememişti kanlı bir tabutu.

The third discourse segment, starting with the clause (k), introduces another event and the elaborative elements (l-o) support the result of this event with further details. The relation between the last discourse segment and the previous two segments can again be considered to be a cause and effect chain. In other words, the event in third discourse segment can be considered to be the result of the previous discourse segments. To put in a nutshell, coordination, the rhetorical relation defined in SDRT, in this paragraph establishes the relations among the event in the first discourse segment, the situation in the second one, and the event in the third one.

The subsets underlined by the participants in both groups (see Appendix 14) indicate that the subordination relation is more dominant in the current data. Considering RFC, the target abstract anaphors attach to a constituent within the same discourse segment, not across the distinct discourse segments. Therefore, to account for the current data, it is more meaningful to mention the implications of subordination, discussed by Asher and Vieu (2005), on the current data.

Subordination, the rhetorical relation that SDRT accounts for, serves to link the elaborative clauses in the paragraph (c-f) to the prior context established by the discourse segment formed by (a) and (b). Subordinate clauses bound to the discourse segment in focus instruct the reader on how to update his or her discourse model with the information in the current sentence alone. In other words, it can be considered that the discourse model was updated as the participants read the elaborative clauses

(c-f). As they updated their discourse models, probably the span of the antecedents underlined by the participants varied.

All in all, considering from the framework of the SDRT, while processing the discourse presented in this text, the readers might have identified (a) and (b) as the discourse segment denoting the event that the target anaphors should attach to. However, as the participants read the following clauses (c-f), they may need to update the discourse structure constituted by the main discourse segment and supported by the elaborative elements. Although the clauses (c-f) do not constitute a new discourse segment, they are the further details of the event in focus i.e., the event in which Mesut was captured and got killed. Some of the participants added (c)³⁷ to the subset A7 or A8, some of them added (c) and (d)³⁸, some of them added (c)-(e)³⁹, some of them added all of the elaborative elements (c)-(f)⁴⁰. Participants might have hesitated to decide how much of these details to involve while specifying the extension of their markings. As a consequence, this might have decreased the scores obtained through α values in this paragraph compared to the previous two paragraphs.

Bu vs. Bu durum:

The number of the subsets is the highest in this paragraph. Therefore, the scores obtained through α value are the lowest ones for this paragraph. This is possibly because many clauses are embedded in each other. Thus, rather than the lexical meaning of the head noun in the label or any kind of complexation because of abstractness, the complexity of discourse structure both in terms of syntactic and semantic elements make the resolution of the target anaphors difficult in this paragraph.

[[[[[[[Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş]_{A8} olmayı]_{A6} hiçbir zaman içine sindirememiş]_{A4} olan]_{A3} ve bu konuda duyduğu ezikliği saklayamayan]_{A2} Turgay'a]_{A1} bu konuda sürekli takılırlar]_A (Aslında şimdi ben de onu yapmış

³⁷ (c) İki kişiymişler.

³⁸ (d) Öbürünün adı Halim.

³⁹ (e) Tanımiyordu.

⁴⁰ (f) Hafif yaralı kurtulmuş o.

oldum.) Çünkü **bu/bu durum**, onun daima kanayan açık yarasıdır, camiada herkesçe bilinir ve kabul edersiniz ki, insan ilişkilerinde en çabuk öğrendiğimiz şeylerden biri, başkalarının yarasını kullanmaktır.

_A[They make fun of _{A1}[Turgay, _{A2}[after being an activist communist who served years in jail, who has never _{A3}[been able to _{A4}[accept _{A6}[that _{A8}[he turned to managing a bar from being a member of a secret organization and who cannot hide the embarrassment he has for this issue.]]]]]]] (In fact, I have done the same thing now.) Because **this/this situation** is his unhealed wound bleeding all the time, it is known by everybody and, as you might accept, one of the things we learn the earliest is to abuse others' wounds.

The possible strategies while specifying the antecedent of the target anaphors applied by the participants can be explained through SDRT and RFC for this paragraph, as well. The constituents of this paragraph can be written in the following linear order to have a better illustration of rhetorical relations in SDRT and the right frontier operations.

- a. Yıllarca hapis yatmış militan bir komünist olarak,
- b. gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı
- c. hiçbir zaman içine sindirememiş
- d. ve bu konuda duyduğu ezikliği saklayamayan
- e. Turgay'a bu konuda sürekli takılırlar
- f. (Aslında şimdi ben de onu yapmış oldum.)
- g. Çünkü **bu/bu durum**, onun daima kanayan açık yarasıdır
- h. camiada herkesçe bilinir
- i. ve kabul edersiniz ki,
- j. insan ilişkilerinde en çabuk öğrendiğimiz şeylerden biri, başkalarının yarasını kullanmaktır.

The clauses (a) and (b) introducing that Turgay was an activist communist who served years in jail and he turned to managing a bar from being a member of a secret organization constitute the first discourse segment in the paragraph. The clauses (c-f) further elaborate on the situation presented in this discourse segment. The target anaphors presented in (g) attach to the discourse segment in focus, i.e. the

discourse segment formed by (a) and (b). The markings of the participants do not violate this constraint, i.e. the most common subsets are A6⁴¹ and A11⁴² (see Appendix 14). These sets involve the discourse segment that should be in focus, which indicates that no one violates RFC.

Rather than being in a sentence-by-sentence successive fashion, different from the previous paragraph, the elaborative clauses together with the clauses constituting the main discourse segment are embedded in each other. One of the clauses is even connected to the discourse segment with the connector've' (Eng. 'and') within the boundaries of the same sentence. This complexity in the syntactic structure of the discourse segment and the elaborative clauses make the specification of the boundaries of the antecedent even a more difficult task compared to the previous paragraphs. Therefore, it is not readily possible to account for the lowest agreement scores in this paragraph merely through the abstractness hierarchy put forward by Consten and Knees (2005), or the lexical meaning of the head noun in the label anaphors.

Going back to the paragraph where the bare anaphor 'bu' and the label anaphor 'bu gerçek' were compared, recall that the best agreement scores were obtained. It was argued that the simplicity of the syntactic structure of the sentence in which the antecedent was presented fostered agreement scores to be the highest. This can be proved within the framework of SDRT and RFC, as well.

- (a) Sulu gözlüydün Nesli,
- (b) ama itiraf etmeye utanıyordun değil mi?
- (c) Ne kadar gizlemeye çalışsan da **bunu/bu gerçeği**, başaramıyordun.

Being the most common antecedent specified by the participants, (a) is alone a discourse segment introducing a state. The clause (b) following the discourse segment in focus is an elaborative element of this state. Compatible with the RFC, the anaphor taking place in the next clause (c) is attached to the discourse segment in focus, i.e. (a). This is also justified by the Passonneau (2005) analysis of this

⁴¹ A6: Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı (Eng. Being an activist communist, turning to managing a bar from being a member of a secret organisation)

⁴² A11: gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı (Eng. turning to managing a bar from being a member of a secret organisation)

paragraph (see Appendix 14). Noone from both groups violate RFC. Besides this, perfect agreement scores through α value indicates that the participants are quite precise in specifying the boundaries of the antecedent to the target anaphors in this paragraph.

However, it is important to note that there seems to be a different phenomenon in the paragraph where 'bu' and 'bu iş' are compared. This distinction was already discussed in the results of the pilot study and Experiment 1.

...Yapılarını önceleri kerpiçten yapıyorlar, ama sonra [[kerpiçle taşı birlikte kullanmayı]_{A6} öğreniyorlar]_{A5}. Mimarlık açısından çok önemli **bu/bu iş**,...

The most common antecedents, as indicated in brackets, are the subsets A6 and A5. In Group 1, A5 is 51,72% and A6 is 31,03% of the total markings. In Group 2, A5 is 19,44% and A6 is 50% of the total markings. As explained in 4.3.3.4, the findings for this paragraph show that participants have a tendency to associate the label anaphor 'bu iş' with the activity of *using* the mud together with stone, whereas they have a tendency to associate the bare anaphor 'bu' with *learning* how to use the mud together with stone. This distinction between 'bu' and 'bu iş' was observed in Experiment 1 as well as the pilot study. This distinction seems to be originating from the lexical meaning of the head noun 'iş' directing readers to look for an activity in the text rather than a syntactic complexity or simplicity.

All in all, Experiment 2 reveals that, as well as the lexical meaning of the head nouns in the label anaphors having an effect on specifying the boundaries of the antecedents, structural factors also play an important role in specifying the boundaries of the antecedents. The more the structural difficulty of the potential antecedents increases in the paragraphs, the lower the α value gets. This is because complexity in syntactic structures makes it difficult to specify the boundaries of the antecedents.

4.4 General Discussion

In both of the experiments the main aim was to observe the differences in the resolution process of the bare anaphor 'bu' (Eng. 'this') and the label anaphors 'bu durum/iş/olay/gerçek' (Eng. 'this situation/activity/event/fact'). Four different

methods were used to analyze the data i.e., consistency analysis on the start and end words, Krippendorff's Alpha, visual illustration of the data through bar graphs and Passonneau's (2004) classification of subsumption, intersection and disjunct relations.

It was observed that the findings in Experiment 2 did not completely match the findings in Experiment 1. In the analysis conducted through α values, both the word interval and word boundary measurements in both experiment revealed that label anaphors had a tendency to produce better agreement scores compared to the bare anaphor 'bu'. However, differences were found regarding the labels used in these experiments. For example, in the text where the label anaphor 'bu olay' (Eng. 'this event) and the bare anaphor 'bu' referring back to an event were used in Experiment 1, α values produced perfect agreement scores both in word boundary and word interval measurements. In Experiment 2, in the text where these anaphors were used, α values produced only moderate agreement scores. Besides this, the lowest agreement scores were found in the text where the bare anaphor and 'bu iş' (Eng. 'this activity') were used in Experiment 1. In the text where these anaphors were used in Experiment 2, α values produced moderate and substantial agreement scores both in word boundary and word interval measurements. Depending on the lexical meaning of the head words in the labels, Consten and Knees' (2005) abstractness hierarchy accounted for the differences obtained through α values in Experiment 1. However, the differences found in Experiment 2 through α values were not possible to be explained with the same theory. This was because, in Experiment 2, the structural complexities in the paragraphs seemed to be the leading factor to obtain α values different than those in Experiment 1. Therefore, in Experiment 2, an account of anaphora resolution was given through SDRT (Asher & Vieu, 2005) and RFC (Webber, 1991).

In the visual illustration of the data for Experiment 1, it was possible to see the alternative interpretations of the participants' while marking the antecedent to the bare anaphor by means of the visually illustrated data. Thus, it was concluded that label anaphors had a tendency to eliminate the competing candidates to be antecedents and minimize the occurrences of the alternative interpretations.

However, such a difference was not observed as a result of the same analysis conducted for Experiment 2.

All in all, label anaphors can be considered to have a disambiguating role in that the lexical meaning of the head nouns helps specifying the boundaries of the antecedents in some cases. However, the lexical meaning of the labels may fall short to resolve the abstract anaphors in cases where the potential antecedents are grammatically complex. Therefore, the grammatical structure cannot be ignored in the resolution process of the anaphors. Such factors like level of embedding can be the leading factor in the accessibility of the antecedents to the abstract anaphors in discourse.

To wrap up, as highlighted by Poesio and Traum (1997) 'language is (joint) action' and the complexity of discourse in practice has resulted in theories of the common ground. However, these theories differ in almost all essential details (p. 2). It does not seem to be possible to account for the complexity of discourse, thus, abstract anaphora resolution from a single point of view. The experiments conducted for this study revealed that not everybody agreed on what the 'information' the anaphors or the antecedents hold. However, depending on the results obtained through the analysis of the data from both experiments, it is at least possible to say that there is a 'common ground' of information to denote the antecedent of the target anaphors. While bare anaphors produced multiple common grounds as antecedents in Experiment 1, therefore, it was concluded that the label anaphors had a tendency to eliminate the competing candidates to be the antecedents. This conclusion was not justified in Experiment 2. It revealed that the common ground of information also depended on the centers that were focus of attention in the discourse model.

CHAPTER 5

SUMMARY, CONCLUDING REMARKS

AND FUTURE WORK

This thesis presented an experimental study conducted with human subjects to gain an insight into the resolution of abstract object anaphors in Turkish discourse. First of all, basic notions (e.g. what anaphora resolution means, anaphors referring to abstract *vs.* concrete objects, types of abstract entity anaphors) providing the background of the study were explained in detail in Chapter 1. The focus was on two different types of abstract object anaphors: (1) the bare anaphor 'bu' (Eng. 'this') with situation reference as discussed by Fraurud (1992, 1992a), and (2) the label anaphors 'bu durum/iş /gerçek/olay' (Eng. 'this situation/activity/fact/event') as discussed by Francis (1994). The main research questions were determined as:

- Is there any difference between the resolution of bare abstract object anaphor 'bu' and label anaphors 'bu+durum/gerçek/olay/iş'?
- Do label anaphors, namely 'bu+durum/gerçek/olay/iş', have a disambiguating role in abstract anaphora resolution in Turkish written discourse?

In Chapter 2, some of the pioneering corpus-based studies conducted in an attempt to understand the phenomenon of abstract anaphora resolution were reviewed in Chapter 2. Among these studies, Botley (2006) was very influential in shaping our decision to determine the types of abstract object anaphors to focus on. Also, the challenges in the specifying the boundaries of the antecedents to abstract anaphors mentioned in this study were one of main discussions in the current study, too. Another important study reviewed in Chapter 2 was Poesio and Artstein (2008). This study was a model in that it deals with the annotation of ambiguous anaphoric expressions, and searches for a reliability measurement of agreement in ambiguous texts. Following their methods, Krippendorff's Alpha was applied in this study, too. Besides these, Consten and Knees' (2005) ideas regarding anaphoric complexation processes and Webber's (1991) elaboration on right frontier operations were summarized. In addition, some of the pioneering studies emphasizing the ambiguous nature of abstract object anaphors, e.g., Botley (2006), Artstein and Poesio (2008), Recasens (2008) and Poesio et al. (2006) were discussed in Chapter 2. These studies were highly influential in the interpretation of the findings in the current study. Finally, the methods i.e., Krippendorff's Alpha and Passonneau (2004) that were used to analyze the data in the current study were explained in detail.

Afterwards, in Chapter 3, various occurrences of abstract object anaphors in Turkish discourse were exemplified with the examples taken from the TDB. The theory of abstractness hierarchy by Consten and Knees' (2005) and right frontier operations by Webber (1991) was adapted to Turkish discourse. Finally, the studies having been conducted on anaphora resolution in Turkish were reviewed.

Finally, in Chapter 4, the pilot study and two off-line experiments conducted to address the main research questions were presented in detail. The participants were given with four texts with the bare anaphor as well as the manipulated texts containing the label anaphors. The qualitative evaluation done as a result of the pilot study indicated that the participants associated the bare anaphor 'bu' (Eng. 'this') and the label anaphor 'bu iş' (Eng. 'this activity') with different entities. The findings in Experiment 1 revealed that the bare abstract object anaphors brought burden on the anaphora resolution process compared to the label anaphors. This is because bare anaphors may not always carry the load of denoting the precise boundaries of their

antecedents. It was observed that label anaphors, on the other hand, could handle this burden in some cases. This is because lexical meaning of their head nouns makes them better equipped in specifying the boundaries of the antecedents.

To understand whether the data set obtained from Experiment 1 was generalizable enough to allow disambiguating role of label anaphors, Experiment 2 was conducted with a new set of texts involving bare anaphor 'bu' and the same label anaphors in Experiment 1. The findings through α value in Experiment 1 matched the abstractness hierarchy suggested by Consten and Knees (2005). However, the scores did not verify this hierarchy in Experiment 2. Rather than the lexical meaning of the head nouns and the ontological categories denoted by these head nouns used in both experiments, the scores in Experiment 2 were accounted for by the structural complexities of the texts used in Experiment 2. Thus, Experiment 2 has revealed, as also discussed by Francis (1994, p.99), that labels do not necessarily refer to a clearly delimited or identifiable stretch of discourse.

All in all, the focus in this study was to gain an insight into the resolution process of bare anaphors and label anaphors in Turkish discourse through off-line methods. The main challenges in the study were caused by (1) using flexible annotation guidelines, (2) ambiguous nature of abstract object anaphors and (3) structural difficulties in the texts in Experiment 2. Ambiguous nature of the abstract object anaphors makes it already difficult for the coders to agree on the span of the antecedents. This challenge was doubled in our preference not to use stricter guidelines. Also, the structural complexities of the texts in Experiment 2 seemed to increase the difficulty for participants to access the antecedents and specify the span of them.

Upon the analysis of data in Experiment 1 and 2 with the help of various methods, the apparent factors having a role in the way humans compute the antecedent of the bare vs. label anaphors are (a) the lexical meaning of the head nouns, (b) the consistency between the ontological category of the anaphor and its antecedent, and (c) the structural complexities like level of embedding. As a result of this study, (a) and (b) are the factors standing out as a difference between the bare and label anaphors in the resolution process. However, when the anaphors are presented in a structurally complex text, especially when the antecedent is presented

in a text involving several levels of embedding, the lexical meaning of the head nouns in the label anaphors falls short to ease resolution process. At this point, it becomes essential to account for the complexity of discourse with the help of various theories like SDRT (Asher and Vieu, 2005) and RFC (Webber, 1991).

For English discourse, the label anaphors and the bare anaphors with reference to abstract entities have recently been comparatively investigated to gain an insight into the distinction during anaphora resolution process by Consten and Knees (2005), and Consten, Knees and Friesel (2007). However, it has not been the focus of attention in Turkish so far. It is important to note that Consten and Knees (2005), and Consten, Knees and Friesel (2007) do not give an empirical account for the distinction between these two types of anaphors.

Considering the contributions this study could provide in anaphora studies in Turkish, we believe, it was a reasonable assumption to separate these two kinds of anaphora. Our assumption was verified by the distinctions observed between the groups in the steps we followed.

Another contribution of the current work can be the avenues it opens for future research. We realize that structural factors can be computationally too expensive in anaphora resolution. It would be interesting to investigate, for instance, how level of embedding influence the abstract anaphora resolution. A new experiment would have to be designed to investigate this. For example, the resolution of the same label anaphor (e.g., 'bu gerçek' (Eng. 'this fact')) in structurally different texts in terms of their complexity can be interesting to study.

Finally, although we are convinced that there are some differences in the resolution process of the anaphors selected for this study, to reach firmer conclusions, it is necessary to measure online processing of abstract object anaphors. We consider focusing on anaphora processing through online procedures like eyetracking. We believe this can give better insight into where exactly people look at while reading a text, how long they fixate on some certain words, which words people skip and whether people have regressive (backwards) eye movements while associating the abstract anaphors with their antecedents.

Another future work for this study would be to discuss the results within the framework of Centering Theory (CT) (Grosz et al., 1995; Walker et al., 1998). CT

accounts for the concrete entities in referential relations. An attempt to gain insight into the current data involving the abstract entity referents can be made through centering.

