SUBJECT AGREEMENT IN -DIK CLAUSES IN THE HISTORY OF TURKISH

BURAK IŞIK

BOĞAZİÇİ UNIVERSITY

2024

SUBJECT AGREEMENT IN -DIK CLAUSES IN THE HISTORY OF TURKISH

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts in

Linguistics

by

Burak Işık

Boğaziçi University

2024

Subject Agreement in -DIK Clauses in the History of Turkish

The thesis of Burak Işık

has been approved by:

Assist. Prof. Metin Bağrıaçık (Thesis Advisor) _____

Prof. Balkız Öztürk Başaran _____

Prof. Mehmet Ölmez (External Member) _____

April 2014

DECLARATION OF ORIGINALITY

I, Burak Işık, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature…….....
Date …….....

ABSTRACT

Subject Agreement in -DIK Clauses in the History of Turkish

In this thesis, subject case and subject agreement in *-DIK* argument subordinate clauses are explored throughout the history of Turkish. Definiteness hierarchy is the determining factor in the genitive marking of the overt internal subjects of complement and non-subject relative clauses in Modern Turkish: specific subjects bear genitive whereas non-specific ones appear in the nominative. This state of affairs was not the same across the history of Turkish. The quantitative investigation of a diachronic corpus that I created manually reveals that, while genitive marking was overall lower, the animacy hierarchy played a crucial role in the subject case of *-DIK* subordinate clauses until the eighteenth century. At this point definiteness hierarchy became the prominent factor. This diachronic change was discussed in the light of an acquisition-based model to language change (Andersen, 1973; Lightfoot, 1976). In particular, it is suggested that due to parametric ambiguity and lack of sufficient data, the animacy hierarchy, which is the predictor in genitive marking, was reanalyzed as definiteness hierarchy by a group of language acquirers around the eighteenth century. In addition, the thesis also investigates diachronic changes in the nominal agreement on the verb of subordinate *-DIK* clauses. The thesis shows how often and since when nominal agreement has been used in all subordinate *-DIK* clause types. For each subordinate *-DIK* clause type, the increase or decrease in the existence of the nominal agreement has been provided numerically.

ÖZET

Türkçenin Tarihi Boyunca –DIK Ekli Yan Tümcelerdeki Özne Uyumu

Bu tezde *-DIK* ekinin kullanıldığı yan tümcelerde özne durumu ve özne uyumu Türkçenin tarihi boyunca incelenmiştir. Modern Türkçede tümleç yan tümcelerinde ve ilgi tümcelerindeki açık öznelerin ilgi eki ile işaretlenmesinde belirlilik hiyerarşisi belirleyici etkendir: Özgül özneler ilgi ekini taşırken, özgül olmayanlar işaretsiz, yalın durumda görünür. Bu durum Türkçenin tarihi boyunca aynı değildir. İlgi ekiyle işaretleme bu dönemde genel olarak daha az olmasına rağmen manuel olarak oluşturduğum art zamanlı derlemin nicel incelemesi, *-DIK* ekinin kullanıldığı yan tümcelerin öznelerinin durumunun belirlenmesinde on sekizinci yüzyıla kadar canlılık hiyerarşisinin çok önemli bir rol oynadığını ortaya koymaktadır. Bu noktada belirlilik hiyerarşisi önde gelen faktör haline gelmiştir. Bu art zamanlı değişim, dil değişimine yönelik edinim temelli model (Andersen, 1973; Lightfoot, 1976) ışığında tartışılmaktadır. Özellikle parametrik bulanıklık ve yeterli verinin eksikliği nedeniyle, ilgi ekiyle işaretlemede yordayıcı olan canlılık hiyerarşisinin, on sekizinci yüzyılda bir grup dil edinicisi tarafından belirlilik hiyerarşisi olarak yeniden analiz edildiği ileri sürülmektedir. Ayrıca bu tezde, *-DI*K ekinin kullanıldığı yan tümcelerin fiillerinde görülen adsıl uyumdaki art zamanlı değişim de incelenmektedir. Tez, *-DIK* ekinin kullanıldığı tüm yan tümce türlerinde adsıl uyumun ne sıklıkta ve ne zamandan beri kullanıldığını gözler önüne sermektedir. Her *-DIK* ekli yan tümce türü için adsıl uyumun varlığındaki artış veya azalış sayısal olarak sunulmuştur.

ACKNOWLEDGMENTS

This thesis could have not been written without the valuable input of many dear scholars of Boğaziçi University. Among them, I would first like to thank my thesis advisor, Metin Bağrıaçık, for believing in me, encouraging me, and showing me the tolerance and patience throughout my thesis writing period. I benefitted a lot from his immense knowledge about diachronic syntax and I could not have finished this thesis without him. Then, Balkız Öztürk Başaran deserves heartfelt thanks for believing in my potential and for encouraging me to pursue a career in linguistics. She has never ceased to support me against every hurdle I encounter. I am also very grateful to Ceyda Arslan Kechriotis for introducing me to linguistics in the first place, and for encouraging me to continue learning more about this field from the first years of my undergraduate education. Elena Guerzoni was always there, and was patiently helping me with all the bureaucratic problems I encountered throughout my MA, for which I am very grateful to her. Mehmet Ölmez provided me with valuable information about Old Turkic and pointed out to me numerous historical texts that I used in the thesis. I would like to thank him for these, and for kindly accepting to be a member of my thesis examination committee. I would like to thank Ümit Atlamaz for guiding me on many academic problems, especially on those pertinent to theoretical syntax, and Ömer Demirok for the moral support he gave me, especially during the COVID period. Didar Akar taught me to look at diachronic data from a broader perspective, always keeping in mind the sociolinguistic ecology of them, which significantly increased the quality of this thesis. Thank you, Didar hocam, for this, and for making me part of your research project. I would like to thank Stefano Canalis for introducing me to historical linguistics during my undergraduate studies,

which paved the way to this thesis. Finally, I would like to thank Onur Keleş for sharing his expertise with me in manipulating and interpreting the numerical data, which has made the thesis a lot more meaningful.

And then this thesis could have not been written if the moral support of the three very dear people had not been there. First, I would like to thank parents, Şükran Işık and Mikdat Işık, for their endless support while I pursue my dreams. Then, I would like to thank my brother, my best friend, and my roommate, İhsan Işık, who was always there for me.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1

1.1 Aim of the thesis 1

1.2 Conclusions in a nutshell 4

1.3 The structure of the thesis 6

CHAPTER 2: PERIODIZATION AND DATA SOURCES 8

2.1 Periodization 8

2.2 Data sources 13

2.3 A note on the choices of texts 26

CHAPTER 3: DATABASE 27

3.1 Extracting the data 27

3.2 Classification 27

3.3 Metadata 30

3.4 Construction-specific criteria 32

3.5 Additional information for criteria 35

3.6 Examples excluded from the data 38

CHAPTER 4: COMPLEMENT CLAUSES 43

4.1 Introduction 43

4.2 Complement clauses in Modern Turkish 43

4.3 Incorporation in complement clauses 44

4.4 Setting the scene for historical investigation: The preamble 46

4.5 The overall numbers 47

4.6 The verbs selecting -*DIK* complement clauses 49

4.7 Other criteria 50

4.8 Sorting out overt internal subjects 50
4.9 Change in the ratio of genitive-subjects 51
4.10 Chapter summary 54
CHAPTER 5: HEADED RELATIVE CLAUSES 55
5.1 Headed relative clause in Modern Turkish: The preamble 55
5.2 Headed relative clauses within historical data: An overview 58
5.3 Historical data: Raw numbers 61
5.4 The overall change in genitive marking. 67
5.5 Chapter summary 68
5.6 A note on headless relative clause 69
CHAPTER 6: DIACHRONIC DISCUSSION 71
6.1 Taking stock 71
6.2 The effect of animacy in genitive-marking in headed relative clauses 73
6.3 The effect of animacy in genitive-marking in complement clauses 76
6.4 Interim summary 78
6.5 What has (not) changed? 87
6.6 Changes in numbers: Summary 93
6.7 Changes in minds 94
6.8 What possibly happened? 95
CHAPTER 7: ADVERBIAL CLAUSES 99
7.1 Overview 99
7.2 Temporal clauses 99
7.3 Reason clauses 109
7.4 Chapter summary 114
CHAPTER 8: CONCLUSION 116

8.1 Findings and summary 116
8.2 Points for further research 120
APPENDIX A: DATA SOURCES ACROSS THE CENTURIES 123
APPENDIX B: THE DISTRIBUTION OF TOTAL NUMBER OF EXAMPLES ACCORDING TO THE TEXTS 125
REFERENCES 127

LIST OF TABLES

Table 1. Total Number of Examples According to Different Clause Types 29

Table 2. The Distribution of Total Number of Examples across Centuries 30

Table 3. Additional Criteria for Complement Clauses 33

Table 4. Additional Criteria for Relative Clauses 34

Table 5. Additional Criteria for Adverbial Clauses 34

Table 6. The Overall Number of Complement Clauses across Centuries 48

Table 7. The Most Frequent Matrix Verbs in Complement Clauses 50

Table 8. The Distribution of the Subject Marking in Complement Clauses across the Centuries 51

Table 9. The Increase of Genitive-Subjects in Complement Clauses (in %) 52

Table 10. The Ratio of Genitive-Subjects in Complement Clauses without SNK.... 53

Table 11. Headed Relative Clauses across the Centuries 61

Table 12. The Distribution of Non-Subject Relative Clauses with Nominal Agreement in Embedded Clauses 62

Table 13. The Percentage of Genitive Case Usage across the Centuries 63

Table 14. The Distribution of Non-Subject Relative Clauses without Nominal Agreement across the Centuries 64

Table 15. Models for Likelihood Ratio Test 74

Table 16. Result of Likelihood Ratio Test (Relatives) 74

Table 17. The lme Result of the Best Model for Relative Clauses 75

Table 18. Likelihood Ratio Test for Complement Clauses 76

Table 19. The lme Result of the Best Model for Complement Clauses 77

Table 20. The Effect of Definiteness Hierarchy on Object-Marking Cross-Linguistically 82

Table 21. The Effect of Animacy Hierarchy on Object-Marking Cross-Linguistically 83

Table 22. The Effect of Definiteness Hierarchy on Subject-Marking Cross-Linguistically 86

Table 23. The Distribution of Temporal Clauses across the Centuries 102

Table 24. The Distribution of Temporal Clauses with Postposition *sonra* and Case Suffix across the Centuries..... 104

Table 25. The Distribution of Temporal Clauses with Noun Head across the Centuries 106

Table 26. The Distribution of Temporal Clauses with *-CA* across the Centuries.... 108

Table 27. The Distribution of Reason Clauses across the Centuries..... 111

Table 28. The Distribution of Reason Clauses with Postposition *üçün>içün/-çUn> için* across the Centuries..... 113

Table 29. The Distribution of Reason Clauses with Case Suffix across the Centuries 114

LIST OF FIGURES

Figure 1. Excel sheet for complement clauses 31

Figure 2. Excel sheet for headed relative clauses 31

Figure 3. Relative proportion of genitive marked subjects to nominative marked ones across the subcorpora in chronological order (thirteenth-seventeenth centuries) 52

Figure 4. Increase in the proportion of GEN-marked subjects from thirteenth to seventeenth centuries. 53

Figure 5. Relative proportion of genitive marked subjects to nominative marked ones across the subcorpora in chronological order (thirteenth-twenty-first centuries). 54

Figure 6. Increase of genitive-marking in relative clauses across the centuries 67

Figure 7. Change in genitive vs nominative in both complements and relatives..... 72

Figure 8. Animacy effect in relative clause for genitive-marking 73

Figure 9. Predicted probabilities of genitive by century and animacy in relative clauses 75

Figure 10. Animacy effect in complement clauses for genitive-marking 76

Figure 11. Predicted probabilities of genitive by century and animacy in complement clauses 78

Figure 12. The proportion of genitive marked subjects of headed relatives when an intervening maximal projection appears between it and the verb 90

Figure 13. The proportion of genitive marked subjects of headed relatives when there is no intervening maximal projection between it and the verb 91

Figure 14. The proportion of genitive marked subjects of complements when an intervening maximal projection appears between it and the verb 92
Figure 15. The proportion of genitive marked subjects of headed relatives when there is no intervening maximal projection between it and the verb. 93

ABBREVIATIONS

1 Abbreviations Used in the Glosses

Ø	zero morpheme
1	first person
2	second person
3	third person
ABL	ablative
ACC	accusative
AGR	agreement
DAT	dative
DO	definite object
FN	finitive marker
FUT	future
GEN	genitive
IMP	imperative
LOC	locative
NEG	negation
NOM	nominative
OPT	optative
PAST	past
PL	plural
POSS	possessive
PRS	present
PROG	progresive

PSS	passive
SG	singular
V	verb
XP	maximal projection

2 Abbreviations Used in the Text

Ø	zero morpheme
Asp P.	Aspectual Phrase
CE	Common Era
cf.	confer
DOM	Diffrential Object Marking
DP	Determiner Phrase
DSM	Differential Subject Marking
e.g.	exempli gratia
etc.	Et cetera
ID	identification
Mood P.	Mood Phrase
NP	Noun Phrase
PLD	Primary Language Data
TAM	Tense- Aspect -Mood
TP	Tense Phrase
V	Verb
VP	Verb Phrase

3 Abbreviations Used in the Manuscripts

AY	Altun Yaruk
BH	Bahrü'l-Hakayık
BHMH	Behcetü'l-Hadâik fî Mev'izati'l-Hâlaik
BQ	Bilge Qaghan Stele
BOUN	BOUN Web Corpus
BögüK	Bögü Qaghan's Acceptance of Maniheism
Dasa	Dasakarmapathavadanamala
EATM	An Ancient Anatolian Text Dated 1343
EM	Edviye-i Müfrede
GB	Grammar-book treebank
GN	Garibname
GT	Gulistan Translation
Hua	Huastuanift
HR	Hilyet'ür-Rical
IB	Irg Bitig
IMST	ITU-METU-Sabancı Treebank
İZG	İkramiyeli Gazete
KBD	Kadı Burhanettin's Divan
KÇ	Küli Çor Stele
KBP	Kitab-ı Baz-name-i Padişahi
KG	Kitab-ı Güzide
KE	Kısas-ı Enbiya
KENET	Turkish WordNet
KT	Kul Tigin Stele

Mai	Maitrisimit
MN	Mecmü'atü'n-nezair
MRB	Marzubanname
MŞ	Muntahab-ı Şifa
O	Ongi Stele
OKD	Harita-ı Kapudanan-I Derya
PENN	Penn Treebank
PUD	Parallel Universal Dependencies
RB	Revnak-ı Bustan
SNB	Süheyl ü Nevbahar
SNK	Süleyman-name-i Kebir: 6th and 7th volume
SVTM	Sultan Veled's Turkish Poems
Şİ	Şerh-i İbarat
Tar	Taryat Stele
Tes	Tes Stele
TMB	Tuhfe-i Mubarizi
TMR	Tuhfe-i Muradi
TOA	Captivity of Osman Agha of Timişoara
TT	Takvimü’t-Tevarih
UD	Universal Dependencies for Turkish
VG	Varka and Gülşah
Xuan	Xuanzang-Biographie
YZ	Yusuf and Zelih

CHAPTER 1

INTRODUCTION

1.1 Aim of the thesis

Modern Turkish, and several Turkic languages in general, flag the internal overt subjects

(a) of situational and propositional complement clauses (in the sense of Wurmbrand & Lohninger, 2019),

(b) of non-subject headed and headless relative clauses, and

(c) of a large array of adverbial clauses, e.g., reason clauses with the postposition *için* 'because' or temporal clauses with the locative *-DA* 'when'

with nominal agreement markers on their necessarily non-finite verb, which, in Modern Turkish may be marked in *-DIK* or *(-y)AcAK*, depending on the temporal specification of the subordinate clause (see Göksel & Kerslake 2005, p. 365-370, 383-387, 404-406; see Johanson, 2021, chapter 55 for an overview of Turkic in general).[1] These nominal agreement markers are identical to the ones in the possessive paradigm. Additionally, while specific/partitive internal subjects of complement and relative clauses are marked with the genitive in Modern Turkish, the subject internal to an adverbial clause appears regularly in the nominative (Underhill, 1972; Hankamer, 1972 and numerous subsequent studies).[2] These facts are illustrated for embedded clauses with *-DIK* in (1)-(3):

[1] Capital vowels denote archephonemes whose surface value is based on well-defined vowel harmony rules. Capital consonants show ether voicing harmony (for coronals) or deletion (for the dorsal /K/).
[2] This, on the surface value, is a robust generalization. The genitive may appear on a number of adverbial clauses, which, however, were ultimately analyzed as hidden relatives (e.g., Aygen 2007).

(1) Complement clause
[sen-in ben-i sev-diğ-in]-I düşün-dü-m.
you-GEN I-ACC love-DIK-2SG-ACC think-PAST-1SG
"I thought you loved me."

(2) (Headed) relative clause
[Muhasebeci-nin gönder-diğ-i] evrak nerede?
accountant-GEN send-DIK-3SG document where
"Where is the document the accountant has sent?"

(3) Adverbial clause (of reason)
[[Kapıcı kapı-yı kilitle-diğ-i] için] bina-ya
doorman door-ACC lock-DIK-3SG because building-DAT
gir-e-me-di-m.
enter-ABL-NEG-PAST-1SG
"I couldn't enter the building because the doorman locked the door."

The nominalizer *-DIK* is claimed to appear on factive subordinate clauses (Lees, 1965; Kornfilt, 1984, as cited in Demirok, 2018). While the nominal agreement appears on other types of subordinate clauses involving different subordinate suffixes, apart from *–(y)AcAK*, this study focuses on *-DIK* clauses only.

A large list of studies have been carried out about the internal and external syntax of the subordinate *-DIK* clauses in Modern Turkish, which are couched upon different theoretical frameworks (e.g. Kornfilt, 2003, 2014; Haig, 1997; Hankamer & Knecht 1976; Barker, Hankamer & Moore 1990), some of them further offering a comparison across different Turkic languages (e.g. Aygen, 2002a; Kornfilt, 2001, 2008; Satik, 2020). Independent of their claim and conclusions, these studies are all analyses of the synchronic facts, namely investigation of the state of affairs coextensive with the time of studies. The only exception to this, as far as we are aware, is Kornfilt (2014), in which the author provides a diachronic syntactic investigation of a subset of these clause types, couched within the framework of the generative grammar of Chomskyan tradition. Leaving aside her conclusions with respect to the evolution of the syntactic structure of the subset of clause types in Turkish, one major flaw in Kornfilt's (2014) study is that she reports individually

selected examples from the historical stages she chooses to verify her conclusions. This creates the risk of the so-called 'confirmation bias' (Gard & McGillivray, 2017; Kroeber & Chrétien, 1937), i.e., the risk that instances contradicting the hypothesis may not be reported. This is a logical outcome of the 'example-based approach' to diachronic syntax she assumes: instead of looking at large historical corpora, she singles out individual examples cited in secondary sources. There are also numerous historical studies available, whose aim is to describe the individual synchronic stages of the agreement pattern in the history of Turkish without focusing on the diachronic change on these (e.g. Erdal, 2004; Timurtaş, 2010; Tekin, 2003). Such studies with philological flavor do not seem to suffer from this problem to the same extent; however, they also offer no link between individual stages, i.e., how the synchronic syntactic facts follow from earlier changes. This, we assume, largely follows from the disinterest of theoretical analyses in such studies, combined with lack of large corpora available for scholars.

In this thesis, I aim to bridge the gap between the philological and the theoretical literature on the phenomena of (a) nominal agreement on *-DIK* subordinate clauses in the history of Modern Turkish and (b) the appearance of the genitive on subtypes of clauses, trying to alleviate the methodological flaws of one and to enrich the other with a theoretical framework. In particular, to understand the diachronic trajectory of the nominal agreement on *-DIK* subordinate clauses and the marking of genitive on their subjects I rely on extensive data I gather from a historical corpus I have created and enriched manually since 2020. This is followed with a qualitative analysis of the data, which by assumption reflects the linguistic facts of their respective era. When necessary, this discussion is enriched with

theoretical observations couched within the generative framework as it is interpreted and developed by Chomsky (1981, 1986b, 1995 a.o.).

1.2 Conclusions in a nutshell

The historical corpus I have created involves a relatively large number of written literary and non-literary prose texts, as well as some poetry, from the seventh century to twenty-first century (see Section 2.2 for the data sources). Obviously, however, it does not include *all* the edited and non-edited texts written in Turkish/Turkic. Notwithstanding, the information I gathered from the texts captures the typical situation for each stage and at least offers sound hypotheses that should –hopefully– be checked and rechecked in the future with larger corpora. As in all historical analyses, therefore, my conclusions should be seen not as absolute claims but hypotheses, which, however — unlike previous studies — were built upon a sound numerical and qualitative analysis. With this caveat in mind, the conclusions reached in this thesis can be summarized as follows:

i) *-DIK* appears on the relevant subordinate clause types since the earliest attestation of Turkic.

ii) The nominal agreement marker on *-DIK* complement clauses and relatives was established very early on, suggesting tentatively that it was a property of Proto-Turkic.

iii) Genitive vs. nominative (Ø) alternation on the subjects of *-DIK* complement clauses and relative clauses is observed since the earliest texts; however, while the proportion of genitive-marked subjects to nominative-marked subjects was about one-thirds from the earliest attestations, there is almost a 10% slope in their increase up to the nineteenth century. In the eighteenth century the number of genitive-marked

subjects increase exponentially, reaching almost 80% in the time, showing us a typical logistic curve. The increase started in relative clauses first and expanded to complement clauses later, with a lapse of about three generations. Genitive-marked subjects increase in both construction types in a constant rate; thereby showing Constant Rate Effect (Kroch, 1989).

iv) Although the total percentage of genitive-marked subjects in complement and relative clauses were low before the eighteenth entury, the genitive/nominative alternation was not stochastic but animacy played a crucial role. A subject NP with an animate referent always had a much higher chance to be genitive marked than a subject NP with an inanimate referent. Therefore, we conclude that animacy hierarchy was a crucial defining factor in genitive-nominative alternation throughout the large portion of the history of Turkish.

v) The animacy factor ceases to be a significant one for genitive vs. nominative alternation in complement clauses around the nineteenth century. Although its effect can still be seen in complement clauses in Modern Turkish today, a major predictor comes into scene then, namely specificity. In other words, while the predictive force of animacy decreases, the predictive force of specificity increases steadily, without, however, replacing the former completely. The role of animacy on genitive vs. nominative alternation can be taken as an atropied property in Modern Turkish, inherited from its earlier stages.

vi) We interpret (v) in the light of an acquisition-based model to language change (Andersen, 1973; Lightfoot, 1976 and much subsequent work). According to our hypothesis, the parametric ambiguity in the input data lead at least a set of acquirers to interpret the crucial factor as specificity, instead of animacy, for marking subjects in the genitive or in the nominative. Simplifying grossly here, some individuals of a

generation of learners probably reinterpreted animacy hierarchy as definiteness hierarchy, the interaction between which is known since at least Comrie (1979).

vii) Adverbial clauses have been a stable configuration type with respect to the case marking on their internal subjects: genitive never appears on their subjects throughout the history, which lends support to the idea that there is a sharp adjunct vs. complement distinction in Turkish (Kornfilt, 2003, 2008; Aygen, 2007; Laszakovits, 2013).

viii) The nominal agreement marker on adjunct clauses, however, is a different issue. By focusing on only two temporal clauses with *-DIK*, namely those that further carry *-DA* 'when' and *-DAn soNra* 'after', we show that agreement markers became obligatory, in the former, or disappeared completely, in the latter, only after the beginning of the twentieth century — a rather recent innovation, which is probably a side-effect of the language reform.

1.3 The structure of the thesis

The structure of this thesis is as follows: Chapter 2 discusses and justifies the periodization of the history of Turkish adopted in this thesis and provides an overview of the textual sources used for the corpus. Chapter 3 provides detailed information about the methodology of annotation adopted in the thesis and offers an overview of the databases.

Chapter 4 is on the complement clauses to verbal predicates in the history of (Modern) Turkish. Here, we present information on the structure of the complement clauses and offer raw numerical numbers on different permutations. Chapter 5 repeats the same for headed relative clauses. In Chapter 6 we discuss the findings of the previous two chapter and offer an analysis based on 'language-change-as-an

acquisition-problem' view. In Chapter 7 we look at the adverbial clauses. Chapter 8 concludes the thesis.

CHAPTER 2

PERIODIZATION AND DATA SOURCES

2.1 Periodization

This chapter is devoted to the periodization of Turkish/Turkic and the sources used in collecting data. It should be noted at the outset that the sources from especially the earlier periods may not fully reflect the complete picture of the language used in the specific era, like all other written languages (Ergin, 2009). However, making the best use of bad data (Labov, 1994, p. 11) is the only way of doing historical linguistics and, we follow the suit.

2.1.1 Old Turkic

The earliest sources of the Turkic language known were written in the Orkhon valley, in today's Mongolia, at the beginning of the eighth century. These are memorial installations, and they belong to the period of the Second Turkic Khanate, whose ruling elite refer to themselves as 'Kokturks (682-744).[3] The language has its own writing system, called the Old Turkic script. The collapse of the Second Turkic Khanate was followed by a number of smaller khanates with Turkic-speaking elites, which were established by the Uyghurs.

The first of these is referred to as the Third Turkic Khanate or the Orkhon Uyghur Khanate (682-742). Some steles, written in the Old Turkic script also survive from this period. It should be noted here that Uyghurs mentioned so far in this thesis is not be confused with the Uyghurs living today in Xinjiang region, of China. These

[3] The First Turkic Khanate, which is the first geopolitical entity that uses the name 'Turk' existed between sixth and seventh centuries. This is kept out of the discussion given no texts in Turkic survive from this period (Ölmez, 2015).

two peoples are ethnolinguistically do not form a homogenous category (Erdal, 2004, p. 7) Texts produced during the Orkhon Uygur Khanate are also written in the Old Turkic script. Apart from the Orkhon Uygur Khanate, two more Uyghur Khanates were established, the first, Qocho Uyghur Khanate in today's Xinjiang region (843-1209) and the second, namely the Ganzhou Uyghur Khanate in today's Ganzhou region (894–1036). Throughout the Uyghur Khanates period, the Turkic speaking population exercised several different religions such as Manichaeism, Buddhism, Nestorianism, and Islam, and importantly they produced a number of religious texts, notably translations into (Uyghur) Turkic. The later of these texts also adopt a new writing system, namely the Old Uyghur alphabet, an adaptation of the Aramaic alphabet.

The period from the first attestation of Turkic in the eighth century until eleventh century, is uniformly referred to as Old Turkic (Tekin, 1976; Gülensoy, 2010; Tezcan, 1978; Gabain, 2007, as cited in Berbercan, 2013).[4] The period is further partitioned into two. The first one is called the Kokturk period, and it covers the eighth and ninth centuries, in which written texts dating back to the Second Turkic Khanate period are found. The second period is the Old Uyghur period, which includes sources belonging to the Uyghurs and covers the ninth and eleventh centuries. Following the literature, we refer to this period as the Old Turkic period.

[4] Some scholars do accept texts from the later period, namely those produced during the Karakhanid khaganate (ninth-thirteenth centuries) also to be part of Old Turkic (Banguoğlu, 1964, 2007; Erdal, 2004; Caferoğlu, 1974, as cited in Berbercan, 2014). We do not take into consideration the texts from this region and the period because there are also a large number of scholars who do not accept this hypothesis.

2.1.2 Anatolian Turkish

Turkic after the Old Turkic period is generally divided into variants of three zones. The first one, namely East Turkic, is claimed to be the Turkic recorded in Central Asia, including Chagatai Turkic. The second, North Turkic, was recorded in the area covering the northern shores of the Caspian Sea to the north of Black Sea; the most notable member of this is Kipchak Turkic. The third one, i.e., West (Oghuz) Turkic refers to Turkic recorded in the Western Asia, all the way to Anatolia and Balkans.

Turkic speaking people began to move towards Anatolia in late twelfth century. The majority of those who settled in Anatolia are known to be Oghuz Turks, a nomenclature referring to a group of Turkic-speaking people, which were already referred to in the history in earlier texts. Following the Turkic influx into Anatolia in the twelfth century, the language used in the majority of the Turkic-language texts produced in this area until the rise of Modern Turkish with the beginning of the twentieth century is referred to as Anatolian Turkish/ Old Turkish/ Ottoman Turkish.

As we mentioned above, texts belonging to the Old Turkic period in the eighth-eleventh centuries were written in present-day Mongolia, the Xinjiang and Guangzhou regions in China, and Eastern Central Asia. Texts belonging to the Anatolian Turkish period are the products of a written language centered in Anatolia. There are no written texts produced in the Western Turkic zone, neither on Anatolia nor in its vicinity before the thirteenth century. It is for this reason that there are varying opinions about whether a parent-child link can safely be established between the language in the last texts in the Old Turkic period (eleventh century) and the language of the earliest texts in the Anatolian Turkish period (thirteenth century). Kornfilt (2009) and Berbercan (2013) state that Old Turkic is the ancestor of East Turkic. Kornfilt further claims that the texts that can be considered the ancestors of

Anatolian Turkish should be accepted as texts written in Anatolia after the thirteenth and fourteenth centuries.

On the other hand, there are many scholars who hold an opposing view. Ercilasun (2004) stated that Old Turkic can be considered the ancestor of Anatolian Turkish and that the language of the sources of the Old Turkic period is the predecessor of the written languages that emerged later. Additionally, Korkmaz (1974) determined in detail the effects of Oghuz language on the written language of the Old Turkic period. She states that, rather than reflecting the language of a single community, Old Turkic, is the common written language of different dialects in a certain area. She claims that the effect of Oghuz people should also be taken into consideration on the Old Turkic written language. Additionally, Ergin (2009) stated that 'Old Turkic' is the common written language used in Central Asia. He argues that the division into different branches emerged after Old-Turkic period. Erdal (2004, p. 10) emphasized that Old Turkic is not the same as Proto-Turkic. He says that none of the Modern Turkic languages have a direct connection with Old Turkic. However, the linguistic features found in Old Turkic are partially similar to the features in all subsequent sub-branch languages. For this reason, Erdal claimed that the ancestors of Modern Turkic languages were not very different from Old Turkic. According to him, although a Modern Turkic language is seen as quite different from Old Turkic, it should be similar to Old Turkic in the early periods. The difference must have occurred in later periods.