REFERENCES

- Artstein, R., & Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4), 555-596.
- Asher, N. (1993). *Reference to abstract objects in discourse*. Dordrecht: Kluwer Academic Publishers.
- Asher, N., & Vieu, L. (2005). *Subordinating and coordinating discourse relations*. *Lingua*, 115(4):591–610.
- Benett, E. M., Alpert, R., & Goldstein, A. C. (1954). Communications through limited questioning. *Public Opinion Quarterly*, 18(3), 303-308.
- Botley, S. P. (2006). Indirect anaphora: Testing the limits of corpus-based linguistics. *International Journal of Corpus Linguistics* 11(1), 73–112.
- Botley, S. & McEnery, A. M. (2001). Proximal and distal demonstratives: a corpus based study. *Journal of English Linguistics*, 29(3), 214-233.
- Byron, D. (2002). Resolving pronominal reference to abstract entities. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 80--87.
- Cohen, J. (1960). A Coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Cokal-Karadas, D. (2005). *A contrastive analysis of the pronominal usages of this and that in academic written discourse*. MA Thesis, Department of English Language Teaching, Middle East Technical University, Ankara, Turkey.
- Consten, M. & Knees, M. (2005). Complex anaphors — ontology and resolution. In

- P. Dekker, editor, *Proceedings of the 15th Amsterdam Colloquium*, Amsterdam: University.
- Consten, M., Knees, M. & Schwarz-Friesel, M. (2007). The Function of Complex Anaphors in Texts. In: Schwarz-Friesel, M., Consten, M. & Knees, M. (eds.), *Anaphors in Texts*
- Dipper, S. & Zinsmeister, H. (2010). *Towards a standard for annotating abstract anaphora*. In Proceedings of the LREC Workshop on Language Resource and Language Technology Standards – state of the art, emerging needs, and future developments, pp. 54-59. Malta.
- Dipper, S. & Zinsmeister, H. (2011). Annotating abstract anaphora. *Language Resources & Evaluation*, Online First, 3. September 2011.
- Dipper, S. & Zinsmeister, H. (2009). Annotating discourse anaphora. In *Proceedings of the Third Linguistic Annotation Workshop, ACL-IJCNLP 2009*, pp. 166-169.
- Eckert, M. & Strube, M. (2000). Dialogue acts, synchronizing units, and anaphora resolution. *Journal of Semantics* 17(1), 51-89.
- Eckert, M. & Strube, M. (1999). Resolving discourse deictic anaphora in dialogues. *EACL '99 Proceedings of the ninth conference on European chapter of the Association for Computational Linguistics*.
- Eckert, M. & Strube, M. (2000). *Dialogue acts, synchronizing units and anaphora resolution*. *Journal of Semantics*, 17(1), pp. 51--89.
- Francis, G. (1994). *Advances in written text analysis*. Malcolm Coulthard (ed.). Labelling discourse: an aspect of nominal-group lexical cohesion. pp.83-101
- Fraurud, K. (1992). *Processing Noun Phrases in Natural Discourse*. PhD Thesis (by Publications), Department of Linguistics, Stockholm University.
- Fraurud, K. (1992a). *Situation Reference: What Does 'It' Refer To?*. GAP Working Paper No 24, Fachbereich Informatik, Universität Hamburg.
- Grosz, J. B. & Sidner, C. L. (1986) *Attention, intentions and the structure of discourse*. *Computational Linguistics*. 12(3): 175-205.
- Grosz, J. B. Joshi, A. K. & Weinstein, S. (1995). *Centering: a framework for modeling the local coherence*. *Computational Linguistics*. 21 (2): 203-225
- Gundel, J. K., Hedberg, N. & Zacharski, R. (2005) Pronouns without NP antecedents: how do we know when a pronoun is referential? In A. Branco, T. McEnery & R. Mitkov (eds.), *Anaphora Processing: Linguistic, Cognitive and Computational Modelling*. Amsterdam: John Benjamins, pp. 351-365.

- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hirst, G. (1981). *Anaphora in natural language understanding: A survey*. Berlin & New York: Springer-Verlag.
- Joshi, A., K. & Kuhn, S. (1979). Centered logic: role of some aspects of discourse structure -centering. *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, Vancouver, 385-7.
- Kılıçaslan, Y., Güner, E.S. & Yıldırım, S. (2009). *Learning-based pronoun resolution for Turkish with a comparative evaluation*. *Computer Speech and Language*, 23(3), pp. 311-331.
- Krippendorff, K. (1995). *On the reliability of unitizing continuous data*. *Sociological Methodology*, 25, 47-76.
- Krippendorff, K. (2004a). *Content analysis : an introduction to its methodology* (2nd ed.). Thousand Oaks, Calif.: Sage.
- Krippendorff, K. (2004b). *Reliability in content analysis: some common misconceptions and recommendations*. *Human Communication Research*, 30(3), 411-433.
- Kroeger, P. R., (2005). *Analyzing grammar: an introduction*. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, Sao Paulo & Delhi: CUP.
- Küçük, D. and Turhan-Yöndem, M. (2007). A knowledge-poor pronoun resolution system for Turkish. In *Proceedings of 22nd International Symposium on Computer and Information Sciences (ISCIS 2007)*, pp. 1--6.
- Küçük, D. and Yazıcı, A. (2008). Identification of Coreferential Chains in Video Texts for Semantic Annotation of News Videos. In *Proceedings of 23rd International Symposium on Computer and Information Sciences (ISCIS 2008)*, pp.1--6.
- Küçük, D. and Yazıcı, A. (2009). Employing Named Entities for Semantic Retrieval of News Videos in Turkish. In *Proceedings of the 24th International Symposium on Computer and Information Sciences (ISCIS 2009)*, pp. 153--158.
- Lakoff, R. (1974). Remarks on ‘this’ and ‘that’. In *Proceedings of the Chicago Linguistics Society 10*, pages 345–356.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *International Biometric Society*, 33(1), 159-174.
- Lee, S. and Song, J. (2010). Annotating Korean demonstratives. In *Proceedings of*

the Fourth Linguistics Annotation Workshop. pp. 162--165.

- Levinson, S. C. (1987). Pragmatics and the grammar of anaphora: A partial pragmatic reduction of Binding and Control phenomena. *Journal of Linguistics*, 23, 379- 434
- Lewis, G. (1967). *Turkish Grammar*. London: Oxford University Press.
- Lyons, J. (1977). *Semantics. Volume 1 and 2*. Cambridge: Cambridge University Press.
- Mann, W.C. and Thompson, S.A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8, pp. 244--277.
- Mungan, M. (2004). *Yüksek topuklar*. İstanbul: Metis.
- Müller, C. (2007). Resolving it, this and that an unrestricted multi-party dialog. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*
- Navaretta, C. and Olsen, S. (2008). Annotating abstract pronominal anaphora in the DAD project. In *Proceedings of LREC 2008*, pp. 2046--2052.
- Navarretta, N. (2007). A contrastive analysis of abstract anaphora in Danish, English and Italian. In A. Branco, T. McEnery, R. Mitkov and F. Silva (eds.) *Proceedings of DAARC 2007 - 6th Discourse Anaphora and Anaphora Resolution Colloquium*, pages 103-109,.
- Navarretta, C. (2001). Identifying situation reference in Danish. *Extended abstracts of the 13th Nordic Conference on Computational Linguistics*, NoDaLiDa '01 Uppsala University, May 21-22.
- Passonneau, R. J. (2004). *Computing reliability for coreference annotation*. In Proceedings of LREC, volume 4, pp. 1503–1506, Lisbon.
- Poesio, M. and Artstein, R. (2008). Anaphoric annotation in the ARRAU corpus. In the Proceedings of the LREC Workshop on Language Resource and Language Technology Standards (LREC 2008), pp. 1170—1174.
- Poesio, M. & di Eugenio, B. (2001). Discourse Structure and Anaphoric Accessibility. *Proc. of the ESSLLI Workshop on Discourse Structure and Information Structure*, Helsinki, August 2001
- Poesio, M., Reyle, U. & Stevenson, R. (2006). "Justified Sloppiness in Anaphoric Reference," In H. Bunt and R. Muskens (eds.), *Computing Meaning 3*, Kluwer
- Poesio, M. (2004). The MATE/GNOME Proposals for Anaphoric Annotation, Revisited. In Proceedings of the 5th SIGDIAL Workshop, pp. 154–162, Boston.

- Poesio, M. & Traum, D. (1997). Conversational Actions and Discourse Situations. *Computational Intelligence*, v. 13, n.3.
- Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of Pragmatics*, 12.
- Prasad, R., Miltsakaki, E., Dinesh, N., Lee, A., Joshi, A. & Webber, B. (2006). *The Penn Discourse TreeBank 1.0. Annotation Manual. IRCS Technical Report IRCS-06-01*. Institute for Research in Cognitive Science, University of Pennsylvania. March 2006.
- Recasens, M. (2008). *Discourse deixis and coreference: evidence from AnCora*. Proceedings of the Second Workshop on Anaphora Resolution (WAR II). Editor: Christer Johansson. NEALT Proceedings Series, Vol. 2 , 73-82
- Scott, W. A. (1955). Reliability of content analysis: the case of nominal scale coding. *Public Opinion Quarterly*, 19(3), 321-325.
- Say, B., Zeyrek, D. Oflazer, K. & Özge, U. (2002). Development of a Corpus and a Treebank for Present-day Written Turkish. In *Proceedings of the Eleventh International Conference of Turkish Linguistics*.
- Sevdik-Çallı, A. (2012). Demonstrative Anaphora in Turkish: A Corpus Based Analysis. In *Proceedings of the of the First Workshop on Language Recourses and Technologies for Turkic Languages*.
- Tın, Erkan and Varol-Akman. 1998. "Situated Processing of Pronominal Anaphora". Bilkent University, Ankara.
- Turan, U.D. (1997). Metin işaret adılları: *Bu, şu* ve metin yapısı (Discourse Deictic pronouns *this, that*, and the structure of discourse). In D. Zeyrek & Ş. Ruhi (Eds.), *XI. Dilbilim Kurultayı Bildiriler (The Proceedings of the XI. Linguistics Conference)*. Ankara: Middle East Technical University, pp. 201--212.
- Tüfekçi, P. & Kılıçaslan, Y. (2005). A computational model for resolving pronominal anaphora in Turkish using Hobbs' naïve algorithm. In *Proceedings of World Academy of Science, Engineering and Technology (WEC)*, pp. 13--17.
- Tüfekçi, P., Küçük, D., Turhan-Yöndem, M. & Kılıçaslan, Y. (2007). Comparison of a syntax-based and a knowledge-poor pronoun resolution systems for Turkish. In *Proceedings of the International Symposium on Computer and Information Sciences (ISCIS 2007)*. Middle East Technical University, p.53.
- Vieira, R., Salmon-Alt, S., Gasperin, C., Schang, E. and Othero, G. (2005). Coreference and anaphoric relations of demonstrative noun phrases in

- multilingual corpus. In A. Branco, T. McEnery, and R. Mitkov (Eds), *Anaphora Processing: Linguistic, Cognitive and Computational Modeling*. Amsterdam: John Benjamins, pp. 385--401.
- Yalçinkaya, Ş., İ. (2011). *An Inter-annotator Agreement Measurement Methodology for the Turkish Discourse Bank (TDB)*. MA Thesis, Department of Cognitive Science, Middle East Technical University, Ankara, Turkey
- Yıldırım, S., Kılıçaslan, Y. and Yıldız, T. (2009). Pronoun resolution in Turkish using decision tree and rule-based learning algorithms. In Z. Vetulani and H. Uszkoreit (Eds.), *Human Language Technology: Challenges of the Information Society, Proceedings of the Third Language and Technology Conference, LTC 2007*. Berlin-Heidelberg: Springer-Verlag, pp. 270--278.
- Yıldırım, S., Kılıçaslan, Y. and Yıldız, T. (2007). A decision tree and rule-based learning model for anaphora resolution in Turkish. In *Proceedings of the 3rd Language and Technology Conference (LTC'07)*. Poznań, Poland, pp. 89--92.
- Yıldırım, S., Kılıçaslan, Y., Aykaç, R.E. (2004). A computational model for anaphora resolution in Turkish via centering theory: an initial approach. In *Proceedings of International Conference on Computational Intelligence*, pp. 124--128.
- Yuksel, O. & Bozsahin, C. (2002). Contextually appropriate reference generation. *Natural Language Engineering*, 8(1), pp. 69--89.
- Walker, M., A., Joshi, A., K. & Prince, E. F. (1998). *Centering theory in discourse*. New York: OUP.
- Webber, B., L. (1988a). Discourse deixis: reference to discourse segments. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, pp. 113--122.
- Webber, B. L. (1988b). *Discourse deixis and discourse processing*. Technical Report MS-CIS-86-74, Department of Computer and Information Science. University of Pennsylvania.
- Webber, B., L. (1991). Structure and ostention in the interpretation of discourse deixis. *Language and Cognition Processes*, (6)2, pp. 107-135.
- Zeyrek, D., Demirsahin, I. n., Sevdik-Çallı, A., Ögel-Balaban, H., Yalçinkaya, I., & Turan, Ü. D. (2010). *The Annotation scheme of the turkish discourse bank and an evaluation of inconsistent annotations*. Paper presented at the Fourth Linguistic Annotation Workshop (LAW IV).

APPENDIX 1

EXPERIMENT 1 TEST MATERIAL FOR GROUP 1

ÖN BİLGİ FORMU

Yaş: Cinsiyet: Bölüm
Sınıf: 5-15 yaşları arasında yaşadığınız şehir(ler):
Dil eğitimi düzeyi: İngilizce Yeterlilik Sınavı'ndan aldığınız puan:

DENEY

Aşağıdaki 4 paragrafta "bu/bunu" ifadesinin gönderimde bulunduğu ifadenin altını çiziniz.

1. Kadınlar her şeyi çok daha kolay kabullenir, değişimlere daha çabuk rıza gösterirler. Nazlı da, komünist günlerini geride bırakıp hiçbir şeyin artık eskisi gibi olmadığını anladığı anda, yeni hayatlarına daha çabuk uyum göstermiş, kocasını da kendi zamanı içinde **bunu** kabulleneceği noktaya gelmesini sessizce beklemeye başlamıştı. Arada bir hapisanedeki eski dostlara, tutuklu ailelerine, derneklere para yardımı yaparak sol vicdanlarını yatıştırmak kalmıştı onlara (Mungan, 2001).

2. Şimdi Kanada'larda sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek doğrusu sevindirmişti beni. Elbet bir gün Kanada'dan dönecekti! **Bunu** açıklarken onun o kırmızıya dönen yüzünü seyretmek hoş olacaktı doğrusu (Mungan, 2001).

3. İlla ki kadın bedeninin gizemi olmalıydı Halil'in gözlerinin önünde. Kadın bedeni çizme tutkusu yaşama amacı haline gelmişti. Çareler düşünüyor, bulamıyordu, çünkü içinde bulunduğu ortamda **bunu** başarabilmesi neredeyse olanaksızdı. Artık bir şey

çizemez olmuştu. Altı ay önce bitirdiği bir resmi uzun süre dayanması ve renklerini koruması için verniklediği bir gece ansızın bir tekme savurarak üst kata çıktı. Hiçbir tablosu değerli değildi onun için (TSB).

4. Kemal'i eliyle yana itip yukarı koştı Ahmet Reşat. Kemal kapıda beti benzi atmış Hakkı Efendi'yi görünce paldır küldür inmeye başladı merdivenleri. Hakkı Efendi, Kemal'e titreyerek az önce yaşadıklarını naklediyordu ki, Ahmet Reşat giyinmiş olarak geri geldi. İşgalciler, Ziya Paşaların konağına el koymuşlardı. Hakkı Efendi'yle Ahmet Reşat alelacele çıkıp Paşa'nın konağına gittiler. **Bu**, bir yıkımdı Ziya Paşa için. Büyük bir hakaretti devlet nazırına bu yapılanlar (Kulin, 2007).

APPENDIX 2

EXPERIMENT 1 TEST MATERIAL FOR GROUP 2

ÖN BİLGİ FORMU

Yaş: Cinsiyet: Bölüm
Sınıf: 5-15 yaşları arasında yaşadığınız şehir(ler):
Dil eğitimi düzeyi: İngilizce Yeterlilik Sınavı'ndan aldığınız puan:

DENEY

Aşağıdaki 4 paragrafta “bu gerçeği/durumu/işi” ifadesinin gönderimde bulunduğu ifadenin altını çiziniz.

1. Kadınlar her şeyi çok daha kolay kabulle nir, değişimlere daha çabuk rıza gösterirler. Nazlı da, komünist günlerini geride bırakıp hiçbir şeyin artık eskisi gibi olmadığını anladığı anda, yeni hayatlarına daha çabuk uyum göstermiş, kocasını da kendi zamanı içinde **bu gerçeği** kabulleneceği noktaya gelmesini sessizce beklemeye başlamıştı. Arada bir hapisanedeki eski dostlara, tutuklu ailelerine, derneklere para yardımı yaparak sol vicdanlarını yatıştırmak kalmıştı onlara.

2. Şimdi Kanada'larda sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek doğrusu sevindirmişti beni. Elbet bir gün Kanada'dan dönecekti! **Bu durumu** açıklarken onun o kırmızıya dönen yüzünü seyretmek hoş olacaktı doğrusu.

3. İlla ki kadın bedeninin gizemi olmalıydı Halil'in gözlerinin önünde. Kadın bedeni çizme tutkusu yaşama amacı haline gelmişti. Çareler düşünüyor, bulamıyordu, çünkü içinde bulunduğu ortamda **bu işi** başarabilmesi neredeyse olanaksızdı. Artık bir şey

çizemez olmuştu. Altı ay 139nce bitirdiği bir resmi uzun 139nce dayanması ve renklerini koruması için verniklediği bir gece ansızın bir tekme savurarak üst kata çıktı. Hiçbir tablosu değerli değildi onun için.

4. Kemal'i eliyle yana itip yukarı koştı Ahmet Reşat. Kemal kapıda beti benzi atmış Hakkı Efendi'yi görünce paldır küldür inmeye başladı merdivenleri. Hakkı Efendi, Kemal'e titreyerek az 139nce yaşadıklarını naklediyordu ki, Ahmet Reşat giyinmiş olarak geri geldi. İşgalciler, Ziya Paşaların konağına el koymuşlardı. Hakkı Efendi'yle Ahmet Reşat alelacele çıkıp Paşa'nın konağına gittiler. **Bu olay**, bir yıkımdı Ziya Paşa için. Büyük bir hakaretti devlet nazırına bu yapılanlar.

APPENDIX 3

ANNOTATION GUIDELINE FOR GROUP 1

1.0 Giriş

Bu deneyin temel amacı, Türkçe'de artgönderim çözümleme (İng. Anaphora resolution) süreçlerini incelemektir. Bu deneyde, size verilen kısa paragraflarda altı çizili söz ya da sözcüklerin göndermelerini (kastedilen nesne ya da kavramı) bulmanız beklenmektedir.