The question, whether there is a direct parent-child link between Old Turkic and Old Anatolian Turkish, in our opinion, can only be resolved after the culmination of thorough diachronic studies of individual linguistic phenomena. In the absence of such works, and historical texts, speculations are inevitable. In this study, we decided

to add sources from the Old Turkic period because the diachronic changes of embedded clauses with the *-DIK* suffix may help in discussion mentioned above.

Anatolian Turkish is divided into Old Anatolian Turkish and the Classical Period. The Old Anatolian Turkish period starts in the thirteenth century and continues up to the fifteenth century, when the Ottoman Empire began to dominate across the entire Anatolian peninsula as well as the Balkans. The period after the fifteenth century can be called the Classical period, which arose in the circles of the Ottoman elites.

An important aspect of the texts from the Old Anatolian Turkic period is that they are claimed to be 'bilingual texts', referred to as *olga-bolga* texts, after the variation in the orthography of the verb 'to be'. The same texts record both *bol-* and *ol-* form of the word; however, *bol-* variant is associated with East Turkic and *ol-* form with the West. Some scholars think that these "mistakes" are made by the scribes that were native speakers of East Turkic, who, however, migrated to Anatolia at that time. While using the Anatolian Turkish written language, due to the lack of standardization in the written language, they carried over features from their native languages (Tekin, 1974, as cited in Korkmaz, 1994). However, Korkmaz (1994, p. 18) argues that these variants are results of a "transition process" from the old written language to the Anatolian written language, which we independently know is gradual and creates pools of variants (Lightfoot, 1991 etc.). Korkmaz argues that such bilingualism emerged when Oghuz people living in the Old Uyghur region in Central Asia created their own written language when they came to Anatolia. Oghuz people reflected the linguistic features of their own language into the written language to a greater extent over time. For this reason, we included these so-called bilingual texts in our study.

2.1.3 Modern Turkish

Modern Turkish or Türkiye Turkish is the period of Turkish that encompasses today. The period from the transition to constitutional monarchy in the Ottoman Empire in 1908 until the proclamation of the Republic (1923) is called Early Modern Turkish. The Arabic alphabet was used during this period. Moreover, since it is a transition period, the features of the Classical Ottoman period are still visible. Especially the influence of Arabic and Persian continues. In addition, the influences of European languages, especially French, are also seen (especially in terms of loanwords). The Turkish spoken since the proclamation of the Republic (1923), especially since the Turkish Language Reform (1938), is called Modern Turkish (Ergin, 2009). The Latin alphabet has been used in this period. In addition, during the language reform, efforts were made to purify Turkish from Arabic-Persian origin words and grammatical structures.

2.2 Data sources

In this section, the texts used in gathering data will be presented in detail and in chronological order. The abbreviations for the title of the texts were given in parentheses throughout in the relevant section titles.

2.2.1 Old Turkic period :Eighth-eleventh centuries

2.2.1.1 Bilge Qaghan (BQ):

It is an inscription erected in the name of Bilge Qaghan after his death in Orkhon Valley, in today's Mongolia. The date of its erection is thought to be 735 CE. As in KT, Bilge Qaghan is the narrator, and the scribe is Yollug Tigin, the nephew of Bilge

Qaghan. The texts of the KT and BQ inscriptions are largely similar. There are some additions in the BQ inscription. The total number of lines is over 70 (Ölmez, 2015).

2.2.1.2 Kul Tigin (KT):

It is an inscription erected in the Orkhon valley, written during the Second Turkic Khaganate period. It was erected by Bilge Qaghan, the ruler of the period, in the name of his brother Kül Tigin. The erection date of the inscription is thought to be 732 CE. The speaker who dictates to the scribe seems to be Bilge Qaghan. The scribe of the inscription is Yollug Tigin. There are in total 66 lines, 40 on the east side of the monument, 13 on the north and 13 on the south (Ölmez, 2015).

2.2.1.3 Ongi (O):

It was found 160 kilometers south of the Orkhon inscriptions, near the Ongi River. The text contains information about the establishment of the Second Turkic Khaganate. It is thought that it was erected in the name of a Bey living in that region. The date of its erection is thought to be the beginning of the eighth century. There are a total of 19 lines in the inscription (Ölmez, 2015).

2.2.1.4 Küli Çor (KÇ):

The inscription describes the life of a person named Koli Çor, starting from the period of Kapgan Qaghan, the uncle of Bilge Qaghan. Therefore, the erection date of the inscription is thought to be the beginning of the eighth century. The inscription consists of 28 lines in total (Ölmez, 2015).

2.2.1.5 Tes (Tes):

The inscription belongs to the Old Uyghur Period. The exact date of erection of the inscription is not known. However, due to the phrase *tavuk yıl* 'chicken year' in the text, it is thought that the date of erection of the inscription was 745. On the other hand, another opinion on the subject is that the date of its erection may have been the year 762, during the reign of Bögü Qaghan. Although some parts of this inscription are damaged, there are 20 lines in total (Ölmez, 2015).

2.2.1.6 Tonyukuk (T 1-2):

It is an inscription consisting of two separate stones erected in the name of Tonyukuk, who is one of the founders of the Second Turkic Khaganate. The inscription was erected in the Orkhon Valley. The text was written from Tonyukuk's perspective. Considering the period in which Tonyukuk lived, it is thought that the inscription was erected in the first half of the eighth century (<750CE). There are a total of 62 lines in both stones (Ölmez, 2015).

2.2.1.7 Taryat (Tar):

It is an inscription from the Old Uyghur Period. It is thought that the inscription was erected in 753. It was found divided into three pieces. It is thought that the inscription was erected in the name of the third Uyghur Qaghan. There are 30 lines in total (Ölmez, 2015).

2.2.1.8 Bögü Qaghan (BögüK):

The text describes Bögü Qaghan's acceptance of Manichaeism. It is a four-page text written in the Uyghur alphabet. It is one of the Turkish Turfan Texts found in the

Turfan region (Yılmaz & Demir, 2007). Considering the period when Bögü Khagan lived and the Uyghurs accepting Manichaeism in Mongolia, it is thought that the text was written in Mongolia and later brought to the Turfan region, in today's Xinjiang region (Tekin, 1997, as cited in Yılmaz & Demir, 2007). Therefore, the date of writing may be the ninth century or tenth century. It is a text of 96 lines in total.

2.2.1.9 Maitrisimit (Mai):

It is a religious work related to Buddhism. The Uyghur text of the work was translated from Tocharian into Turkic by Prajnaksita. The first of the copies we used in our study was found by the Second Turfan Archeology Committee in the Twelve Burkan Monastery Libraries buried in the rocks, on the west bank of the river of the same name in the town of Sengim. It is thought that these monasteries were active in the eighth and eleventh centuries. The second copy was found in the monasteries on the banks of a river called Murtuk, five kilometers north of Sengim. The third copy is another copy found in Sengim. This copy is only two pages long and a facsimile of only one is provided (Tekin, 1976).

2.2.1.10 Huastuanift (Hua):

It is a prayer of repentance belonging to the Manichaean Uyghurs. It was discovered by Aurel Stein in Turfan region (1900). There are more than twenty copies. Although the exact date of writing of the copies is not known, it is thought that the copies were hidden in a cave by people to protect themselves from the Tangut attack in 1028 (Özbay, 2019). Therefore, it is thought that the text was written in the early eleventh century.

2.2.1.11 Altun Yaruk (AY 2-3-5-6-10):

It is a translation effort of Buddhist Uyghurs. The work has been translated from Chinese. Its translator is Şınko Şeli Tutuŋ. The translation date is not exactly known. However, as a result of Gabain's (Gabain, 1935, as cited in Kaya, 2021) study on other works of the translator, it is estimated that Şınko Şeli Tutuŋ lived in the tenth century. In his study on AY, which we benefited from, Kaya (2021) stated that the copy he used belongs to a much later period, but the original language of the text was meticulously preserved. It is a very comprehensive text. It consists of ten books in total. We used the second, third, fifth, sixth and tenth books in our study for convenience reasons.

2.2.1.12 Irk Bitig (IB):

It is a fortune telling book, thought to belong to the Manichaean Uyghurs. It was found by Stein in a cave in the city of Dunghuang (Uluç et al., 2013). The work was written in 930. The date of writing is written at the end of the text. Most of the text is written in Orkhon script.

2.2.1.13 Xuanzang-Biographie (2-3-6-9-10):

The work is a translation of a work written in Chinese by a Buddhist traveler named Xuanzang in the seventh century. Its translator is Şınko Şeli Tutuŋ. Since the translator is thought to have lived in the tenth century, it is thought that the translation was made in this period (Tezcan, 1975, as cited in Ağca, 2014). It consists of ten books in total. Among these, we used the second, thirrd, sixth, ninth and tenth books in our study.

2.2.1.14 Dasakarmapathavadanamala (Dasa):

The text is a translation from a religious textbook. It is a work of Buddhist Uyghurs. The language of the work is Old Uyghur. The person who translated it into Uyghur is Şilazin Praşnike. The translation date of the work is not mentioned in the documents belonging to the work, but it is thought to have been translated in the eleventh or twelfth centuries. In the edition of Elmalı (2016), which we used in our study, the copies found in Petersburg, Berlin and Urumqi were combined comparatively. Many of the copies were obtained during the studies carried out in the Turfan region.

2.2.2 Old Anatolian Turkic

We would like to remind that there is imperfection in the textual records between the eleventh and the thirteenth centuries. The following texts are used for the Old Anatolian period and were separated according to the centuries they were written in.

2.2.2.1 Thirteenth century

2.2.2.1.1 Sultan Veled's Turkish Poems (SVTM):

In the source we used, Mansuroğlu (1958) collected Turkish couplets from the works of Sultan Veled, the son of Mevlana Celaleddin-i Rumi, in İbtidaname (1291), Rebabname (1301) and his Divan. There are 76 Turkish couplets in İbtidaname, 162 in Rebabname, and 129 in his Divan.

2.2.2.1.2 Behcetü'l-Hadaik Fi Mev'izati'l-Halayık (BHMH):

The work is a text containing sermons on religious subjects. It is thought to have been written in the Anatolian region at the end of the twelfth or the beginning of the thirteenth century. It is a work written in a very early period, as it is thought to

belong to the Seljuk period in Anatolia. The language of the text shows both Old Anatolian Turkish and Eastern Turkic features fitting into the description of a 'bilingual text', discussed in section 2.1.2. It is seen as a transitional work to Anatolian Turkish in the period when the written language was being formed (Canpolat, 2018).

2.2.2.2 Fourteenth century

2.2.2.2.1 An Ancient Anatolian Text Dated 1343 (EATM):

The text was translated from Persian and is on inheritance issues. Its main author is Fakih Yakut Arslan. The copy used by Tekin (1974) was copied by Müskin Abbas. The copying date of the text is 1343. It is a 'bilingual text'.

2.2.2.2.2 Varka and Gülşah (VG):

It is a story of love and heroism. Its author is Yusuf-ı Meddah. It is understood that the author was born in the Azerbaijan region and later came to Konya. He wrote the text in Sivas (Nikoosokhan Sedghi, 2015). The date of writing of the text is 1342.

2.2.2.2.3 Kadı Burhanettin's Divan (KBD):

The text is a poetry book containing mostly poems about love. Kadı Burhanettin, the author of the work, served as a vizier in Anatolia and later founded a state centered in Sivas. Kadı Burhanettin lived from 1345 to 1398. In the source we use, the only copy in the British Museum was used. Halil bin Ahmed copied the used copy in 1393. For this reason, it is thought that the copy was also seen by Kadı Burhanettin (Ergin, 1980).

2.2.2.2.4 Tuhfe-i Mubarizi (TMB):

It is a medical book. It is a work written by Hekim Bereket, containing also some sections from Ibn Sina's 'The Law'. He firstly wrote in Arabic and then translate into Turkish. It is thought that the text goes back to the early thirteenth century or the fourteenth century. In the source we used, Doğuer (2013) edited the text by comparing two different copies of the text. While the date of incorporation of one of them is unknown, it is thought that the other was written in the early fifteenth century.

2.2.2.2.5 Süheyl ü Nevbahar (SNB):

The work is a masnavi[5] about love. It was translated from Persian to Turkish by Hoca Mesud and his nephew Izzettin Ahmed in the fourteenth century. Dilçin (1991) states that the author lived within the borders of Germiyanid Beylik between 1300-1370. The source we use contains the text prepared by Dilçin (1991) based on two copies. However, these copies do not have a copying date.

2.2.2.3 Fifteenth century

2.2.2.3.1 Gulistan Translation (GT):

The work is the translation of Sadi's work of the same name into Turkish. The translator of the work is Mahmud bin Kadı-ı Manyas. It is a translation that was left unfinished in 1430. It is thought that this translation started with the order of Murad The Second (Özkan, 2014, as cited in Anaçoğlu, 2019).

[5] Masnavi is a type of poetry used in Arabic, Persian and Turkish literature. It is written using couplets. It is mostly used in narrative works that are longer than a poem.

2.2.2.3.2 Mecmü'atü'n-nezair (MN):

The work is a collection of poems compiled in 1437. Its compiler is Ömer bin Mezid. There are a total of 397 poems from 84 poets in the work (Canpolat 1995).

2.2.2.3.3 Marzubanname (MRB):

It consists of stories and fables created to give advice on specific ethical subjects. Marzubanname was written in Persian by an author named Marzuban and translated into Turkish by Şeyhoğlu Sadraddin Mustafa. While the exact translation date is unknown, since the translation was made by the order of Germiyan Bey Suleyman Shah, it is thought to have been made in the mid-fourteenth century. The date of the copy of the work is 1445 (Özavşar, 2009).

2.2.2.3.4 Garib-name (GN):

It was written by Aşık Pasha in 1330. It is a moral and mystical masnavi. There are over ten thousand couplets. Aşık Pasha served in the state during the establishment of the Ottoman Empire during the reign of Osman Gazi, and later died in Kırşehir. In our study, the Suleymaniye Copy of the work was used. The date of copying is 1477 (Yavuz, 2000).

2.2.2.4 Sixteenth century

2.2.2.4.1 Kitab-ı Güzide (KG):

The work is a religious book. It was translated into Anatolian Turkish by Mehmet bin Bali in the fourteenth century. The copy in the study we used was copied in 1565 (Erdem, 1992).

2.2.2.4.2 Tuhfe-i Muradi (TMR):

The subject of the work is about precious stones, beautiful subjects, medicine, and objects that increase sexual potency. It was written in Bursa by Muhammed bin Mahmud-ı Şirvani in 1430. It was presented to the Ottoman ruler Murad The Second. The source we used (Argunşah, 1999) used three copies. The first of these was copied by Kasım in 1523. The date of copying of the second copy is unknown. The third copy was copied in 1571.

2.2.2.4.3 Süleyman-name-i Kebir: Sixth and seventh volume (SNK):

it is a religious, historical and epic work. It was written by Firdevsî-i Tavîl or Firdevsi-i Rumi. The author's date of birth is 1453. For this reason, it is thought that the work was written after this date, at the end of the fifteenth century or the beginning of the sixteenth century. In the source we used, the sixth and seventh volumes of SNK are worked. The copy used was copied in 1544.

2.2.2.4.4 Yusuf and Zeliha (YZ):

The work is an masnavi. It was written by Sayyad Hamza. It is thought that Seyyad Hamza died in 1285 and lived in Akşehir (Taş, 2017). Therefore, it is thought that this work was written in the first half of the thirteenth century. The only existing copy of the text was copied in 1545. The copyist is Kasımoğlu Abd'ur-rahman.

2.2.2.4.5 Edviye-i Müfrede (EM):

The work is a medical book. It was written by İshak bin Murad in 1389. In the source we used, Malatya Copy was used. The copying date of the copy is 1547 (Aydeniz, 2008).

2.2.2.4.6 Muntahab-ı Şifa (MŞ):

The work is a medical book. Written by Celalüddin Hızır. There is no information about its writing date and author. In the source we used, Malatya Copy and Paris Copy were used. Malatya Copy was copied by Muhammed el Zeceri in 1527 (Önler, 1981). The date of copying of the Paris Copy is unknown.

2.2.2.4.7 Revnak-ı Bustan (RB):

The work was on agricultural practices. In terms of its linguistic features, it is thought to have been written in the second half of the sixteenth century. Among the three existing copies, the Topkapı Copy is dated 1577. Although the date of copying of the Hagia Sophia Copy is unknown, its language is similar to the Topkapı Copy. The source we use (Önler, 2000) is based on the Hagia Sophia Copy.

2.2.2.4.8 Bahr'ül-Hakayık (BH):

It is the poetic translation of Hacı Bektaş Veli's work ‘Makalat’. The original work is in Arabic. It was written by Hatipoğlu in 1410. It is thought that the author lived in the Germiyanid Beylik. The source we use (Çiftçi, 1997) is based on Manisa Copy. The date of copying of this copy is 1565.

2.2.2.4.9 Hilyet'ür-Rical (HR):

It is a work that is about the lives of saints and their relationships with each other. It was written by Gelibolulu Mustafa Ali in 1541. The copy date of the copy used in the source we use (Özezen, 1994) is 1577.

2.2.2.4.10 Kısas-ı Enbiya (KE):

It is a masnavi written about the lives of the prophets. The author, Hindi Mahmud, wrote the work in 1579 and it was presented to Murad The Third. The date of copy is 1585 (Karataş & Coşkun, 2013).

2.2.2.5 Seventeenth century

2.2.2.5.1 Kitab-ı Baz-name-i Padişahi (KBP):

It is a book about the capture, raising, care, diseases, and treatments of birds of prey. Although its author is unknown, it was written at the beginning of the seventeenth century (Ergüzel, 2009).

2.2.2.5.2 Takvimü't-Tevarih (TT):

This is a history book. Additionally, information about the birth of the calendars used until that period is also given. It was written by Katip Çelebi in Istanbul in 1648 (Başar, 2021).

2.2.2.6 Eighteenth century

2.2.2.6.1 Captivity of Osman Agha of Timişoara (TOA):

The work is about the captivity period of an Ottoman soldier named Osman Agha of Timişoara. It is an autobiography written in Osman Agha 's own words. It was written in 1724 (Koç, 2021).

2.2.2.6.2 Şerh-i İbarat (Şİ):

The work is about the Mevlevi order and Sufism. It is a commentary on a work written in Arabic in the same century. Written by Müstakîmzâde Süleymân

Sa‘deddîn. The author also gave biographical information about the people he mentioned. It was written in Istanbul in 1789 (Karagöz, 2019).

2.2.2.7 Nineteenth century

2.2.2.7.1 Harita-I Kapudanan-I Derya (OKD):

The book is about the biographies of people who served as high admirals in the Ottoman Empire. It was written by Râmizpaşazâde Mehmed İzzet Bey in 1808. However, additional sections were written until 1853 (Sağlam, 2021).

2.2.2.7.2 Tarih-i Askeri-i Osmani: The fourth book (TAOCP):

The book we use as a source is about military innovations made in the Ottoman Empire in the eighteenth and nineteenth centuries. It also mentions some of the wars of the period. The book was written by Ahmet Cevat Pasha and published in Istanbul in 1881 (Çoban, 2009).

2.2.2.8 Early twentieth century

2.2.2.8.1 İkramiyeli Gazete (İZG):

İkramiyeli Gazete is a newspaper about agriculture. In the study we used, the first 13 issues of the newspaper published in 1909 were evaluated (Yılmaz, 2020).

2.2.2.9 Late twentieth century and early twenty-first century

2.2.2.9.1 Universal Dependencies for Turkish (UD):

Universal Dependencies (UD) is a framework for consistent annotation of grammar across different human languages including Turkish. It contains different Turkish UD treebanks. These are Parallel Universal Dependencies (PUD) (Türk et al., 2019),

ITU-METU-Sabancı Treebank (IMST) (Sulubacak, Pamay & Eryiğit, 2016, Türk et al., 2019), Boun Web Corpus (BOUN) (Türk et al., 2022), Grammar-book treebank (GB) (Çöltekin, 2015), KENET (Bakay et al., 2021), FrameNet (Marşan, 2021), Penn Treebank (PENN) (Kuzgun et al., 2020). Turkish examples are from 1980 to 2020.

The distribution of sources across centuries can be seen in Appendix A.

2.3 A note on the choices of texts

We took particular care to choose textual sources with different subjects. Naturally, different genres involve different types of words, hence, different grammatical subjects. It therefore creates a pool of variability in the grammatical subjects and the verbs of the subordinate clauses. This choice will prove to be useful to determine whether this differentiation changes the embedded clause structure. Additionally, there is a difference between the date of composition and the date of copying of many sources we use. We cannot be sure whether it was written in the language of the date in which it was written or in the language of the date in which it was copied. The person who copied the work may have adapted the original language of the work to the language in his copy if there was a language change. However, we can simply assume that the copied version was understandable for that period. We accepted the copy dates as the main date when listing them chronologically. If definitive judgments were made regarding the preservation of the language, we listed it according to its original writing date.

CHAPTER 3

DATABASE

In this chapter, we explain how we collected the data, the conventions we employed while annotating the data and the general numerical description of the data.

3.1 Extracting the data

Data for this study were taken mostly from transliterated sources, described in detail in Chapter 2. Transliterations were preferred simply out of convention: reading and scanning from the original texts would be beyond the limits of an MA thesis. Some of these works are available in pdf format yet many of them were old works and were not searchable. Yet in many ones that seem to be searchable, the search function did not yield reliable results. To avoid any doubt, all the sources were scanned manually over a two-year-period. Some of the resources are not available in pdf format. Hard copies of these were accessed at libraries and they were scanned manually. All instances of the sentences that involve the *-DIK* suffix in the sources were first identified manually.[6]

3.2 Classification

The sentences containing *-DIK* were then classified under five different categories. These categories are;

[6] An interesting avenue to pursue in the future is to transcribe and annotate these texts, using e.g., Transkribus, an AI powered handwritten text recognition. This, however, cannot be undertaken for MA research.

a. complement clauses to verbal predicates

b. adverbial clauses of reason

c. adverbial clauses of time

d. headed relative clauses

e. headless relative clauses.

This classification is based on structural and semantic criteria. The presence vs. absence of a matrix verbal predicates whose internal argument is a *-DIK* clause cut across complement clauses vs. others. A *-DIK* clause modifying an overt nominal head as syntactic adjectives are taken to be headed relative clauses. In the absence of an overt head, we took a *-DIK* marked clause functioning as a NP with definite or universal reading as a headless relative clause. When in doubt, we relied on further semantic criteria to distinguish headless relative clauses from complement clauses to verbal predicates. To illustrate, what we mean, consider (4) and (5) from Modern Turkish. Although both are the same on the surface, they may translate into two underlying structures, as a complement clause (4) or a headless relative clause (5), respectively[7].

(4) ye-diğ-in-i gör-dü-m.
eat-DIK-3SG-ACC see-PAST-1SG
"I saw that you ate."

(5) ye-diğ-in-Ø-i gör-dü-m
eat-DIK-3SG-Ø-ACC see-PAST-1SG
"I saw what you ate."

In such ambiguous cases, we looked for whether a referent of the bracketed clause is made salient in the context. If no clear judgments could be made, the -DIK clause was discarded from the database.

[7] Kornfilt (2005) claims that the headless relative in (2) is in actual fact a sort of light-headed relative– so light that it is zero. One would be tempted to think that they involve silent sortal heads, as is done by Cinque (2020). We simply follow Kornfilt's assumption here and show the light head as Ø.

Adverbial clauses are first collected based on the selecting adpositions, such as *üçün>içün/çUn, so*ŋ/*nra, evvel*... Since we know from Modern Turkish that an adposition can be a complementizer or a true adposition selecting a NP, we read all the potential adverbial sentences one by one in their own context. Any *-DIK* clause that is potentially both a relative clause or an adverbial clause was discarded from the database, as illustrated from Modern Turkish in (6)-(7):

(6) sen-in gel-diğ-in zaman çok geç-ti.
you-GEN come-DIK-2SG time too late-PAST
"the time when you came was too late."

(7) sen gel-diğ-in zaman çok geç-ti.
you come-DIK-2SG whe too late-PAST
"It was too late when you came."

While we are aware that the exact criteria that distinguish different clause types in Modern Turkish may not be applicable fully to the historical stages of Turkish, we are confident that the ultimate data points capture the typical clause types in each era.

A total of 8818 tokens were found in sum from all the texts. Of these, a total of 1052 examples were discarded from analysis, because we could not classify them precisely into one or more of the categories above. Of the remaining 7766 examples, 1717 are unambiguously complement clauses to predicates, 778 are reason clauses, 2038 are temporal clauses, 2126 are headed relative clauses, and 1107 are headless relative clauses. For ease of exposition, the numbers are also given in Table 1.

Table 1. Total Number of Examples According to Different Clause Types

Complement Clauses	Reason Clauses	Temporal Clause	Headed Relative Clause	Headless Relative Clause	Examples Excluded from the Data	Total
1717	778	2038	2126	1107	1052	8818

The distribution of these examples across the centuries are given in Table 2:

Table 2. The Distribution of Total Number of Examples across Centuries

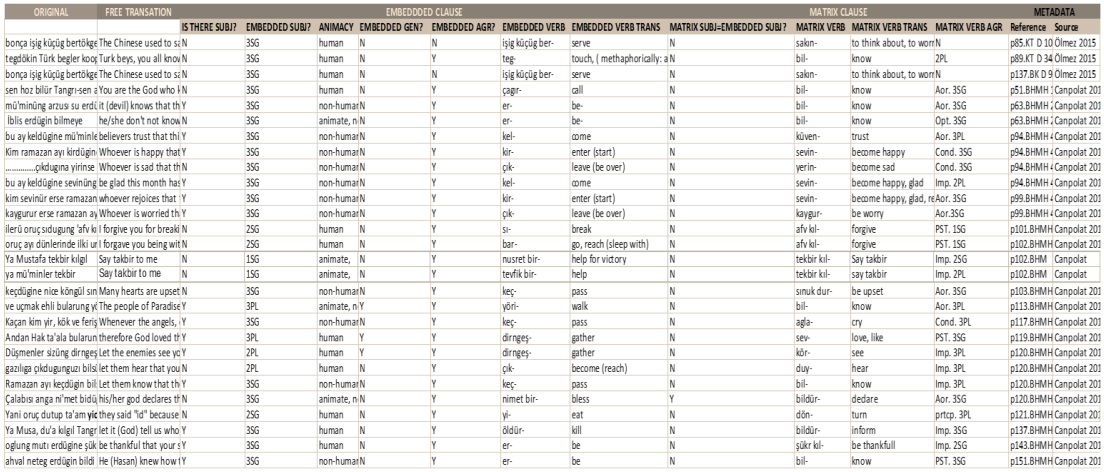
Century	Complements	Reason	Temporal	Headed Relatives	Headless Relatives
Seventh-Eleventh Centuries	8	58	133	31	6
Thirteenth-Fourteenth Centuries	307	112	127	217	271
Fifteenth Century	234	37	156	160	208
Sixteenth Century	367	53	599	442	335
Seventeenth Century	27	8	79	66	36
Eighteenth Century	146	45	441	333	127
Nineteenth Century	100	309	311	264	55
Twentieth Century	130	99	147	248	69
Twenty-first Century	398	57	45	365	0
Total	1717	777	2038	2126	1107

The distribution of these according to the texts are given in Appendix B:

As the tables suggest, there are only very few or no tokens from specific stages or authors. When necessary, these stages/authors were therefore taken out from numerical analysis.

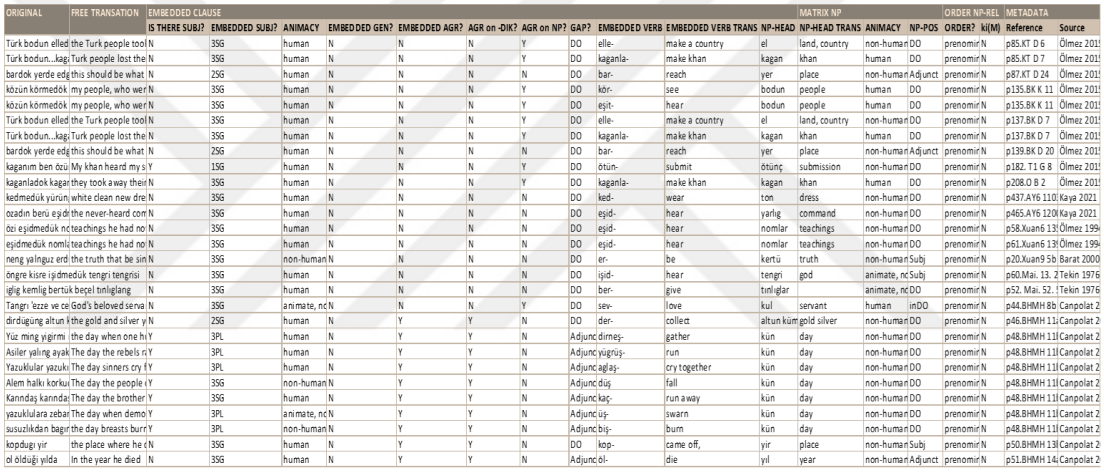
3.3 Metadata

The data points were collected in a database created using Microsoft Excel. For every clause type, a distinct sheet was created. In every sheet, a token is given a Sentence ID. This was followed by the original token given in broad transcription. For every sentence, a free translation into English was also provided. Metadata involved the following fields: (a) year of writing, (b) century it belongs to, (c) source ID (abbreviations in chapter 2), (d) secondary source (citation), (e) page-line number. An illustration is given in Figure 1 - 2.