1.1 Söylem Bağlıları

Aralarında belli bağlantılar olan tümcelerin biraraya gelmesiyle oluşan dilsel birime *söylem* (İng. discourse) denir. Örneğin (1a)'da ilk tümce ikinci tümcenin nedeni olduğundan, (1a) neden-sonuç içeren bir söylemdir. Buna karşın, (1b)'de ilk tümcenin anlamıyla (Ahmet'in susaması) ikinci tümcenin anlamı (global ekonomik krizin tüm dünya ülkelerini etkilemesi) arasında herhangi bir ilgi kurulamamaktadır.⁴³

- (1a) Ahmet çok susamıştı. Kalkıp bir bardak su içti.
(1b) # Ahmet çok susamıştı. Global ekonomik krizden tüm ülkeler etkileniyor.

Söylemlerde kastedilen nesne ya da kavramların belirlenme işlemine *artgönderim çözümlemesi* (İng. anaphora resolution) denir. Örneğin,

- (2) Ali görüşmeden önce bir bardak viski içti. Bu, sinirlerinin yatışması için ona çok iyi geldi.

şeklindeki ardışık iki cümleden oluşan bir söylemde, ikinci cümlede bulunan "bu" adılı, birinci cümlede *Ali'nin görüşmeden önce bir bardak viski içmesi* ifadesine gönderimde bulunmaktadır. Burada, "bu" artgörmüsel (İng. anaphoric), *Ali'nin görüşmeden önce bir bardak viski içmesi* ise gönderim yapılan (İng. referent) ifadedir.

Bazen, gönderimde bulunan ifade bir ya da birkaç cümleyi birden içeriyor olabilir. Örneğin,

- (3) Sitemin hakkından en iyi daha baskın bir sitem gelir: "Sayenizde kapıma ilk defa polis dayandı. Bütün apartman beni biriyle basıldı sandı hanımefendi. Ne

⁴³ Bağdaşık olmayan metinler # işaretiyle gösterilmektedir.

hakkınız var bana **bunu** yaşatmaya?" dedim. Namusu konusunda uyananşüphelerden gururu incinmiş mağrur bir kadın gibi tonladım bu cümleyi. Tahmin ettiğim gibi, sözlerim çok dokundu ona, üst üste özürler diledi (Mungan, 2001, p.88).

şeklindeki bir söylemde, altıçizi "bu" artgörümsele ifadesi, "*Sayenizde kapıma ilk defa polis dayandı. Bütün apartman beni biriyle basıldı sandı*" cümlelerine göndermede bulunmaktadır.

1.2 Yeterlik İlkesi

Deney sırasında artgörümsele ifadelerle kastedilen nesne ya da kavramları belirlerken dikkat edilecek nokta, gönderim yapılan ifadenin anlaşılması için yeterli olan metin aralığının *üye* olarak seçilmesi gereğidir. Buna *yeterlik ilkesi* diyebiliriz. Üyeye dahil edilmesine gerek olmayan bölümlerin dahil edilmemesi önemlidir. Örneğin,

- (4) Bu yaşta bir kadın olarak niye hala çocuk sahibi olmadığım yolundaki bütün soruları sahtekarca yanıtlamıştım demek. **Bunu** şimdi anlıyordum; ne hiç evlenmemiş olmam bir nedendi, ne de karşıma evlenmek için uygun biri çıkmamış olması... (Mungan, 2001, p.17).

şeklindeki bir söylemde "bu" artgöndermesi, *Bu yaşta bir kadın olarak niye hala çocuk sahibi olmadığım yolundaki bütün soruları sahtekarca yanıtlamış olma* durumuna gönderimde bulunmaktadır.

Böyle bir söylemde gönderim yapılan ifade aşağıda altı çizili olan söz öbeği olarak belirlenmelidir:

- (4) Bu yaşta bir kadın olarak niye hala çocuk sahibi olmadığım yolundaki bütün soruları sahtekarca yanıtlamıştım demek. **Bunu** şimdi anlıyordum; ne hiç evlenmemiş olmam bir nedendi, ne de karşıma evlenmek için uygun biri çıkmamış olması... (Mungan, 2001, p.17).

İlk cümlelele son kelimesi olan "demek" sözcüğünü üyeye dahil etmekle, yeterlilik ilkesinin gerekliliklerine uyulmamış olunur.

1.3 Artgönderme ve Gönderge Arasındaki Mesafe

Artgönderimsel ifade ile gönderimde bulunan ifade arasında mesafe olması doğal bir durumdur.

- (5) Masallar, efsaneler bedel ödemekten bahsederler. Geçmişte yaptığınız kötülükler için, kötülük yaptığınız bir insanın ruhu sizi yer yüzünde, bir sonraki hayatınızda bulur ve siz **bunun** bedelini ödersiniz. Musluğu iyice açıyorum. Su öylesine şiddetli çarpıyor ki yüzüme, sıksa bedenim küvete yığılıverecekmiş gibi sarsılıyor. Ağlıyorum. Ağladığımı da yüzüme şiddetle çarpan suyun parçaladığı göz yaşlarımdan anlamıyorum. Göğsümde saat başı vuran bir gong beliriyor. Ağladığımı ancak öyle fark edebiliyorum (ODTÜ).

Yukarıdaki söylemde, altıçizili "bu" artgöndermesi, birinci satırda geçen "geçmişte yaptığınız kötülükler" ifadesine gönderimde bulunmaktadır.

1.4 Örnek

a. Aşağıdaki paragrafta "bu" ifadesinin gönderimde bulunduğu ifadenin altını çiziniz.

- (6) Akılcıca onu bilinçlendirmeye çalışıyordum galiba. Birbirine bu biçimde benzemeye çalışan kadınların, yalnızca mutsuzlukları birbirine benziyordu sonuçta. Belli ki, benden çok daha önce görmüştü **bunu** Aysel (Mungan, 2001, p.95).

APPENDIX 4

ANNOTATION GUIDELINE FOR GROUP 2

1.0 Giriş

Bu deneyin temel amacı, Türkçe'de artgönderim çözümü (İng. Anaphora resolution) süreçlerini incelemektir. Bu deneyde, size verilen kısa paragraflarda altı çizili söz ya da sözcüklerin göndermelerini (kastedilen nesne ya da kavramı) bulmanız beklenmektedir.

1.1 Söylem Bağlıları

Aralarında belli bağlantılar olan tümcelerin biraraya gelmesiyle oluşan dilsel birime *söylem* (İng. discourse) denir. Örneğin (1a)'da ilk tümce ikinci tümcenin nedeni olduğundan, (1a) neden-sonuç içeren bir söylemdir. Buna karşın, (1b)'de ilk tümcenin anlamıyla (Ahmet'in susaması) ikinci tümcenin anlamı (global ekonomik krizin tüm dünya ülkelerini etkilemesi) arasında herhangi bir ilgi kurulamamaktadır.⁴⁴

(1a) Ahmet çok susamıştı. Kalkıp bir bardak su içti.

(1b) # Ahmet çok susamıştı. Global ekonomik krizden tüm ülkeler etkileniyor.

Söylemlerde kastedilen nesne ya da kavramların belirlenme işlemine *artgönderim çözümü* (İng. anaphora resolution) denir. Örneğin,

(2) Ali görüşmeden önce bir bardak viski içti. Bu, sinirlerinin yatışması için ona çok iyi geldi.

şeklindeki ardışık iki cümleden oluşan bir söylemde, ikinci cümlede bulunan "bu" adlı, birinci cümlede *Ali'nin görüşmeden önce bir bardak viski içmesi* ifadesine gönderimde bulunmaktadır. Burada, "bu" artgörmüsel (İng. anaphoric), *Ali'nin görüşmeden önce bir bardak viski içmesi* ise gönderim yapılan (İng. referent) ifadedir.

Bazen, gönderimde bulunan ifade bir ya da birkaç cümleyi birden içeriyor olabilir. Örneğin,

(3) Sitemin hakkından en iyi daha baskın bir sitem gelir: "Sayenizde kapıma ilk defa polis dayandı. Bütün apartman beni biriyle basıldı sandı hanımefendi. Ne

⁴⁴ Bağdaşık olmayan metinler # işaretiyle gösterilmektedir.

hakkınız var bana **bu durumu** yaşatmaya?" dedim. Namusu konusunda uyanan şüphelerden gururu incinmiş mağrur bir kadın gibi tonladım bu cümleyi. Tahmin ettiğim gibi, sözlerim çok dokundu ona, üst üste özürler diledi (Mungan, 2001, p.88).

şeklindeki bir söylemde, altıçizi "bu durumu" artgörümsel ifadesi, "*Sayenizde kapıma ilk defa polis dayandı. Bütün apartman beni biriyle basıldı sandı*" cümlelerine göndermede bulunmaktadır.

1.2 Yeterlik İlkesi

Deney sırasında artgörümsel ifadelerle kastedilen nesne ya da kavramları belirlerken dikkat edilecek nokta, gönderim yapılan ifadenin anlaşılması için yeterli olan metin aralığının *üye* olarak seçilmesi gereğidir. Buna *yeterlik ilkesi* diyebiliriz. Üyeye dahil edilmesine gerek olmayan bölümlerin dahil edilmemesi önemlidir. Örneğin,

- (4) Bu yaşta bir kadın olarak niye hala çocuk sahibi olmadığım yolundaki bütün soruları sahtekarca yanıtlamıştım demek. **Bu gerçeği** şimdi anlıyordum; ne hiç evlenmemiş olmam bir nedendi, ne de karşıma evlenmek için uygun biri çıkmamış olması...

şeklindeki bir söylemde "bu gerçeği" artgöndermesi, *Bu yaşta bir kadın olarak niye hala çocuk sahibi olmadığım yolundaki bütün soruları sahtekarca yanıtlamış olma* ifadesine gönderimde bulunmaktadır.

Böyle bir söylemde gönderim yapılan ifade aşağıda altı çizili olan söz öbeği olarak belirlenmelidir:

- (4) Bu yaşta bir kadın olarak niye hala çocuk sahibi olmadığım yolundaki bütün soruları sahtekarca yanıtlamıştım demek. **Bu gerçeği** şimdi anlıyordum; ne hiç evlenmemiş olmam bir nedendi, ne de karşıma evlenmek için uygun biri çıkmamış olması...

İlk cümlelerin son kelimesi olan "demek" sözcüğünü üyeye dahil etmekle, yeterlilik ilkesinin gerekliliklerine uyulmamış olunur.

1.3 Artgönderme ve Gönderge Arasındaki Mesafe

Artgönderimsel ifade ile gönderimde bulunan ifade arasında mesafe olması doğal bir durumdur.

- (5) Masallar, efsaneler bedel ödemekten bahsederler. Geçmişte yaptığınız kötülükler için, kötülük yaptığınız bir insanın ruhu sizi yer yüzünde, bir sonraki hayatınızda bulur ve siz **bu işin** bedelini ödersiniz. Musluğu iyice açıyorum. Su öylesine şiddetli çarpıyor ki yüzüme, sıksa bedenim küvete yığılıverecekmiş gibi sarsılıyor. Ağlıyorum. Ağladığımı da yüzüme şiddetle çarpan suyun parçaladığı göz yaşlarımdan anlamıyorum. Göğsümde saat başı vuran bir gong beliriyor. Ağladığımı ancak öyle fark edebiliyorum.

Yukarıdaki söylemde, altıçizili "bu işin" artgöndermesi, birinci satırda geçen "geçmişte yaptığınız kötülükler" ifadesine gönderimde bulunmaktadır.

1.4 Örnek

a. Aşağıdaki paragrafta "bu durum" ifadesinin gönderimde bulunduğu ifadenin altını çiziniz.

- (6) Aklımca onu bilinçlendirmeye çalışıyordum galiba. Birbirine bu biçimde benzemeye çalışan kadınların, yalnızca mutsuzlukları birbirine benziyordu sonuçta. Belli ki, benden çok daha önce görmüştü **bu durumu** Aysel.

APPENDIX 5

START-END WORD CONSISTENCY ANALYSIS FOR THE PILOT STUDY

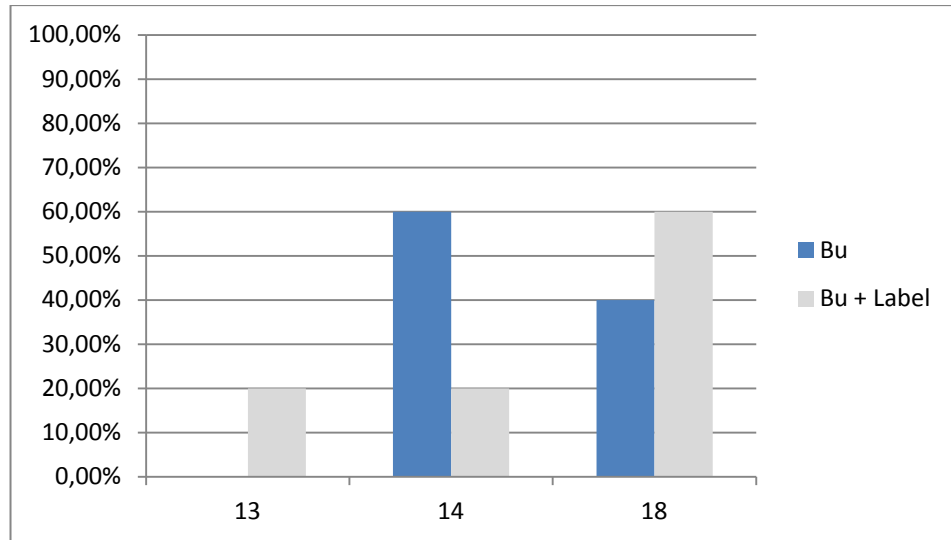


Figure 1. The frequency of start words of the participants in Group 1 and Group 2 for paragraph 1 with the most common start words being the 14th and 18th ones.

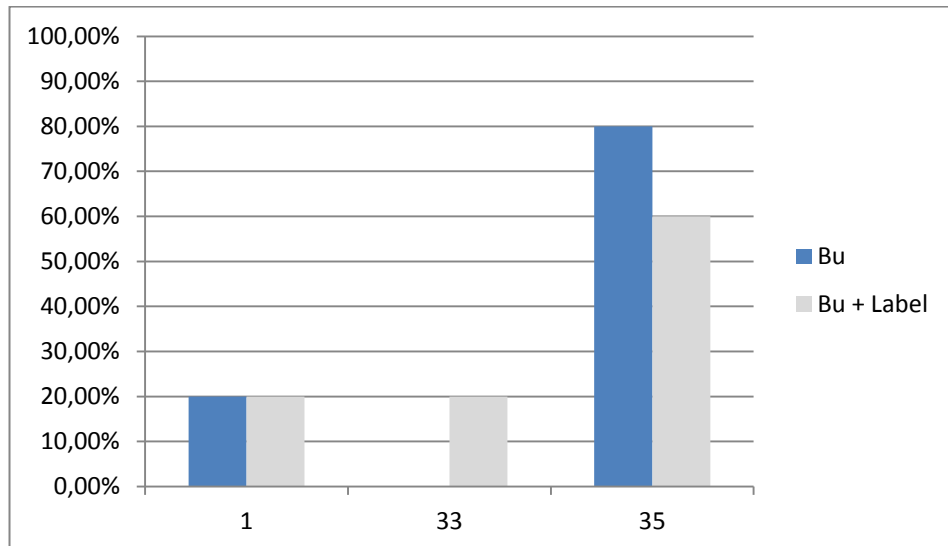


Figure 2. The frequency of start words of the participants in Group 1 and 2 for paragraph 2 with the most common start words being the 33rd and 35th ones.

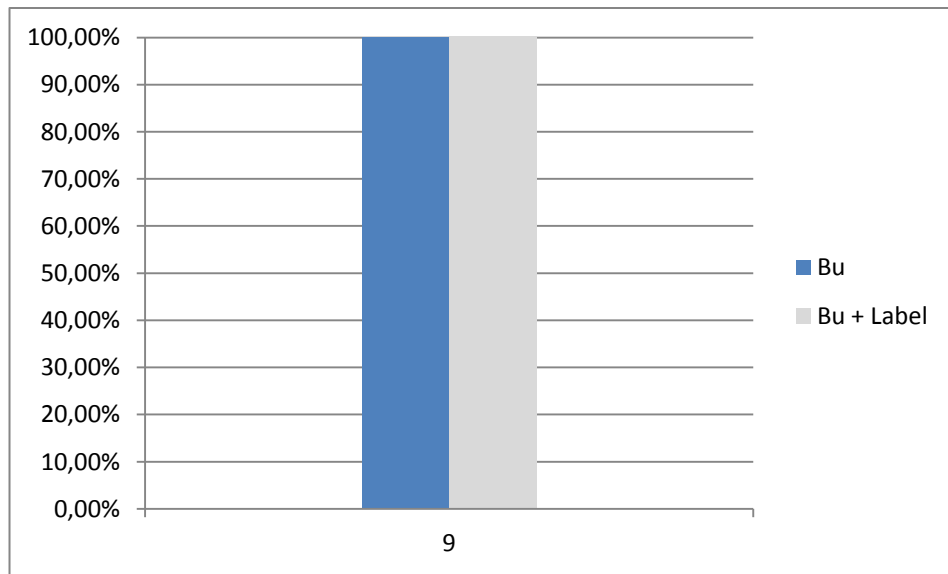


Figure 3. The frequency of start words of the participants in Group 1 and 2 for paragraph 3 with the most common start word being the 9th one.

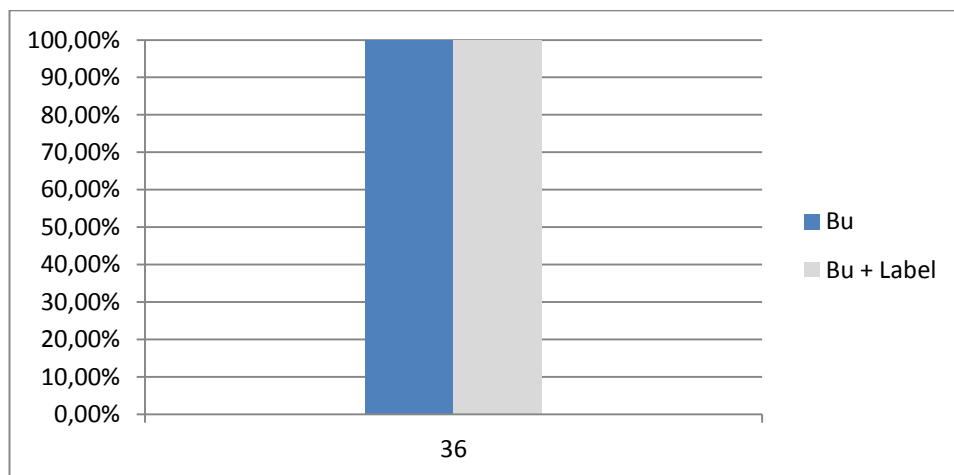


Figure 4. The frequency of start words of the participants in Group 1 and 2 for paragraph 4 with the most common start word being the 36th one.