ORIGINAL	FREE TRANSATION	EMBEDDED CLAUSE							MATRIX CLAUSE				METADATA	
		IS THERE SUBJ?	EMBEDDED SUBJ?	ANIMACY	EMBEDDED GEN?	EMBEDDED AGR?	EMBEDDED VERB	EMBEDDED VERB TRANS	MATRIX SUBJ–EMBEDDED SUBJ?	MATRIX VERB	MATRIX VERB TRANS	MATRIX VERB AGR	Reference	Source
bonça işig küçüg bertökge	The Chinese used to sı	N	3SG	human	N	N	işig küçüg ber-	serve	N	sakın-	to think about, to worr	N	p85.KT D 10	Ölmez 2015
tegdökin Türk begler koo	Turk beys, you all knov	N	3SG	human	N	Y	teg-	touch, (methaphorically: a	N	bil-	know	2PL	p89.KT D 34	Ölmez 2015
bonça işig küçüg bertökge	The Chinese used to sı	N	3SG	human	N	N	işig küçüg ber-	serve	N	sakın-	to think about, to worr	N	p137.BK D 9	Ölmez 2015
sen hoz bilür Tangrı-sen z	You are the God who i	N	3SG	human	N	Y	çagır-	call	N	bil-	know	Aor. 3SG	p51.BHMH 1	Canpolat 201
mü'minüng arzus su erdi(	it (devil) knows that th	Y	3SG	non-human	N	Y	er-	be-	N	bil-	know	Aor. 3SG	p63.BHMH 2	Canpolat 201
İblis erdügin bilmeye	he/she don't not know	N	3SG	animate, n	N	Y	er-	be-	N	bil-	know	Opt. 3SG	p63.BHMH 2	Canpolat 201
bu ay keldügine mü'minle	believers trust that thi	Y	3SG	non-human	N	Y	kel-	come	N	küven-	trust	Aor. 3PL	p94.BHMH 4	Canpolat 201
Kim ramazan ayı kirdügin	Whoever is happy that	Y	3SG	non-human	N	Y	kir-	enter (start)	N	sevin-	become happy	Cond. 3SG	p94.BHMH 4	Canpolat 201
.....çıkdugına yirinse	Whoever is sad that th	N	3SG	non-human	N	Y	çık-	leave (be over)	N	yerin-	become sad	Cond. 3SG	p94.BHMH 4	Canpolat 201
bu ay keldügine sevinüng	be glad this month has	Y	3SG	non-human	N	Y	kel-	come	N	sevin-	become happy, glad	Imp. 2PL	p94.BHMH 4	Canpolat 201
kim sevinür erse ramazan	whoever rejoices that	Y	3SG	non-human	N	Y	kir-	enter (start)	N	sevin-	become happy, glad, re	Aor. 3SG	p99.BHMH 4	Canpolat 201
kaygurur erse ramazan ay	Whoever is worried th	Y	3SG	non-human	N	Y	çık-	leave (be over)	N	kaygur-	be worry	Aor.3SG	p99.BHMH 4	Canpolat 201
ilerü oruç sıdugung 'afv kı	I forgive you for breaki	N	2SG	human	N	Y	sı-	break	N	afv kıl-	forgive	PST. 1SG	p101.BHMH	Canpolat 201
oruç ayı dünlerinde ilki ur	I forgave you being wit	N	2SG	human	N	Y	bar-	go, reach (sleep with)	N	afv kıl-	forgive	PST. 1SG	p102.BHMH	Canpolat 201
Ya Mustafa tekbir kılgıl	Say takbir to me	N	1SG	animate,	N	Y	nusret bir-	help for victory	N	tekbir kıl-	Say takbir	Imp. 2SG	p102.BHM	Canpolat
ya mü'minler tekbir	Say takbir to me	N	1SG	animate,	N	Y	tevfik bir-	help	N	tekbir kıl-	say takbir	Imp. 2PL	p102.BHM	Canpolat
keçdügine nice köngül sın	Many hearts are upset	N	3SG	non-human	N	Y	keç-	pass	N	sınuk dur-	be upset	Aor. 3SG	p103.BHMH	Canpolat 201
ve uçmak ehli bularung yi	The people of Paradise	Y	3PL	animate, n	Y	Y	yörü-	walk	N	bil-	know	Aor. 3PL	p113.BHMH	Canpolat 201
Kaçan kim yir, kök ve feriş	Whenever the angels,	Y	3SG	non-human	N	Y	keç-	pass	N	agla-	cry	Cond. 3PL	p117.BHMH	Canpolat 201
Andan Hak ta'ala bularun	therefore God loved th	Y	3PL	human	Y	Y	dirngeş-	gather	N	sev-	love, like	PST. 3SG	p119.BHMH	Canpolat 201
Düşmenler sizüng dirngeş	Let the enemies see yo	Y	2PL	human	Y	Y	dirngeş-	gather	N	kör-	see	Imp. 3PL	p120.BHMH	Canpolat 201
gazlıga çıkdugunguzı bilsı	let them hear that you	N	2PL	human	N	Y	çık-	become (reach)	N	duy-	hear	Imp. 3PL	p120.BHMH	Canpolat 201
Ramazan ayı keçdügin bil	Let them know that th	Y	3SG	non-human	N	Y	keç-	pass	N	bil-	know	Imp. 3PL	p120.BHMH	Canpolat 201
Çalabısı anga ni'met bidü	his/her god declares th	N	3SG	animate, n	N	Y	nimet bir-	bless	Y	bildür-	declare	Aor. 3SG	p120.BHMH	Canpolat 201
Yani oruç dutup ta'am yıg	they said "id" because	N	2SG	human	N	Y	yi-	eat	N	dön-	turn	prtcp. 3PL	p121.BHMH	Canpolat 201
Ya Musa, du'a kılgıl Tangr	let it (God) tell us who	Y	3SG	human	N	Y	öldür-	kill	N	bildür-	inform	Imp. 3SG	p137.BHMH	Canpolat 201
oglung mutı erdügine şük	be thankful that your s	Y	3SG	human	N	Y	er-	be	N	şükr kıl-	be thankfull	Imp. 2SG	p143.BHMH	Canpolat 201
ahval neteg erdügin bildi	He (Hasan) knew how t	Y	3SG	non-human	N	Y	er-	be	N	bil-	know	PST. 3SG	p151.BHMH	Canpolat 201

Figure 1. Excel sheet for complement clauses



ORIGINAL	FREE TRANSATION	EMBEDDED CLAUSE										MATRIX NP				ORDER NP-REL	METADATA	
		IS THERE SUBJ?	EMBEDDED SUBJ?	ANIMACY	EMBEDDED GEN?	EMBEDDED AGR?	AGR on -DIK?	AGR on NP?	GAP?	EMBEDDED VERB	EMBEDDED VERB TRANS	NP-HEAD	NP-HEAD TRANS	ANIMACY	NP-POS	ORDER?	ki(M)	Reference	Source
Türk bodun elled	the Turk people tool	N	3SG	human	N	N	N	Y	DO	elle-	make a country	el	land, country	non-human	DO	prenomir	N	p85.KT D 6	Ölmez 201
Türk bodun...kag	Turk people lost the	N	3SG	human	N	N	N	Y	DO	kaganla-	make khan	kagan	khan	human	DO	prenomir	N	p85.KT D 7	Ölmez 201
bardok yerde edg	this should be what	N	2SG	human	N	N	N	N	DO	bar-	reach	yer	place	non-human	Adjunct	prenomir	N	p87.KT D 24	Ölmez 201
közün körmedök	my people, who wer	N	3SG	human	N	N	N	Y	DO	kör-	see	bodun	people	human	DO	prenomir	N	p135.BK K 11	Ölmez 201
közün körmedök	my people, who wer	N	3SG	human	N	N	N	Y	DO	eşit-	hear	bodun	people	human	DO	prenomir	N	p135.BK K 11	Ölmez 201
Türk bodun elled	the Turk people tool	N	3SG	human	N	N	N	Y	DO	elle-	make a country	el	land, country	non-human	DO	prenomir	N	p137.BK D 7	Ölmez 201
Türk bodun...kag	Turk people lost the	N	3SG	human	N	N	N	Y	DO	kaganla-	make khan	kagan	khan	human	DO	prenomir	N	p137.BK D 7	Ölmez 201
bardok yerde edg	this should be what	N	2SG	human	N	N	N	N	DO	bar-	reach	yer	place	non-human	Adjunct	prenomir	N	p139.BK D 20	Ölmez 201
kaganım ben özü	My khan heard my s	Y	1SG	human	N	N	N	Y	DO	ötün-	submit	ötünç	submission	non-human	DO	prenomir	N	p182. T1 G 8	Ölmez 201
kaganladok kagar	they took away their	N	3SG	human	N	N	N	Y	DO	kaganla-	make khan	kagan	khan	human	DO	prenomir	N	p208.O B 2	Ölmez 201
kedmedük yürün	white clean new dre	N	3SG	human	N	N	N	N	DO	ked-	wear	ton	dress	non-human	DO	prenomir	N	p437.AY6 110	Kaya 2021
ozadın berü eşidi	the never-heard com	N	3SG	human	N	N	N	N	DO	eşid-	hear	yarlıg	command	non-human	DO	prenomir	N	p465.AY6 120	Kaya 2021
özi eşidmedük nc	teachings he had no	N	3SG	human	N	N	N	N	DO	eşid-	hear	nomlar	teachings	non-human	DO	prenomir	N	p58.Xuan6 13	Ölmez 199
eşidmedük nomlı	teachings he had no	N	3SG	human	N	N	N	N	DO	eşid-	hear	nomlar	teachings	non-human	DO	prenomir	N	p61.Xuan6 13	Ölmez 199
neng yalnguz erdi	the truth that be sin	N	3SG	non-human	N	N	N	N	DO	er-	be	kertü	truth	non-human	Subj	prenomir	N	p20.Xuan9 5b	Barat 2000
öngre kisre işidmedük tengri tengrisi		N	3SG	human	N	N	N	N	DO	işid-	hear	tengri	god	animate, nc	Subj	prenomir	N	p50.Mai. 13. 2	Tekin 1976
işlig kemlig bertük beçel	tınlıglang	N	3SG	human	N	N	N	N	DO	ber-	give	tınlıglar		animate, nc	DO	prenomir	N	p52. Mai. 52. 1	Tekin 1976
Tangrı 'ezze ve ce	God's beloved serva	N	3SG	animate, nc	N	N	N	Y	DO	sev-	love	kul	servant	human	inDO	prenomir	N	p44.BHMH 8b	Canpolat 2
dirdügüng altun k	the gold and silver y	N	2SG	human	N	Y	Y	N	DO	der-	collect	altun küm	gold silver	non-human	DO	prenomir	N	p46.BHMH 11	Canpolat 2
Yüz ming yigirmi	the day when one hi	Y	3PL	human	N	Y	Y	N	Adjunc	dirneş-	gather	kün	day	non-human	DO	prenomir	N	p48.BHMH 11	Canpolat 2
Asiler yalıng ayak	The day the rebels ri	Y	3PL	human	N	Y	Y	N	Adjunc	yügrüş-	run	kün	day	non-human	DO	prenomir	N	p48.BHMH 11	Canpolat 2
Yazuklular yazuk	The day sinners cry f	Y	3PL	human	N	Y	Y	N	Adjunc	ağlaş-	cry together	kün	day	non-human	DO	prenomir	N	p48.BHMH 11	Canpolat 2
Alem halkı korku	The day the people o	Y	3SG	non-human	N	Y	Y	N	Adjunc	düş	fall	kün	day	non-human	DO	prenomir	N	p48.BHMH 11	Canpolat 2
Kanndaş kannda	The day the brother	Y	3SG	human	N	Y	Y	N	Adjunc	kaç-	run away	kün	day	non-human	DO	prenomir	N	p48.BHMH 11	Canpolat 2
yazuklulara zebar	The day when demo	Y	3PL	animate, nc	N	Y	Y	N	Adjunc	üş-	swarn	kün	day	non-human	DO	prenomir	N	p48.BHMH 11	Canpolat 2
susuzlıkdan bagır	the day breasts burn	Y	3PL	non-human	N	Y	Y	N	Adjunc	biş-	burn	kün	day	non-human	DO	prenomir	N	p48.BHMH 11	Canpolat 2
kopdugı yir	the place where he c	N	3SG	human	N	Y	Y	N	DO	kop-	came off,	yir	place	non-human	Subj	prenomir	N	p50.BHMH 13	Canpolat 2
ol öldüği yılda	In the year he died	N	3SG	human	N	Y	Y	N	Adjunc	öl-	die	yıl	year	non-human	Adjunct	prenomir	N	p51.BHMH 14	Canpolat 2

Figure 2. Excel sheet for headed relative clauses

3.3.1 General criteria for annotation

Besides metadata, every token was further annotated according to general as well as construction-specific criteria. The annotation criteria that are common to all datasets are given as follows:

a. Whether the subject internal to the embedded clause is overt or not. “Yes” (Y) if overt, “No” (N) if not.

b. Whether genitive case is expressed on the subject of the embedded clause: Y/N

c. Person and number specification of subjects, regardless of whether it is understood or overt, in the embedded clause: multiple criteria (1SG, 2SG, 3SG… etc).

d. Animacy properties of subjects in the embedded clause: three criteria "human", "animate, non-human", "inanimate."

e. Whether there is nominal agreement on the verb of the embedded clause: Y/N

f. The lexical verb/copula of the embedded clause and its translation.

3.4 Construction-specific criteria

Depending on the construction type under investigation, tokens were annotated according to numerous construction-specific criteria. In the next subsections, these criteria are given.

3.4.1 Additional criteria for complement clauses

a. Matrix verb and its translation

b. TAM specification on the matrix verb

c. Whether the subject in the complement and the matrix clauses are the same or not: Y/N

d. Distance of the embedded subject to the embedded verb quantized regarding the number of intervening XPs.

3.4.2 Additional criteria for temporal and reason clauses

a. Case suffix on the verb of the embedded clause. Y/N, if Y, which one

b. Postposition/complementizer used after case.

c. Whether the subject in the adverbial clause and the matrix clause is the same or not: Y/N

d. verb in the matrix clause and its translation.

3.4.3 Additional criteria for relative clauses

a. whether the nominal agreement is on the head or on the embedded clause: Y/N

b. gap position inside relative clauses.

c. The head of the relative clause and its translation.

d. Animacy property of the head of the relative clause ("human", "animate, non-human", "inanimate").

e. The grammatical status of the relative clause in the sentence (subject, direct object, indirect object, adjunct).

f. Distance of the embedded subject to the embedded verb quantized regarding the number of intervening XPs. Examples of the annotation according to the additional criteria are given in Tables (3 - 5)

Table 3. Additional Criteria for Complement Clauses

Complement Clauses	
Original	
Free Transation	
Embedded Clause	Is There Subject?
	Embedded Subject?
	Animacy
	Embedded Genitive?
	Embedded Agreement?
	Embedded Verb
	Embedded Verb Trans.
Matrix Clause	Matrix Subject=Embedded Subject?
	Matrix Verb
	Matrix Verb Trans.
	Matrix Verb Agreement
Meta Data	Reference
	Source
	Text
	Century

Table 4. Additional Criteria for Relative Clauses

Temporal Clauses	
Original	
Free Transation	
Embedded Clause	Is There Subject?
	Embedded Subject?
	Animacy
	Embedded Genitive?
	Embedded Agreement?
	Embedded Verb
	Embedded Verb Trans.
	Dative or Locative?
Matrix Clause	Matrix Subject=Embedded Subject?
	Matrix Verb
	Matrix Verb Trans
	Matrix Verb Agreement
	Atttechment Site?
Meta Data	Reference
	Source
	Postposition
	Century

Table 5. Additional Criteria for Adverbial Clauses

Headed Relative Clauses	
Original	
Free Transation	
Embedded Clause	Is There Subject?
	Embedded Subject?
	Animacy
	Embedded Genitive?
	Embedded Agreement?
	Agr on -DIK?
	Agr on NP?
	Gap?
	Embedded Verb
	Embedded Verb Trans.
Matrix NP	NP-Head
	NP-Head Trans.
	Animacy
	NP-Position
Order NP-Relatives	Order?
	ki(M)
Meta Data	Reference
	Source
	Text
	Century

Before moving on to the next sections, it should be noted that one of the main subjects of our thesis is the change in subject markings in complement and headed relative clauses in the history of Turkish. Throughout our study, we also annotated headless relative clauses in order to look at the change from a broader perspective. However, we did not take them into consideration in this thesis. We will reconsider and evaluate them in our future studies.

3.5 Additional information for criteria

The next sections (Section 3.4.1 to 3.4.3) elaborate on the criteria of the subject marking on the embedded clause because it requires additional information.

3.5.1 Subject of embedded clauses:

Turkish throughout its history has always been a consistently null-subject language (Özsoy, 1987b). Subjects of embedded clauses or matrix clauses may not be overt. For this reason, whether the subjects of embedded clauses had genitive cases or not was evaluated only through examples where the subjects were overt.

Some problems were encountered in determining whether the subject of the embedded clause is overt or not. For example, *ol* is the nominative form of third person pronoun in the pre-Modern Turkish period. When any case suffix is added, this pronoun becomes *an-*. Additionally, *ol* can also be understood as a demonstrative pronoun or as a demonstrative adjective. Although the internal subjects of complement and relative clauses have genitive cases in the Modern Turkish period, there are examples in the pre-Modern Turkish period where the internal subjects did not have genitive. There are examples in which we are undecided about the function of the *ol* morpheme in some sentences. Illustrations are given in (8).

(8) a. Ol ḥarāmī-ler ile cenk it-düg-in di-di
That/he bandit-PL with war do-DIK-3SG+ACC say-PAST.3SG
"He said he fought with that bandits."

(VG 1303)

b. ol öl-düg-i yıl-da
this/he die-DIK-3SG year-LOC
"in the year he died."

(BHMH 14a 10)

In cases where the subject in the complement clause and matrix clause is identical and in the third person singular, as in the example in (8a), if one of the subjects is not overt, the existing pronoun *ol* can be considered as belonging to the matrix or complement clause. In such cases, we accepted all *ol* pronouns as the subject of the matrix clause, and in these examples, we accepted that the subject is not overt in the embedded clauses. For this reason, these samples were excluded from our evaluation of the presence of the genitive case. *ol* morpheme in (8b) is understood as a demonstrative adjective. However, in cases where the agreement in the embedded clause is third person singular, morpheme *ol* can also be understood as a pronoun. In all such examples, we considered the morpheme *ol* as a demonstrative adjective and excluded it from our evaluation of the presence of the genitive case.

A similar situation applies to first and second person subjects. If there is a first- and second-person pronoun in nominative form in the sentence and the subject of the embedded clause and matrix clause is the same, this pronoun can be considered both the subject of the embedded clause and the subject of the matrix clause. For all such examples, these pronouns were accepted as the subject of the matrix clause.

3.5.2 Person, number, animacy

Properties such as number person or animacy of the subjects of the embedded clause, which may affect the existence of genitive case or nominal agreement, are annotated. In examples where there is no nominal agreement in the embedded clause and the subject is not overt, it is difficult to determine the number and person properties of the subject of the embedded clause. In such examples, the subject is annotated according to the context to determine the person, number and animacy properties. In addition, some parts of some examples in sources from the Old Turkic period are missing. For these examples, if there was nominal agreement in the embedded clauses, the features were annotated according to the agreement suffix or context. However, if it cannot be understood from the context, these sections are left blank. Subjects in the complement clause are classified as human, inanimate, animate-but-non-human. The categories human and inanimate, as their names suggest, are created for humans and inanimate objects, respectively. The animate-but-non-human category is created for examples whose subject referent is animate yet not human. Supernatural beings such as giants, gods, fairies, and angels were all classified in the animate-but-non-human category. In addition, since there are literary texts among the sources, some animals are personified and used with human characteristics. These were also considered animate-but-non-human. Therefore, we remained agnostic to whether there are honorary human categories in the history of Turkish or not.

As we briefly mentioned, some parts of the sources, especially those belonging to the Old Turkic period, are missing. These parts were broken in the inscriptions, and there was wear on the edges of the books. For this reason, these parts could not be read in the transcriptions. In some of the examples we evaluated, apart from the parts where the embedded clause is written, the parts where we will

evaluate the criteria related to the matrix clause are also missing. For this reason, we could not annotate the parts related to the matrix clause for such examples and left them blank. We had to exclude these examples from the relevant evaluations. However, the number of such examples is quite small. We do not think it will affect the overall evaluation. In short, when we are in doubt, we excluded the examples from the calculations.

3.5.3 Postposition *üçün>içün/-çUn*

The postposition *üçün>içün/-çUn* used in reason clause structures has undergone phonological change over time. As seen in (9), the form *-çUn* is used as a clitic and written as attached. Since vowels are mostly not written in Old Anatolian Turkish written with the Arabic alphabet, it may be difficult to detect the presence of nominal agreement in examples containing *-çUn*. In general, in these structures with *-çUn*, the agreement is assumed to be present and read. We accepted the presence of the agreement marker in these examples. *-çUn* was used in 61 examples.

(9) Talut Şah Davud-a ḳız-ın vir-me-düg-i -çün
Talut shah Davud-DAT girl-3SG+ACC give-NEG-DIK-3SG because
"Because Shah Talut did not give his daughter to David"
(SNK 40b 24)

3.6 Examples excluded from the data

It was difficult to determine the category of some samples found in the data. These examples could potentially be in different categories. Since evaluating these examples in one category would affect our evaluation if they were in the other category, we excluded these examples from our data. These examples will be mentioned in this section.

3.6.1 Subordinate clauses used as subjects of the main clause

In our data, there are subordinate clauses used as subject clauses in main clauses that are not headless relative clauses. These examples could be considered as complement clauses. However, since these structures were outside the VP, they were not accepted as complement clauses and these examples were excluded from the data. You can see an example of this structure in (10).

(10) ag mercan-a tolaş-dug-ı bil-in-di
net coral-DAT entangle-DIK-3SG know-PASS-PAST.3SG
"the net has been known to become entangled in the coral"
(TMR 87a 8)

3.6.2 Headed relative clause or adverbial clause

It is not clear whether some examples are an adverbial clause or a headed relative clause. The examples given in (6) and (7) are given again below.

(=6) sen-in gel-diğ-in zaman çok geç-ti.
you-GEN come-DIK-2SG time too late-PAST
"the time you came is too late."

(=7) sen gel-diğ-in zaman çok geç-ti.
you come-DIK-2SG when too late-PAST
"it is too late when you came."

For Modern Turkish, the main clue in distinguishing these two structures is the genitive case in (6) in the subject. In modern Turkish, while genitive case is used in subjects in relative clauses, it is not used in adverbial clauses (Kornfilt, 2003, 2008; Aygen, 2007). However, in the pre-modern period, the use of genitive cases in headed relative clauses was not obligatory. Therefore, it is impossible to determine the category of these samples.

3.6.3 Main clause examples

In the sources we scanned, it seems that some examples with *-DIK* suffix are in the TP or Asp P., Mood P. in the main clause, not in the subordinate clause. As seen in example (11), there are differences of opinion on how the structure with *-DIK* should be read. Tekin (1998) thinks that the suffix *-madok* is the negative form of the reported past tense suffix *-mIş*. The word *erinç* seen in the example means 'undoubtedly' and is not a postposition. Ölmez (2015) read the sentence as a reason clause, but in our opinion, he gave this reading according to the context. In the example, there is no morphological clue to read this part as a reason clause.

(11) ini-si eçi-sin teg kıl-ın-ma-dok
younger. brother-3SG older.brother-3SG+ACC like create-PASS-NEG-DIK
erinç
undoubtedly

(KT E5)

a. "Undoubtedly, because the younger brother was not created like his older brother,"
(Ölmez, 2015)

b. "Undoubtedly, the younger brother was not created like his older brother."
(Tekin, 1998)

As seen in example (12), in some examples the tense suffix *-DI*, is used after *er-* or *bol->ol-* verbs. Considering that *er-* and *bol-> ol-* are an insertion to attach the tense suffix, the *-DIK* suffix is placed under TP in the structure as Asp P. or Mood P. Since the main subject of our thesis is subordinate clauses, we excluded such examples from our categories. There are a total of 41 examples in this structure.

(12) yazuk kıl-ma-duk er-dü- ŋüz
sin do-NEG-DIK be-PAST-2PL
"you have never sinned"

(BHMH.112b 23)

3.6.4 Disease names

Among the sources we scanned, some disease names in medical sources were consistently established with the *-DIK* suffix. In an example like (13), we are

undecided about which category we should evaluate these disease names under. Since we could not predict the structure of the embedded clause, we excluded such examples from our evaluation. There are a total of 60 examples of this structure being used as a disease name.

(13) gönül dön-düg-in sakin ed-er
heart turn-DIK-POSS.3SG+ACC calm make-PRS
"it calms dizziness."

(MŞ M78a 6)

3.6.5 Titles in KBP

Structures with *-DIK* suffix are used as titles of chapters in KBP. There are 147 examples in total. Since there is a contents section at the beginning of the book, one example is actually written in two places. As can be seen in example (14), since there is a *-DA* locative case in such examples, it can be read as both a temporal clause and a headless relative clause. Due to this uncertainty, we excluded these samples from our evaluation.

(14) yitmiş sekizinci bab sovuk-dan kus-dug-ın-da-dur.
seventy eighth chapter cold-ABL vomit-DIK-3SG-LOC-PRS

(KBP 4a 9)

a. "Chapter seventy-eight is about what to do when vomiting due to cold."
b. "Chapter seventy-eight is about treating vomiting due to cold."

3.6.6 Adverbial clauses that are not temporal or reason clauses

In some examples, it gives the meaning of 'as much as possible' in adverbial clauses where verbs such as *güci yit-* 'afford' or *mümkin ol-* 'be possible' are used, formed with the *-CA* suffix or *kadar* proposition, and where the *-DIK* suffix is present. Since we could not include these examples in the temporal or reason clause categories, we excluded such examples from evaluation.

Apart from these examples, the examples we excluded from evaluation will be given according to their categories in following chapters, along with their reasons.

CHAPTER 4

COMPLEMENT CLAUSES

4.1 Introduction

In this section, we offer a diachronic description of complement clauses to lexical verbal predicates that bear the *-DIK* morpheme on their embedded verb. We will evaluate these *-DIK* clause structures according to the categories mentioned in Section 3.2. In what follows, we first provide the structure with examples from Modern Turkish. Then, the historical change in those structures will be revealed with the examples we obtained from our data.

4.2 Complement clauses in Modern Turkish

According to standard grammar sketches of Modern Turkish (Goksel & Kerslake, 2005; Kornfilt, 2003), the subordinating suffix *-DIK* alternates with *-(y)AcAK*. While both are obligatorily followed by a nominal agreement suffix (possessive suffix according to Goksel & Kerslake, 2005), which refer to the subject of the noun clause (15a) vs. (15b). *-DIK* expresses relative present or relative past tense and *-(y)AcAK* relative future tense. We will only focus on *-DIK* complement clauses. Kornfilt (2003) refers to *-DIK* clauses as factive nominalized embedded clauses. These complement clauses are largely selected (in terms of Grimshaw, 1979) by non-control semi-factive verbs (commmunication, cognitive), emotive factives, non-factiver communication and non-factive cognitive verbs (Sheehan & Hinzen, 2011).

(15) a. [Ben-im gel-diğ-im]-e sevin-di.
I-GEN.1SG come-DIK-1SG-DAT be happy-PAST.3SG
"He was happy I came."

b. [Ben-im gel-eceğ-im]-e sevin-di.
I-GEN come-DIK-DAT be happy-PAST
"He was happy I will come."

4.3 Incorporation in complement clauses

The subject internal to the *-DIK*-embedded clause may bear genitive if it is overt, as seen in (15a). This is so because of the specificity feature or lack of pseudo-incorporation. In the literature, it is accepted that genitive marking in Modern Turkish indicates a higher level of specificity (Erkman-Akerson & Özil, 1992; Barker, Hankamer & Moore, 1990 a.o.). Also, Öztürk (2009) claims that agents can freely undergo incorporation in transitive and ergative constructions. When these agent incorporation structures are used as complement clauses, the subjects of the embedded clause do not have genitive cases. In these cases, the subject occurs immediately preverbal position. As can be seen in example (16), the subject of the embedded clause must be preverbal and not specific to be used without a genitive case.

(16) a. [köy-e doktor gel-diğ-in]-I söyle-di-ler.
village-DAT doctor come-DIK-3SG-ACC say-PAST-3PL
"They said a doctor came to the village"

b.*[doktor köy-e gel-diğ-in]-I söyle-di-ler.
doctor village-DAT come-DIK-3SG-ACC say-PAST-3PL
"They said a doctor came to the village"

Pronominal subjects, being inherently specific, must bear genitive, compare (17) with (15).

(17) *Ben gel-diğ-im-e sevin-di.
I come-DIK-1SG-DAT be happy-PAST.3SG
"He was happy I came."

Kennelly (1997) observes that there is no genitive case on the nonspecific subjects of embedded unaccusative, existantial and passive verbs. The author further

suggests that these must be in the preverbal focus position and that the genitive case is obligatory when they are located anywhere else. Kennelly also stated that when it takes a genitive case in the preverbal focus position, the subject loses the nonspecific interpretation.

On the other hand, genitive case is not observed on incorporated themes of passives in Modern Turkish. To illustrate, in (18a), the predicate of the embedded clause is passive and its theme is specific, in which configuration the subject has genitive case. In (18b), on the other hand, there is an incorporated theme. Unlike (18a), genitive case is not used after the word *kitap* 'book'. Öztürk (2007) states that the passivation of incorporated themes yields impersonal passive constructions. Additionally, she states that incorporated themes are not visible as syntactic objects but are rather part of the predicate. For this reason, in an example like (18b), the theme *kitap* 'book' does not take a genitive case because it is not visible for case computation, also, as a syntactic subject.

(18) a. Mehmet [kitab-ın oku-n-duğ-un]-u öğren-di.
Mehmet book-GEN read-PASS-DIK-3SG-ACC learn-PAST.3SG
"Mehmet learned the the book was read."

b. Mehmet [kitap oku-n-duğ-un]-u öğren-di.
Mehmet book read-PASS-DIK-3SG-ACC learn-PAST.3SG
"Mehmet learned that book-reading was done."

In short, the properties of *-DIK* complement clauses in Turkish can be seen in (19).

(19) Basic facts about the *-DIK* complements in Modern Turkish:
a. They are non-finite clauses selected by non-control verbs (assertive or factive),
b. They encode their internal subject with agreement/possessive markers on the verb
c. Their subject appears in the genitive case if the subject is specific and/or not incorporated.
d. They are case-marked by the selecting verb.

4.4 Setting the scene for historical investigation: The preamble

The state of affairs presented in sections 4.2-4.3 did not hold uniformly across all the historical stages of the language. To illustrate, consider the examples in (20). The complement clauses, including the case marking, are provided in square brackets.

(20) a. [Talut-uŋ leşker çek-üb gel-düg-in] anla-dı.
Talut-GEN soldier gather-FN come-DIK-3SG+ACC understand-PAST.3SG
"He realized that Talut had gathered soldiers and came."
(SNK 46b 4)

b. [bonça iş-ig küç-üg bėr-dök-gerü] sakın-matı
so many work-ACC power-ACC give-DIK- DAT think-FN
türk bodun ölör-eyin urugsırat-ayin te-r er-miş)
turkic people kill-IMP+1SG destroy- IMP+1SG say-PRS be-PAST
"(They said they would kill and destroy the Turkic people) without considering that they had served them so much."
(KT E10, BQ E9)

c. [Musa öldür-düg-in-i] hep bil-ür-ler.
Musa kill-DIK-3SG-ACC all know-PRS-3PL
"They all know that Moses killed"
(KE 155bBL 9)

The three examples illustrate three possibilities from different stages of the language. Notice at this point that, examples (20) do not provide the typical situation of their era. They are simply given here to illustrate the variability seen in *-DIK* complement clauses in the history of Turkish. While (20a) virtually has no difference from what it would be like in modern Turkish, in that its subject appears in the genitive and the same subject is encoded on the verb, which also bears case marking,[8] (20b-c) depict a different picture each. In (20b), there is neither an overt

[8] In Turkish, the 3rd person possessive suffix is *-(s)I(n).* The uses in parentheses are used to show the change in the suffix according to different conditions. If the morpheme to which the 3rd person possessive suffix is attached ends in a vowel and is followed by another suffix, the suffix is used as *-sIn*. If the morpheme before the suffix ends with a consonant, the (s) at the beginning of the suffix is not used. If no other suffix is used after the suffix, the (n) at the end of the suffix does not surface. As in the example in (20a), in the Old Anatolian Turkish period, when the accusative case suffix comes after the 3rd person possessive suffix, the accusative suffix -(y)I can be dropped. However, the 3rd person possessive suffix is used with the -(s)In form.

(i) a. on-un ev-I
he-GEN house-3SG
"his house"

b. on-u araba-sı
he-GEN car-3SG
"his car"

c. on-un araba-sın-da
he-GEN car-3SG-LOC
"in his car"

internal subject nor nominal agreement on the verb of the embedded clause. In (20c), we see that the subject of the embedded clause, while it is overt, does not take genitive case. As in Modern Turkish, the genitive suffix is used after the subject in complement clause and the agreement suffix is used after *-DIK* suffix in the embedded clause in example (20a). However, we see that the sentences in examples (20b) and (20c), which would be ungrammatical for Modern Turkish, are possible. (Example in (20b) is from the eighth century, example in (20c), from sixteenth century)

The comparison between (20a) on one hand and (20b-c) on the other, point to a number of differences observed in the history of Turkish, the notable of which is that the genitive on the subject has become obligatory in Modern Turkish, as long as the referent of the subject is specific, whereas it was not in older stages.

To understand this and many other changes, we need to provide a through historical account, in other words, a diachronic corpus analysis. We acknowledge that diachronic analysis is making the best use of bad data—bad because textual records provide a biased picture of the facts, i.e., they only provide what written texts looked alike, and in so doing, they are fragmentary. With a careful analysis however, which combines quantitative data with interpretation, we can overcome this hurdle. In the rest of the chapter, we offer numerical information on the *-DIK* clauses in the history of Turkish.

4.5 The overall numbers

In total, 1962 *-DIK* clauses that could potentially be taken as complement clauses were extracted from the written texts. Of these, 210 were taken out because they seem to behave as non-finite sentential subjects, as given in (21).

(21) Süheyl an-ı öldür-düg-i adl i-di.
Süheyl he-ACC kill-DIK-3SG fair be-PAST
"It was fair for Suheyl to kill him."

(SNB 5176)

Among the remaining, 35 examples involve subordinate subjects with the accusative case, as in (22).

(22) Hiç kimesne [an-ı ne ol-dug-ın] bil-me-z-di
anyone it-ACC what be-DIK-3SG+ACC know-NEG-PRS-PAST.
"No one knew what it was."

(TMR 52a 12)

While these structures are interesting by themselves, we leave them for future research.

Once the data were sorted out there are in total 1717 *-DIK* complement clauses, spreading over the entire corpus. The breakdown of the numbers is given in Table 6:

Table 6. The Overall Number of Complement Clauses across Centuries

	Overt Internal SUBJ		Non-Overt Internal SUBJ	Total Number
CENTURIES	GEN on SUBJ	No GEN on SUBJ		
Seventh-Eleventh Centuries	1	0	7	8
Thirteenth-Fourteenth Centuries	23	93	191	307
Fifteenth Century	15	117	102	234
Sixteenth Century	54	116	197	365
Seventeenth Century	4	13	10	27
Eighteenth Century	14	43	89	146
Nineteenth Century	33	14	53	100
Twentieth Century	66	19	44	129
Twenty-first Century	350	48	0	398
Total	560	463	693	1717

Crucially, only two tokens that do not carry nominal agreement on the *-DIK* marked clause in our corpus. Both of these, which are found in the BQ and KT inscriptions respectively, are from the eighth century and they are identical. This sentence, which was cited as (20b) above, is given here again:

(=20b) [bonça iş-ig küç-üg bėr-dök-gerü] sakın-matı (türk
so many work-ACC power-ACC give-DIK- DAT think-FN turkic
bodun ölör-eyin urugsırat-ayin te-r er-miş)
people kill-IMP+1SG destroy- IMP+1SG say-PRS be-PAST
"(They said they would kill and destroy the Turkic people) without considering that they had served them so much."

(KT E10, BQ E9)

That an identical sentence is used in two separate inscriptions may point out to the possibility that we are dealing with a formulaic expression. Furthermore, as can be seen in example (20b), the *-gArU* morpheme is used after the *-DIK* suffix. This morpheme may be taken as the spell out of a lexical case assigned by the verb *sakın-* 'to worry, to think about'. That it is phonologically heavy (bisyllabic) may have an effect on the lack of nominal agreement. There are no other examples of complement clauses from this era with the with the *-gArU* suffix.

In any event, even when we accept the above two examples not as exceptions but as some possible variant, reflecting a possible grammar of their era, we must conclude that the obligatory appearance of nominal agreement marker on the embedded verb established very early on, definitely before the eleventh century, as the rest of the examples from the following eras indicate.

4.6 The verbs selecting *-DIK* complement clauses

DIK-complement clauses appear with a large number of verbs in the corpus, notably cognitive factives such as *bil-* 'know' (GN 69, p. 24) *aŋla-* (SNK 39a 12, p. 285), *unut-* 'to forget' (EK 108bBL 2), (emotive) factives such as *sevin-* 'to be happy' (SNB 1563, p. 132), *ħavf it-* 'to be afraid' (SNK 36a 9, p. 419), *peşiman ol-* 'to regret' (MRB 46a 1, p. 93), p. 418), semi-factives *ma'lum eyle-* 'to expose' (GN 7080, p. 362), *ayan id-* 'to expose' (SNK 30a 26, p. 404), non-factives (communication or cognitive) such as *bildür-* (BHMH 4a 1, p. 39) *eyd-* 'say' (GN

4942, p. 255), *di-* ‘say’ (KG 86b 13, p. 141). They also appear with non-assertive ones such as *iste-* ‘to want (GN 5824, p. 302) One illustration is given in (23):

(23) kamu-lar-uŋ hırmen-I köyün-düg-in iste-r
all-PL-GEN harvest-3PL burn-DIK-3SG+ACC want-PRS.
“He (devil) wants everybody's harvest burned.”

(BHMH 161b 5)

The table 7 summarizes the six most frequent matrix verbs:

Table 7. The Most Frequent Matrix Verbs in Complement Clauses

Token #	Verb
143	Know: *bil-*
56	See/realize: *kör-, gör-*
29	Say: *ayd-, eyd- de-, di-, eyit-*
19	Understand: *anla, aNgla-*
24	Hear: *duy- işit, tuy-*
18	Declare: *beyan it/et-/kil-*

There seem to be no restrictions on the nature of the embedded verb; verbs of any category combine with the *-DIK* suffix to form a complement clause.

4.7 Other criteria

There are no constrains on the number, person or animacy specification of the subject internal to the complement clause.

4.8 Sorting out overt internal subjects

The total number of the examples in which the referent of the subject of the complement clause and the referent of the subject of the matrix clause are the same are 134. None involves an overt subject at the same time both in the matrix and the embedded clause, suggesting that Turkish has been a well-behaved consistently null-subject language throughout its history.

Of the 1717 complement clauses, 1023 have an overt internal subject. Of these 560 have a subject in the genitive, while 463 have a subject in the bare/nominative form. The distribution of the subject types across the centuries are given in Table 8.

Table 8. The Distribution of the Subject Marking in Complement Clauses across the Centuries

	Overt Internal SUBJ	
CENTURIES	GEN on SUBJ	No GEN on SUBJ
Seventh-Eleventh Centuries	1	0
Thirteenth-Fourteenth Centuries	23	93
Fifteenth Century	15	117
Sixteenth Century	54	116
Seventeenth Century	4	13
Eighteenth Century	14	43
Nineteenth Century	33	14
Twentieth Century	66	19
Twenty-first Century	350	48
Total	560	463

As seen in Table 8, the number of complement clauses with *-DIK* is rather few in the Old Turkic period when compared to the following periods. Moreover, there is only one example with an overt subject. Hereafter, we are obliged to leave this stage aside.

4.9 Change in the ratio of genitive-subjects

In table (9), the proportional change in genitive cases in complement clauses can be seen. While there is a small decrease in the occurrence of the overt genitives between thirteenth -fourteenth centuries and fifteenth century, there is a good increase of over 50% between the fifteenth century and the sixteenth century.

Table 9. The Increase of Genitive-Subjects in Complement Clauses (in %)

CENTURIES	Percentage change %
Seventh-Eleventh Centuries	19.83%
Thirteenth-Fourteenth Centuries	11.36%
Fifteenth Century	31.76%
Sixteenth Century	23.53%
Seventeenth Century	24.56%
Eighteenth Century	70.21%
Nineteenth Century	77.65%
Twentieth Century	87.94%

When we plot the gen/non-gen proportion across the subcorpora (between thirteenth and seventeenth century) the figure 3 emerges (basis for normalization of the data in this thesis is 100.

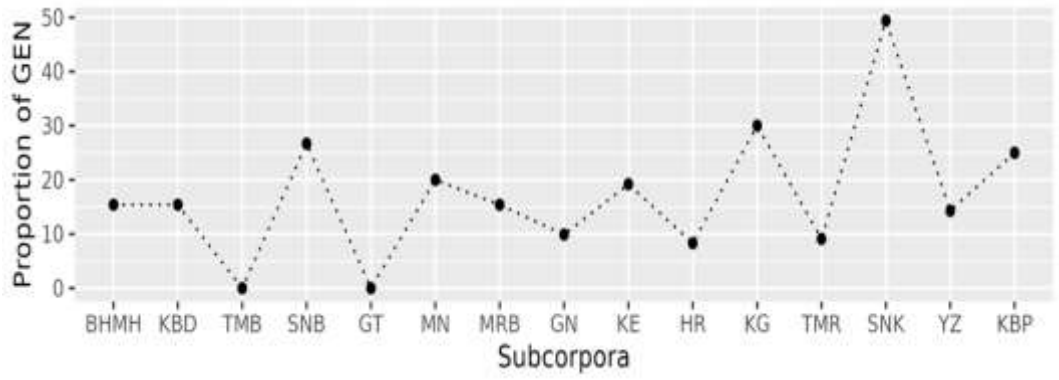


Figure 3. Relative proportion of genitive marked subjects to nominative marked ones across the subcorpora in chronological order (thirteenth-seventeenth centuries)

However, this should be read with caveat. A text in the sixteenth century, namely SNK, is an outlier in that its genitive subjects are 35.47 percent more than the other texts from the same century. We do not really know why this text behaves rather spuriously both here and in relative clauses (chapter 5). When we take this text out and proportions the changes, the table 10 emerges:

Table 10. The Ratio of Genitive-Subjects in Complement Clauses without SNK

CENTURIES	Percentage change %
Seventh-Eleventh Centuries	19.83%
Thirteenth-Fourteenth Centuries	11.36%
Fifteenth Century	15.56%
Sixteenth Century	23.53%
Seventeenth Century	24.56%
Eighteenth Century	70.21%
Nineteenth Century	77.65%
Twentieth Century	87.94%

Even without the SNK, there seems to be an increase in the sixteenth century in terms of overt subjects with the genitive.

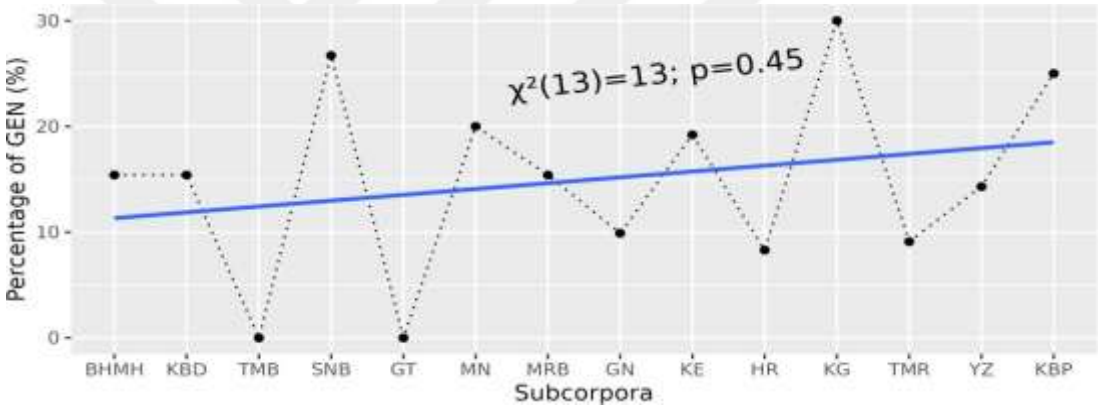


Figure 4. Increase in the proportion of GEN-marked subjects from thirteenth to seventeenth centuries.

The upward trend of genitive-marked subjects across the centuries (between thirteenth and seventeenth centuries) can also be visually demonstrates. As can be seen in Figure 4, Kruskal-Wallis test (Kurskal & Wallis 1952) shows a moderately significant increase of the genitive (p=0.45).

As the Figure 5 shows, although genitive-marking increases very moderately since the first attestation of Turkic, it reaches a sudden increase around the nineteenth century. Crucially, it never reaches a 100%, as expected, because nominative-marked

subjects are still available in Modern Turkish as non-spesific, incorporated ones. The plot in Figure 5 also fits in well with the well-established logistic curve (S-curve) in diachronic syntactic changes: a change starts slowly, then makes a sudden increase in a short time span and then tapers off again (Kroch, 1989; Denison, 2002). We will return to the plot in Chapter 6.

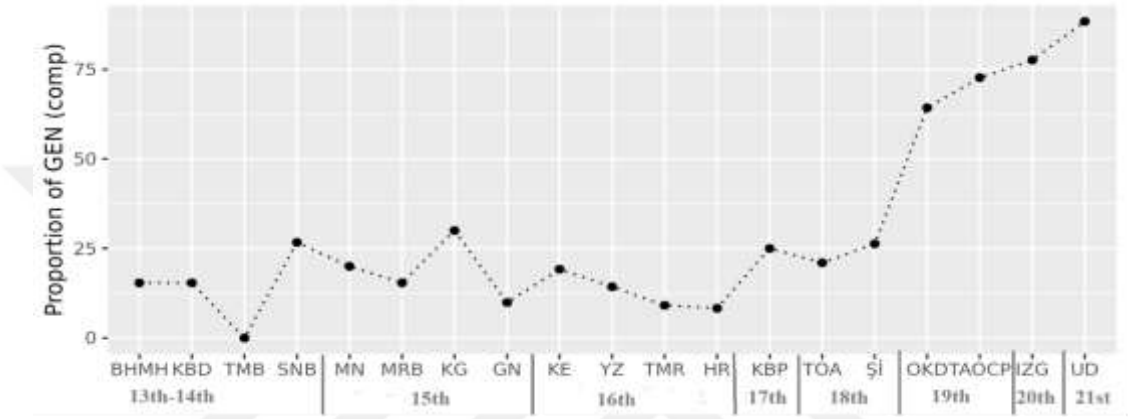


Figure 5. Relative proportion of genitive marked subjects to nominative marked ones across the subcorpora in chronological order (thirteenth-twenty-first centuries).

4.10 Chapter summary

Nominal agreement is obligatory from very early on the *-DIK*-marked verb of the in complement clauses, which suggests that it may be a feature inherited from a previous stage. On the other hand, we cannot say the same for the genitive marking on the internal subject of complement clauses. From the early periods until the nineteenth century, the proportion of genitive-marked subjects to those that are marked in the nominative was low although there was a steady increase. A sudden increase was observed around the nineteenth century. The change in this ratio, although other factors seem to be constant, indicates that we are facing at that point in time a diachronic change relevant to genitive- vs. nominative-marking of the internal subjects. We will return to this in chapter 6

CHAPTER 5

HEADED RELATIVE CLAUSES

5.1 Headed relative clause in Modern Turkish: The preamble

In Modern Turkish, subject relative clauses and non-subject relative clauses with non-future temporal reference differ in the relativization strategy (Underhill, 1972; Hankamer & Knecht, 1976; Csato, 1985; Barker, Hankamer & Moore, 1990): While both are prenominal with a gap, *-(y)An* suffix is chosen when the nominal relativized on is the subject internal to the relative clause. In this case, there is also no nominal agreement on the *(-y)An*-marked verb as seen in (24):

(24) kitab-ı ver-en çocuk
book-ACC give-AN child
"the boy who gave the book"

Non-subject relative clauses, on the other hand, are established with the *-DIK* suffix. Furthermore, the subject of the relative clause is indexed with a nominal agreement marker on the *-DIK*-marked verb (Kornfilt, 1997; Hankamer, 1972; Underhill, 1972; Hankamer & Knecht, 1976; Csato, 1985; Barker, Hankamer & Moore, 1990). Finally, if the specific internal subject is overt, it bears genitive case (25a). Relative clauses become ungrammatical when the nominal agreement is absent, as can be seen in example (25b), or if it is placed on the Head (25d), or when the specific internal subject does not bear genitive (25c).

(25) a. Ben-im git-tiğ-im ev
I-GEN go-DIK-1SG home
"the house I go to."

b. *Benim git-tik ev
I-GEN go-DIK home
"the house I go to."

c. *Ben git-tiğ-im ev
I go-DIK-1SG home
"the house I go to."

d. *Ben-im git-tik ev-im
I-GEN go-DIK home-1SG
"the house I go to."

While the above presentation captures the typical situation, it does not reveal the complete situation. In the next two subsections, we provide more details on the headed relative clauses.

5.1.1 Incorporation in non-subject relative clauses in Modern Turkish

In non-subject relative clauses where agent- or theme-subject incorporation happens within the relative clause, *-(y)An* suffix is used instead of *-DIK*. As in subject relatives, in this case as well, there is no nominal agreement (Hankamer & Knecht, (1976); Barker, Hankamer & Moore, 1990; Haig, 1997). To illustrate this, consider the examples in (26)-(27), which are all non-subject relatives. Examples (26a) and (26b) have specific agent- and theme- subjects respectively. As discussed in the previous section, they bear genitive morphology. Furthermore, these subjects can be separated by the verb by an intervening XP, say a time adverbial, such as *dün* 'yesterday'. On the other hand, in examples (26b) and (27b) the NP subjects receive necessarily a non-specific interpretation in the absence of the genitive. Nominative-marked subjects in non-subject relative clauses must be immediately preverbal, witness the ungrammaticality caused by the intervening *dün* 'yesterday'. This is the reason why we refer to (26b) as agent incorporation and (27b) as theme incorporation. In these configurations, relativization does not proceed with *-DIK*, rather, it is established with the *–(y)An* suffix, and there is no nominal agreement, as Barker, Hankamer and Moore (1990) and Haig (1997) already pointed out.

(26) a. Doktor-un (dün) gel-diğ-i köy
doctor-GEN yesterday come-DIK-3SG village
"the village that the doctor came yesterday"

b. Doktor (*dün) gel-en köy
doctor yesterday come-AN village
"the village that a doctor came"
Literally= "the village that doctor-coming happened"

(27) a. Kitab-ın (dün) oku-n-duğ-u ev
book-GEN yesterday read-PASS-DIK-3SG house
"the house that the book was read"

b. Kitap (*dün) oku-n-an ev
book yesterday read-PASS-AN house
"the house that book-reading happened"

Additionally, passive structures that involve theme incorporation results in impersonal reading. (27b) is incompatible with a *by-phrase*. Incorporated themes will therefore be visible as syntactic objects. Additionally, agent incorporations are not visible as syntactic arguments (Öztürk, 2009). In Öztürk (2004, 2005, 2008), it is argued that such subjects do not have to raise into [Spec, TP] but can remain in their base-position.

When we compare incorporation facts with that in complement clauses (section 4.3), we see that the syntactically non-visible arguments only affect genitive case-marking in the complement clause, while a completely different structure is used in relative clauses. In complement clause structures with incorporated agent or theme, there is a nominal agreement with the *-DIK* suffix and the predicate agrees with the incorporated subject. However, this agreement is not seen in the relative clause with incorporated subjects, which are instead formed by relativizer *-(y)An*.

5.1.2 Subject relative clauses with *-DIK* in Modern Turkish

While it was maintained in section 5.1, that subject relatives are formed with the *-(y)An* suffix, this statement should be relaxed. Modern Turkish does also allow the

formation of subject relatives with the *-DIK* suffix, albeit in a very limited usage/genre. Almost all subject relativels with *-DIK* have a verbal predicatemarked in the passive suffix *-Il,- In*, and almost all of them carry the negation marker (28a). As seen in example (28b), if the negation suffix *-mA* does not occur on the verb of the relative clause, the structure becomes unacceptable. Additionally, as seen in example (28c), if the verb of the relative clause is in the active voice, the sentence will also become ungrammatical. There are very few examples as fixed expressions that go against these, one of which is given in (28d). This observation was already made by Kornfilt (1997), who states that these are frozen expressions. We therefore conclude that subject relative clauses in the ordinary case is formed with the *-(y)An* suffix.

(28) a. Kır-ıl-ma-dık kemik
break-PASS-NEG-DIK bone
"unbroken bone"

b. *Kır-ıl-dık kemik
break-PASS-NEG-DIK bone
"unbroken bone"

c. *Kır-ma-dık kemik
break-NEG-DIK bone
"unbroken bone"

d. Ol-ma-dık iş-ler
happen-NEG-DIK work-PL"
"unusual works"

5.2 Headed relative clauses within historical data: An overview

Non-subject relative clauses have been formed in a large array of syntactic structures in the history of Turkish. The most common headed relative clause structure we encounter involves nominal agreement of the subject on the verb of the embedded clause, which can be seen in examples (29a) and (29b). Of 2125 examples 2036 have

nominal agreement in embedded clause. Moreover, as seen in examples (29a) and (29b), genitive case may be present on the subject of embedded, but it does not have to, regardless of the specificity of the referent of the subject NP.

(29) a. kişi-nüŋ otur-dug-ı yir
person-GEN sit-DIK-3SG place
"the place the person sit

(EM 58b 1)

[[Subj-GEN V-DIK-Agr] Head]

b. Musa di-düg-i öküz
Moses mention-DIK-3SG ox
"the ox Moses mentioned"

(KE 190aBL 8)

[[Subj-NOM V-DIK-Agr] Head]

In another relative clause structure nominal agreement is missing on the relative verb (30). In this group one possibility is to locate the nominal agreement suffix on the Head of the relative clause (30a-b). In this configuration, the internal subject may bear genitive case, as in (30a) but it should be noted that the number of such examples is only two and interestingly, both examples appear in GN, from the fifteenth century. Additionally, the internal subject may also be left in the nominative, as in (30b). Finally, a relative clause may also involve neither genitive case on the internal subject, nor nominal agreement on the verb or on the head (30c). It is very difficult to tell apart structurally the configurations in (30b) and (30c), simply because in (30b) it is nearly impossible to determine in a written text whether the agreement suffix on the head of the relative clause is a possessive marker, or and agreement marker flagging the subject of the relative clause.

(30) a. Taŋrı-nuŋ sev-dük kul-ı
God-GEN like-DIK slave-3SG
"God's beloved servant"

(GN 7658)

[[Subj-GEN V-DIK] Head- Agr]

b. Ben öz-üm bilge tonyukuk ötün-tök ötünç-üm
I self-1SG wise Tonyukuk propound-DIK propoundation-1SG
"the propoundation that I myself, Tonyukuk The Wise, propound."
(T15)

[[Subj-NOM V-DIK] Head-Agr]

c. göz-i gör-dük yir-e
eye-3SG see-DIK place-DAT
"to the place his eye saw"
(GN1467)

[[Subj-NOM V-DIK] Head]

In any event, all of the examples cited in (30) are non-subject relative clause structures. Apart from these, there are also subject relative clause structures established with the *-DIK* suffix, whic has mentioned in section 5.1.2, are also possible for Modern Turkish, albeit in a marginal form in the relevant Modern Turkish example in (28a) is cited here again.

(=28a) a. kır-ıl-ma-dık kemik
break-PASS-NEG-DIK bone
"unbroken bone"

Unlike Modern Turkish, there are numerous examples from the historical corpusin which the verb of subject rleatives with *-DIK* is in the active form and there is no negation. An example is cited in (31). This points out to the possibility that these structures had a wider distribution historically.

(31) yumurtla-ma-duk bülüc
lay.egg-NEG-DIK chick
"the chick that didn't lay egg" egg"
(MŞ M110b 15)

[[V-NEG-DIK] Head]

It should be noted here that we did not focus on subject relative clauses in this thesis.

Headed relative clauses with *-DIK* across the centuries are given in raw numbers in Table 11.

Table 11. Headed Relative Clauses across the Centuries

	Overt		Non-overt	[DIK+AGR] (YES)	[DIK+AGR] (NO)
CENTURIES	Gen	No Gen			
Seventh-Eleventh Centuries	0	1	30	1	30
Thirteenth-Fourteenth Centuries	6	78	133	203	14
Fifteenth Century	9	31	120	138	22
Sixteenth Century	35	129	278	422	20
Seventeenth Century	6	14	46	64	2
Eighteenth Century	76	34	223	333	0
Nineteenth Century	88	21	154	263	1
Twentieth Century	93	17	138	248	0
Twenty-first Century	362	1	0	364	0
Total	675	327	1122	2037	89

There are no exclusive criteria that affect the headed relatives; virtually verbs, and NP heads or internal subjects with all person, number, animacy specifications appear in relatives.

What appears from the presentation so far are a number of diachronic changes in the headed relative clauses in the history of Turkish. To summarize these, in modern Turkish, in non-subject relative clauses with *-DIK*, genitive case appears on the internal subject (and the subject is necessarily specific), and an agreement marker is obligatory after *-DIK*. However, it is seen that in older periods of Turkish, that none were strictly obligatory. Quantitative data regarding this diachronic change will be given in the following sections.

5.3 Historical data: Raw numbers

5.3.1 Non-subject relative clauses with nominal agreement after *-DIK*

In 989 of the total 2037 examples in our database, the subject of the relative clause is overt, and in 669 of them, the subject has a genitive case. There is only one example from the Old Turkic period with a nominal agreement in the embedded clause, and in this example the subject is not overt. For this reason, we will exclude the Old Turkic

period from our evaluation. On the other hand, the use of genitive case is less in every period in the post-Old Turkic period until the eighteenth century. We see that for the first time in the eighteenth century, the number of subjects with the genitive case becomes higher than those without. The incidence of genitive cases has increased rapidly since this century. Table 12 summarizes the above information.

Table 12. The Distribution of Non-Subject Relative Clauses with Nominal Agreement in Embedded Clauses

	Overt		Non-overt	[DIK+AGR] (YES)
CENTURIES	Gen	No Gen		
Seventh-Eleventh Centuries	0	0	1	1
Thirteenth-Fourteenth Centuries	6	74	123	203
Fifteenth Century	6	25	107	138
Sixteenth Century	32	132	258	422
Seventeenth Century	6	14	44	64
Eighteenth Century	76	34	223	333
Nineteenth Century	88	21	154	263
Twentieth Century	93	19	138	250
Twenty-first Century	362	1	0	363
Total	669	320	1048	2037

Table 13 offers the percentage of genitive case usage across the centuries. While relatives with genitive-marked subjects consitute 7.50% of all relatives in seventh-eleventh centuries, it increases roughy to 20% in the thirteenth-fourteenth centuries. In the seventeenth century, it reached 30%. Crucially, it is seen that the rate has reached 70% since the eighteenth century, 80% in the nineteenth and twentieth centuries, and almost 100% today. In sum, a steady rise has been observed since the thirteenth century and since the eighteenth century, the rate has increased rapidly.

Table 13. The Percentage of Genitive Case Usage across the Centuries

CENTURIES	Percentage
Seventh-Eleventh Centuries	7.50%
Thirteenth-Fourteenth Centuries	19.35%
Fifteenth Century	19.51%
Sixteenth Century	30.00%
Seventeenth Century	69.09%
Eighteenth Century	80.73%
Nineteenth Century	83.04%
Twentieth Century	99.45%

5.3.1.1 A note on agent-incorporation

Before moving on with data presentation, we would like to point out structures in the date which may involve agent-incorporation in data. In section 5.1.1, we mentioned that *-(y)An* suffix is used in non-subject relative clauses with agent and theme incorporation in the embedded clause in Modern Turkish. As seen in example (29b), genitive case is not always used in even in cases where the subject of the embedded clause is specific. However, when we look at the example in (32) and the context in the text from which this example is taken, it can be said that *it* 'dog', which is the subject of the embedded clause, is not specific. Therefore, one may conclude that this structure also ınvolves agent-incorporation. Although it is quite difficult to conclude from a written text whether this is indeed so, we may still conclude that the rule of using the *-(y)An* suffix in non-subject relative clauses with incorporation in the embedded clause in Modern Turkish may not have existed in much earlier periods. *-DIK* suffix and nominal agreement may have been used in previous periods even in this configuration. The true evaluation of this awaits a full-fledged diachronic study on the relative clauses with *–(y)An*, which we cannot undertake in this thesis and we will leave this issue for further research.

(32) it ısır-dug-ı yer-e
dog bite-DIK-3SG place-DAT
"the place a dog bites"

(MŞ M22B 9)

5.3.2 Non-subject relative clauses without nominal agreement after *-DIK*

In the non-subject relative clause examples, we will talk about in this section, agreement does not appear after *-DIK*. The internal subject is overt in only 14 out of a total of 50 samples. Only two of them have a genitive in the internal subject. In the remaining 12 examples, the internal subject is in nominative form. On the other hand, in 16 examples, agreement appears after the head as in (33.c). In the remaining 34 examples like (33a-b), nominal agreement did not exist. In (33), different non-subject relative clause structures in which agreement did not appear after *-DIK* in data are schematized. Additionally, the distribution of these structures across the centuries can be seen in table 14.