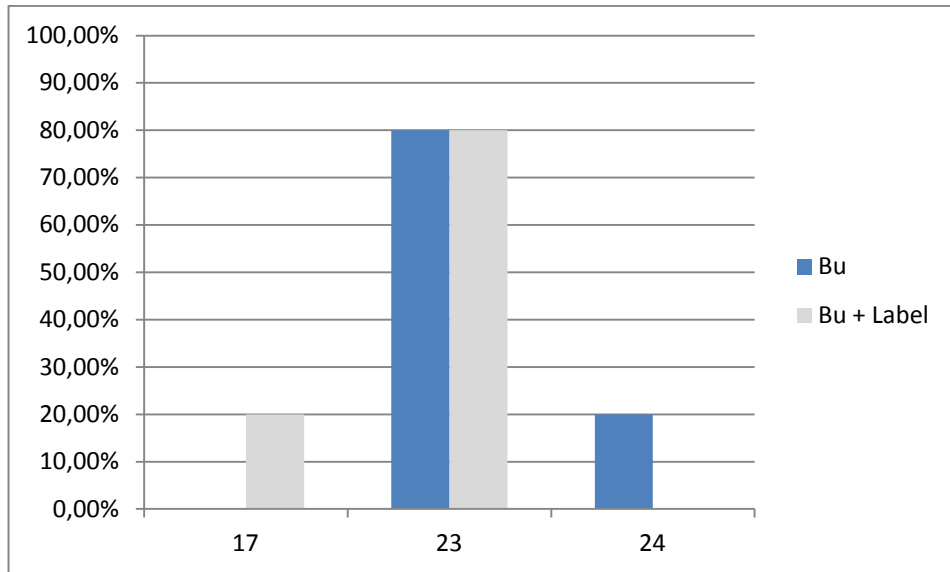


Figure 5. The frequency of end words of the participants in Group 1 and 2 for paragraph 1 with the most common start word being the 23rd and 24th ones.

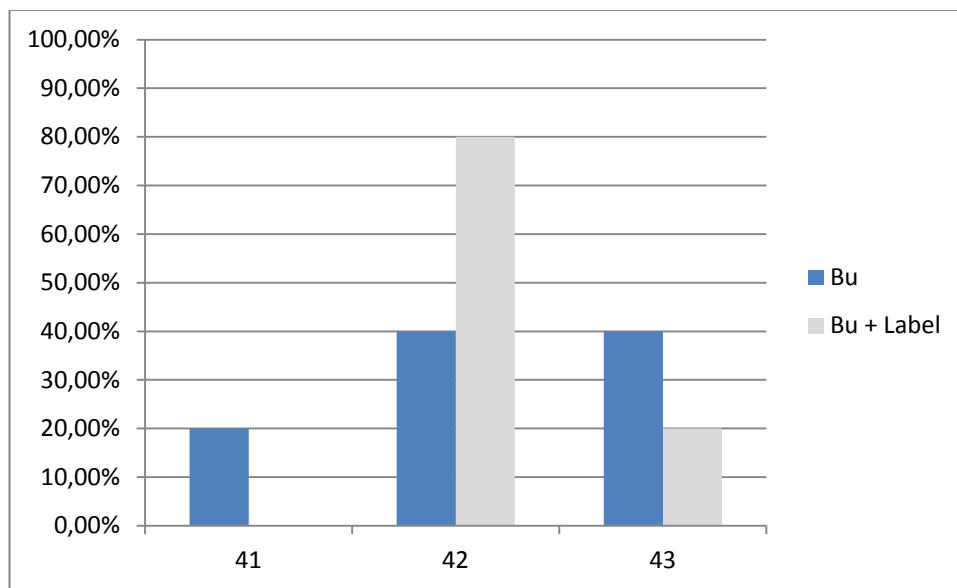


Figure 6. The frequency of end words of the participants in Group 1 and 2 for paragraph 2 with the most common start word being the 42nd and 43rd ones.

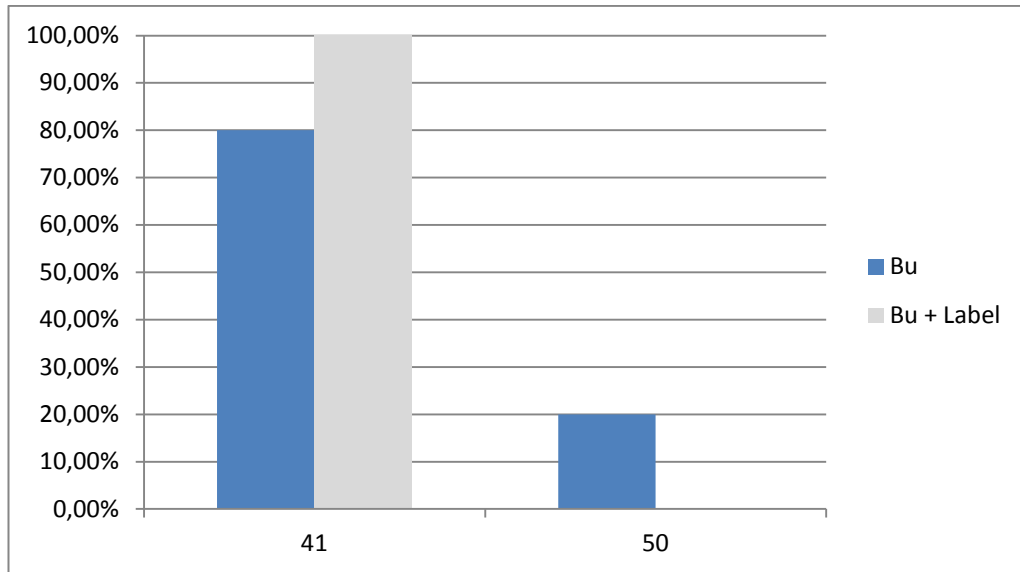


Figure 7. The frequency of end words of the participants in Group 1 and 2 for paragraph 4 with the most common start word being the 41st one.

APPENDIX 6

START WORD CONSISTENCY ANALYSIS FOR EXPERIMENT 1

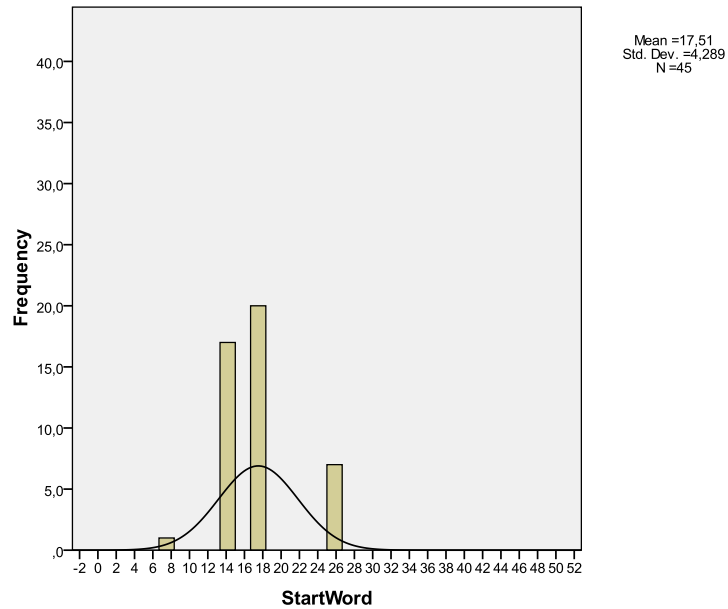


Figure 8. The frequency of start words of the participants in Group 1 for paragraph 1 with the most common start words being the 14th and 18th ones.

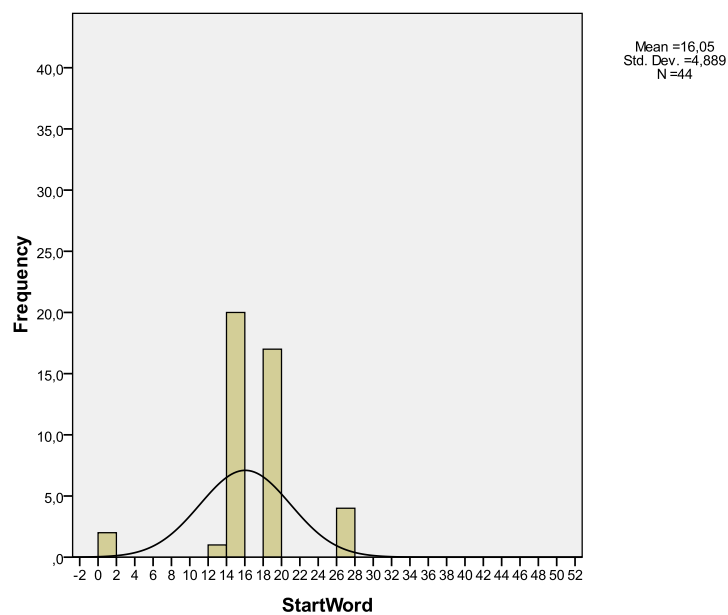


Figure 9. The frequency of start words of the participants in Group 2 for paragraph 1 with the most common start words being the 14th and 18th ones.

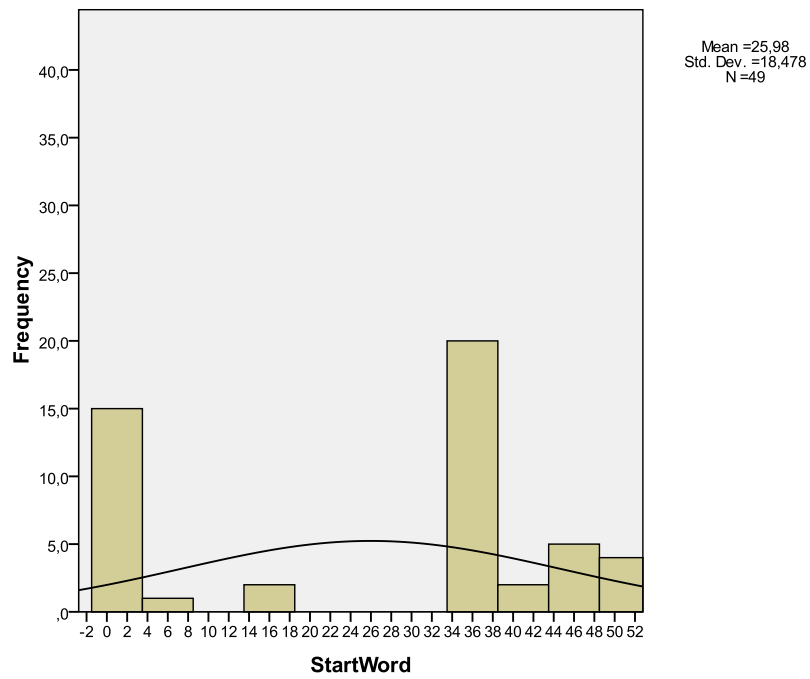


Figure 10. The frequency of start words of the participants in Group 1 for paragraph 2 with the most common start words being the 1st and 35th ones.

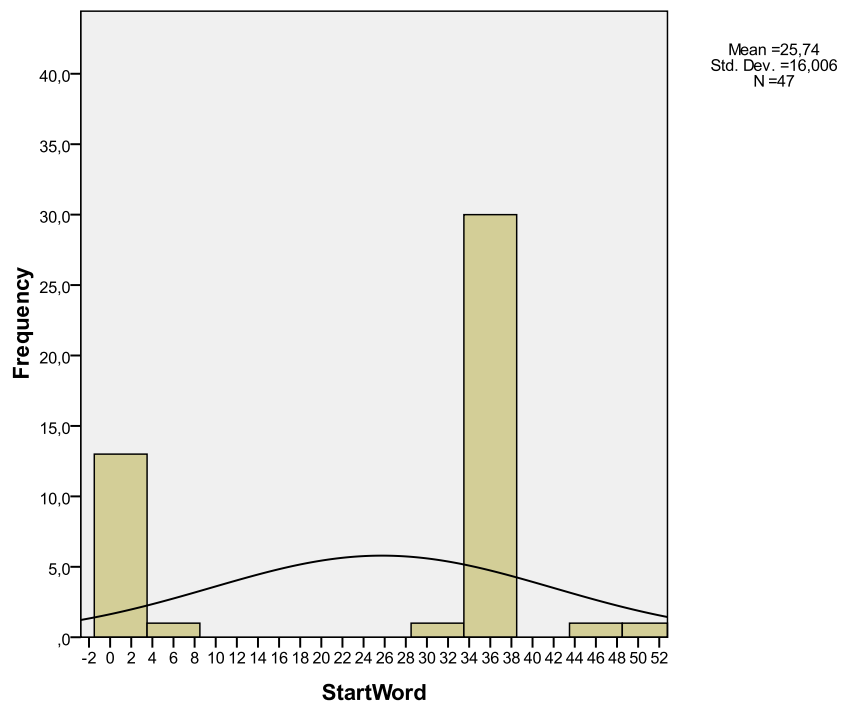


Figure 11. The frequency of start words of the participants in Group 2 for paragraph 2 with the most common start words being the 1st and 35th ones.

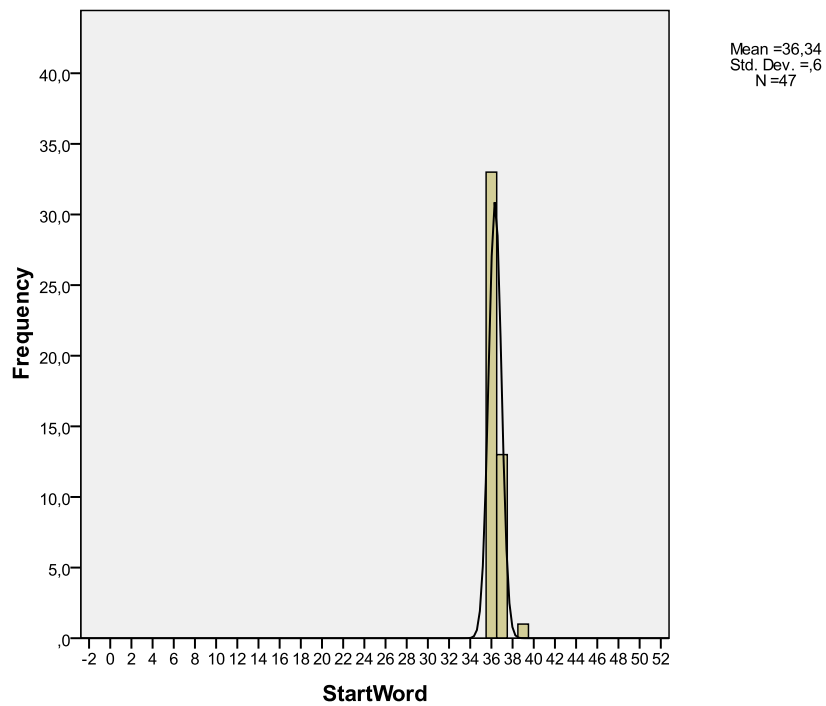


Figure 12. The frequency of start words of the participants in Group 1 for paragraph 4 with the most common start words being the 36th and 37th ones.

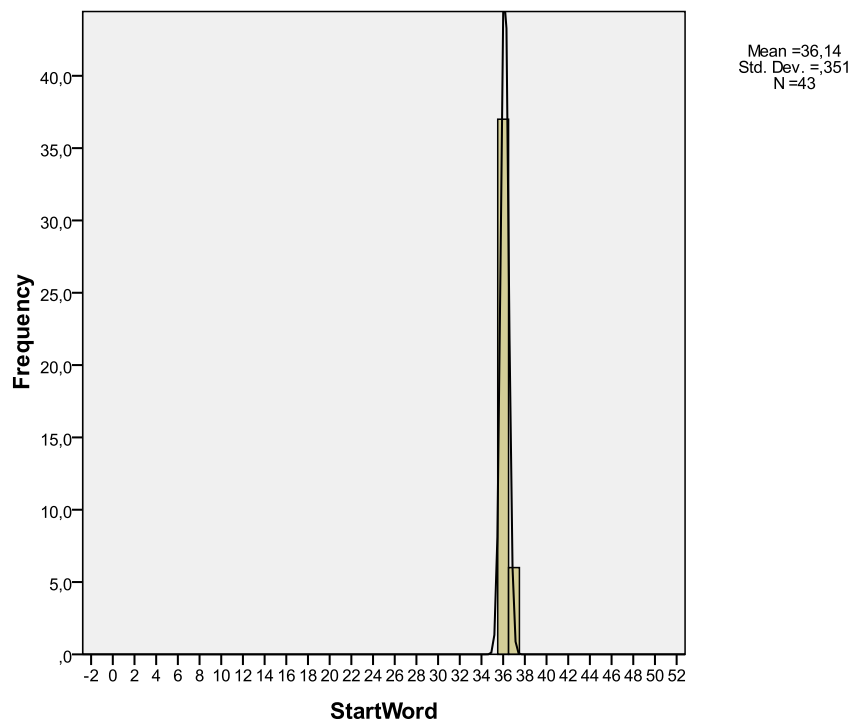


Figure 13. The frequency of start words of the participants in Group 2 for paragraph 4 with the most common start words being the 36th and 37th ones.

APPENDIX 7

END WORD CONSISTENCY ANALYSIS FOR EXPERIMENT 1

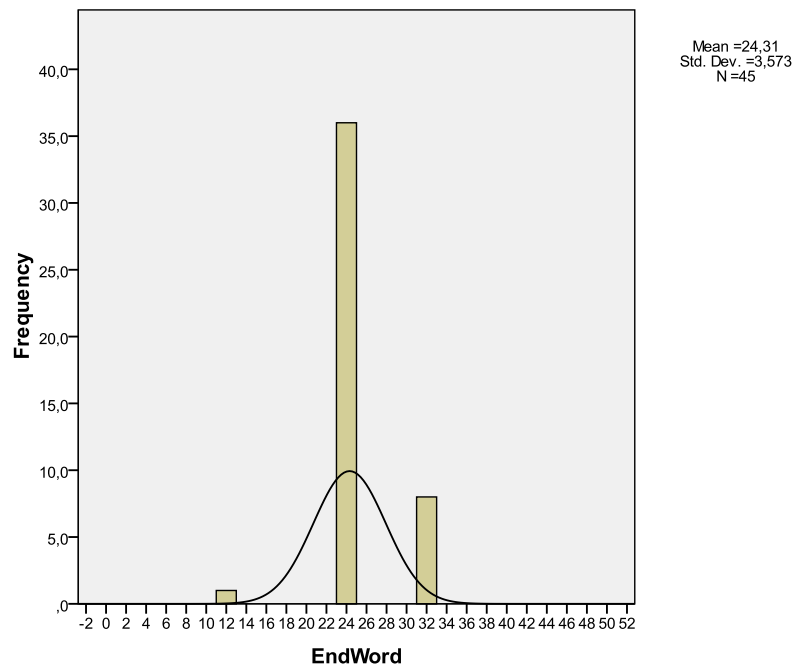


Figure 14. The frequency of end words of the participants in Group 1 for paragraph 1 with the most common end word being the 23rd one.