(33) a. [[Subj-NOM V-DIK] Head]

b. [[Subj-NOM V-DIK-Agr] Head-AGR]

c. [[Subj-GEN V-DIK] Head-Agr]

Table 14. The Distribution of Non-Subject Relative Clauses without Nominal Agreement across the Centuries

	Overt subject		Non-overt subject	[DIK+AGR] (NO)	[HEAD+AGR] (YES)	[HEAD+AGR] (NO)
CENTURIES	Gen	No Gen				
Seventh-Eleventh Centuries	0	1	20	21	8	13
Thirteenth-Fourteenth Centuries	0	4	5	9	3	6
Fifteenth Century	2	7	10	19	5	14
Sixteenth Century	0	0	0	1	0	1
Seventeenth Century	0	0	0	0		
Eighteenth Century	0	0	0	0		
Nineteenth Century	0	0	0	0		
Twentieth Century	0	0	0	0		
Total	2	12	35	50	16	34

As we mentioned in 5.3.1, the number of non-subject relative clauses where nominal agreement appears after *-DIK* in the Old Turkic period is only one. However, as can be seen in Table 14, nominal agreement did not appear after *-DIK* in 21 non-subject relative clauses in Old Turkic period. Therefore, the structures seen in (33) are the most common structures in Old Turkic. In Post-Old Turkic period, structures in (33) are very rare. In the Post-Old Turkic. The structures in (33) are found in only 29 of 2086 non-subject relative clauses. In only two of them, the internal subeject also has a genitive, as in (33.c), and agreement exists after the head. This example is from GN in the fifteenth century. In other examples, overt subjects are in nominative form, as in (33a-b). Only eight of these have agreement after the head. These structures disappeared in the sixteenth century, and nominal agreement became obligatory in embedded clauses (see also section 5.3.1)

5.3.2.1 Non-subject relatives without agreement in other Turkic languages

In this section, we would like to briefly talk about non-subject relatives in other Modern Turkic languages. Kornfilt (2012) mentions that there are three main non-subject relative clause structures in Turkic languages. The first of these is the examples with nominal agreement in the embedded clause, as in Modern Turkish. The second possibility is forming relative clauses in which there is no nominal agreement and no genitive case, as seen in (34). This is observed, among others, in Uzbek, Uyghur, Azerbaijani and Kyrgyz.

(34) a. [men qil-gan] isloh
I do-P reform(s)
"(the) reforms which I have made" (Uzbek)

b. [sen iste-yen] sey
you want-P thing
"the thing that you want" (Azerbaijani)
(Kornfilt, 2012, p. 515 (original reference omitted))

The third structure is the relative clause structures, whose subject is marked in the genitive and addiditionally a agreement marker marks the head of the relative clause, as in (33c). As seen in (35) This structure is found in Tuvan, Uzbek, Turkmen, Kazakh and Uyghur, among others (Kornfilt, 2012).

(35) Uzbek:
[men-in gapir-gan] gap-im
I-GEN say-P word-1SG
"the word(s) I said"
(Kornfilt, 2012, p. 516 (original reference omitted))

Such structures, which are used extensively in other contemporary Turkic languages, were also present in the Old Anatolian Turkish period, and continued to exist until the sixteenth century in our corpus. Observe further that the suffix *-(y)An/GAn* is used in non-subject relative clauses in numerous contemporary Turkic languages (Kornfilt, 2012; Satik, 2020). This suffix is the same as *-(y)An* suffix in Modern Turkish. [g], especially found in suffixes, disappeared in the Old Anatolian Turkish period (Timurtaş, 1973). The *-(y)An* suffix is used only in subject relative clauses in Turkish and exceptionally in non-subject relative clauses with incorporated subject in Modern Turkish. The striking point is that although they are non-subject relative clauses, the agreement is not used after the *-(y)An/-GAn* suffix in most Modern Turkic languages (Kornfilt, 2003, 2012 a.o). On the other hand, Haig (1997) mentions that there are examples in East Middle Turkic where the relativizer *-(y)An/-GAn* is used along with agreement on the predicate, as seen in example (36).

(36) oltur- ɣan-ım yär-dä monçuk-lar-ım-nı bul-du-m
sit-AN-1SG place-LOC pearl-PL-1SG-ACC find-PAST-1SG
"I found my pearls at place I had been sitting.
(Haig,1997, p. 200 (original reference omitted))

The relationship between embedded *-(y)An* clause and *–DIK* clause is quite interesting and is a subject that should be evaluated across Turkic languages, both diachronicly and synchronically. We will however leave it to future studies.

5.4 The overall change in genitive marking.

As stated in section 5.3.1, the number of genitive-marked subjects in relative clauses with *-DIK* increases throughout the centuries. In Figure 6, we plot this increase (SNK is excluded as in 4.9).

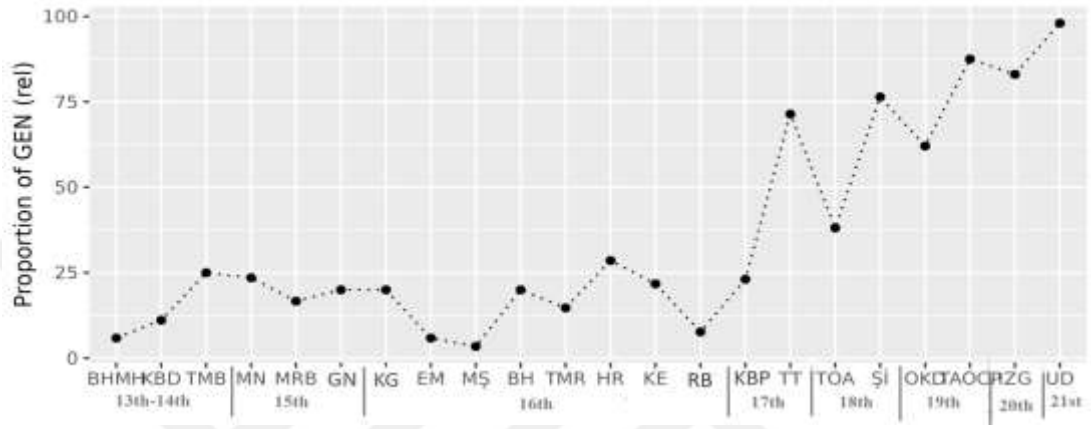


Figure 6. Increase of genitive-marking in relative clauses across the centuries

As figure 6 makes it clear, the change in the genitive-marked subjects is insignificant until the late seventeenth-eighteenth century, when the change speeds up. In the twenty-first century, we reach to a point where almost 100% of *-DIK* relative clauses bear a subject marked in the genitive. The few examples that do not have genitive marking in the subcorpus from the twenty-first century are due to mistakes in the annotation. One example is given in (37).

(37) [Başesgioğlu ver-diğ-i önerge-yle]……düzenleme-ler-in
Başesgioğlu give-DIK-3SG proposion-INS arrangement-PL-GEN
devam-ı sağla-n-dı.
contuniation-3SG provide-PASS-PAST
"with proposion that Başesgioğlu gave.....the arrangement were continued"

This example is ungrammatical for Modern Turkish under the reading that the *Başesgioğlu* is the relative clause-internal subject (indicated in square brackets) but this is how that NP is annotated. We think a mistake has been made here. The author

may have been confused because the headed relative clause is very far from the matrix cluster's predicate, which is in passive form. The passive may be a slip-of-the -tongue. *Besesgioğlu*, the so-called subject of the headed relative clause is to be understood as the subject of the matrix clause, and in this configuration the subject of the headed relative clause would not be overt.

We will return back to the interpretation of the curve in Figure 6 in Chapter 6

5.5 Chapter summary

Examples of non-subject relative clauses with no nominal agreement in the embedded clause can be seen until the fifteenth century (see also 5.5.1.2). However, the number of these examples is quite small compared to the examples in which nominal agreement existed, and these examples have disappeared after the sixteenth century. Therefore, we can argue that the examples where nominal agreement did not exist for the early periods of Old Anatolian Turkish are exceptional or archaic. While the number of examples where nominal agreement exists is 1674, the number of examples where nominal agreement does not exist is only 48. Therefore, nominal agreement is obligatory for non-subject relative clauses from the sixteenth century onwards, and it is also the highly preferred variable between the twelfth and sixteenth centuries.

When we look at encoding subjects of the headed relative clauses, there are large differences proportionally over the centuries. For non-subject relative clauses, a greater proportion of subjects were in the nominative until the late seventeenth-eighteenth century. Although there were small increases in the genitive case rate until these centuries, it is proportionally less than the nominative form. However, the rate of existence of genitive cases has increased rapidly since the eighteenth century.

Because; We can argue that the turning point of change was the eighteenth century for relative clause.

5.6 A note on headless relative clause

The grammatical structure of non-subject headless relative clauses in Turkish are similar to the headed ones, and both can be taken as headed relatives, albeit the headless ones involving a silent head (Cinque 2020). The subject of the embedded clause has genitive case and embedded clause necessarily bears nominal agreement.

As mentioned in Section 5.3.2, there are also examples in which the agreement marker appears on the head in the pre-Modern Turkish period. In such examples, if the head of the relative clause is not overt, the location of the agreement cannot be determined in the headless relative clause, as given in (38).

(38) a. ben-im ver-diğ-Ø-im-i getir-in
I-GEN give-DIK-Ø-3SG-ACC bring-IMP
"bring what I give."

b. ben-im ver-diğ-im-Ø-i getir-in
I-GEN give-DIK-3SG-Ø-ACC bring-IMP
"bring what I give."

Additionally, in many examples it is difficult to distinguish between the complement clause and the headless relative clauses. As can be seen from (4) and (5) are identical on the surface, although they have different underlying structures. To distinguish these, we benefited from context. However, we excluded from the data those we could not decide on.

(=4) ye-diğ-in-i gör-dü-m.
eat-DIK-3SG-ACC see-PAST-1SG
"I saw that you ate."

(=5) ye-diğ-in-Ø-i gör-dü-m
eat-DIK-Ø-3SG-ACC see-PAST-1SG
"I saw what you ate."

Additionally, a similar surface ambiguity appears between headless relative clauses and adjunct clauses. Relevant examples are given in (39). (39a) illustrates an adjunct clause and (39b) a headless relative clause.

(39) a. sen gel-diğ-in gibi gid-eceğ-iz
you come-DIK-3SG as go-FUT-1PL
"We will go as you come"

b. sen-in gel-diğ-in gibi gid-eceğ-iz.
you-GEN come-DIK-3SG like go-FUT-1PL
"We will go like how you come"

The only difference between these two examples is the presence of genitive on the internal subject of the headless relative clause. However, as seen in the previous discussion, genitive case is not obligatory in pre-Modern Turkish. For this reason, it is very difficult to identify the categories of such examples in the pre-modern period. As far as can be understood from the context, we have identified what we think are headless relative clauses. We removed the ones we were in doubt about from the data. Since in the end it turned out to be difficult to securely identify headless relative clauses, we decided to exclude them from the evaluation and focused on headed relative clauses

CHAPTER 6

DIACHRONIC DISCUSSION

6.1 Taking stock

In Chapters 4 and 5 revealed two important changes in the history of Turkish with respect to two types of embedded clauses with *-DIK*, namely headed relative clauses and complement clauses to predicates. First, although we established that nominal agreement has been obligatory in the complement clauses since their earliest attestations, examples of non-subject relative clauses without nominal agreement in the embedded clause existed until the sixteenth century (see also 5.3.2), albeit in small numbers. Therefore, nominal agreement has beenobligatory for non-subject relative clauses since the sixteenth century onwards, and it was the highly preferred variant between the twelfth and sixteenth centuries. In complement clauses, nominal agreement was there throughout.

Second, when we look at the genitive-marking on the subject of complement and relative clauses, there are large differences proportionally over the centuries. For non-subject relative clauses, a greater proportion of subjects were in the nominative form until the eighteenth century. In complement clauses, this situation continues until the nineteenth century. Although there was a small increase in the genitive case rate until these centuries, it is proportionally less than the nominative form. However, the rate of genitive cases has increased rapidly since the centuries we mentioned for both clauses. We can argue that the beginning of the change was the late seventeenth century, the early eighteenth century for headed relative clauses, and the nineteenth century for the complement clauses. The plots in figures 5 and 6 are brought together in figure 7, to visualize the time difference.

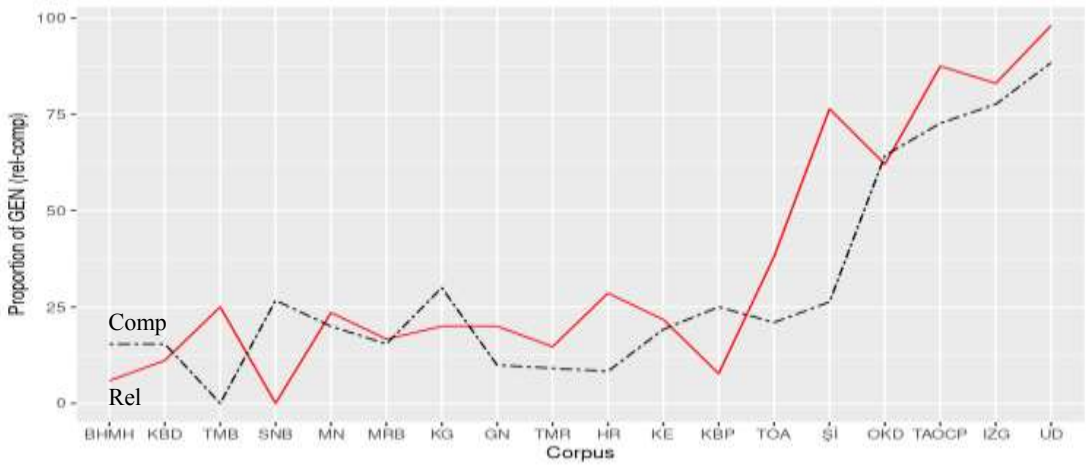


Figure 7. Change in genitive vs nominative in both complements and relatives

Figure 7 reveals that the change in genitive marking on relatives and complements happened with a lapse of roughly three generations (the time between TOA (1724) and OKD (1808). Furthermore, this change happens at a constant rate in both relative clauses and complement clauses, fitting well with the hypothesis that when some reason leads to a change in the syntactic configuration, the rate of replacement is the same in all contexts affected by the change, known as the Constant Rate Hypothesis (Kroch, 1989; Kauhanen & Walkden, 2018).

In this section, we discuss how this change, namely the increase in genitive marking of the internal subject came into being, asking in particular the question, what happened shortly before 1974. Anticipating the discussion, we claim that around the eighteenth century (for headed relatives) and nineteenth century (for complement clauses), animacy, which we claim was one defining factor in the ratio of genitive marking until then, was reanalyzed as definiteness, possibly because of parametric ambiguity in the input data to the acquirers. We therefore assume a view to syntactic change as an acquisition problem (Lightfoot, 1976; Lightfoot, 2006; Roberts, 2021; Yang, 2002).

6.2 The effect of animacy in genitive-marking in headed relative clauses

As seen in figure 6, and elaborated on chapter 5, the ratio of genitive-marked subjects to nominative-marked subjects is rather low until the eighteenth century. This being the case, we claim that the occurrence of genitive-marking is not stochastic. Crucially, in this stage, whether the referent of the subject is animate or not makes a difference. As figure 8 shows below, around 75% of the genitive-marked subjects have a referent that is animate, suggesting that animacy, although not a categorical predictor, affects whether genitive marker appears on a subject or not.

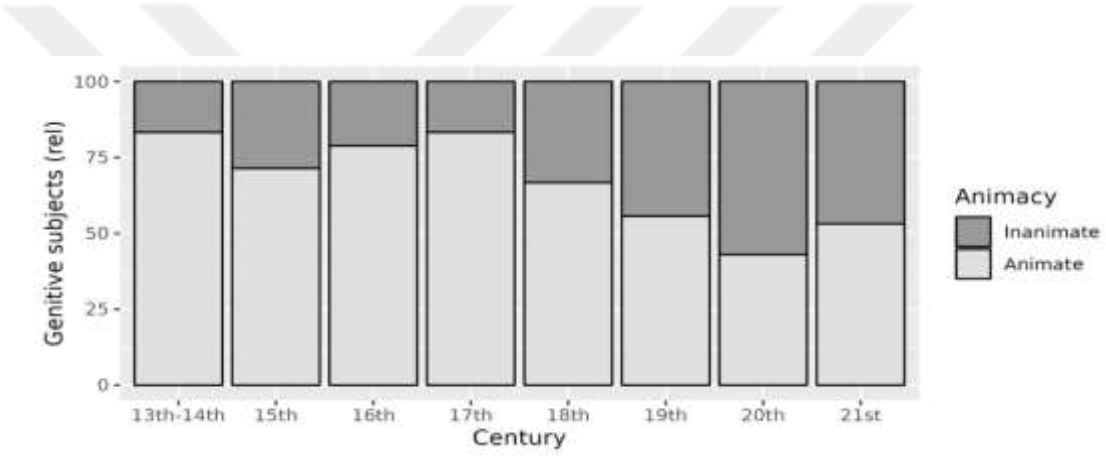


Figure 8. Animacy effect in relative clause for genitive-marking

Crucially, when we look at figure 8, we see ther beginning of a steady decrease in the effect of animacy on genitive marking around the eighteenth century. This is precisely the time when the genitive-marking increases. We interpret this as follows: while animacy still plays a role in genitive marking in and after the eighteenth century, its effect has been reduced gradually since then. To understand the effect of animacy on genitive marking, we have fitted a number of Linear Mixed-effect Models with binominal data, using lm4e package (Bates et al., 2012) and language R (Baayen 2008), ranging from simplex to complex. In the simplest case, we took

animacy and the century as the only predictors of genitive and further added into the models the interactions between animacy and century as separate predictor. The models given in Table 15

Table 15. Models for Likelihood Ratio Test

model1	Gen ~ Animacy + scale(Century) + (1 | ID)
model2	Gen ~ Animacy * scale(Century) + (1 | ID)
model3	Gen ~ Animacy + scale(Century) + (1 | ID) + (Animacy | Century)
model4	Gen ~ Animacy * scale(Century) + (1 | ID) + (Animacy | Century)

To compare the models we run a Likelihood Ratio Tests, whose results are given in Table 16.

Table 16. Result of Likelihood Ratio Test (Relatives)

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
model1	4	685.45	705.10	-338.72	677.45			
model2	5	687.11	711.67	-338.55	677.11	0.3415	1	0.5589
model3	7	658.88	693.27	-322.44	644.88	32,2292	2	1,004e-04***
model4	8	660.74	700.04	-322.37	644.74	0.1371	1	0.7112
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1								

Model3 emerges as the most appropriate model based on these comparisons. It significantly improves upon both the basic and interaction models by incorporating random slopes for Animacy, suggesting that the effect of Animacy on Gen varies in a complex manner across centuries. The lme result of the best model is in Table 17 below. Simply put, the results of the best model states that while the plays temporal scale (Century) plays definitely a significant role in the change of the rate of genitive, it improves when animacy as a significant factor is added.

Table 17. The lme Result of the Best Model for Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.3217	0.4598	0.700	0.4841
Animacy1	0.7565	0.3361	2.251	0.0244*
scale(Century)	1.9464	0.2492	7.810	5.74e-15***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

The predicted probabilities of genitive by century and animacy in relative clauses are given in Figure 9. While the error margin is wide for especially the initial stages for the relatives— probably because of fewer examples with inanimate subjects—the difference between animates and inanimates reduces after the eighteenth century.

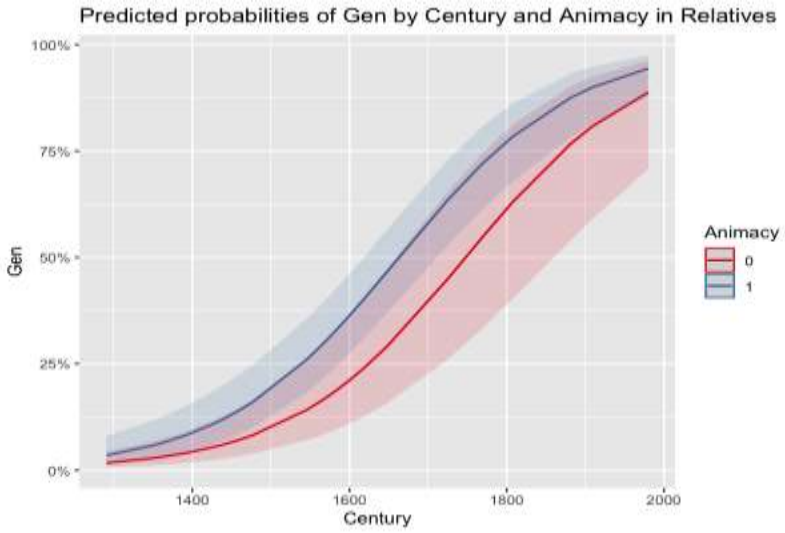


Figure 9. Predicted probabilities of genitive by century and animacy in relative clauses

To conclude, we have established here beyond impression that animacy had a significant effect on whether a subject NP of a relative is marked in the genitive or not before the eighteenth century. We see that this affect has reduced somewhat after the eighteenth century.

6.3 The effect of animacy in genitive-marking in complement clauses

Similar to the case of headed relative clauses, animacy plays a crucial role in whether the subject internal to the complement clause will be marked in the genitive or not. Figure 10 illustrates this point:

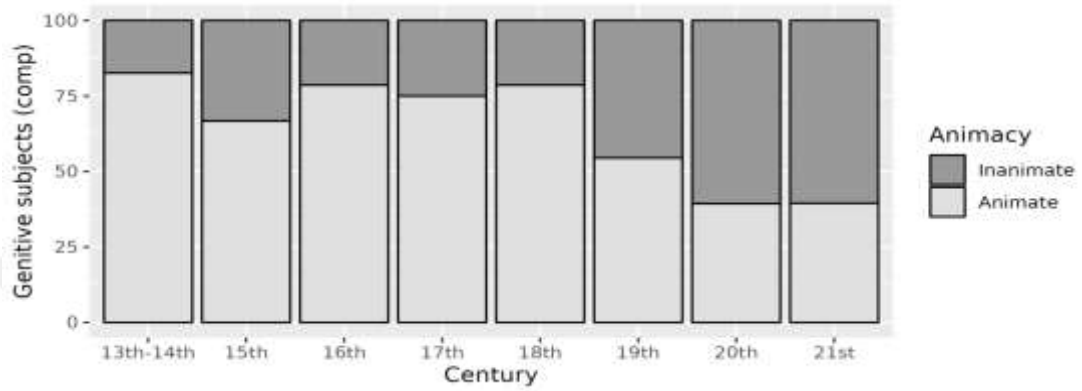


Figure 10. Animacy effect in complement clauses for genitive-marking

We notice a rapid decrease around the nineteenth century in the ratio of genitive-marked animate subjects with respect to inanimate ones, almost 100 years after. To understand the effect of animacy on genitive marking on the subject of complement clauses, we have relied on the same type of binominal mixed-effect models as in relative clauses (see Table 15). The Likelihood test yield the results in Table 19.

Table 18. Likelihood Ratio Test for Complement Clauses

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
model1	4	915.09	934.80	-453.54	907.09			
model2	5	915.75	940.38	-452.87	905.75	1.3399	1	0.2471
model3	7	887.26	921.75	-436.63	873.26	324.847	2	8,832e-05***
model4	8	889.00	928.42	-436.50	873.00	0.2615	1	0.6091
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1								

Model3 is the preferred model based on these results. It significantly improves upon Model1 and Model2 by incorporating random slopes for Animacy, capturing the

variability of Animacy's effect across centuries. The superiority of Model3 indicates that the effect of Animacy on the outcome variable Gen varies significantly across different levels of Century. This variability is best captured by allowing the slope of Animacy to change, rather than assuming it remains constant. It suggests that the relationship between Animacy and Gen is not uniform but depends on the context defined by century. For instance, the impact of being animate versus inanimate on the likelihood of Gen may increase or decrease over different centuries.

By incorporating random slopes, Model3 provides a more accurate and nuanced understanding of how Animacy influences Gen across Century. This results in a better fit to the observed data, as evidenced by the significant improvement in likelihood compared to models without random slopes. The lme results of the best model is given in Table 18 below:

Table 19. The lme Result of the Best Model for Complement Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.7399	0.1919	-3.855	0.000116***
Animacy1	14.550	0.2817	5.165	2.41e-07***
scale(Century)	17.100	0.2063	8.288	< 2e-16***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

That the effect of animacy decreases in the overt realization of genitives around the nineteenth century is also shown on the plot in Figure 11:

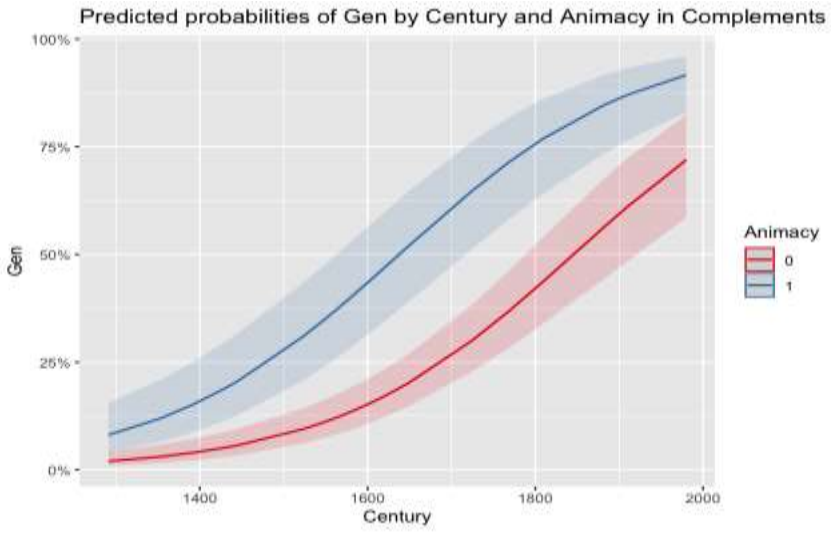


Figure 11. Predicted probabilities of genitive by century and animacy in complement clauses

To summarize, both complements and relatives exhibit an increasing trend in the predicted probability of genitive marking on the subject over the centuries, regardless of Animacy. More importantly, NPs with animate referents consistently show a higher probability of Genitive than inanimates in both configuration types, which suggests that Animacy *was* a significant predictor of genitive marking. The rate in predicted probability of genitives appears to be more pronounced in the relative data for animate NPs starting around the eighteen hundreds compared to the complement data. The confidence bands (shaded areas) for the relative data seem to be wider than for the complement data, suggesting more uncertainty or variability in the predictions for relatives.

6.4 Interim summary

As discussed in chapters 4 and 5, Modern Turkish does not mark its embedded subjects with the genitive if they are non-specific. In other words, in Modern Turkish, specificity/definiteness hierarchy plays a crucial role in embedded subject

marking (as well as direct object marking, yielding the so-called Differential Object Marking). The definiteness hierarchy is given in (40):

(40) Definiteness Hierarchy
pronoun > proper noun > definite > specific > non-specific
(Aissen 2003, p. 437)

Any subject that is specific or definite bears genitive case, as illustrated with complement clauses in (41b).

(41) a. köy-e doktor gel-diğ-in-i söyle-di-ler.
village-DAT doctor come-DIK-3SG-ACC say-PAST-3PL
"They said a doctor came to the village"

b. köy-e doktor-un gel-diğ-in-i söyle-di-ler.
village-DAT doctor-GEN come-DIK-3SG-ACC say-PAST-3PL
"They said a doctor came to the village"

In relative clauses, while specific subjects are encoded in the genitive, non-specific ones are not. Additionally, in the latter case, the verb bears *–(y)An* instead of *-DIK*. As in (42)

(=26) a.doktor-un gel-diğ-i köy
doctor-GEN come-DIK-3SG village
"the village that the doctor came"

b.doktor gel-en köy
doctor come-AN village
"the village that doctor came"
Literally= "the village that doctor-coming happened"

As the historical data so far has shown, this was not the case: specific/definite subjects could too appear in the nominative (20c, 29b), with the proviso that whether their referent was animate or not seems to have affected it, as discussed in the previous section.

(=20c) [Musa öldür-düg-in-i] hep bil-ür-ler.
Musa kill-DIK-3SG-ACC all know-PRS-3PL
"They all know that Moses killed"
(KE 155bBL 9)

(=29b) [[Musa di-düg-i] öküz]
Moses mention-DIK-3SG ox
"the ox Moses mentioned"

(KE 190aBL)

In what follows, we will argue that a change occurred in the most defining factor of the genitive-marking, which started around the eighteenth century; namely the predictor, which was the animacy hierarchy was reevaluated as the definiteness hierarchy. To proceed with this story, we will first look at the close relationship between animacy and definiteness.

6.4.1 The interaction between animacy and definiteness

Cross-linguistically, features such as definiteness, animacy and specificity are observed to be determining criteria in Differential Subject Marking (DSM) and Differential Object Marking (DOM). These can be summed up in two monotonic hierarchies, as proposed by Silverstein (1976): Definetness Hierarchy and Animacy Hierarchy. The first one, Definiteness Hierarchy is given below again.

(=40) Definiteness Hierarchy
pronoun > proper noun > definite > spesific > non-spesific

(Aissen, 2003, p. 437)

According to this view, if a language marks an argument with a feature situated on the scale, it also marks arguments with the features that are hierarchically higher than this feature (Aissen, 2003) too. In the Modern Turkish example, indefinite specific direct objects are marked with accusative case. whereas non-specific/generic objects are marked in the nominative case (42).

(42) a. (Ben) kitab-ı oku-du-m. (definite)
(I) book-ACC read-PAST-1SG
"I read the book."

b. (Ben) bir kitab-ı oku-du-m. (indefinite specific)
(I) a book-ACC read-PAST-1SG
"I read a certain book" resp.
"I read one of the books."

c. (Ben) bir kitap oku-du-m. (indefinite nonspecific)
(I) a book read-PAST-1SG
"I read a book."

d. (Ben) kitap oku-du-m. (bare/generic)
(I) book read-PAST-1SG
"I was (book) reading."

(Laszakovits, 2013, p. 45)

Since Modern Turkish marks indefinite specific objects with the accusative –*(y)I*, it marks all types of objects that are higher in the hierarchy and have features such as definite, name, pronoun. Marking type and the cutting point for marking varies cross-linguistically. As can be seen in (43), only pronominal objects are marked in English. Pronouns are at the top of the definiteness hierarchy, and other features below them are not marked.

(43) English (only pronoun)
*Leyla saw hi-**m**/Yusuf/the boy/a boy.*

(Haspelmath, 2008, p. 2)

Similarly, as can be seen in (44), proper noun and pronouns are marked in Old German when they are direct objects.

(44) older German (pronouns, proper nouns)
*Friedrich sah ih-**n**/Gertrud-**en**/das Kind/ein Kind.*
"Friedrich saw him/Gertrud/the child/a child."

(Haspelmath, 2008, p. 2)

Pronouns, proper names or common NPs that are definite feature are all marked in Persian (45).

(45) Persian (pronouns, proper nouns, and definite)
Hasan u-ra/ Ali-ra/ ketab-ra/ ketab did.
Hasan he-ACC Ali-ACC book-ACC book saw
"Hasan saw him/Ali/the book/a book."

(Haspelmath, 2008, p. 2)

There are also languages, such as Hungarian, which mark all the object types throughout, or languages, such as Vietnamese, in which no object type is marked. In every case however, marking or its absence proceed monotonically on the hierarchy in (40). This can also be visualized in Table 20.

Table 20. The Effect of Definiteness Hierarchy on Object-Marking Cross-Linguistically

	pronoun	proper n.	definite	specific	Non-specific
Vietnamese	-	-	-	-	-
English	m	-	-	-	-
Old German	m	m	-	-	-
Persian	m	m	m	-	-
Turkish	m	m	m	m	-
Hungarian	m	m	m	m	m

(Haspelmath, 2008, p. 3)

Besides definiteness, animacy is also a criterion that operates on argument-marking. The relevant hierarchy is given in (46).

(46) Animacy Hierarchy
human > animate > inanimate

(Aissen 2003, p. 437)

As in definiteness hierarchy, argument-marking based on the Animacy Hierarchy also proceeds monotonically. Thus, a language that marks arguments with any feature in the hierarchy in (46), will also mars arguments with features higher in the hierarchy likewise (Aissen 2003). To illustrate, only objects with human referents are marked in Spanish with the dative *a*—non-humans do not carry this marker (47) (this statement about Spanish will be slightly revised below, around the example (50)).

(47) Spanish (only human)
*El director busca el carro/el perro/**a** su hijo.*
"The director is looking for the car/the dog/his son."

(Haspelmath, 2008, p. 2)

In Russian, on the other hand, objects with the human or animate feature specifications are marked in overt accusative but inanimates are not (48).

(48) Russian (animate and human)
*Miša uvidel dom/kot-**a**/brat-**a**.*
"Misha saw the house/the cat/the brother."

(Haspelmath, 2008, p. 2)

Finally, Haspelmath (2008) reports that in Hungarian all objects are marked. (49)

(49) Hungarian (inanimate, animate and human -- no split!)

(Haspelmath, 2008, p. 2)

Examples of cross-linguistically different possibilities are given in Table 21.

Table 21. The Effect of Animacy Hierarchy on Object-Marking Cross-Linguistically

	human	animate	inanimate
Vietnamese	-	-	-
Spanish	m	-	-
Russian	m	m	-
Hungarian	m	m	m

(Haspelmath, 2008, p. 2)

Crucially, it is not obligatory for languages to mark their arguments according to only one of the two hierarchies. There are also languages that perform object-marking according to different combinations of them. For example, Spanish marks its objects differentially not only based on animacy; it also refers to definiteness. As can be seen in (50a-b), inanimate objects in Spanish encode definiteness with definite articles and in neither case is there the differential marker *a*. While neither does an indefinite NP with a human referent has the differential marker (50c), a definite one does (50d). Therefore, only human definite NPs are marked differentially in Spanish.

(50) a. El director busca un automóvil.
"The director is looking for a car."

b. El director busca el automóvil.
"The director is looking for the car."

c. El director busca un empleado.
"'The director is looking for a clerk (i.e. any-clerk)

d. El director busca al (= a el) empleado.
"The director is looking for the clerk.'

(Comrie, 1979, p. 15)

To provide another example from Mongolian, the accusative suffix *-(ii)g* in this language is used with NP objects of the following specifications (Poppe, 1970; as cited in Comrie, 1979): if the direct object is human, using the accusative suffix *-(ii)g* is obligatory (51). If the direct object is non-human and in immediately preverbal position, it is not marked (52a). If it is not in the immediately preverbal position, it is marked in the accusative case (52b).

(51) a. Dorž bagš -iig zalav
Dorj teacher DO invited
"Dorj invited the teacher."

b. bid nar olan xün -iig üzsen.
we many people DO saw
"We saw many people."

(Comrie 1979, p. 18)

(52) a. Čoidog zurag zurav.
Choidog picture painted
"Choidog painted a picture."

b. Zurag -iig Čoidog zurav
picture DO Choidog painted
"Choidog painted the picture; as for the picture, it was Choidog that painted it."

(Comrie 1979, p. 19)

We therefore conlcude that animacy and definiteness should not be thought as isolated categories and the interaction between them manifests itself in many languages.

6.4.2 Motivation for DOM and DSM

To emphasize the interaction between animacy and definiteness, Comrie (1979) claimed that examples where definiteness was claimed to be the defining factor in many descriptive grammars he examined could also be explained with Animacy Hierarchy (e.g., in Hindi and Tagalog). So, what is exactly the nature for this relationship? What is the underlying motivation behind the Definetness Hierarcy and the Animacy Hierarcy?

Caldwell (1856, as cited Haspelmath, 2008) mentions that it is more natural for rational beings to act than to act upon them. Therefore, when a NP with a referent that can be claimed to be a rational being is used as an object, case-marking is necessary to avoid missapprehension. Comrie (1989) further emphasized that it is natural for subjects to be high in the hierarchy of animacy and definiteness in a transitive contruction, and for objects to be low in these hierarchies. Being agent, human, animate, definite or specific are the characteristics that are expected to be seen in subjects. Being theme and having inanimate or non-specific properties are expected to be seen in objects. A hearer, so the functional analysis goes, expects this, and it is only natural. However, if there is an object with human feature or definite feature, it is natural to mark the object because this object can cause ambiguity compared to the subject. When the opposite happens, that is, if subject is inanimate or non-specific, this subject can be marked so that it does not cause ambiquity with the object.

DOM and DSM are actually about case-marking subjects and objects that have features that are opposite to this tendency. It seems that subjects and objects naturally tend towards opposite properties. For this reason, Silverstein (1976, as cited in Aissen, 2003) mentions that there is a mirror-image in the Definiteness and

Animacy Hierarchies seen in the DOM in DSM. That is, these hierarchies work in reverse in the DSM and DOM. Table 22 shows how subjects are marked according to different features in languages with DSM. When we compare table 22 with tables 21 and 20, a mirror-image is seen.

Table 22. The Effect of Definiteness Hierarchy on Subject-Marking Cross-Linguistically

	1s/2nd	3rd	proper	human	inanimate
Lezgian	M	m	m	m	m
Dyirbal	-	m	m	m	m
Guugu Yimidhirr	-	-	m	m	m
Gumbainggir	-	-	-	m	m
Lakhota	-	-	-	-	m
Hungarian	-	-	-	-	-

(Haspelmath, 2008, p. 10)

6.4.3 Parallelism between DSM and DOM in Modern Turkish

Modern Turkish seems to be an exception to the conlusion in the previous section. Contrary to the generalization given there, Kornfilt (2008) states that there is a parallelism between DSM and DOM in Turkish (see also Kennelly, 1997). This parallelism is seen in (53)-(54): specificity and not being in the immediately preverbal position are the determining factors in both DSM and DOM. Simply put, embedded subjects without the genitive case must be in preverbal position and non-specific, just like direct objects without the accusative case (see also (42)).

(53) a. köy-e doktor-(un) gel-diğ-in-i söyle-di-ler.
village-DAT doctor-GEN come-DIK-3SG-ACC say-PAST-3PL
"They said the/a doctor came to the village"

b. doktor-*(un) köy-e gel-diğ-in-i söyle-di-ler.
doctor-GEN village-DAT come-DIK-ACC say-PAST-3PL
"They said the/a doctor came to the village"

(54) a. Mehmet bu oda-da kitab-(ın) oku-n-duğ-un-u
Mehmet this room-LOC book-GEN read-PASS-DIK-3SG-ACC
öğren-di
learn-PAST.3SG
"Mehmet learn that the book was read in this room"

b. Mehmet kitab-*(ın) bu oda-da oku-n-duğ-un-u
Mehmet book-GEN this room-LOC read-PASS-DIK-3SG-ACC
öğren-di.
learn-PAST.3SG
"Mehmet learn that the book was read in this room"

6.5 What has (not) changed?

Sections 6.2-6.3 established that animacy was a crucial factor in the history of Turkish with respect to genitive marking on embedded subjects. In actual fact, it still is; in other words, its effect has never reduced to zero. Consider for instance the problem posed by the wh-indefinites as given in (55):

(55) a. . ne-(yi) gör-dü-n?
what-(ACC) see-PAST-2SG
"What did you see?"

b.. kim-*(i) gör-dü-n?
who-(ACC) see-PAST-2SG
"Who did you see?"

(Laszakovits 2013, p. 54)

While in (55a) the genitive cuts across the specificity line if the wh-indefinite refers to a non-human entity (*ne* vs *ne-yi*), the non-specific counterpart of the animate *kim* does not exist. Krause and von Heusinger (2019) further show that animacy is still a factor in accusative marking in Turkish. In their research, they selected objects according to three animacy categories (human, animal, inanimate) and then divided them into two sub-animacy categories ((i.e., Human (Relation Terms & Profession Terms), Animal (Wild Animals & Domestic Animals)), and Inanimate Entities (Concrete Entities & Abstract Entities). Additionally, all of these objects are indefinite. They showed their participants both examples where objects

were marked with accusative case and examples where they were not. They asked to what extent these were acceptable. Participants found the versions of indefinite objects with human features marked with the accusative case significantly more acceptable than their inanimate counterparts. Therefore, accusative marking is more acceptable on objects with the indefinite feature located just above specificity in the Definitness Hierarcy, and on humans located higher in the Animacy Hierarcy. Zero-marking is more acceptable for inanimate object located lower in the hierarchy.

We can conclcude that definiteness alone cannot probably affect genitive case marking. While inanimate, non-specific, indefinite subject are easily possible in nominative in complement clauses (56b), its animate counterpart are not so (56a). These examples further show that genitive marking on embedded subjects is still sensitive to animacy.

(56) a.*/? Köy-ü bir haydut bas-tığ-ın-ı duy-du-m.
village-ACC a robber raid-DIK-3SG-ACC hear-PAST-1SG
" I heard thatsome robber raided the village"

b. yol-dan bir araba geç-tiğ-in-i gör-dü-m.
road-ABL a car pass-DIK-3SG-ACC see-PAST-1SG
"I saw that a/some car went by on the road."

(Laszakovits 2013, p. 54)

We claim here however that this crucial factor (animacy) reduced in power after the eighteenth century stage although it never reached to zero, and specificity took over the leading stage.

Unfortunately, it is near impossible to define with security whether an embedded subject is specific or not in the historical corpora. Trying to figure this out simply from the context of the sentence is obliged to result in erroneous tagging. However, we know from Modern Turkish that specificity interacts closely with linear order, i.e., in pseudo-incorporation: As discussed in various sections in this thesis,

non-specific subjects of complement clauses in Modern Turkish appear immediately preverbally. A minimal pair is reprouduced below:

. (=16) a. Köy-e doktor gel-diğ-in-i söyle-di-ler.
village-DAT doctor come-DIK-3SG-ACC say-PAST-3PL
"They said a doctor came to the village"

b.*doktor köy-e gel-diğ-in-i söyle-di-ler.
doctor village-DAT come-DIK-3SG-ACC say-PAST-3PL
"They said a doctor came to the village"

If we take it that pseudo-noun incorporation is a syntactic reflex of specificity of the subject in Turkish, then we can provide real (but indirect evidence) for the absence of specificity effects of genitive marker, and the appearance of these effects, in the historical corpora. In what follows, we will undertake this task. In doing so, we follow the following axioms in (57), true observations for Turkish:

(57) a. if a subject is immediately preverbal and is not genitive marked then it is non-specific.

b. if a subject is separated from the verb by at least one intervening maximal projection, it is necessarily interpreted as specific/kind-referring, and bears genitive obligatorily.

6.5.1 Specificity in headed relative clauses

With (57) in mind, we have looked at the distance of the subject from the verb in relative clauses in the historical corpus. Figure 13 below provides the proportion of genitive marked subjects when an intervening maximal projection appears between them and the verb. As seen in the figure, the marking of genitive jumps from below 50% to around 70% in the eighteenth century and it reaches to a maximum of 100% in the twenty-first century, verifying the axioms in (57).

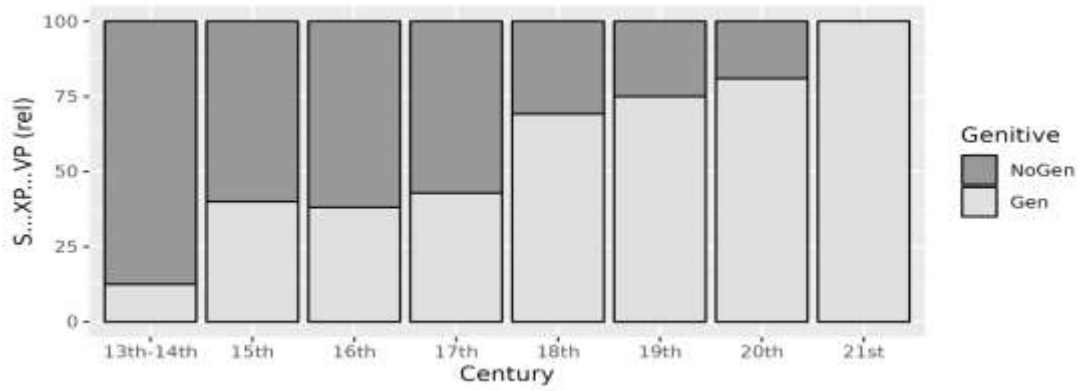


Figure 12. The proportion of genitive marked subjects of headed relatives when an intervening maximal projection appears between it and the verb

The histograms show that specificity (a function of linear distance between subject and the verb) does not play a crucial role until the eighteenth century (see the <50% genitive up to then), for which stage we claimed that animacy was a more decisive factor. After then, the effect of specificity (again measured as a function of linear distance between subject and the verb) increased. Interestingly until the twenty-first century, subjects could still have nominative marking when in this configuration, which, from a Modern Turkish perspective is unexpected because this configuration would translate to a non-specific subject of the relative clause and in this context, the relativizer would appear as *–(y)An*. One example from the corpus is given in (58a), and its Modern Turkish counterpart is as (59):

(58) a. tiyatro bulun-ma-dığ-ı yer-de
theatre exist-NEG-DIK-3SG place-LOC
"in the place theatre didn't exist"
(İZG)(Early- twentieth century)

b. tiyatro bulun-ma-yan yer-de
theatre exist-NEG-(Y)AN place-LOC
"in the place theatre didn't exist" (Modern Turkish)

To this, we would add the observation by Erdal (1981), Nilson (1985), Haig (1997). In particular, Erdal (1997, as cited in Haig, 1997) and Nilson (1985, as cited in Haig, 1997) point out that non-specific internal subjects (semi-subject according to

Haig) could be used in the nominative form with *-DIK* relativizer as in (59). Considering the writing dates of the mentioned articles, the structure in (59) is possible in the nineteen eighties.

(59) dondurma sat-ıl-dığ-ı yer
ice-cream sell-pass-DIK-3SG place
"the place where ice cream is sold"
(Haig, 1997, p. 196 (original reference omitted))

Haig (1997, p. 197) sees this example marginal for Modern Turkish (of 1997, according to the writing date). He mentions that he could not find an example similar to (59) in the corpus with over 1000 relative clauses.

To conclude so far with data from our corpus and from the observation by Haig (1997), although *–(y)An* may not be a recent innovation in the history of Turkish, its complete obligatoriness on non-subject relative clauses with a non-specific subject is a recent phenomenon. According to Haig's (1997) observations, we can say that it was not fully completed at least until the nineteen nineties. It could be interesting to look at the historical evolution of *–(y)An* clauses, which we leave for future research. Figure 14 below, provides the histogram for gentivie/no-genitive marking in headed relatives in which case there is no intervening XP between the subject and the verb. In other words, the subject and the V are adjacent:

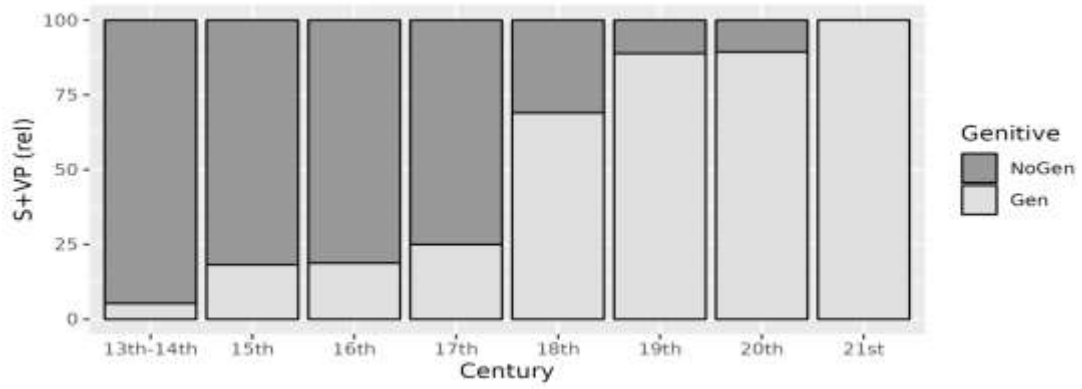


Figure 13. The proportion of genitive marked subjects of headed relatives when there is no intervening maximal projection between it and the verb

The histogram shows that until the twenty-first century, non-specific subjects in the nominative were tolerated in non-subject headed relative clauses with *-DIK*, suggesting that *–(y)An* in this configuration became obligatory only recently. The increase in the genitive around the eighteenth century should be read as a reflex of the general increase of genitive subjects in the headed relative clauses in this century.

6.5.2 Specificity in complement clauses

Again keeping (57) in mind, we have looked at the distance of the subject from the verb in complement clauses. Figure 14 below provides the proportion of subjects in the genitive to subjects in the nominative when an intervening maximal projection appears between them and the verb. As seen in the figure, the marking of genitive jumps from below 40% to around 70% in the nineteenth century and reaches to a maximum of ~100% in the twenty-first century, verifying the axioms in (57). We interpet this as not only the overall increase in the genitive-marked subjects in the corpus but we also claim that specificity (read as a function of linear distance between a subject and the verb) comes into play around the nineteenth century in complement clauses.

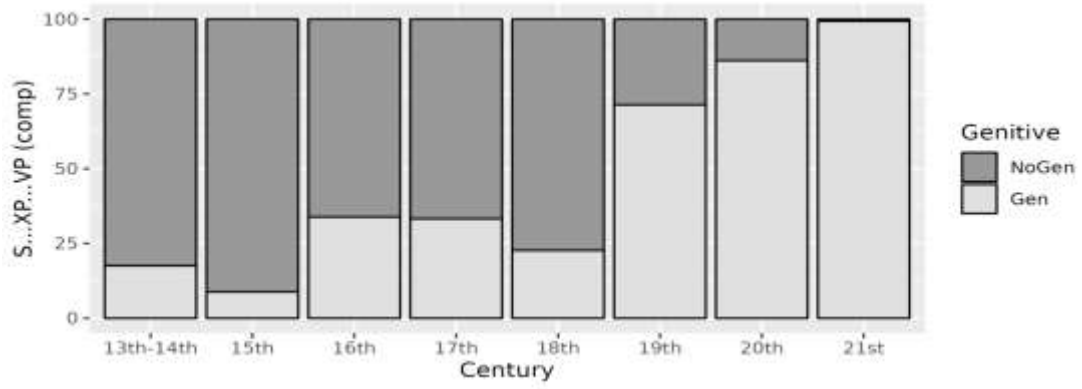


Figure 14. The proportion of genitive marked subjects of complements when an intervening maximal projection appears between it and the verb

Similarly, when the subject and the verb are not separated by an intervening XP, the total ratio of genitive marking still increases in the nineteenth century; however, unlike in the headed relatives, it does not reach 100% as given in figure 16. This is obviously expected because non-specific subjects are tolerated with a zero (nominative) marking as long as they are immediately preverbal. This only proves our conclusion indirectly: subjects are prone to be more specific in general and hence they will appear more in the genitive regardless of their linear position.

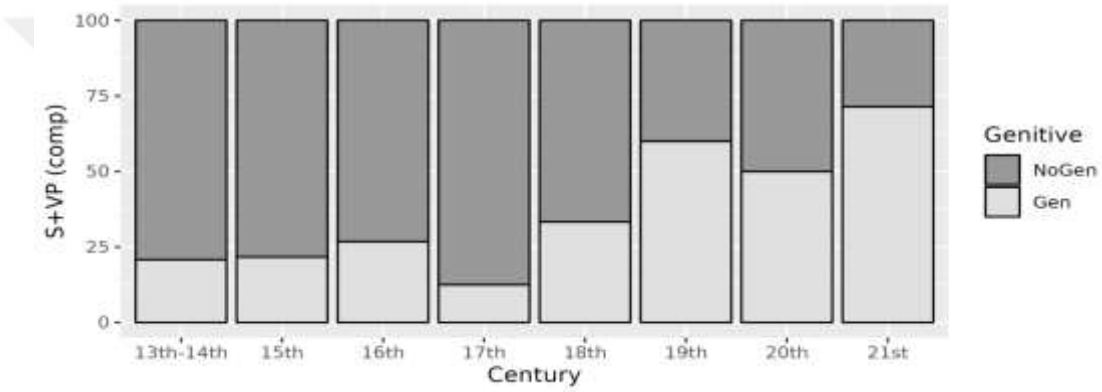


Figure 15. The proportion of genitive marked subjects of headed relatives when there is no intervening maximal projection between it and the verb.

6.6 Changes in numbers: Summary

Animacy played a crucial role in subject-marking for pre-modern Turkish. This dominant effect continued until the eighteenth century for headed relative clauses and until the nineteenth century for complement clauses. The binomial mixed-effect models and the Likelihoood tests showed how subjects with the animate feature were prone to take genitive cases Later, however this effect lost its power in the respective configrations, without however reaching to zero-point even today. When we look at the effect of specificity on genitive marking, we see the emergence of this effect around the eighteenth century for headed relative clauses and around the nineteenth

century for complement clauses and we conclude that definiteness hierarchy took the foreground in these stages of Turkish.

6.7 Changes in minds

In what follows, we speculate on how the numerical changes above can be captured within a framework of historical change based on acquisition.

6.7.1 Language change as an acquisition problem

In line with much previous work in language acquisition, we think change begins in the I-language yet, as a function of language acquisition, is always triggered by E-language. I-language is the language internally in the mind of each individual (Comsky 1986). I-language is cognitive. This is linguistic competence. However, the performed version of this I-language is E-language, which we refer to as English, Turkish etc. E-language is external and is about performance. E-language is performed with the I-language: Two speakers can communicate with E-language produced by their respective I-languages. A direct connection cannot be made between each person's I-languages. For this reason, a child develops their I-language only with data from the E-language.

This difference between I-language and E-language and the lack of a direct connection of speakers and hearers with I-languages can cause a grammar abducted in the acquisition stage (Andersen, 1973; Lighfoot, 1979; Roberts, 2021). According to this view, when a language acquirer creates his own I-language, he/she acquires the language by the E-language performed by the speakers of that language around him/her. In other words, the previous generation creates the E-language externally with their own I-language. The new generation collects these inputs and creates their

I-language. These are called Primary Language Data (PLD). Lightfood (2006) stated that I-languages are systems that emerge in children according to the dictates of the language capacity and to the demands of the external language to which they are exposed. Since the new generation does not have a direct connection with the I-language of the previous generation an ambiguity in PLD may cause the new generation abduct a grammar as a result of them not being able to access sufficient data, especially minimal pairs, to create the grammar.

When the new generation abducts the grammar, they can build their I-languages that differ underlying from the grammar of the previous generation that created the PLD. That is, they may reanalyze the data, and as a result, the new generation acquires a different grammar than the previous generation. This is the Z-model of language change, proposed by Andersen (1973), and adapted and modified by Lightfoot (1979), Roberts (2021), and Kodner (2020) among others. Andersen (1973) sees this as a cycle of error-prone abductive and inductive learning. For this reason, change does not occur suddenly. Language change is the observable state of the differences that occur from generation to generation at the end of a process.

We mentioned above that language change can occur as a result of ambiguity or limited input in E-language. So, what could have caused this abduction or reanalysis in our data so that definiteness hierarchy assumes the prominent role?

6.8 What possibly happened?

We claim that while Animacy Hierarchy played an important role in subject-marking in complement and relative clauses in the pre-eighteenth century Turkish period, the importance of Animacy Hierarchy decreased then and the Definiteness Hierarchy began to play a major role. Animacy Hierarchy does not have an obligatory effect in

the pre-Modern Turkish period. That is, the fact that a subject is higher in Animacy Hierarchy does not automatically cause it to be marked with a genitive case; we present this as a significant predictor, not categorical. In actual fact, examples were cited in the literature that animacy may optionally determine case-markiong. For example, in Sinhalese only animate direct objects can be marked, but this is optional (Aissen, 2003, p. 456). Additionally, in Yiddish, only direct objects that are human are case- marked; however, while only some of them are obligatory case marked, some are optionally case-marked (three masculine nouns which are obligatorily case-marked (*tate* 'father', *zeyde* 'grandfather' and *rebe* 'teacher'), two masculines which are optionally case-marked (*Yid* 'Jew' and *mentsh* 'person') and three feminine nouns, which are optionally case marked (*mame* 'mother', *bobe* 'grandmother', *mume* 'aunt') (Aissen, 2003, p. 456)) . Therefore, it is normal from a cross-linguistic perspective that Animacy hierarchy optionally affects case-marking. Yet, this is enough to introduce the culmination of ambiguity in the PLD.

What we speculate is as follows: animacy and definiteness features of the subjects certainly overlap in numerous instances. From an acquirer's perspective it is possible in many cases to take the genitive marking as a function of definiteness rather than of animacy. The potential to do this 'misinterpretation' grows as the number of genitive marked subjects grow until the eighteenth century. We think up to then while this reanalysis of lexical specifications must have happened in individuals, it was never at a point to alter the grammar: animacy was always the first defining factor. Once, however, the number of such 'genitive = definite' examples reached a tipping point (Yang 2002), the change was inevitable: one generation abducted the grammar, starting with relative clauses. In this slow change from animacy to definiteness, very clear role was played by inanimate NP subjects.

Consider (60) and (61), which are are almost the same in everything else but the subject marking.: while genitive case was not used in (60), it was used in (61). These two sentences are taken from the same source, that is, they appeared as input in the E-language by the same person. When encountered by such an ambiguous data, driven by the Third Factor, the child tries to seek a pattern. Although animate NPs provide them a clear rule (Rule 1: "if NP is animate, mark it in the genitive"), inanimate ones blur the picture. The child seeks a pattern that can potentially be common to both NP types in the given context. The commonality in most cases is a feature from the Definiteness Hierarchy. The child then adds a new rule as "(Rule 2: "if NP is high in the Definiteness scale, mark it in the genitive"). If the examples that prove the newer rule reaches a certain threshold, then the new acquirers simply do away with the former rule.

(60) Evvelgi fasl zümürrüd neden ol-dug-ın ve .....beyan id-er.
previous chapter emerald how be-DIK-3SG.ACC and declare-PRS
"the previous chapter explains what the emerald is made of."
(TMR 43b 11)

(61) zümürrüd-üŋ aslda neden ol-dug-ın Tıgışı Belinas-dan
emerald-GEN *actually* how be-DIK-3SG.ACC Tıgışı Belinas-ABL
nakl id-er
convey-PRS
"He narrates from Tıgışı-Belanas what the emerald actually consists of."
(TMR 44a 5)

It is once again important to note the close interaction between animacy and definiteness here. Comrie (1979) makes the following observation when talking about the defineness-animacy interaction in object marking in some languages:

> Indeed, when we turn to descriptive grammars of some languages, it is very difficult to decide whether we are dealing with definite direct object marking or animate direct object marking often different feature of direct objects, often the same descriptive grammar will present an account which, if taken, is confused or self-condradictory. (Comrie, 1979, p.16)

We claim that the tipping point in the PLD was reached around the late seventeenth century, which, we claim wass a point of no-return. The number of gentivie marked NPs that can be analyzed as specific reached to a number that the acquirers developed a grammar with the rule 2 as the most prominent rule and treated cases where non-specific animates marked in the genitivie as exceptions. The atrophied effect of animacy on genitive marking we see today is in actual fact Rule 1 that was once operant in Turkish grammar saliently. The S-curves only verify the gradualness of this change. Why the tipping point was reached in the eighteenth century, we do not really know. It may be the case that more L2-speakers were coined in the society before the late sixteenth and seventeenth century. Notice that the seventeenth century is known to be a time in the history of Turkish in which many changes appear. In particular, the phonological and morphological changes seen during the transition from Old Anatolian Turkish to Ottoman Turkish were completed only in the seventeenth century (Timurtaş, 2010, p. 331; Ergin, 2009, p. 19). This is mere speculation, but the L2-speaker population is shown to affect change (Kauhannen, 2022).

CHAPTER 7

ADVERBIAL CLAUSES

7.1 Overview

To complete the picture on our diachronic investigation on nominal agreement and genitive-marked subjects, we would like to report our findings on adverbial clauses with *-DIK* as well. Our presentation in this chapter is largely expository. We do not have data from many sub-types of adverbial clauses from the twenty-first century; therefore, we rely on our own judgements for this era.

7.2 Temporal clauses

According to the corpus, there are a total of five temporal clause structures with *-DIK* in the pre-Modern Turkish period. The first of these is a structure in which only an oblique case suffix is used after the embedded clause (63a).

(63) a. bodun elig ekegü bol-tok-ın-ta izgil bodun birle
people state two.peace become-DIK-3SG-LOC İzgil poeople with
süngüş- tü-müz
fight-PAST-1PL
"When the people and the state became two parts, we fought with the Izgil people"

(KT N 3)

The second type of examples additionally involve a postposition (63b).

(63) b. yakut od-a gir-dük-den sonra hüsn-i ve nur-ı
ruby fire-DAT enter-DIK-ABL after beauty-3SG and light-3SG
art-ma-ya
increase-NEG-OPT
"The beauty and light of a ruby do not increase after it enters the fire."

(TMR 35b 9)

The third type involves no case suffixes but certain sortal nouns, such as *vakt, zaman* 'time' are used, (63c).