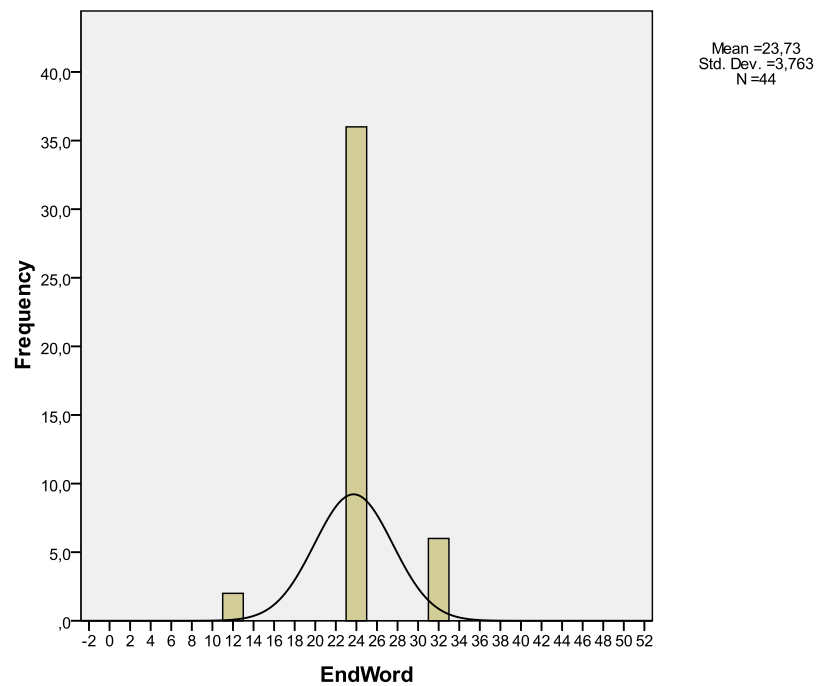


Figure 15. The frequency of end words of the participants in Group 2 for paragraph 1 with the most common end word being the 23rd one.

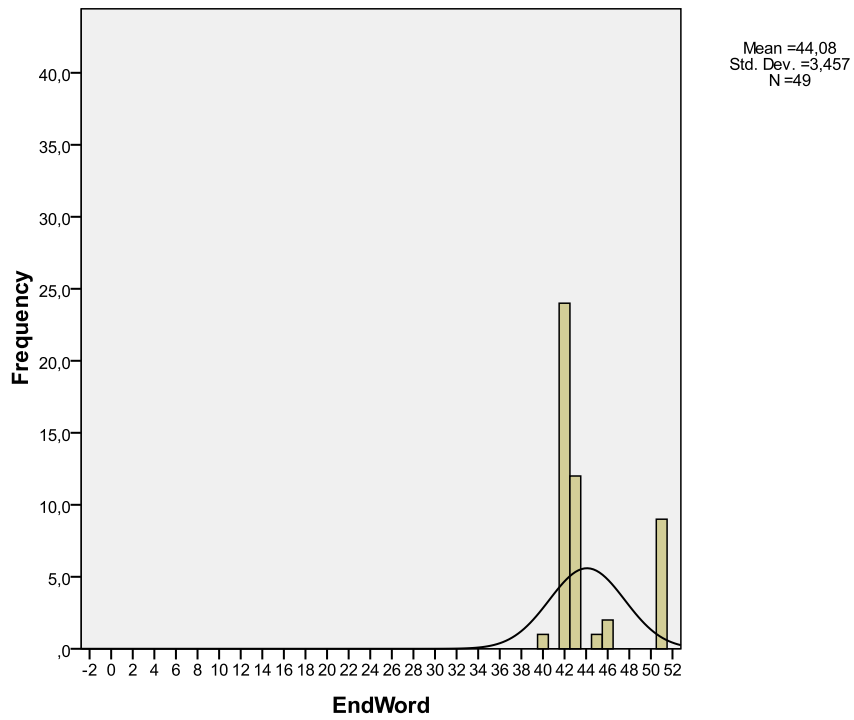


Figure 16. The frequency of end words of the participants in Group 1 for paragraph 2 with the most common end words being the 42nd, 43rd and the 50th.

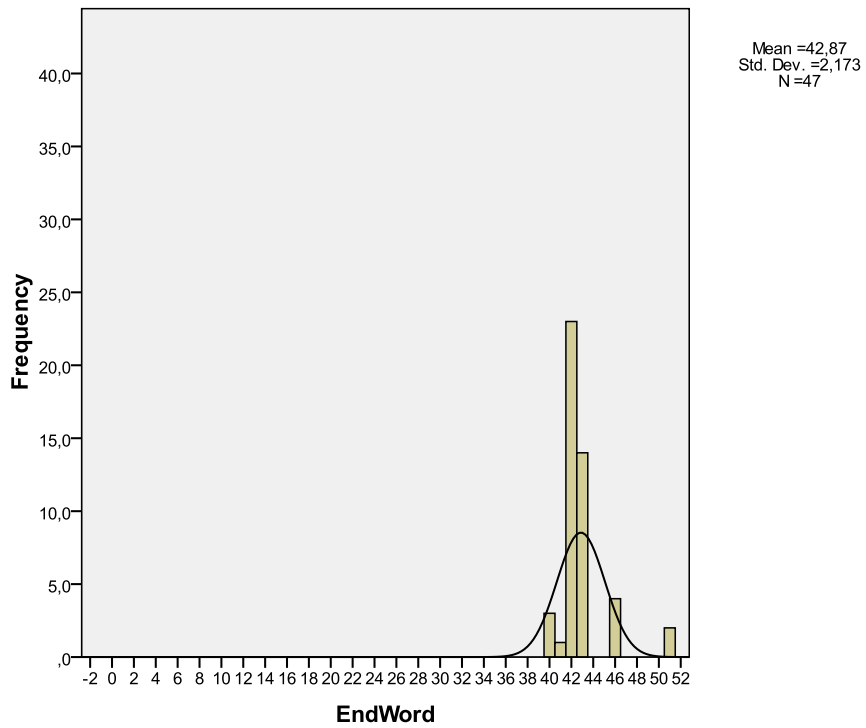


Figure 17. The frequency of end words of the participants in Group 2 for paragraph 2 with the most common end words being the 42nd and 43rd.

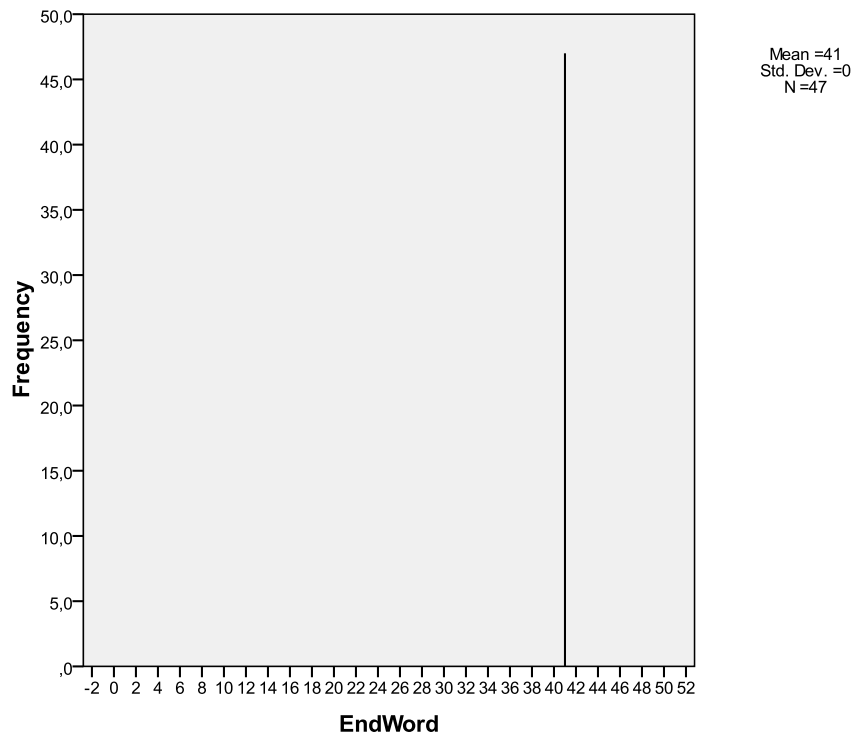


Figure 18. The frequency of end words of the participants in Group 1 for paragraph 4 with the most common end word being the 41st.

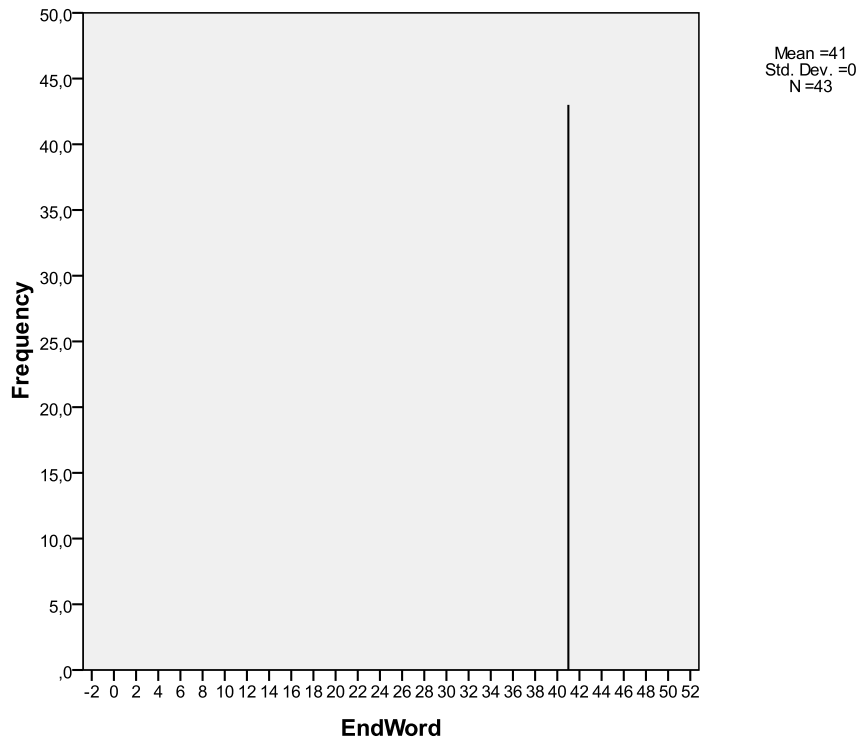


Figure 19. The frequency of end words of the participants in Group 2 for paragraph 4 with the most common end word being the 41st.

APPENDIX 8

PASSONEAU (2005) ANALYSIS FOR EXPERIMENT 1

Experiment 1 - Paragraph 1

A: Nazlı da, komünist günlerini geride bırakıp hiçbir şeyin artık eskisi gibi olmadığını anladığı anda, yeni hayatlarına daha çabuk uyum göstermiş

A1: komünist günlerini geride bırakıp hiçbir şeyin artık eskisi gibi olmadığını anladığı anda yeni hayatlarına daha çabuk uyum göstermiş

A2: hiçbir şeyin artık eskisi gibi olmadığını anladığı anda, yeni hayatlarına daha çabuk uyum göstermiş

A3: komünist günlerini geride bırakıp hiçbir şeyin artık eskisi gibi olmadığını anladığı

A4: komünist günlerini geride bırakıp hiçbir şeyin artık eskisi gibi olmadığını

A5: hiçbir şeyin artık eskisi gibi olmadığını

A6: hiçbir şeyin artık eskisi gibi olmadığını anladığı

A7: yeni hayatlarına daha çabuk uyum göstermiş

A8: hayatlarına daha çabuk uyum göstermiş

B: Kadınlar her şeyi çok daha kolay kabullenir, değişimlere daha çabuk rıza gösterirler

B1: değişimlere daha çabuk rıza gösterirler

No Match:

- komünist günlerini geride bırakıp hiçbir şeyin artık eskisi gibi olmadığını anladığı anda
- komünist günlerini
- Nazlı da, komünist günlerini geride bırakıp hiçbir şeyin artık eskisi gibi olmadığını anladığı anda, yeni hayatlarına daha çabuk

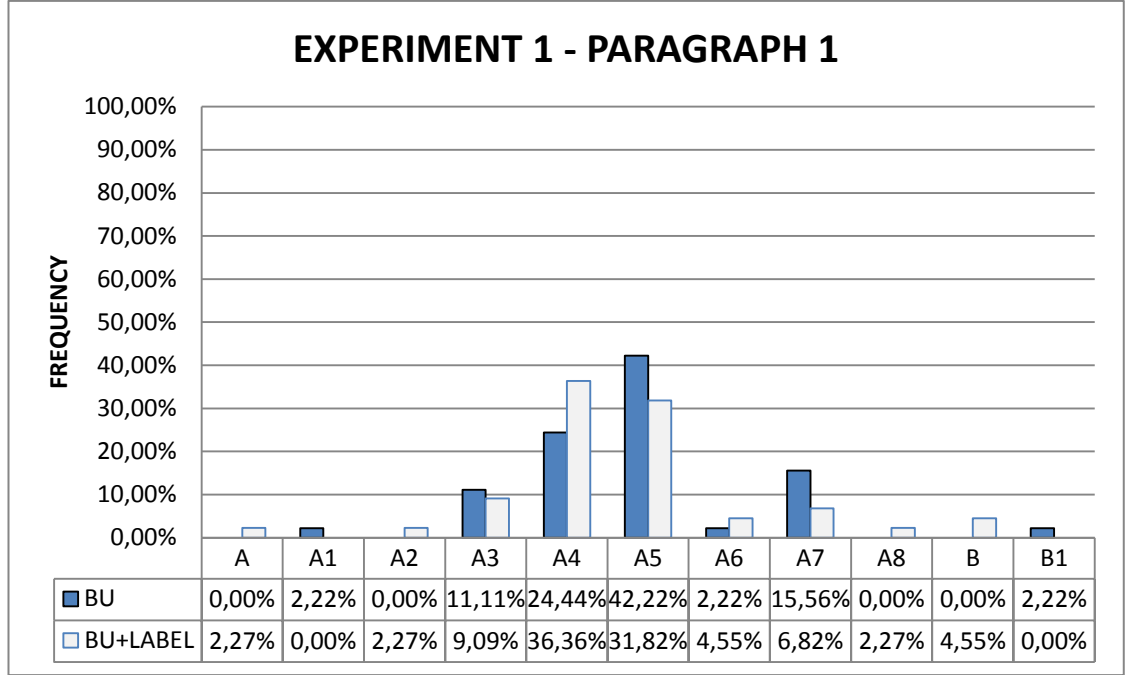


Figure 20. Pasonneau's classification of main set, subsumption, intersection and disjunct relations are illustrated. This is the whole data for Paragraph 1 from both groups. Y axis shows the frequency of the classifications. Below the graph, the percentages of each set's occurrence in Group 1 (bu) and Group 2 (Bu + label) are provided.

Experiment 1 - Paragraph 2

A: dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek doğrusu sevindirmişti beni

A1: dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu okuyan biri olduğunu öğrenmek doğrusu sevindirmişti

A2: dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek

A3: kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek doğrusu sevindirmişti beni

A4: dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A5: boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek doğrusu sevindirmişti beni

A6: dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri

A7: çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A8: boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek

A9: kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A10: boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A11: gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek

A12: boş zamanlarında gizli gizli fotoromanlar okuyan

A13: gizli gizli fotoromanlar okuyan biri olduğunu

A14: fotoromanlar okuyan biri olduğunu

B: Elbet bir gün Kanada'dan dönecekti

B1: bir gün Kanada'dan dönecekti

B2: Kanada'dan dönecekti

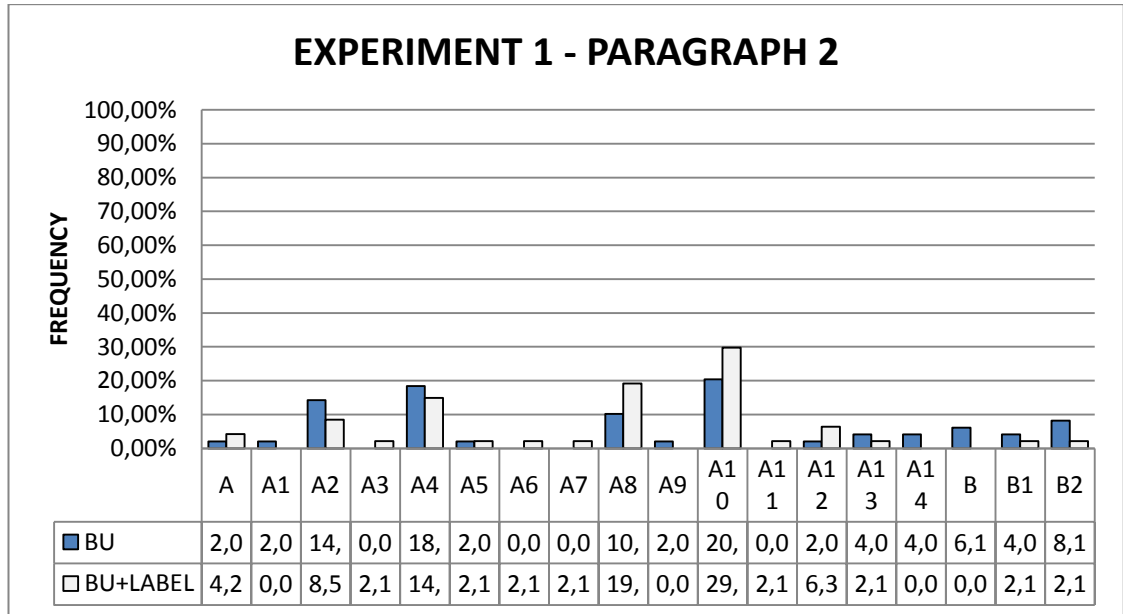


Figure 21. Pasonneau's classification of main set, subsumption, intersection and disjunct relations are illustrated. This is the whole data for Paragraph 2 from both groups. Y axis shows the frequency of the classifications. Below the graph, the percentages of each set's occurrence in Group 1 (bu) and Group 2 (Bu + label) are provided.

Experiment 1 - Paragraph 4

A: İşgalciler, Ziya Paşaların konağına el koymuşlardı

A1: Ziya Paşaların konağına el koymuşlardı

A2: konağına el koymuşlardı

No Match:

- Ahmet Reşat alelacele çıkıp Paşa'nın konağına gittiler
- Kemal'i eliyle yana itip yukarı koştu Ahmet Reşat. Kemal kapıda beti benzi atmış Hakkı Efendi'yi görünce paldır küldür inmeye başladı merdivenleri. Hakkı Efendi, Kemal'e titreyerek az önce yaşadıklarını naklediyordu ki, Ahmet Reşat giyinmiş olarak geri geldi. İşgalciler, Ziya Paşaların konağına el koymuşlardı. Hakkı Efendi'yle Ahmet Reşat alelacele çıkıp Paşa'nın konağına gittiler
- İşgalciler, Ziya Paşaların konağına el koymuşlardı. Hakkı Efendi'yle Ahmet Reşat alelacele çıkıp Paşa'nın konağına gittiler

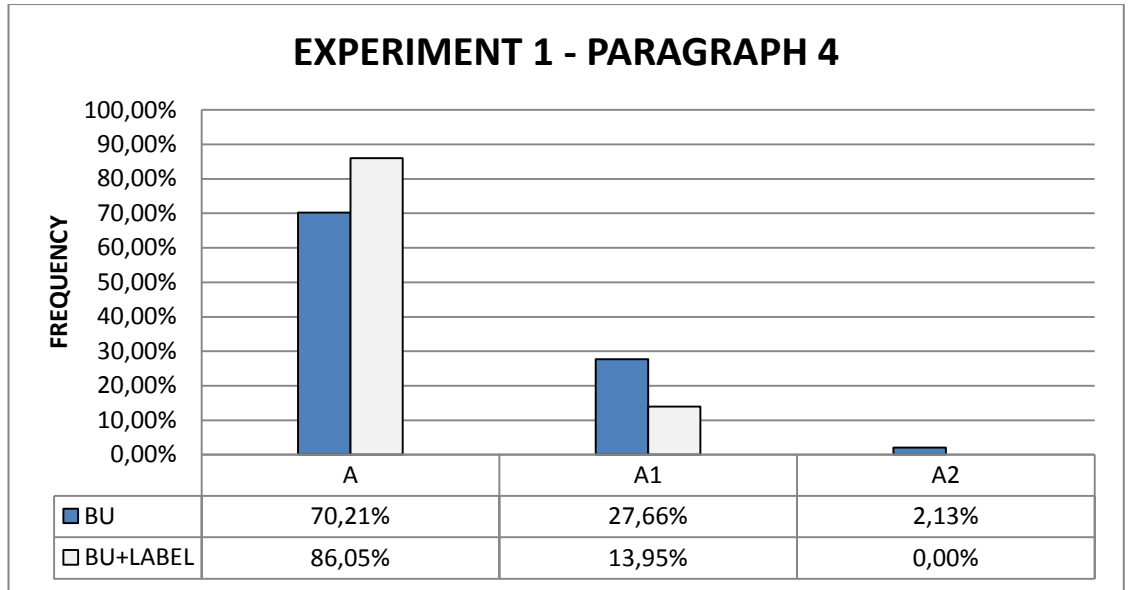


Figure 22. Pasonneau's classification of main set, subsumption, intersection and disjunct relations are illustrated. This is the whole data for Paragraph 4 from both groups. Y axis shows the frequency of the classifications. Below the graph, the percentages of each set's occurrence in Group 1 (bu) and Group 2 (Bu + label) are provided.

APPENDIX 9

PASSONEAU (2005) ANALYSIS WITHOUT ADJUNCTS FOR EXPERIMENT 1

Experiment 1 - Paragraph 2

A: Şimdi Kanada'larda sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek doğrusu sevindirmişti beni

A1: Şimdi Kanada'larda sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu okuyan biri olduğunu öğrenmek doğrusu sevindirmişti

A2: Şimdi Kanada'larda sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek

A3: Şimdi Kanada'larda sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A4: Kanada'larda sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek

A5: Kanada'larda sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A6: Kanada'larda sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri

A7: sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek

A8: sosyalbilimci sosyalbilimci sürten, aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A9: aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek

A10: aşk ve erkekler için önemli değilmiş gibi yapan; başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A11: başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin,

çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek

A12: başkalarının yaşadığı duygusal sarsıntılara, hep bir böcek inceliyormuş gibi nesnel ve kuru gözlerle, hafif de küçümseyerek bakan o dişlek, çirkin, çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A13: kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek doğrusu sevindirmişti beni

A14: boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek doğrusu sevindirmişti beni

A15: çokbilmiş kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A16: boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek

A17: kızın, boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A18: boş zamanlarında gizli gizli fotoromanlar okuyan biri olduğunu

A19: gizli gizli fotoromanlar okuyan biri olduğunu öğrenmek

A20: boş zamanlarında gizli gizli fotoromanlar okuyan

A21: gizli gizli fotoromanlar okuyan biri olduğunu

A22: fotoromanlar okuyan biri olduğunu

APPENDIX 10

EXPERIMENT 2 TEST MATERIAL FOR GROUP 1

ÖN BİLGİ FORMU

Yaş: Cinsiyet: Bölüm
Sınıf: 5-15 yaşları arasında yaşadığınız şehir(ler):
Dil eğitimi düzeyi: İngilizce Yeterlilik Sınavı'ndan aldığınız puan:

DENEY

Aşağıdaki 4 paragrafta "bu/bunu" ifadesinin gönderimde bulunduğu ifadenin altını çiziniz.

1. Taş kaynaklarına baktığınızda, doğal olarak çok taş malzemesi var; bazalt, kireçtaşı, tüfler, volkanik taşlar çok fazla var. Aşıkli'ya yerleşenler ilk geldiklerinde kerpiç mimariyi biliyor. Yapılarını önceleri kerpiçten yapıyorlar, ama sonra kerpiçle taşı birlikte kullanmayı öğreniyorlar. Mimarlık açısından çok önemli **bu**, çünkü bir yapı malzemesini başka bir malzemeyle beraber kullanmayı, ilk defa burada görüyoruz. Evler kerpiçten, ama tapınağın temelinde taş var; çevre duvarı neredeyse tamamen taştan. (TSB).

2. Ahmet, bir gün gazeteyi bir açmıştı, Mesut yok! Pusuya düşürülmüş oracıkta can vermişti. İki kişiymişler. Öbürünün adı Halim. Tanımıyordu. Hafif yaralı kurtulmuş o. **Bu** derinden sarsmıştı Ahmet'i. Evden ayrılmıştı. Okula gitmiyordu artık. Kimseyle görüşmüyordu. Köyüne gömmüşlerdi Mesut'u. 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 (TSB).

3. Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı hiçbir zaman içine sindirememiş olan ve bu konuda

duyduđu ezikliđi saklayamayan Turgay'a, bu konuda s¼rekli takılırlar. (Aslında Őimdi ben de onu yapmıŐ oldum.) ¼nk¼ **bu**, onun daima kanayan aık yarasıdır, camiada herkese bilinir ve kabul edersiniz ki, insan iliŐkilerinde en abuk ¼đrendiđimiz Őeylerden biri, baŐkalarının yarasını kullanmaktır (Mungan, 2001).

4. D¼rt kiŐiydik, t¼renden sonra deniz kenarında bir kahvede oturmuŐtuk. KıŐtı, kar yerden kalkmamıŐtı. Sen ađlamıŐtın. Mutluluktan, yalnızca mutluluktandı g¼z yaŐların. Sulu g¼zly¼d¼n Nesli, ama itiraf etmeye utanıyordun deđil mi? Ne kadar gizlemeye alıŐsan da **bunu**, baŐaramıyordun. Biz yasalar karŐısında evli sayılacak, ama gerekte evli iki insan gibi deđil de (evlilikler sıradanlaŐıyordu ¼nk¼, tekd¼ze ve sıkıcıydı; biz farklı olacaktık), aynı evi paylaŐan iki ¼đrenci gibi yaŐayacaktık (TSB).

APPENDIX 11
EXPERIMENT 2 TEST MATERIAL FOR GROUP 2

ÖN BİLGİ FORMU

Yaş: Cinsiyet: Bölüm
Sınıf: 5-15 yaşları arasında yaşadığınız şehir(ler):
Dil eğitimi düzeyi: İngilizce Yeterlilik Sınavı'ndan aldığınız puan:

DENEY

Aşağıdaki parağraflarda "bu durum/olay/iş/gerçek" ifadesinin gönderimde bulunduğu ifadenin altını çiziniz.

1. Taş kaynaklarına baktığımızda, doğal olarak çok taş malzemesi var; bazalt, kireçtaşı, tüfler, volkanik taşlar çok fazla var. Aşıklı'ya yerleşenler ilk geldiklerinde kerpiç mimariyi biliyor. Yapılarını önceleri kerpiçten yapıyorlar, ama sonra kerpiçle taşı birlikte kullanmayı öğreniyorlar. Mimarlık açısından çok önemli **bu iş**, çünkü bir yapı malzemesini başka bir malzemeye beraber kullanmayı, ilk defa burada görüyoruz. Evler kerpiçten, ama tapınağın temelinde taş var; çevre duvarı neredeyse tamamen taştan.

2. Ahmet, bir gün gazeteyi bir açmıştı, Mesut yok! Pusuya düşürülmüş oracıkta can vermişti. İki kişiymişler. Öbürünün adı Halim. Tanımıyordu. Hafif yaralı kurtulmuş o. **Bu olay** derinden sarsmıştı Ahmet'i. Evden ayrılmıştı. Okula gitmiyordu artık. Kimseyle görüşmüyordu. Köyüne gömmüşlerdi Mesut'u. 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.

3. Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı hiçbir zaman içine sindirememiş olan ve bu konuda duyduğu ezikliği saklayamayan Turgay'a, bu konuda sürekli takılırlar. (Aslında şimdi

ben de onu yapmış oldum.) Çünkü **bu durum**, onun daima kanayan açık yarasıdır, camiada herkesçe bilinir ve kabul edersiniz ki, insan ilişkilerinde en çabuk öğrendiğimiz şeylerden biri, başkalarının yarasını kullanmaktır.

4. Dört kişiydik, törenden sonra deniz kenarında bir kahvede oturmuştuk. Kıştı, kar yerden kalkmamıştı. Sen ağlamıştın. Mutluluktan, yalnızca mutluluktandı göz yaşların. Sulu gözlüydün Nesli, ama itiraf etmeye utanıyordun değil mi? Ne kadar gizlemeye çalışsan da **bu gerçeği**, başaramıyordun. Biz yasalar karşısında evli sayılacak, ama gerçekte evli iki insan gibi değil de (evlilikler sıradanlaşıyordu çünkü, tekdüze ve sıkıcıydı; biz farklı olacaktık), aynı evi paylaşan iki öğrenci gibi yaşayacaktık.

APPENDIX 12

START WORD CONSISTENCY ANALYSIS FOR EXPERIMENT 2

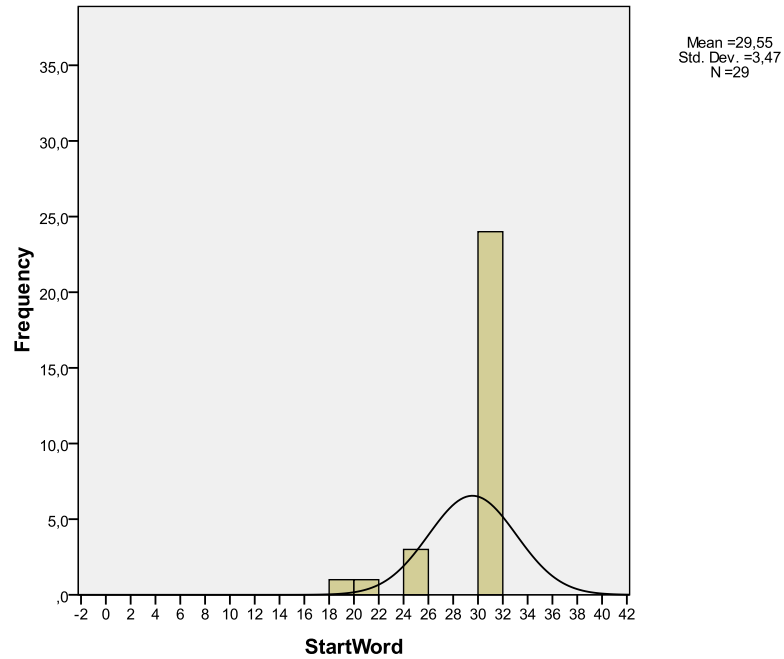


Figure 23. The frequency of start words of the participants in Group 1 for paragraph 1 with the most common start word being the 31st one.

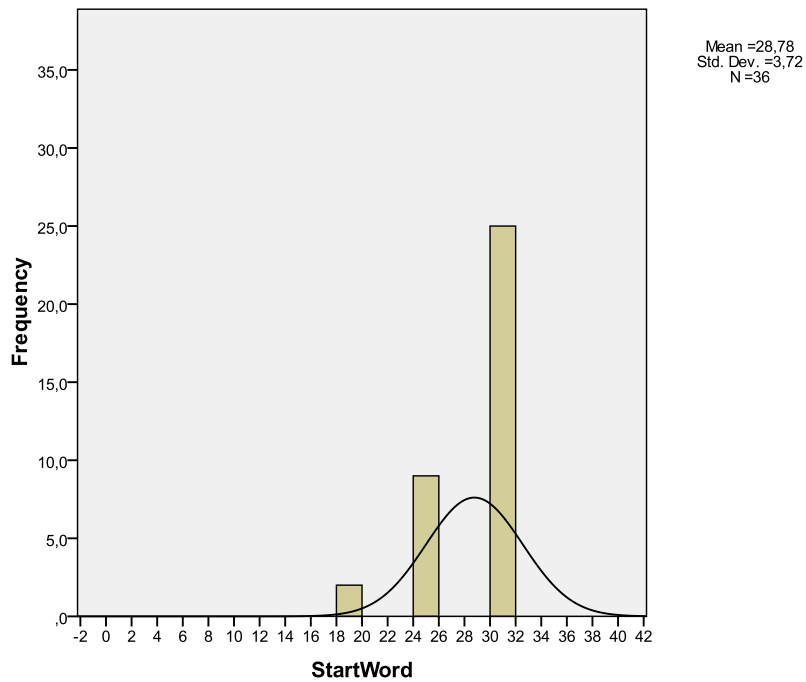


Figure 24. The frequency of start words of the participants in Group 2 for paragraph 1 with the most common start words being the 31st one.

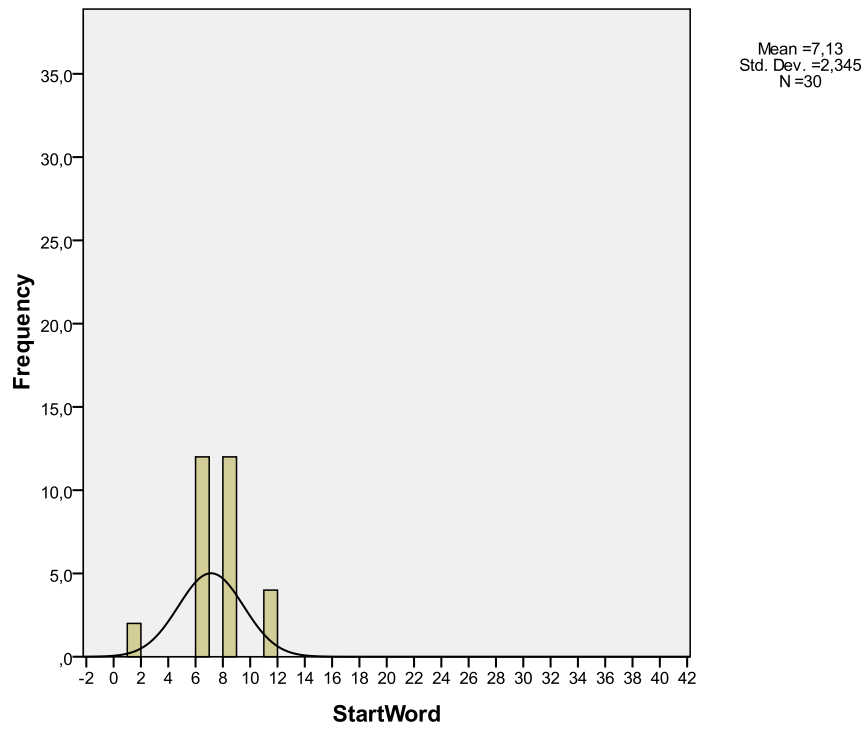


Figure 25. The frequency of start words of the participants in Group 1 for paragraph 2 with the most common start words being the 6th and 8th words.

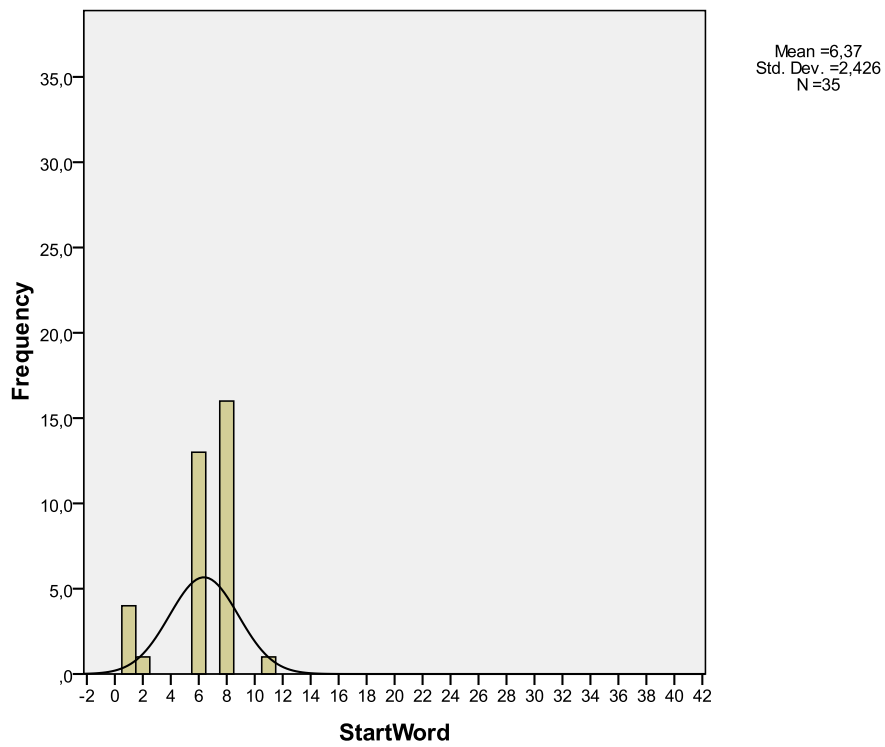


Figure 26. The frequency of start words of the participants in Group 2 for paragraph 2 with the most common start words being the 6th and 8th words.

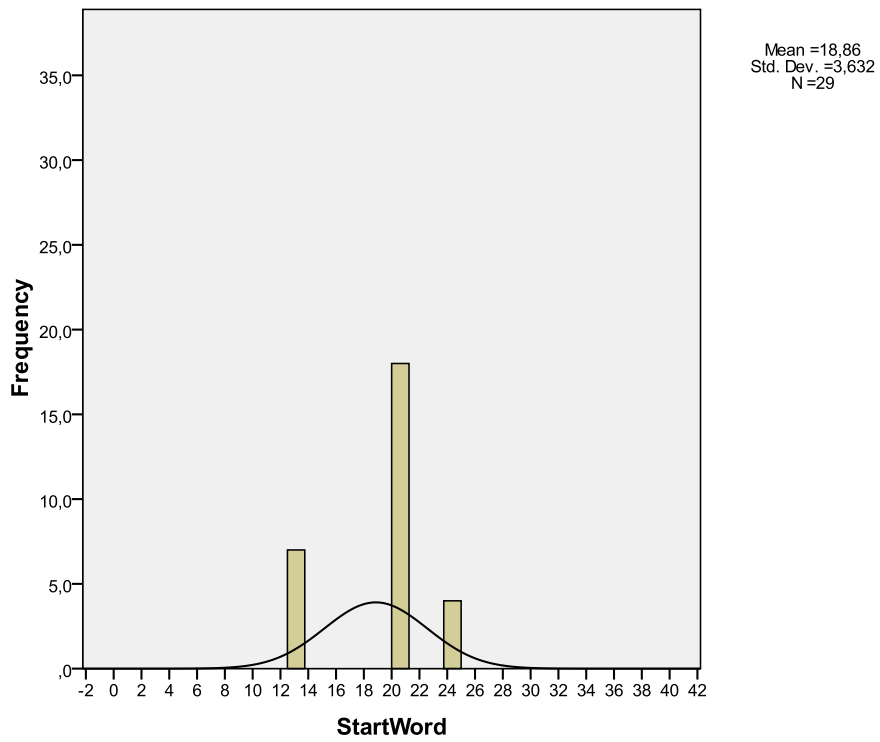


Figure 27. The frequency of start words of the participants in Group 1 for paragraph 4 with the most common start words being 20th and 13rd ones.