(63) c. hava ısı ol-dug-ı vakt güneş-e koy-a-lar
weather hot become-DIK-3SG time sun-DAT put-OPT-3PL
"When the weather is hot, put them in the sun."
(MŞ M164b 12)

In the final type, the *-CA* suffix is used after the *-DIK* suffix (63d).

d. Ömer gül yüz-üm gör-dük-çe ög-er bülbül-i güya ben-i
Ömer rose face-1SG see-DIK-CA praise-PRS nightingale singing I-ACC
"When Ömer sees my rose face, the singing nightingale praises me."
(MN 191b 7)

The structures are schematically given in (64).

(64) a. [Subj-NOM V-DIK-AGR/Ø -LOC]

b. [[Subj-NOM V-DIK-AGR/Ø -ACC/LOC/ABL] soŋra> sonra/ berü>beri]

c. [[Subj-NOM V-DIK-AGR] vakt/ vakit/ vaktın/ an/ dem/ zaman]

d. [Subj-NOM V-DIK-AGR/Ø-CA]

Similar structures are also seen in Modern Turkish; however, there is a difference in the use of nominal agreement in embedded clauses. We will specifically evaluate each structure we mentioned above. In the following sections, a comparison will be made between Modern Turkish and pre-Modern Turkish for each structure.

7.2.1 Temporal clauses with case suffix

The structure in (64a) is the one with the most examples among the temporal clause: 1167 of 2052 temporal clauses have this structure. 1101 of them contain the locative case *-DA*. Only 30 of them contain the ablative case *-DAn*, as in (65).

(65) süd-den ayrıl-duk-dan binü-ler-i yazı-lar-da at katır eşşek
milk-ABL wean-DIK-ABL riding-PL-3SG plain-PL-LOC horse hinny donkey
er-di
be-PAST
"After he was weaned, his mounts were horses, donkeys and mules."
(BHMH 67b 26)

This structure can be confused with the similar structure in the reason clause (see section 7.3.2. Examples containing the *-DAn* suffix, which we consider as a

temporal clause, are examples that can only be understood with a temporal meaning. Additionally, 26 of these examples are from the BHMH book. For this reason, we can say that the examples with the *-DAn* ablative case in the embedded clause are exceptions. Because, apart from the BHMH book, only three examples come from different sources.

Modern Turkish also allows the structure with the locative *-DA*, on adverbial clauses as well as on certain temporal nouns (Yenice, 2024). As can be seen in (66a), nominal agreement appears in this configuration, and the internal subject never bears genitive; see the ungrammaticality of (66b). If the subject appears with genitive case, the embedded clause can only be read as a headless relative clause. Finally, if there is no nominal agreement in the embedded clause, the sentence becomes ungrammatical again (66c).

(66) a. Ben gel-diğ-im-de sen git-ti-n.
I come-DIK-1SG-LOC you go-PAST-2SG
"When I came you left"

b. *Ben-im gel-diğ-im-de sen git-ti-n
I-GEN come-DIK-1SG-LOC you go-PAST-2SG
"When I came you left"

c. *Ben gel-dik-de sen git-ti-n
I come-DIK-LOC you go-PAST-2SG
"When I came you left"

In the pre-modern Turkish period, however, there are examples where there is no nominal agreement, as can be seen in (67).

(67) Bu kelâm-ı Sâre Ana işit-dük-de taʿaccüb ile gül-er-ler.
this word-ACC Sare mother hear-DIK-LOC surprise with laugh-PRS-3PL
"When Mother Sare hears this word, they laugh in surprise."
(KE 7bBL 4)

Table 23 provides the distribution of temporal clauses with the locative *-DA* across the centuries. In no example is there genitive case on the subject. On the other hand, only 255 out of a total of 1167 samples have nominal agreement. The rate of of

nominal agreement is consistently until the twenty-first century. The data is interesting in that the nominal agreement is obligatory in temporal clauses of this structure in Modern Turkish. Even in the IZG published in 1909, the rate of existence of nominal agreement is rather low. This shows us that nominal agreement has become obligatory very quickly in the last century.

Table 23. The Distribution of Temporal Clauses across the Centuries

	Overt internal SUBJ		Non-overt internal SUBJ	Total number	[DIK+AGR] (YES)	[DIK+AGR] (NO)
CENTURIES	Gen	No Gen				
Seventh-Eleventh Centuries	0	40	85	125	19	107
Thirteenth-Fourteenth Centuries	0	10	27	37	15	22
Fifteenth Century	0	7	12	19	2	17
Sixteenth Century	0	131	201	332	79	253
Seventeenth Century	0	28	16	44	5	39
Eighteenth Century	0	123	232	355	63	292
Nineteenth Century	0	57	130	187	42	149
Twentieth Century	0	27	15	42	4	38
Twenty-first Century	0	26	26	26	26	0
Total	0	423	718	1167	255	917

7.2.2 Temporal clauses with a case suffix + postposition

Recall that in the second type of temporal adverbial clauses there is a proposition after the case suffix (64b). In total, 276 out of 388 examples have the postposition *soŋra/sonra* 'after'. In examples with *soŋra*, ablative case *-DAn* appears on the embedded clause (68a). Apart from this postposition, *kėsre* 'after' and, *berü* 'since' are also attested. The postposition *kėsre* is found only in examples from the Old Turkic period, in which period the *-DA* suffix expresses syncretically both locative-

ablative (Ercilasun, 2004) (68b). Finally, the postposition *berü* also governs ablative case -*DAn,* on the embedded clause (68c).

(68) a. tennur kız-duk-dan sonra tezeg-i tennur-a ol kadar
tandoor get.hot-DIK-ABL after dung-ACC tandoor-DAT that much
koy-a-lar
put-OPT-3PL
"After the tandoor gets hot, put the dung into the tandoor."
(TMR 84b 8)

b. ol kan yok bol-tok-da kėsre el yit-miş
that khan disappear-DIK-LOC/ABL after state lost-PAST
"After that khan died, the state was lost"
(O B1)

c. geyig-in-den al-duk-lar-ın-dan berü
deer-3SG-ABL take-DIK-PL-3SG-ABL since
"Since he took it from his deer."
(TMR 118b 5)

While *kėsre* is not recognized as a postposition in Modern Turkish, postpositions *soŋra>sonra* and *berü>beri* are still in use; however, the structures in which these two postpositions occur are not the same. In modern Turkish, in temporal clauses with the postposition *soŋra> sonra*, there is no nominal agreement in the embedded clause and the internal subject is in the nominative form (69a). In the presence of nominal agreement, the sentence becomes ungrammatical (69c). On the other hand, in structures with the *berü>beri* postposition, nominal agreement is the norm (69b), and the absence of it results in ungrammaticality (69d).

(69) a. Sen söyle-dik-ten sonra yap-tı-m.
you say-DIK-ABL after do-PAST-1SG
"I did it after you said"

b. Köpeğ-i gör-düğ-üm-den beri koş-uyor-um.
dog-ACC see-DIK-1SG-ABL since run-PROG-1SG
"I've been running since I saw the dog"

c. *Sen söyle-diğ-in-den sonra yap-tı-m
you say-DIK-ABL after do-PAST-1SG
"I did it after you said"

d. *Köpeğ-i gör-dük-ten beri koş-uyor-um.
dog-ACC see-DIK-1SG-ABL since run-PROG-1SG
"I've been running since I saw the dog"

In the pre-modern Turkish period, however, nominal agreement on the embedded clause appears with postposition *soŋra* 'after', as given in (70)

(70) ta'am ma'de-de hazm ol-dug-ın-dan sonra ed-e-ler
food stomach-LOC digest-DIK-3SG-ABL after make-OPT-3PL
"Let them do this after the food is digested in the stomach
(MŞ M7a 2)

The distribution of temporal clauses with postposition *sonra* 'after' can be seen in table 24

Table 24. The Distribution of Temporal Clauses with Postposition *sonra* and Case Suffix across the Centuries

	Overt internal SUBJ		Non-overt internal SUBJ	Total number	[DIK+AGR] (YES)	[DIK+AGR] (NO)
CENTURIES	Gen	No Gen				
Seventh-Eleventh Centuries	0	2	4	6	3	3
Thirteenth-Fourteenth Centuries	0	6	9	15	12	3
Fifteenth Century	0	1	3	4	3	1
Sixteenth Century	0	53	93	146	61	85
Seventeenth Century	0	6	10	16	4	12
Eighteenth Century	0	18	38	56	5	51
Nineteenth Century	0	35	39	74	1	74
Twentieth Century	0	23	31	54	0	54
Twenty-first Century	0	17	0	17	0	17
Total	0	161	227	388	89	300

As seen in Table 24, the subjects of the embedded clause in this structure are in nominative form in every period. 363 of the 388 examples date from the sixteenth to twentieth centuries. We can say that the use of this structure is not frequent in earlier periods. The interesting part in table 24 is that although there is no nominal agreement in structures with postposition *soŋra>sonra* in Modern Turkish, there are

examples of nominal agreement in these structures in the pre-Modern Turkish period. Moreover, although there were fewer examples of this structure from the thirteenth century to the fifteenth century, nominal agreement was more common in this period. For this reason, while the tendency to use nominal agreement was higher in the early periods in temporal clause with *soŋra>sonra* postposition, this tendency started to decrease later. Especially in nineteenth century, we see that the change has been completed and nominal agreement has completely disappeared.

7.2.3 Temporal clauses with noun head

In this structure, certain nouns are used after the embedded clause without case suffix. The embedded clause looks like it is embedded under a DP, as can be seen in example (71) with the noun *vaktın* 'time'.

(71) yakut bul-un-dug-ı vaktın yay-ı artuk çek-er-ler
ruby find-PASS-DIK-3SG time bow-ACC more pull-PROG-3PL
"When rubies are found, they pull the bow harder."

(TMR 37b 8)

There are in total 83 examples of this structure among temporal clauses. The nouns used in these structures are *vaktın, dem, vakt, an, zaman, lahza,* all of which can be translated as 'time/moment'. Certain nouns were preferred in certain sources. These structures could in principle be considered as headed relative clauses. We excluded from the data the examples that we could not decide were not temporal clauses.

There are similar structures in Modern Turkish. As can be seen in example (72a), the internal subject is in the nominative form and nominal agreement exists in the embedded clause. As seen in example (72b), if there is no nominal agreement in the embedded clause, the sentence becomes ungrammatical. If internal subject is marked in the genitive case, as in example (72c), there is ambiguity in the sentence.

If we accept the example in example (72c) as headed relative clause, the sentence becomes grammatical. However, if we accept this structure as adjunct clause, the sentence will be ungrammatical.

(72) a. O gel-diğ-i zaman yemek yi-yeceğ-iz
he/she come-DIK-3SG time food eat-FUT-1PL
"We will eat when he comes."

b. *O gel-dik zaman yemek yi-yeceğ-iz
he/she come-DIK-3SG time food eat-FUT-1PL
"We will eat when he comes."

c.*O-nun gel-diğ-i zaman yemek yi-yeceğ-iz
he/she-GEN come-DIK-3SG time food eat-FUT-1PL
"We will eat when he comes."

As seen in table 25, there is no temporal clause in this structure in the Old Turkic period. They appear in the later period. In post-Old Turkic period, the internal subject is in the nominative form and nominal agreement exists for each century. Therefore, there have been no diachronic changes in temporal clauses of this structure throughout the history of Turkish.

Table 25. The Distribution of Temporal Clauses with Noun Head across the Centuries

	Overt internal SUBJ		Non-overt internal SUBJ	Total number	[DIK+AGR] (YES)	[DIK+AGR] (NO)
CENTURIES	Gen	No Gen				
Seventh-Eleventh Centuries	0	0	0	0	0	0
Thirteenth-Fourteenth Centuries	0	8	21	29	29	0
Fifteenth Century	0	3	4	7	7	0
Sixteenth Century	0	19	24	43	43	0
Seventeenth Century	0	0	4	4	4	0
Eighteenth Century	0	0	0	0	0	0
Nineteenth Century	0	10	4	14	14	0
Twentieth Century	0	17	4	25	25	0
Twenty-first Century	0	11	0	11	11	0
Total	0	68	61	133	133	0

7.2.4 Temporal clauses with *-CA*

In this structure, *-CA* gives the sentence a meaning like 'as long as/when'. A total of 350 examples are in this structure. In all examples, the subject of the embedded clause is in the nominative form. However, as seen in examples (73a) and (73b), in some examples of this structure there is a nominal agreement and there are not in others.

(73) a. su-yı eksil-düg-in-ce karşu su koy-a-lar
water-3SG decrease-DIK-3SG-CA instead water put-OPT-3PL
"When water runs out, let them replace it with water."
(MŞ M107b 9)

b. Ömer gül yüz-üm gör-dük-çe ög-er bülbül-i güya ben-i
Ömer rose face-1SG see-DIK-CA praise-PRS nightingale singing I-ACC
"When Ömer sees my rose face, the singing nightingale praises me."
(MN 191b 7)

A similar structure exists in modern Turkish. As seen in example (74a), in these structures, the internal subject is in the nominative form and there is no nominal agreement. As seen in example (74b), if there is nominal agreement in the embedded clause, these sentences become ungrammatical. Notice that there are also examples in which the internal subject has genitive case and nominal agreement exists, as illustrated in (74c). In such contexts, the embedded clause can also be understood as a headless relative clause. We removed the examples which is understood as a headless relative clause from this category.

(74) a. O ban-a bak-tık-ça sinirlen-di-m
he/she I-DAT look-DIK-CA get.angry-PAST-1SG
"I got angry when he/she looked at me.

b. *O ban-a bak-tığ-ın-ca sinirlen-di-m
he/she I-DAT look-DIK-3SG-CA get.angry-PAST-1SG
"I got angry when he/she looked at me.

c. Ben-im dile-diğ-im-ce yap-tı
I-GEN wish-DIK-1SG-CA do-3SG
"He did as I wished."

As seen in Table 26, the subject of the embedded clause is in the nominative form for each period. This structure is not seen in the Old Turkic period. The interesting part in the data is that in this adjunct clause structure that we see in Old Anatolian Turkish, nominal agreement is proportionally more common in the thirteenth and fourteenth centuries than in other periods. In total, 33 out of 46 examples contain nominal agreement. The rate is 71.73%. In the fifteenth century, the nominal agreement rate in these structures becomes 36.50%. In the sixteenth and seventeenth centuries, this rate is 42.85%. Considering that nominal agreement is not used in these structures in Modern Turkish, it is interesting that nominal agreement is found at a higher level in the earliest periods. Over time, the nominal agreement rate in such structures began to decrease. Especially since the eighteenth century, nominal agreement has almost completely disappeared, and the change was completed in the nineteenth century.

Table 26. The Distribution of Temporal Clauses with -*CA* across the Centuries

	Overt internal SUBJ		Non-overt internal SUBJ	Total number	[DIK+AGR] (YES)	[DIK+AGR] (NO)
CENTURIES	Gen	No Gen				
Seventh-Eleventh Centuries	0	0	0	0	0	0
Thirteenth-Fourteenth Centuries	0	17	29	46	33	13
Fifteenth Century	0	57	69	126	46	80
Sixteenth Century	0	27	46		32	41
Seventeenth Century	0	5	10	15	4	11
Eighteenth Century	0	12	18	30	1	29
Nineteenth Century	0	16	15	31	0	31
Twentieth Century	0	14	13	27	0	27
Twenty-first Century	0	2	0	2	0	2
Total	0	148	200	348	116	234

7.3 Reason clauses

According to our corpus, four different structures with *-DIK* suffix are seen in the corpus which carry the semantics of a reason clause. In one, illustrated in (75a), a postposition *üçün>içün/çUn>için* 'because' is used after the embedded clause, which carries no case marking. In another, illustrated by (75b) and (75c), postpositions appear as *ötürü* and *bina'en* 'because' are used after the embedded clause with case suffixes. In a structure such as (75d), there is no postposition after the embedded clause but there is simply a case marker.

(75) a. Boz-up gel-düg-iŋ içün sav-ıŋ-ı sen saŋ-a
break-FN come-DIK-POSS.2SG because fast-POSS.2SG-ACC you you-DAT
on gün dahı emr eyle-dü-m ben
ten day more order give-PAST-1SG I
"Since you broke your fast and came, I ordered ten more days"
(KE 182bBL 3)

b. tuzlu balgam bun-lar-da çok mütevellid ol-duğ-ın-dan ötürü
salty phlegm this-PL-LOC frequently emerge- DIK-3SG-ABL because
"Because salty phlegm occurs frequently in these cases."
(MŞ M130a 15)

c. fahr -ı alem salla'llahu ve sellem-e muhabbet it-düg-in-e
boast of world sallallahu alayhi wa sallam-DAT conversation do-DIK-3SG
bina'en
because
"Because of his conversation with the boast of the world"
(HR T20a 12)

d. Yusuf ana-sı öl-düg-in-den Ya'kup bir karavaş satun al-dı
Joseph mother-3SG die-DIK-3SG-ABL Jacob a slave buy-PAST.3SG
"Jacob bought a female slave because Joseph's mother died."
(BHMH 98b 12)

We can then divide the reason clauses in the pre-Modern Turkish period into three main groups. The first is those that contain only postposition, the second is those that contain only case, and the third is those that contain both case and postposition. Their structures can be seen in (76).

(76) a. [[Subj-NOM V-DIK-AGR/Ø] üçün>içün/-çUn>için]

b. [Subj-NOM V-DIK-AGR/Ø -ABL]

c. [[Subj-NOM V-DIK-AGR/Ø -DAT/LOC/ABL] ötrü/ötürü, bina'en]

The structures in (76) are also found in reason clauses in Modern Turkish

(77) a. Ben iste-diğ-im için Ali gel-di
I want-DIK-1SG because Ali come-PAST
"Ali came because I wanted" (=76a)

b. Ayağ-ı kır-ıl-dığ-ın-dan gel-me-di
foot-3SG break-DIK-3SG-ABL come-NEG-PAST
"He didn't come because his leg was broken" (=76b)

c. Sen ev-den git-tiğ-in-den dolayı çok kız-dı-m
You home-ABL go-DIK-3SG-ABL because so get angry-PAST-1SG
"I was so angry because you left home." (76c)

We see that internal subjects of the reason clause in Modern Turkish are in the nominative form (78a). As seen in example (78b), when the subject appears with the genitive case, the embedded clause is interpreted as a headless relative clause. This is also valid for all other structures in reason clauses. For this reason, we can say that in Modern Turkish, the subjects of reason clauses are used obligatorily in the nominative form (Kornfilt, 2003, 2008; Aygen, 2007; Laszakovits, 2013). As seen in example (78c), if there is no nominal agreement, the sentence becomes ungrammatical again. This is also valid for all structures of reason clauses. For this reason, we can say that in Modern Turkish reason clauses, nominal agreement is obligatory.

(78) a. ben gel-diğ-im için o git-ti
I-GEN come-DIK-1SG because he/she go-PAST
"He/she left because I came"

b. *ben-im gel-diğ-im için o git-ti
I-GEN come-DIK-1SG because he/she go-PAST
"He/she left because I came"

c. *ben gel-dik için o git-ti.
I come-DIK because he/she go-PAST
"He/she left because I came"

However, according to the data, nominal agreement is not seen in every type of reason clause for the pre-Modern Turkish period, as can be seen in example (79). Table 27 shows the distribution of reason clauses across the centuries

(79) Tengri yarlıka-dok üçün
God bless-DIK because
"by the will of the God"

(BQ E13)

Table 27. The Distribution of Reason Clauses across the Centuries

	Overt internal SUBJ		Non-overt internal SUBJ	Total number	[DIK+AGR] (YES)	[DIK+AGR] (NO)
CENTURIES	Gen	No Gen				
Seventh-Eleventh Centuries	0	30	28	58	23	35
Thirteenth-Fourteenth Centuries	0	48	55	103	103	0
Fifteenth Century	0	16	18	34	33	1
Sixteenth Century	0	28	37	65	65	0
Seventeenth Century	0	4	4	8	8	0
Eighteenth Century	0	18	27	45	45	0
Nineteenth Century	0	188	121	309	309	0
Twentieth Century	0	75	22	99	99	0
Twenty-first Century	0	58	0	58	58	0
Total	0	465	312	778	743	36

In the 465 examples in our corpus, the subject of the reason clause is overt. In all of these, the subject is in the nominative form. There is no example of these with a genitive case in the subject. Therefore, we can conclude that it obligatory for the subjects of the embedded clauses to be in nominative form in a reason clause, and this rule has not changed over time. In 743 examples, there is nominal agreement in the embedded clause. As seen in table 27, there are more examples where nominal agreement did not exist in the Old Turkic. Of the 58 examples in this period, 23 had

nominal agreement (=39.66%). In the following period, this rate increased significantly, almost to a total of 100%. We can therefore conclude that nominal agreement in this type of adverbial clauses became obligatory immediately in the in post-Old Turkic period.

We must note that in reason clauses, the change in the nominal agreement refers only to the structure with the postposition *üçün>içün/-çUn> için*. The examples nominal agreement doesn't exist are found only in the Old Turkic period. we identified and annotated all reason clauses in our data. However, since we could not detect any change, we did not split them according to different reason clause structures after the sixteenth century. In the following sections, only raw numbers up to the sixteenth century are given.

7.3.1 Reason Clause with postposition *üçün>içün/-çUn> için*

In reason clauses formed with the postposition *üçün* in the Old Turkic period, if there is a nominal agreement in the embedded clause and if the subject of the embedded clause is the third person, accusative case is used after the agreement marker. There are three types of accusative cases in Old Turkic: *-ug, -In, -nI*. Of these, *-In* is used after a agreement marker (Ercilasun, 2004). As seen in example (80), there is an accusative case (third person suffix "-I" in Old Turkic). In the examples where there is no agreement marker, there are no examples where the accusative case exists alone.

(80) Tengri yarlıka-dok-ın üçün
God bless-DIK-3SG+ACC because
"by the will of the God"

(BQ N7)

As seen in Table 28, there is no nominal agreement in only 27 out of 207 examples in the reason clauses where the *üçün>içün/-çUn* postposition is used until

sixteenth century. 26 of these are seen in sources from the Old Turkic period. The one example belongs to GN. It is found exceptionally in the early periods of Old Anatolian Turkish. For this reason, we can say that the use of nominal agreement in reason clauses with the *üçün>içün/-çUn* postposition was obligatory in the Old Anatolian Turkish period, with just one exception.

Table 28. The Distribution of Reason Clauses with Postposition *üçün>içün/-çUn>* için across the Centuries

	Overt internal SUBJ		Non-overt internal SUBJ	Total number	[DIK+POSS] (YES)	[DIK+POSS] (NO)
CENTURIES	Gen	No Gen				
Seventh-Eleventh Centuries	0	25	21	46	20	26
Thirteenth-Fourteenth Centuries	0	48	56	104	103	1
Fifteenth Century	0	10	18	28	28	0
Sixteenth Century	0	15	14	29	29	0
Total	0	98	109	207	180	27

7.3.2 Reason clauses with case suffix

There are 42 examples in this pattern until sixteenth century. As seen in Table 29, there is no reason clause in the Old Turkic period in which postposition does not exist. We have mentioned in section 7.2.1 that there is no *-DAn* suffix in Old Turkic and the *-DA* suffix is used as locative-ablative case (Ercilasun, 2004). As we mentioned in the section 7.2.1, the *-DA* suffix is used with temporal clauses without postposition in the Old Turkic period. This may be the reason why there are no examples without postposition and with case suffix in Old Turkic. In the post-Old Turkic period, there are no examples where nominal agreement did not exist, as seen in table 29. Therefore, the nominal agreement in the reason clause with case suffix is

obligatory. This suggests that there have been no diachronic changes in reason clauses with only case suffix throughout the history of Turkish.

Table 29. The Distribution of Reason Clauses with Case Suffix across the Centuries

	Overt internal SUBJ		Non-overt internal SUBJ	Total number	[DIK+AGR] (YES)	[DIK+AGR] (NO)
CENTURIES	Gen	No Gen				
Seventh-Eleventh Centuries	0	0	0	0	0	0
Thirteenth-Fourteenth Centuries	0	4	4	8	8	0
Fifteenth Century	0	4	5	9	9	0
Sixteenth Century	0	11	14	25	25	0
Total	0	19	23	42	42	0

7.4 Chapter summary

As can be seen in sections 7.2 and 7.3, the subjects of adverbial clauses have been in nominative form throughout the history of Turkish. There has been no change in this regard. Moreover, in reason clauses, the difference only appears in examples with the preposition *üçün>içün/-çUn>* için 'for' proposition. However, this change was completed in the early period of Old Anatolian Turkish. There is no change in other structures of reason clauses compared to modern Turkish. When it comes to temporal clauses, the examples are different from modern Turkish. Temporal clauses in various structures have undergone different change processes. For example, nominal agreement in temporal clauses with the *-DA* locative case marker was much rarer until the beginning of the twentieth century. It seems than this change was completed only in the last century, perhaps even more recently, and nominal agreement became necessary in Modern Turkish. There is no nominal agreement in modern Turkish in adverbial clauses estalished with *-DAn son/ŋra* 'after' and *-CA* 'as long as/when'.

However, according to our data, there is nominal agreement in these structures in the early periods of Old Anatolian Turkish. Over time, the number of examples with nominal agreement decreased and eventually disappeared in the eighteenth century. The situation is quite interesting no note because, as we discussed in chapters 4, 5 and 6, the presence of genitive cases in the subjects of headed relative clauses and complement clauses increased rapidly around the same period.

CHAPTER 8

CONCLUSION

8.1 Findings and summary

In this thesis, subject agreement in subordinate *-DIK* clauses are discussed diachronically throughout the history of Turkish. It has been revealed how often and since when nominal agreement has been used in all subordinate *-DIK* clause types. For each subordinate *-DIK* clause type, the increase or decrease in the existence of the nominal agreement has been determined. Moreover, the diachronic change in genitive marking of the internal subjects of complement and relative clauses in Modern Turkish was determined. This diachronic change was discussed theoretically, and an account of this change was put forward based on the "language-change-as-an-acquisition-problem" hypothesis.

There are many studies in which subordinate clauses in Modern Turkish are analyzed theoretically (e.g. Kornfilt, 2003, 2014; Haig, 1997; Hankamer & Knecht, 1976; Barker, Hankamer & Moore, 1990). Some studies offer comparative accounts across living Turkic languages (e.g. Aygen, 2002a; Kornfilt, 2001, 2008; Satik, 2020) There are also studies in the literature that examine the different stages of Turkish from a philological perspective (e.g. Erdal, 2004; Timurtaş, 2010; Tekin 2003). Yet, as far as we can tell, there are no diachronic studies that combine a theoretical analysis with a large historical corpus. In order to fill the gap, I manually created a large historical corpus as much as the thesis time allowed, with the sources I collected since 2020. I annotated the *-DIK* clauses in this historical corpus according to a number of criteria deemed relevant.

In the light of the data I obtained this way, I first determined that the *-DIK* suffix has been used in all subordinate clause types since the earliest periods of Turkic. Moreover, I showed that nominal agreement started to be seen in all subordinate *-DIK* clause types in every period of Turkish, including Old Turkic. However, in each subordinate *-DIK* clause types and subtypes, nominal agreement occurrence rates and the increase and decrease in these rates vary. For example, except for two identical examples, nominal agreement is constantly seen on complement clauses since the eighth century. On the other hand, examples of relative clauses where nominal agreement does not exist can be found until the sixteenth century but their number decreased b since the thirteenth century till then, when they completely disappeared. In other words, while agreement was the most preferred variant until the sixteenth century, it became obligatory in this century.

In adverbial clauses, the probability of appearance of the nominal agreement varies according to each adverbial clause type. For example, in temporal adverbial clauses established with the *-DA* case marker, the ratio of nominal agreement to its absence until the early twentieth century was only one-thirds (Section 7.2.1). For now, we think that this change is a result of the language reform or standardization in the Republic Era. However, for a more accurate evaluation, the last century of Modern Turkish should be divided into certain stages and examined in more detail. On the other hand, although there is no nominal agreement in adverbial *-DIK* clauses with *-DAn sonra* 'after' or *-CA* 'as long as/when' in Modern Turkish, nominal agreement was seen in these structures around the thirteenth century and its existence continued down to the the eighteenth century (albeit with a decrease) and disappeared completely in this century. In other subordinate *-DIK* clauses, nominal agreement emerged but, in these ones, the nominal agreement disappeared over time

instead of increasing. This shows us that if we accept the nominal agreement as an innovation, diachronic change does not always have to move in the same direction. There are examples where innovation appears and then disappears over time (section 7.2).

As a second finding, the genitive vs. nominative alternation in the internal subjects of complement clauses and relative clauses has been seen since the very early periods of Turkic. Until the eighteenth century for relative clauses and the nineteeth century for complement clauses, the ratio of internal subjects marked with genitive case to those in nominative form was significantly lower, roughly about one-thirds. Until the mentioned centuries, only a ten percent increase in the proportion of subjects marked with genitive was observed but then a sudden increase appears, and the rate of genitive marked subjects reaches to about 80% first in relatives in the eighteenth century and then in the complement clauses in the nineteenth century, and this rate has continued to increase until today. This shows us a typical logistic curve. In addition, the proportion of internal subjects marked in the genitive increased in similar rates in these two clause types. This shows us the Constant Rate Effect (Kroch, 1989). In addition, sudden increases are first seen in relative clauses.

While for most of the pre-Modern Turkish period, the rate of internal subjects marked with genitive was less than the rate of internal subjects in nominative form, this distribution was not stochastic. Our statistical analysis showed that the rate of animate subjects bearing genitive was higher than the inanimate subjects. This shows us that the animacy hierarchy had a crucial role in genitive marking in the pre-Modern Turkish period: animate internal subjects had a greater chance of being marked in the genitive than inanimates. This chance was around 75% in the early stages. However, this chance started to decrease proportionally after

the eighteenth century for relative clauses and after the nineteenth century for complement clauses. As we get closer to today, it has approached the stochastic 50% level, as expected, but did not completely disappear.

In modern Turkish, however, specificity and genitive-marking interact closely. Internal non-specific subjects must be in preverbal position, and they are in the nominative form. Hence we think there was a change from the Animacy Hierarchy to Definiteness Hierarchy. We looked separately for examples in which the internal subjects were in the immediately preverbal position and those in which the internal subject was far from the verb. The genitive marking rate, which was below 50% in relative clauses where the internal subject is not in a preverbal position, increased to over 70% after the eighteenth century, and reached 100% in the twenty-first century. Similar results were obtained in complement clauses, albeit the leap happened in the nineteenth century. These show us that the internal subjects of relative clauses or complement clauses, which were not in preverbal position until the twenty-first century, could be in nominative form. Additionally, the use of the *-(y)An* relativizer in non-subject relative clauses with a non-specific internal subject was not obligatory until the twentieth century. For complement clauses, a similar increase occurs after the nineteeth century. Subjects far from the verb have an increased chance of being marked with genitive. However, until the twentieth century, it did not reach 100%. For complement clauses with internal subject in the nominative form, this means incorporation is not obligatory until twentieth century.