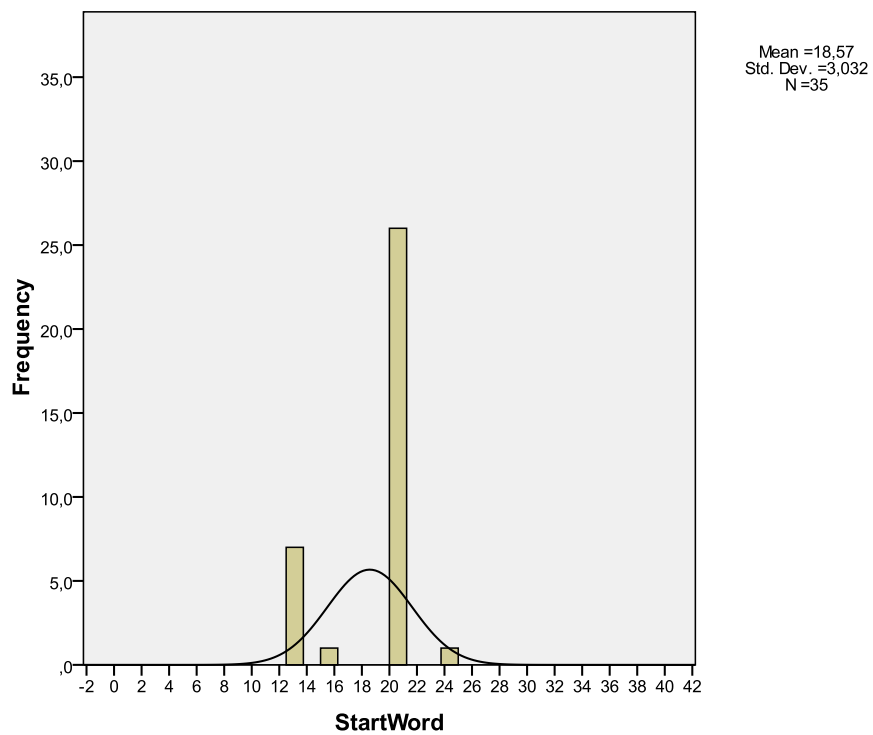


Figure 28. The frequency of start words of the participants in Group 2 for paragraph 4 with the most common start words being 20th and 13rd ones.

APPENDIX 13

END WORD CONSISTENCY ANALYSIS FOR EXPERIMENT 2

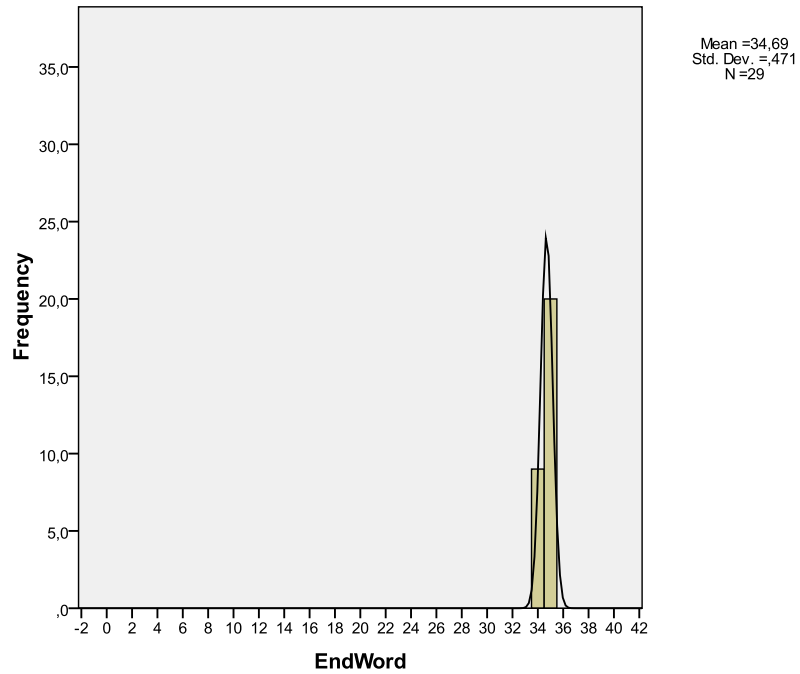


Figure 29. The frequency of end words of the participants in Group 1 for paragraph 1 with the most common end word being 34th and 35th words.

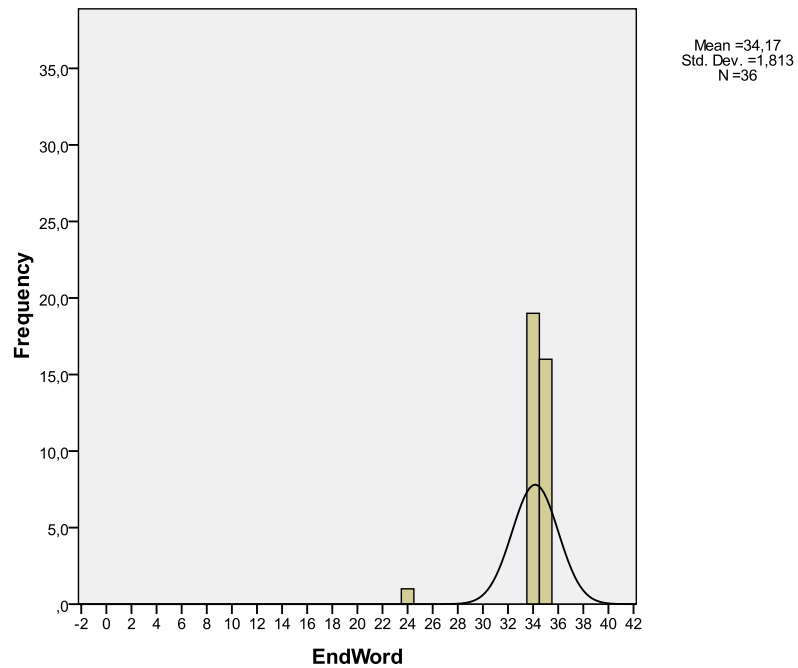


Figure 30. The frequency of end words of the participants in Group 2 for paragraph 1 with the most common end word being 34th and 35th words.

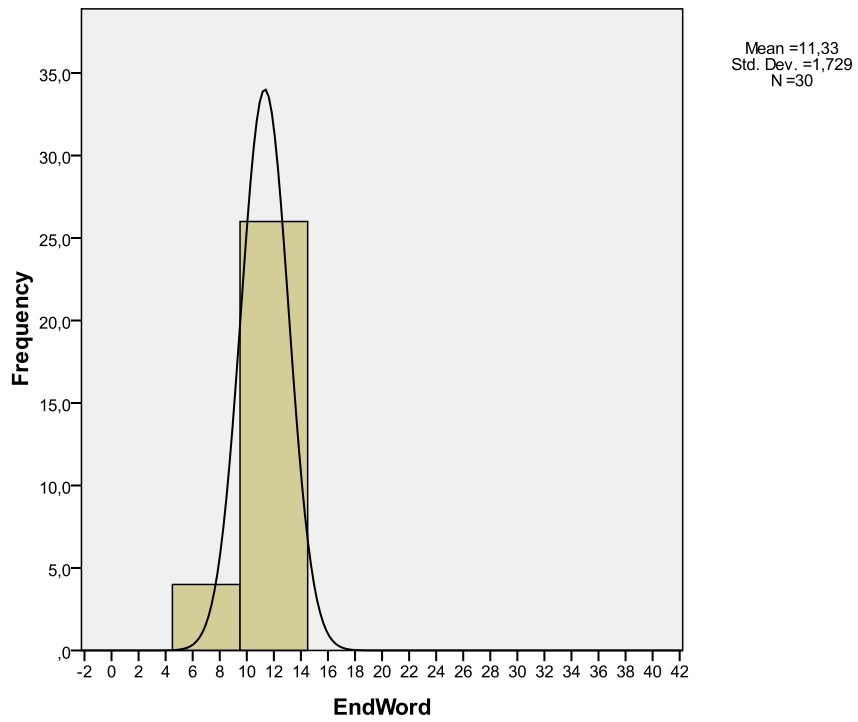


Figure 31. The frequency of end words of the participants in Group 1 for paragraph 2 with the most common end word being the 12th word.

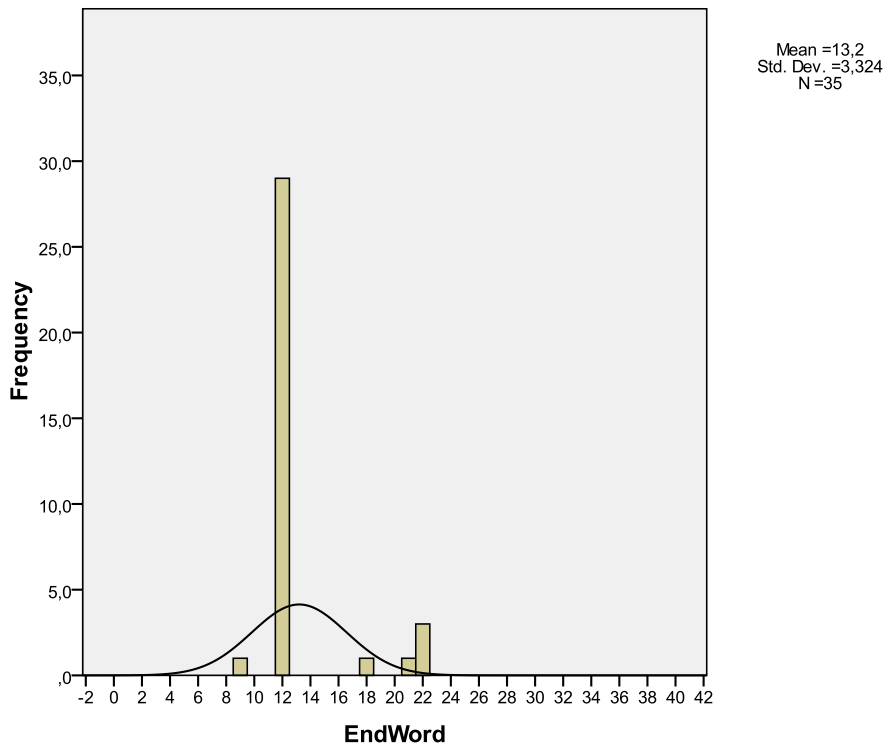


Figure 32. The frequency of end words of the participants in Group 2 for paragraph 2 with the most common end word being the 12th word.

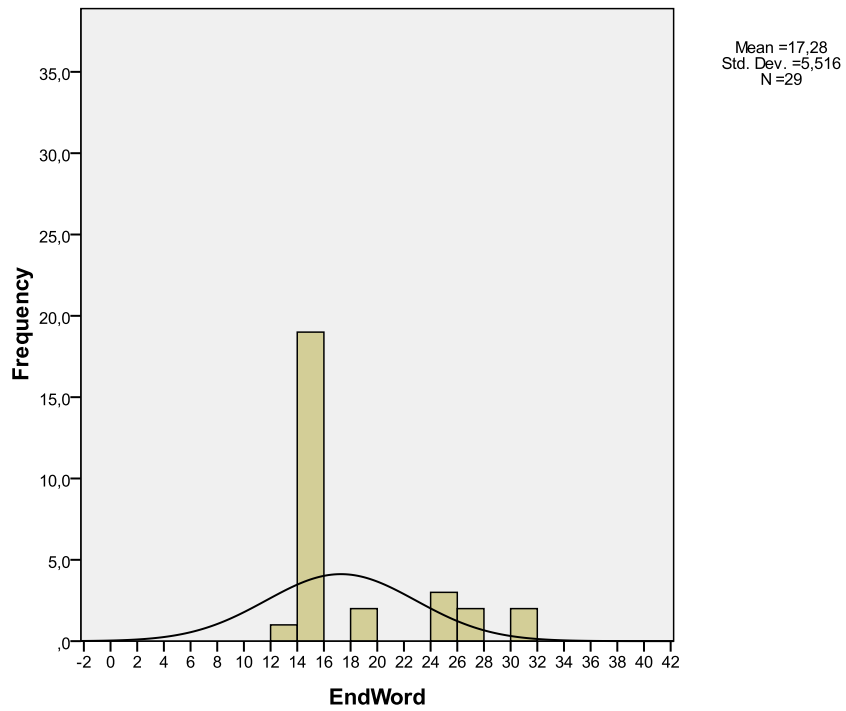


Figure 33. The frequency of end words of the participants in Group 1 for paragraph 3 with the most common end word being the 12th word.

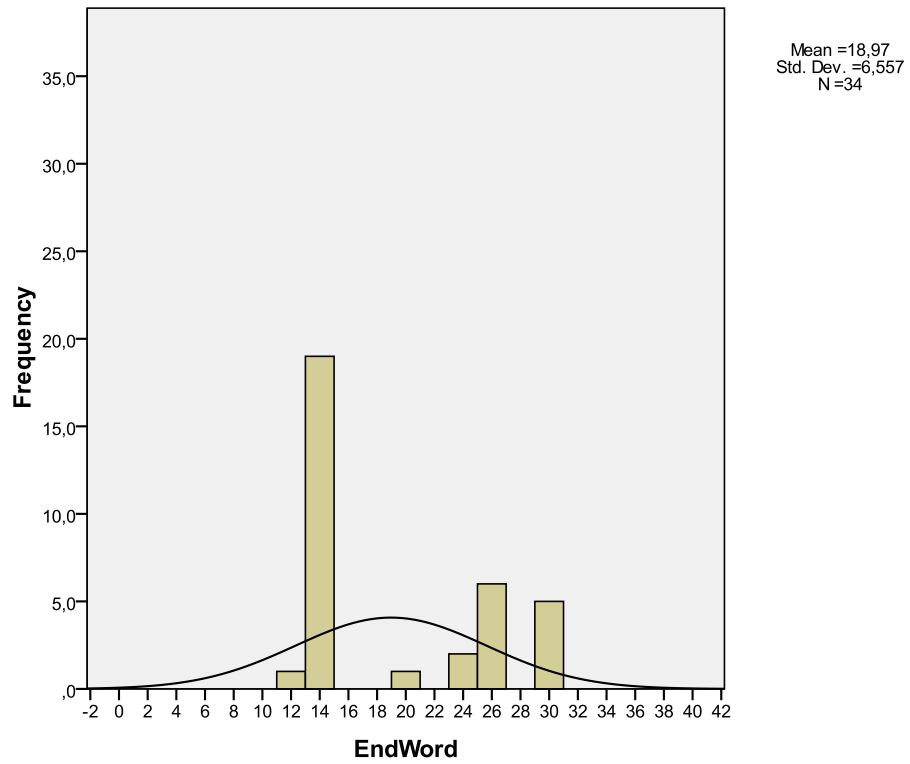


Figure 34. The frequency of end words of the participants in Group 2 for paragraph 3 with the most common end word being the 12th word.

APPENDIX 14

PASSONEAU (2005) ANALYSIS FOR EXPERIMENT 2

Experiment 2 - Paragraph 2

A: Ahmet, bir gün gazeteyi bir açmıştı, Mesut yok! Pusuya düşürülmüş oracıkta can vermişti. İki kişiymişler. Öbürünün adı Halim. Tanımıyordu. Hafif yaralı kurtulmuş o.

A1: Ahmet, bir gün gazeteyi bir açmıştı, Mesut yok! Pusuya düşürülmüş oracıkta can vermişti. İki kişiymişler. Öbürünün adı Halim. Tanımıyordu.

A2: Mesut yok! Pusuya düşürülmüş oracıkta can vermişti. İki kişiymişler. Öbürünün adı Halim. Tanımıyordu. Hafif yaralı kurtulmuş

A3: Pusuya düşürülmüş oracıkta can vermişti. İki kişiymişler. Öbürünün adı Halim. Tanımıyordu. Hafif yaralı kurtulmuş o.

A4: Ahmet, bir gün gazeteyi bir açmıştı, Mesut yok! Pusuya düşürülmüş oracıkta can vermişti.

A5: bir gün gazeteyi bir açmıştı, Mesut yok! Pusuya düşürülmüş oracıkta can vermişti.

A6: Ahmet, bir gün gazeteyi bir açmıştı, Mesut yok!

A7: Mesut yok! Pusuya düşürülmüş oracıkta can vermişti.

A8: Pusuya düşürülmüş oracıkta can vermişti.

A9: Pusuya düşürülmüş

A10: can vermişti.

A10: Mesut yok!

No Match:

- Pusuya düşürülmüş oracıkta can

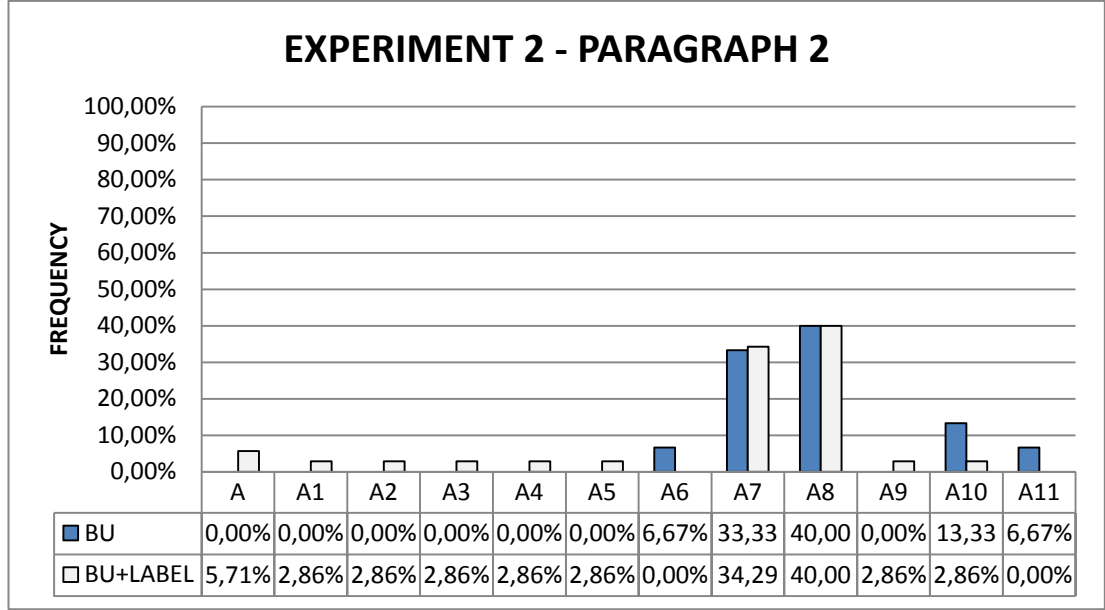


Figure 35. Pasonneau's classification of main set, subsumption, intersection and disjunct relations are illustrated. This is the whole data for Paragraph 2 from both groups. Y axis shows the frequency of the classifications. Below the graph, the percentages of each set's occurrence in Group 1 (bu) and Group 2 (Bu + label) are provided.

Experiment 2 - Paragraph 3

A: Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı hiçbir zaman içine sindirememiş olan ve bu konuda duyduğu ezikliği saklayamayan Turgay'a bu konuda sürekli takılırlar.

A1: Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı hiçbir zaman içine sindirememiş olan ve bu konuda duyduğu ezikliği saklayamayan Turgay'a

A2: Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı hiçbir zaman içine sindirememiş olan ve bu konuda duyduğu ezikliği saklayamayan

A3: Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı hiçbir zaman içine sindirememiş olan

A4: Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı hiçbir zaman içine sindirememiş

A5: gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı hiçbir zaman içine sindirememiş olan ve bu konuda duyduğu ezikliği

A6: Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı

A7: bar işletmeciliğine geçmiş olmayı hiçbir zaman içine sindirememiş olan ve bu konuda duyduğu ezikliği

A8: Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine geçmiş

A9: Yıllarca hapis yatmış militan bir komünist olarak, gizli örgüt üyeliğinden bar işletmeciliğine

A10: gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı hiçbir zaman içine sindirememiş olan

A11: gizli örgüt üyeliğinden bar işletmeciliğine geçmiş olmayı

A12: bu konuda duyduğu ezikliği

No Match:

- Yıllarca hapis yatmış militan bir komünist
- Turgay'a, bu konuda sürekli takılırlar.

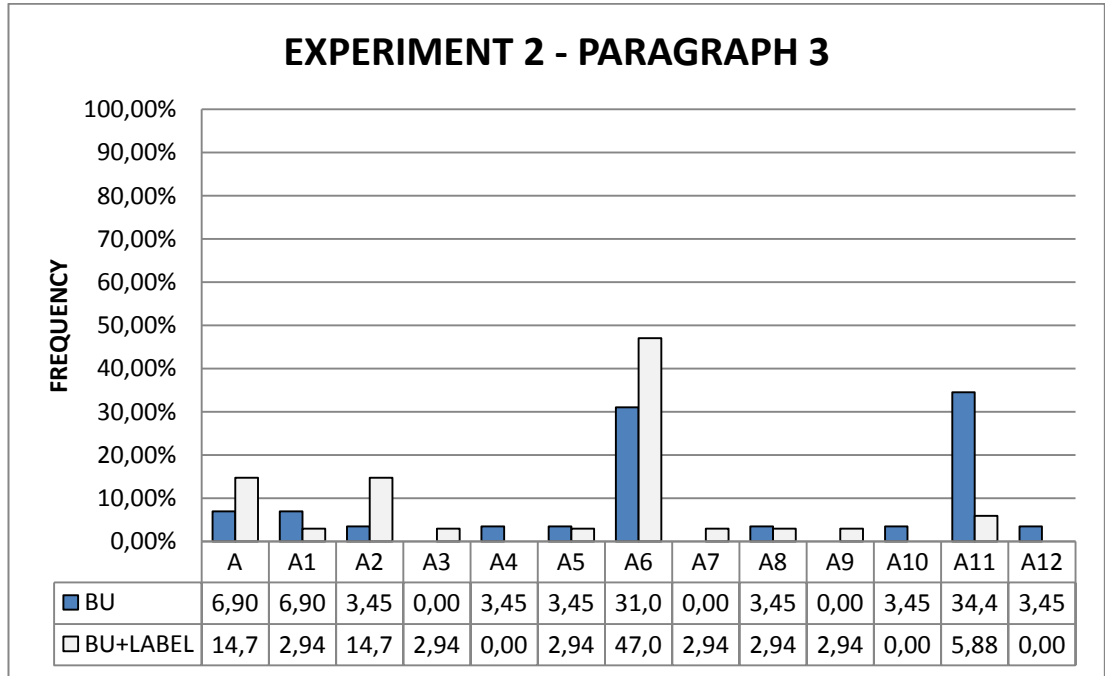


Figure 36. Pasonneau's classification of main set, subsumption, intersection and disjunct relations are illustrated. This is the whole data for Paragraph 3 from both groups. Y axis shows the frequency of the classifications. Below the graph, the percentages of each set's occurrence in Group 1 (bu) and Group 2 (Bu + label) are provided.

Experiment 2 - Paragraph 4

A: Sen ağlamıştın. Mutluluktan, yalnızca mutluluktandı göz yaşların. Sulu gözlüydün Nesli, ama itiraf etmeye utanıyordun değil mi?

A1: Sen ağlamıştın. Mutluluktan, yalnızca mutluluktandı göz yaşların.

A2: Sulu gözlüydün Nesli, ama itiraf etmeye utanıyordun değil mi?

A3: Sulu gözlüydün Nesli ama itiraf etmeye utanıyordun

A4: Mutluluktan, yalnızca mutluluktandı göz yaşların.

A5: itiraf etmeye utanıyordun

A6: Sulu gözlüydün Nesli

A7: Su gözlüydün

A8: itiraf etmeye

A9: Sen ağlamıştın

No Match:

- gizlemeye

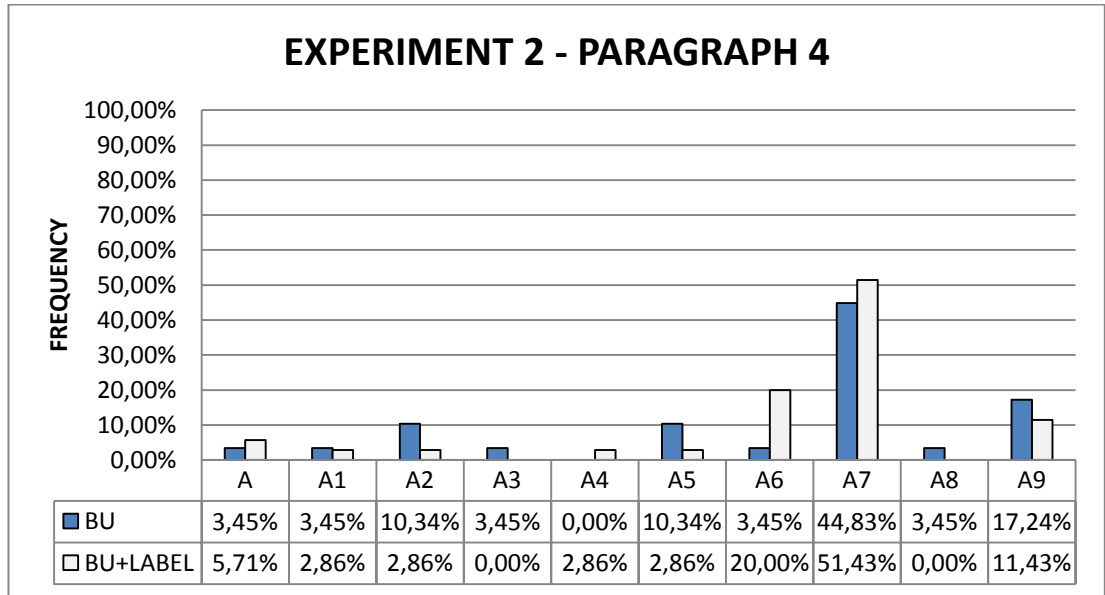


Figure 37. Pasonneau's classification of main set, subsumption, intersection and disjunct relations are illustrated. This is the whole data for Paragraph 4 from both groups. Y axis shows the frequency of the classifications. Below the graph, the percentages of each set's occurrence in Group 1 (bu) and Group 2 (Bu + label) are provided.

APPENDIX 15

FREQUENCY TABLE OF START&END WORDS IN EXPERIMENT 1

Table 18 Frequency of Start Words in Experiment 1**

Paragraph #	Group 1		Group 2	
	(Bare Anaphor)		(Label Anaphor)	
	start word	frequency	start word	frequency
1	1	0,00%	1	4,55%
	8	2,22%	8	0,00%
	13	0,00%	13	2,27%
	14	37,78%	14	45,45%
	18	44,44%	18	38,64%
	26	15,56%	26	6,82%
	27	0,00%	27	2,27%
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	start word	frequency	start word	frequency
	1	22,45%	1	17,02%
2	2	4,08%	2	6,38%
	3	4,08%	3	4,26%
	6	2,04%	6	2,13%
	14	4,08%	14	0,00%
	33	0,00%	33	2,13%
	34	2,04%	34	2,13%
	35	34,69%	35	57,45%
	37	4,08%	37	4,26%
	39	4,08%	39	0,00%
	47	6,12%	47	0,00%
	48	4,08%	48	2,13%
	50	8,16%	50	2,13%
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	start word	frequency	start word	frequency
	1	2,17%	1	8,51%
3	2	4,35%	2	4,26%
	9	67,39%	9	85,11%
	11	0,00%	11	2,13%
	13	2,17%	13	0,00%
	17	21,74%	17	0,00%
	19	2,17%	19	0,00%
	Paragraph #	(Bare Anaphor)		(Label Anaphor)
start word		frequency	start word	frequency
36		70,21%	36	86,05%
4	37	27,66%	37	13,95%
	39	2,13%	39	0,00%

Table 19 Frequency of End Words in Experiment 1

	Group 1		Group 2	
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	end word	frequency	end word	frequency
1	12	2,22%	12	4,55%
	23	66,67%	23	68,18%
	24	13,33%	24	13,64%
	31	17,78%	31	13,64%
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	end word	frequency	end word	frequency
2	40	2,04%	40	6,38%
	41	0,00%	41	2,13%
	42	48,98%	42	48,94%
	43	24,49%	43	29,79%
	45	2,04%	45	0
	46	4,08%	46	8,51%
	51	18,37%	51	4,26%
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	end word	frequency	end word	frequency
3	5	2,17%	5	4,26%
	8	2,17%	8	6,38%
	11	43,48%	11	57,45%
	12	19,57%	12	25,53%
	14	4,35%	14	0,00%
	16	2,17%	16	6,38%
	19	26,09%	19	0,00%
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	end word	frequency	end word	frequency
4	41	100,00%	41	100,00%

APPENDIX 16

FREQUENCY TABLE OF START&END WORDS IN EXPERIMENT 2

Table 20 Frequency of Start Words in Experiment 2

	Group 1		Group 2	
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	start word	frequency	start word	frequency
1	18	3,45%	18	5,56%
	20	3,45%	20	0,00%
	25	10,34%	25	25,00%
	31	82,76%	31	69,44%
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	start word	frequency	start word	frequency
2	1	6,67%	1	11,43%
	2	0,00%	2	2,86%
	6	40,00%	6	37,14%
	8	40,00%	8	45,71%
	11	13,33%	11	2,86%
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	start word	frequency	start word	frequency
3	1	55,17%	1	88,24%
	8	41,38%	8	8,82%
	11	0,00%	11	2,94%
	21	3,45%	21	0,00%
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	start word	frequency	start word	frequency
4	13	24,14%	13	20,00%
	15	0,00%	15	2,86%
	20	62,07%	20	74,29%
	24	13,79%	24	2,86%

Table 21 Frequency of End Words in Experiment 2

	Group 1		Group 2	
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	end word	frequency	end word	frequency
1	24	0,00%	24	2,78%
	34	31,03%	34	52,78%
	35	68,97%	35	44,44%
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	end word	frequency	end word	frequency
2	7	13,33%	7	0,00%
	9	0,00%	9	2,86%
	12	86,67%	12	82,86%
	18	0,00%	18	2,86%
	21	0,00%	21	2,86%
	22	0,00%	22	8,57%
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	end word	frequency	end word	frequency
3	12	0,00%	12	2,94%
	13	3,45%	13	2,94%
	14	65,52%	14	52,94%
	18	3,45%	18	0,00%
	19	3,45%	19	2,94%
	24	6,90%	24	5,88%
	25	3,45%	25	14,71%
	26	6,90%	26	2,94%
	30	6,90%	30	14,71%
Paragraph #	(Bare Anaphor)		(Label Anaphor)	
	end word	frequency	end word	frequency
4	14	17,24%	14	11,43%
	19	3,45%	19	5,71%
	21	44,83%	21	51,43%
	22	3,45%	22	20,00%
	25	3,45%	25	0,00%
	26	13,79%	26	2,86%
	28	13,79%	28	8,57%

APPENDIX 17

ANALYSIS OF THE MOST FREQUENT TYPES OF THE PHRASES MARKED BY THE PARTICIPANTS IN EXPERIMENT 1

Table 22 Types of the Most Frequent Phrases Marked by the participants for Pargraph 1 in Experiment 1

Experiment 1			
Paragraph 1 (Bu vs. Bu gerçek)			
		Group 1	Group 2
A	VP	0,00%	2,27%
A1	VP	2,22%	0,00%
A2	VP	0,00%	2,27%
A3	NP	11,11%	9,09%
A4	NP	24,44%	36,36%
A5	NP	42,22%	31,82%
A6	NP	2,22%	4,55%
A7	VP	15,56%	6,82%
A8	NP	0,00%	2,27%
B	VP	0,00%	4,55%
B1	VP	2,22%	0,00%

Table 23 Types of the Most Frequent Phrases Marked by the participants for Pargraph 2 in Experiment 1

Experiment 1			
Paragraph 2 (Bu vs. Bu durum)			
		Group 1	Group 2
A	VP	2,04%	4,26%
A1	VP	2,04%	0,00%
A2	NP	14,29%	8,51%
A3	VP	0,00%	2,13%
A4	NP	18,37%	14,89%
A5	VP	2,04%	2,13%
A6	NP	0,00%	2,13%
A7	NP	0,00%	2,13%
A8	NP	10,20%	19,15%
A9	NP	2,04%	0,00%
A10	NP	20,41%	29,79%
A11	NP	0,00%	2,13%
A12	AdjP	2,04%	6,38%
A13	NP	4,08%	2,13%
A14	NP	4,08%	0,00%
B	VP	6,12%	0,00%
B1	VP	4,08%	2,13%
B2	VP	8,16%	2,13%

Table 24 Types of the Most Frequent Phrases Marked by the participants for Pargraph 3 in Experiment 1

Experiment 1			
Paragraph 3 (Bu vs. Bu iş)			
		Group 1	Group 2
A	VP	2,17%	0,00%
A1	NP	0,00%	2,13%
A2	VP	2,17%	6,38%
A3	VP	0,00%	6,38%
A4	NP	2,17%	0,00%
A5	VP	2,17%	4,26%
A6	NP	19,57%	23,40%
A7	NP	43,48%	55,32%
A8	NP	2,17%	0,00%
A9	NP	0,00%	2,13%
B	VP	21,74%	0,00%
B1	VP	2,17%	0,00%
A4 \cap B	VP	2,17%	0,00%

Table 25 Types of the Most Frequent Phrases Marked by the participants for Pargraph 4 in Experiment 1

Experiment 1			
Paragraph 4 (Bu vs. Bu olay)			
		Group 1	Group 2
A	VP	70,21%	86,05%
A1	VP	27,66%	13,95%
A2	VP	2,13%	0,00%

APPENDIX 18

ANALYSIS OF THE MOST FREQUENT TYPES OF THE PHRASES MARKED BY THE PARTICIPANTS IN EXPERIMENT 2

Table 26 Types of the Most Frequent Phrases Marked by the participants for Pargraph 1 in Experiment 2

Experiment 2			
Paragraph 1 (Bu vs. Bu iş)			
		Group 1	Group 2
A	VP	3,45%	5,71%
A1	VP	3,45%	2,86%
A2	VP	10,34%	2,86%
A3	NP	0,00%	2,86%
A4	VP	0,00%	2,86%
A5	VP	51,72%	2,86%
A6	NP	31,03%	0,00%

Table 27 Types of the Most Frequent Phrases Marked by the participants for Pargraph 2 in Experiment 3

Experiment 2			
Paragraph 2 (Bu vs. Bu olay)			
		Group 1	Group 2
A	VP	0,00%	14,71%
A1	VP	0,00%	2,94%
A2	VP	0,00%	14,71%
A3	VP	0,00%	2,94%
A4	VP	0,00%	0,00%
A5	VP	0,00%	2,94%
A6	VP	6,67%	47,06%
A7	VP	33,33%	2,94%
A8	VP	40,00%	2,94%
A9	VP	0,00%	2,94%
A10	VP	13,33%	0,00%
A11	VP	6,67%	5,88%

Table 28 Types of the Most Frequent Phrases Marked by the participants for Pargraph 3 in Experiment 2

Experiment 2			
Paragraph 3 (Bu vs. Bu durum)			
		Group 1	Group 2
A	VP	6,90%	14,71%
A1	PP	6,90%	2,94%
A2	AdjP	3,45%	14,71%
A3	AdjP	0,00%	2,94%
A4	AdjP	3,45%	0,00%
A5	NP	3,45%	2,94%
A6	NP	31,03%	47,06%
A7	NP	0,00%	2,94%
A8	VP	3,45%	2,94%
A9	PP	0,00%	2,94%
A10	AdjP	3,45%	0,00%
A11	NP	34,48%	5,88%
A12	NP	3,45%	0,00%

Table 29 Types of the Most Frequent Phrases Marked by the participants for Pargraph 4 in Experiment 2

Experiment 2			
Paragraph 4 (Bu vs. Bu gerçek)			
		Group 1	Group 2
A	VP	3,45%	5,71%
A1	VP	3,45%	2,86%
A2	VP	10,34%	2,86%
A3	VP	3,45%	0,00%
A4	VP	0,00%	2,86%
A5	VP	10,34%	2,86%
A6	VP	3,45%	20,00%
A7	VP	44,83%	51,43%
A8	NP	3,45%	0,00%
A9	VP	17,24%	11,43%

TEZ FOTOKOPİ İZİN FORMU

ENSTİTÜ

Fen Bilimleri Enstitüsü	<input type="checkbox"/>
Sosyal Bilimler Enstitüsü	<input type="checkbox"/>
Uygulamalı Matematik Enstitüsü	<input type="checkbox"/>
Enformatik Enstitüsü	<input checked="" type="checkbox"/>
Deniz Bilimleri Enstitüsü	<input type="checkbox"/>

YAZARIN

Soyadı : ERGİN SOMER
Adı : RABİYE
Bölümü : BİLİŞSEL BİLİMLER

TEZİN ADI (İngilizce) : **AN EXPERIMENTAL STUDY ON ABSTRACT ANAPHORA RESOLUTION IN TURKISH WRITTEN DISCOURSE**

TEZİN TÜRÜ : Yüksek Lisans Doktora

1. Tezimin tamamı dünya çapında erişime açılsın ve kaynak gösterilmek şartıyla tezimin bir kısmı veya tamamının fotokopisi alınsın.
2. Tezimin tamamı yalnızca Orta Doğu Teknik Üniversitesi kullanıcılarının erişimine açılsın. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.)
3. Tezim bir (1) yıl süreyle erişime kapalı olsun. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.)

Yazarın imzası

Tarih