Animacy started to lose its effect since the eighteenth century for the relative clauses and the nineteenth century for the complement clause. After the mentioned centuries, specificity began to have a critical role. In other words, the predictive force of animacy decreased and the predictive force of specificity increased steadily. It

should however be mentioned that the effect of animacy in subject-marking continues in Modern Turkish in a limited sample (Section 6.5). The role of animacy in Modern Turkish can be taken as an atrophied property, inherited from its earlier stages.

In the light of an acquisition-based model to language change (Andersen, 1973, Lightfoot, 1976 and much subsequent work), we claimed that the defining criteria were reinterpreted as specific hierarchy rather than animacy hierarchy. According to our hypothesis, the parametric ambiguity in the input data lead at least a set of acquirers to interpret the crucial factor as specificity, instead of animacy, for marking subjects in the genitive or in the nominative. While this ambiguity may be caused by the interaction between animacy hierarchy and definiteness hierarchy, it may also be caused by the fact that the animacy factor optionally affects genitive-marking in the pre-modern period. Acquirers who could not access sufficient data abducted grammar due to this ambiguity.

8.2 Points for further research

In this thesis, we had to leave some of the theoretical analysis to future studies. First. the change in the nominal agreement in adverbial clauses is only given as raw numbers. What underlying factors caused the changes have not been investigated. Second, as mentioned in section 4.5, there were examples with accusative case on the subject of the complement clause as in (22). While their numbers were very few and, they have disappeared quickly, the reasons behind the emergence of this variant should be properly investigated.

(=22) Hiç kimesne [an-ı ne ol-dug-ın] bil-me-z-di
anybody it-ACC what be-DIK-3SG+ACC know-NEG-PRS-PAST
"No one knew what it was"

Third, in section 6.5.1, we concluded that the use of *-(y)An* as a relativizer in non-subject relatives with an incorporated non-specific subject has recently become obligatory. Crucially, however, I did not look at the history of *-(y)An* relativizer. To provide a better picture of the changes, the diachrony of the *-(y)An* relativizer should be properly investigated.

Fourth, as we mentioned in section 5.3.1.2, the *-DIK* and *-(y)An* asymmetry found in Modern Turkish is not seen in other Turkic languages. Especially in languages such as Turkmen and Azerbaijani in the Oghuz language group, which includes Modern Turkish, the relative clause structures established with the *-DIK* and *–(y)An* suffix can be compared synchronically and diachronically in a future study. The difference that arises from Old Oghuz to Modern Oghuz languages can be discussed diachronically.

Fifth, the effect of specificity is very well known in Differential Object Marking in Turkish. It would be interesting to look at the history of the emergence of differential object marking in Turkish, to investigate whether animacy was also a bigger defining factor in the history for it, similar to the case of gentive vs. nominative alternation in embedded subjects of complement and non-subject relatives.

Finally, in Modern Turkish, genitive case has always been associated with specificity in the literature (Erkman-Akerson & Özil, 1992; Barker, Hankamer & Moore, 1990 a.o.). For instance, specificity has a determining role in genitive marking in possessive constructions as seen in (81) (Öztürk & Taylan, 2016).

(81) a. Kadın-ın doktor-u
women-GEN doctor-3SG
"The doctor of the women"

b. Kadın doktor-u
women doctor-3SG
"Women's doctor (gynecologist)"

(Öztürk & Taylan 2016)

Similarly, genitive vs. nominative alternation appears in subordinate clauses established with other subordinators, such as *-mA, -ACAK*, as well (82)-(83).

(82) a. köy-e doktor gel-eceğ-in-i söyle-di-ler.
village-DAT doctor come-ECEK-3SG-ACC say-PAST-3PL
"They said that a doctor will come to the village."

b. köy-e doktor-un geleceğ-in-i söyle-di-ler.
village-DAT doctor-GEN come-ECEK-3SG-ACC say-PAST-3PL
"They said that the doctor will come to the village."

(83) a. köy-e doktor gel-me-sin-i isti-yor-uz.
village-DAT doctor come-MA-3SG-ACC want-PRS-3PL
"We want a doctor to come to the village.

b. köy-e doktor-un gel-me-sin-i isti-yor-uz.
village-DAT doctor-GEN come-ECEK-3SG-ACC want-PRS-3PL
"We want a doctor to come to the village."

A larger corpus should be established to include all sorts of embedded propositional and situational complement clauses and non-subject relatives, and possessive-genitive constructions, to assess whether specificity was always a crucial factor, or whether animacy also played a crucial role in overt marking in these configurations.

APPENDIX A

DATA SOURCES ACROSS THE CENTURIES

Text	Year	Century	Edition
BK	735	Eighth.	Ölmez 2015
KT	732	Eighth.	Ölmez 2015
O	-	Eighth.	Ölmez 2015
KÇ	-	Eighth.	Ölmez 2015
Tes	745	Eighth.	Ölmez 2015
Tar	753	Eighth.	Ölmez 2015
Bögük	762	Eighth.	Özbay 2019
T1	-	Eighth.	Ölmez 2015
T2	-	Eighth.	Ölmez 2015
Mai	-	Eighth.	Tekin 1976
AY2	-	Tenth	Kaya 2021
AY3	-	Tenth	Kaya 2021
AY5	-	Tenth	Kaya 2021
AY6	-	Tenth	Kaya 2021
AY10	-	Tenth	Kaya 2021
IB	-	Tenth	Uluç et al. 2013
Xuan2	-	Tenth	Ölmez 2002
Xuan3	-	Tenth	Ölmez and Röhrborn 2001
Xuan6	-	Tenth	Ölmez 1994
Xuan9	-	Tenth	Barat 2000
Xuan10	-	Tenth	Barat 2000
Hua	-	Eleventh	Özbay 2019
Dasa	-	Eleventh	Elmalı 2016
SVTM	1291 and 1301	Thirteenth- Fourteenth	Mansuroğlu 1958
BHMH	1303	Thirteenth- Fourteenth	Canpolat 2018
EATM	1343	Thirteenth- Fourteenth	Tekin 1974
VG	1342/43	Thirteenth- Fourteenth	Nikoosokhan Sedghi 2015
KBD	1393/94	Thirteenth- Fourteenth	Ergin 1980
TMB	-	Fourteenth	Doğuer 2013
SNB	-	Fourteenth	Dilçin 1991
GT	1430	Fifteenth	Anaçoğlu 2019
MN	1437	Fifteenth	Canpolat 1995
MRB	1445	Fifteenth	Özavşar 2009
GN	1446- 1477	Fifteenth	Yavuz 2000
KG	1565	Sixteenth	Erdem 1992

TMR	1523 and 1571	Sixteenth	Argunşah 1999
SNK	1544/45	Sixteenth	Babür 2013
YZ	1545	Sixteenth	Taş 2017
EM	1547	Sixteenth	Aydeniz 2008
MŞ	1547	Sixteenth	Önler 1981
RB	1557	Sixteenth	Önler 2000
BH	1565/66	Sixteenth	Çiftçi 1997
HR	1577	Sixteenth	Özezen 1994
KE	1585	Sixteenth	Karataş 2013
KBP	-	Seventeenth	Ergüzel 2009
TT	1648	Seventeenth	Başar 2021
TOA	1724	Eghteenth	Koç, 2021
Şİ	1769	Eghteenth	Karagöz 2019
OKD	1808	Nineteenth	Sağlam 2021
TAOCP	1881/1882	Nineteenth	Çoban 2009
İZG	1909	Twentieth	Yılmaz 2020
UDG	1980- 2020	Twenty-First	(references too many; see section 2.2.2.9.1)

APPENDIX B

THE DISTRIBUTION OF TOTAL NUMBER OF EXAMPLES

ACCORDING TO THE TEXTS

Text	Complements	Reason	Temporal	Headed Relatives	Headless Relatives
BK	1	18	10	5	0
KT	2	9	8	3	0
O	0	2	4	1	0
KÇ	3	0	9	0	1
Tes	0	0	1	0	0
Tar	0	2	1	0	0
Bögük	0	2	0	0	0
T1	0	2	0	1	1
T2	0	5	0	0	0
Mai	0	5	10	4	0
AY2	0	0	2	3	0
AY3	0	1	10	3	1
AY5	0	2	5	4	0
AY6	0	0	7	2	0
AY10	0	1	10	0	0
IB	0	0	0	0	2
Xuan2	0	0	1	0	0
Xuan3	1	1	5	0	1
Xuan6	0	0	7	2	0
Xuan9	1	1	20	1	0
Xuan10	0	2	5	1	0
Hua	0	5	3	1	0
Dasa	0	0	15	0	0
SVTM	0	0	0	1	4
BHMH	80	79	80	111	88
EATM	0	0	3	2	0
VG	7	2	4	4	6
KBD	30	18	5	33	50
TMB	17	4	15	19	18
SNB	173	9	20	47	105
GT	9	1	7	12	5
MN	18	17	98	33	26
MRB	25	7	13	36	38
GN	182	12	38	79	139
KG	19	1	10	23	30
TMR	54	23	93	91	26
SNK	164	6	26	69	27
YZ	31	1	4	12	17
EM	3	2	14	30	28
MŞ	5	10	69	84	58
RB	3	4	109	37	14
BH	2	0	5	8	18
HR	22	2	40	31	22
KE	64	4	229	57	95
KBP	25	3	23	51	27
TT	2	5	56	15	9

TOA	61	16	274	120	57
Şİ	85	29	167	213	69
OKD	47	150	199	108	10
TAOCP	53	159	112	156	46
İZG	130	99	147	248	69
UD	398	57	45	365	0
Total	1717	778	2052	2126	1107

REFERENCES

Ağca, F. (2014). Eski Uygurca metinlerin tarihlendirişi ile ilgili tespitler. *Türkoloji Dergisi, 21*(1), 1-14.

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory 21*, 435-483.

Akerson-Erkman, F. & Ozil, Ş. (1998). *Türkçede Niteleme Sıfat İşlevli Yan Tümceler.* İstanbul: Simurg Yayınevi.

Anaçoğlu, E. (2019). *Mahmud bin Kadi-i Manyas'ın manzum Gülistan tercümesi (Vr. 1b-100b): Metin-dil incelemesi-sözlük-tıpkıbasım* (MA thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=vjszP7PzV0HebcjFEvDfwMq2XvhJuaGfzpEyfo2BO0M7KqJj0vFgEKdMxgDBuaKm

Andersen, H. (1973). Abductive and deductive change. *Language,49*(4), 765–793.

Argunşah, M. (Ed.). (1999). *Tuhfe-i Muradî: İnceleme-metin-dizin.* Ankara: Türk Dil Kurumu Yayınları.

Aydeniz, S. (2008). *Edviye-i müfrede: Metin, indeks, sözlük, dil incelemesi* (Phd thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=fS4sqEZr79C_n60Rk6MjFc_4zPTbSt2Um_kovKIAPL6G_YCt4oz1S1EKDJAvTS 8xB

Aygen, G. (2002a), Subject case in Turkic subordinate clauses: Kazakh, Turkish and Tuvan. *North East Linguistics Society 32* (2), 17.

Aygen, N. G. (2007). Syntax and semantics of genitive subject-case in Turkic. *California Linguistic Notes 32*(2).

Babür, Y. (2013). *Firdevs-i Rumi Süleyman-name-i kebir: 6-7. ciltler/ inceleme-transkripsiyonlu metin* (MA thesis). Retrieved from: https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=iTkOhwevEenJZ3onUvs52uWaC2HCfpPAkhZCnCK-Cb3eAcJfE7UxIXZgQ3zQtKQY

Bakay, Ö., Ergelen, Ö., Sarmış, E., Yıldırım, S., Arıcan, B. N., Kocabalcıoğlu, A., …Yıldız, O. T. (2021). Turkish WordNet KeNet. In *Proceedings of the 11th Global Wordnet Conference* (pp. 166–174). University of South Africa (UNISA): Global Wordnet Association.

Başar, O. S. (2021). *Katip Çelebi'nin Takvimü't-Tevarih'i: Metin-inceleme* (MA thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=RjZwH00oMG4iNa5SgvlggwwUBt6hB-CBzFzfY4oB5OMnjmDW82Bn2H70F0LIFFbH

Barat, K. (Ed.). (2000). *The Uygur-Turkic biography of the seventh-century Chinese Buddhist pilgrim Xuanzang, ninth and tenth chapters.* Blooming, Indiana: Sinor Research Institute of Inner Asian Studies.

Barker, C. & Hankamer, J. & Moore, J. (1990). Wa and Ga in Turkish. In Dziwirek, K. & Farell, P. & Mejías-Bikandi, E. (eds.), *Grammatical relations: A cross-theoretical perspective* (pp. 21-43). Stanford, California: Center for Study of Language and Information.

Berbercan, M. T. (2013). Türk yazı dilinin tarihi dönemleri ve Türkçenin yeri meselesi. *Journal of History School, 2014*(XVII), 765-783.

Canpolat, M. (Ed.). (1995). *Mecmû'atü'n-nezâir: Metin-dizin-tıpkıbasım.* Ankara: Türk Dil Kurumu Yayınları.

Canpolat, M. (Ed.). (2018). *Behcetü'l-hadâik fî mev'izati'l-hâlaik.* Ankara: Türk Dil Kurumu Yayınları.

Chomsky, N. (1981). *Lectures on government and bindings.* Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language*. New York: Praeger.

Chomsky, N. (1986b). *Knowledge of language: Its nature, origin and use.* New York: Praeger.

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

Cinque, G. (2020). *The Syntax of relative clauses: A unified analysis.* Cambridge: Cambridge University Press.

Comrie, B. (1979). Definite and animate direct objects: A natural class, *Linguistica silesiana 3*, 13–21.

Csató, É. Á. 1985. A syntactic analysis of participant nominal constructions in Modern Turkish. In *Beşinci Milletler Arası Türkoloji Kongresi, 23–28 Eylül 1985, Tebliğler 1,* Türk dili. İstanbul: İstanbul Üniversitesi, 39-56.

Çiftçi, S. (1997). *Bahrü'l-hakayık: Transkripsiyon, dil özellikleri, sözlük* (MA thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/tezDetay.jsp?id=R_BkKR9aqtom6IUuZWxtMQ&no=R_BkKR9aqtom6IUuZWxtMQ

Çoban, A. H. (2009). *Ahmet Cevat Paşa'nın Tarih-i Askeri-i Osmani (Kitab-ı Rabi) adlı eserinin transkripsiyonlu metni* (MA thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=CwVIqqBuz1VkysVpueogAZep4zRvs3-fXIcQSyyfi3CjAZA4PXsejk5oiphlkJJ2

Çöltekin, Ç. (2015). A grammar-book treebank of Turkish. In Dickinson, M., Hinrichs, E., Patejuk, A., & Przepiórkowski, A. (Eds.), *Proceedings of the 14th workshop on Treebanks and Linguistic Theories (TLT 14)* (pp. 35–49).

Demirok, Ö. (2019). A semantic characterization of Turkish nominalizations. In Stockwell et al. (Ed.), *Proceedings of the 36th west coast conference on formal linguistics* (pp. 132-42).

Denison, D. (2002, September). History of the sort of construction family. In *Second International Conference on Construction Grammar, Helsinki* (Vol. 7, pp. 279-304).

Dilçin, C. (1991). *Süheyl ü Nev-bahar: İnceleme-metin-sözlük.* Ankara: Atatürk Kültür Merkezi Yayınları.

Elmalı, M. (Ed.). (2016). *Daśakarmapathâvadânamâlâ: Giriş-metin-çeviri-notlar-dizin-tıpkıbasım*. Ankara: Türk Dil Kurumu Yayınları.

Ercilasun, A. B. (2004). *Başlangıcından yirminci yüzyıla Türk Dili Tarihi.* Ankara: Akçağ Yayınları.

Erdağı Doğuer, B. (Ed.). (2013). *Tuḥfe-i Mübârizî: Metin-sözlük.* Ankara: Türk Dil Kurumu Yayınları.

Erdal, M. (2004). *A grammar of old Turkic*. Leiden-Boston: Brill.

Erdağı, B. (2001). Anadolu'da yazılmış ilk Türkçe tıp kitabı. *Türkbilig, 2*(2), 46-54.

Erdem, M. (1992). *Kitab-ı güzide (76a-134a): İnceleme-metin-sözlük* (MA thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=ePX_SaJ0b35Gq45swKG3lOnDoiCKFacQSLRIZqq-72hVdcvuDR6FhPxL98wKuNN-

Ergin, M. (1980). *Kadı Burhaneddin divanı.* İstanbul: İstanbul Üniversitesi Edebiyat Fakültesi Yayınları.

Ergin, M. (2009). *Türk dil bilgisi*. İstanbul: Boğaziçi Yayınları.

Ergüzel, M. M. (Ed.). (2009). *Kitab-ı bâz-name-i padişahi: İnceleme-metin-dizin.* Ankara: Türk Dil Kurumu Yayınları.

Göksel, A. & Kerslake, C. (2005). *Turkish: A comprehensive grammar.* London and New York: Routledge.

Haig, G. (1997). Turkish relative clauses: A tale of two participles. *Turkic languages 1*(2), 184-209.

Hankamer, J. (1972). On the nonexistence of mirror image rules in syntax. In *Syntax and Semantics Volume 1* (pp. 199-212). Brill.

Hankamer, J., & Knecht, L. (1976). The role of subject-object distinction in determining the choice of relative clause participle in Turkish. *Harvard Syntax and Semanties*, *2*, 197-219.

Haspelmath, M. (2008). Object marking, definiteness and animacy. *Handout in Leipzig Spring School on Linguistic Diversity.*

Jenset, G. & McGillivray, B. (2017). *Quantitative historical linguistics: A corpus framework.* Oxford: Oxford University Press.

Johanson, L. (2021). *Turkic.* Cambridge: Cambridge University Press.

Karagöz, E. (2019). *Şerh-i ibarat: Mevlevilik, musiki ve sema* . İstanbul: Türkiye Yazma Eserler Kurumu Başkanlığı Yayınları.

Karataş, A., & Coşkun, H. (Eds.). (2013). *Kısas-ı enbiya: Peygamber kıssaları.* İstanbul: Türkiye Yazma Eserler Kurumu Başkanlığı.

Kauhanen, H. (2022). A bifurcation threshold for contact-induced language change. *Glossa: a journal of general linguistics, 7*(1).

Kauhanen, H., Walkden, G. (2018). Deriving the constant rate effect. *Natural Language & Linguistic Theory, 36*, 483-521.

Kaya, C. (2021). *Uygurca Altun Yaruk: Giriş, metin ve dizin* (Rev. ed.). Ankara: Türk Dil Kurumu Yayınları.

Kennelly, S. (1997). Nonspecific external arguments in Turkish. *Dilbilim Araştırmaları Dergisi, 8*, 58-75.

Koç, U. (2021). *Bir Osmanlı Türk askerinin maceralı esirlik hikayesi: Temeşvarlı Osman Ağa'nın Esaretnamesi'nin orjinal ve sadeleştirilmiş Latin harfleriyle transkripsiyonu.* Berlin: Epubli.

Kodner, J. (2020). *Language acquisition in the past.* (PhD thesis). University of Pennsylvania, Philadelphia, United State of America.

Kormaz, Z. (1974). Eski Türkçe'deki Oğuzca belirtiler. *Türkoloji Dergisi, 6*(1), pp. 15-30.

Korkmaz, Z. (1994). Eski Türk yazı dilinden yeni yazı dillerine geçiş devri ve özellikleri. *Türk Dili Araştırmaları Yıllığı- Belleten, 39*(1991), pp. 55-64.

Kornfilt, J. (1997). *Turkish.* London: Routledge.

Kornfilt, J. (2001). On the syntax and morphology of clausal complements and adjuncts in the Turkic languages. *Aspects of typology and universals, 1*, 63-82.

Kornfilt, J. (2003). Subject case in Turkish nominalized clauses. In U. Junghanns, & L. Szucsich, *Syntactic Structures and Morphological Information* (pp. 129-215). Berlin and New York: Mouton de Gruyter.

Kornfilt, J. (2006). Free relatives as light-headed relatives in Turkish. In H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz & J. Koster (Ed.), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk* (pp. 340-349). Berlin, Boston: De Gruyter Mouton.

Kornfilt, J. (2008). DOM and two types of DSM in Turkish. In *Differential subject marking* (pp. 79-111). Dordrecht: Springer Netherlands.

Kornfiilt, J. (2009). Turkish and Turkic languages. In B. Comrie (Ed.), *The World's major languages* (2nd ed., pp. 519-544). London: Routledge.

Kornfilt, J. (2012). Agreement and its placement in Turkic nonsubject relative clauses. In Guglielmo Cinque, and Richard S. Kayne (eds). The Oxford Handbook of Comparative Syntax.

Kornfilt, J. (2014). Free adjuncts and non-free relatives in Turkish. In Arslan Kechriotis, C., Akar, D., Kelepir, M., Öztürk, B., (Eds.). *Dilbilim Araştırmaları* (pp. 117-29). Istanbul: Boğaziçi University Publications.

Krause, E. & von Heusinger, K. (2019). Gradient effects of animacy on differential object marking in Turkish. *Open Linguistics, 5*(1), 171-190.

Kroch, A. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change, 1,* 199–244.

Kroeber, A, L., & Chretien, C.D. (1937). Quantitative classification of Inda-European languages. *Linguistic of Society of America, 13*(2), 83-103.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association, 47*(260), 583-621.

Kuzgun, A., Cesur, N., Arıcan, B. N., Özçelik, M., Marşan, B., Kara, N., ... & Yıldız, O. T. (2020, October). On building the largest and cross-linguistic Turkish dependency corpus. In *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)* (pp. 1-6). IEEE.

Laszakovits, S. (2013). *Specificity and differential object and subject marking in Turkish.* (PhD thesis). University of Vienna, Vienna, Austria.

Lightfoot, D. (1976). Trace theory and twice-moved NPs. *Linguistic Inquiry, 7(*4), 559-582.

Lightfoot, D. (1991). *How to set parameters: Arguments from language change.* Cambridge: The MIT Press.

Lightfoot, D. (1979). *Principles of diachronic syntax.* Cambridge: Cambridge University Press.

Lightfoot, D. (2006). *How new languages emerge.* Cambridge: Cambridge University Press.

Mansuroğlu, M. (1958). *Sultan Veled'in Türkçe manzumeleri.* İstanbul: İstanbul Üniversitesi Edebiyat Fakültesi Yayınları.

Marşan, B., Kara, N., Özçelik, M., Arıcan, B. N., Cesur, N., Kuzgun, A., … & Yıldız, O. T. (2021, January). Building the Turkish FrameNet. In *Proceedings of the 11th Global Wordnet Conference* (pp. 118-125). University of South Africa (UNISA): Global Wordnet Association.

Nikoosokhan Sedghi, E. (2015). *Ayyuki ile Yusuf-ı Meddah'ın Varka ve Gülşah mesnevilerinin karşılaştırılması* (MA thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=Br_XTptK8CZ70f0JGX9xEhlNvx1vYU3URmNGCRJ0YT3_NvyFxNxGxUCFqDUIsFSJ

Ölmez, M. (1994). *Hsüan-tsang'ın Eski Uygurca yaşam öyküsü VI. bölüm* (PhD thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=EEdeQgIdFRxX5NbvVau-Arl3g348fuNW3vEHR6LEQl7DyuVDqXGU_ugM_1KOVF32

Ölmez, M. (2002). Die alttürkische Xuanzang-biographie II. In M. Ölmez, & S.-C. Raschmann (Eds.), *Splitter aus der gegend von Turfan. Festschrift für Peter Zieme anläßlich seines 60. geburtstags.* (pp. 199-242). İstanbul-Berlin: İstanbul-Berlin 2002.

Ölmez, M. (Ed.). (2015). *Orhon-Uygur Hanlığı dönemi Moğolistan'daki eski Türk yazıtları: Metin-çeviri-sözlük* (3rd ed.). Ankara: BilgeSu Yayıncılık.

Ölmez, M & Klaus Röhborn, K. (2001). *Die alttürkische Xuanzang-Biographie III.* (Gabain, Ed.). Wiesbaden: Harrassowitz Verlag.

Önler, Z. (1981). *Celalüddin Hızır (Hacı Paşa) müntahab-ı şifa: İnceleme-metin-dizinler* (PhD thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/tezDetay.jsp?id=yL_BQdegmn0_utcjx19qUg&no=cDkMUdh6K6gFajNzpzuJNw

Önler, Z. (Ed.). (2000). *Revnak-ı bustan.* Ankara: Türk Dil Kurumu Yayınları.

Özavşar, R. (2009). *Marzubanname tercümesi: Metin, çeviri, art zamanlı anlam değişimleri, dizin* (MA thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=NtBAevXNhYaNqJFoAcdBdsUIjKFP92jPMAGrSEOp_9WyhpXsr33HN5Nq470FFDPI

Özbay, B. (2019). *Huastuanift: Manihaist Uygurların tövbe duası* (Rev. ed.). Ankara: Türk Dil Kurumu Yayınları.

Özezen, M. (1994). *Gelibolulu Mustafa Ali'nin Hilyetü'r- Rical'i* (MA thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=48XPj7KKQhKUgntkUiKO3KD1DTIGQM-H2eCeyi0Gq9Km6MhPGwzJQshY_G-FLESc

Özsoy, S. (1987b). The null subject parameter in Turkish. In H. Boeschoten & L. Verhoeven (Eds.), *Studies on Modern Turkish: Proceedings of the 3rd Conference on Turkish Linguistics* (pp. 82-90). Tilburg: Tilburg University Press.

Öztürk, B. (2007). Asymmetry in double-object constructions in Turkish. In M. Kelepir & B. Öztürk (Eds.), *Proceedings on Workshop on Altaic Formal Linguistics 2*, Boğaziçi University, Istanbul, October 2004. MITWPL, MIT, Cambridge. (pp. 305-316)

Öztürk, B. (2008). Relativization strategies in Turkish. In C. Boeckx & S. Ulutaş (Eds.), *Proceedings of Workshop on Altaic Formal Linguistics 4*, Harvard University, Cambridge, May 2007. MITWPL, MIT, Cambridge. (pp. 241-253).

Öztürk, B. (2009). Incorporating agents. *Lingua, 119*(2), 334-358.

Öztürk, B., & Taylan, E. E. (2016). Possessive constructions in Turkish. *Lingua*, *182*, 88-108.

Roberts, I. (2021). *Diachronic syntax* (2nd ed.). Oxford, UK ; New York, NY: Oxford University Press.

Sağlam, C. (2021). *Harita-i kapudanan-ı derya: Osmanlı kaptanıderyaları (1352-1853).* İstanbul: Türkiye Yazma Eserler Kurumu Başkanlığı Yayınları.

Satık, D. (2020). *Turkic genitive case and agreement asymmetries.* Manuscript, Harvard University.

Sheehan, M., & Hinzen, W. (2011). Moving towards the edge. *Linguistic Analysis, 37*(3-4), 405-458.

Silverstein, M. (1976). Hierarchy of features and ergativity. In R. M. W. Dixon, (Eds.), *Grammatical Categories in Australian Languages* (pp. 112-171). New Jersey: Humanities Press.

Sulubacak U., Pamay, T., & Eryiğit, G. (2016). Universal dependencies for Turkish. In *Proceedings of the 26th International Conference on Computational Linguistics* (pp. 3444-3454). Osaka, Japan: The COLING 2016 Organizing Committee.

Taş, İ. (Ed.). (2017). *Yusuf ve Zeliha: Giriş-metin-çeviri-notlar-sözlük.* Ankara: T.C. Kültür ve Turizm Bakanlığı Kütüphaneler ve Yayınlar Müdürlüğü.

Tekin, Ş. (1974). 1343 tarihli bir Eski Anadolu Türkçesi metni ve Türk dili tarihinde 'Olġa-bolġa' sorunu. *Türk Dili Araştırmaları Yıllığı - Belleten, 21-22*(1973-1974), 59-157.

Tekin, Ş. (1976). *Uygurca metinler II: Maytrısimit.* Ankara: Atatürk Üniversitesi Yayınları.

Tekin, T. (1998). *Orhon Yazıtları: Kül Tigin, Bilge Kağan, Tunyukuk* (2nd ed.). (M. Ölmez, Ed.) İstanbul: Simurg Yayınevi.

Tekin, T. (2003). *Orhon Türkçesi Grameri.* İstanbul: Türk Dilleri Araştırmaları Dizisi.

Timurtaş, F. (2010). Küçük Eski Anadolu Türkçesi grameri. *Türkiyat Mecmuası*, 18, 331-368.

Türk, U., Atmaca, F., Özateş, Ş.B., Köksal, A., Ozturk Basaran, B. Gungor, T., & Özgür, A. (2019). Turkish treebanking: Unifying and constructing efforts. In *Proceedings of the 13th Linguistic Annotation Workshop* (pp. 166–177). Florence, Italy: Association for Computational Linguistics.

Türk, U., Atmaca, F., Özateş, Ş. B., Öztürk Başaran, B., Güngör, T., & Özgür, A. (2019). Improving the annotations in the Turkish Universal Dependency Treebank. In *Proceedings of the Third Workshop on Universal Dependencies* (pp. 108–115). Paris, France: Association for Computational Linguistics.

Türk, U., Atmaca, F., Ozates, Ş. B., Berk, G., Bedir, S. T., Koksal, A., … Ozgur, A. (2022). Resources for Turkish dependency parsing: introducing the BOUN Treebank and the BoAT annotation tool. *Language Resources and Evaluation, 56*(1), 259–307.

Uluç, G., Hüseynli, D., Mehdiyeva, D., Qır, Ö., Sərxanbəyova, C., & İbrahimova, İ. (2013). *Irk Bitig: Qədim Uyğur dilində yazılmış fal kitabı.* Baku: Türkçəsi Varkən Topluluğu.

Underhill, R. (1972). Turkish participles. *Linguistic Inquiry 3*, 87-99.

Wurmbrand, S., & Lohninger, M. (2019). An implicational universal in complementation. In Hartmann, Jutta M. & Angelika Wöllstein (eds.), *Propositional arguments in cross-linguistic research.* Tübingen: Gunter Narr Verlag.

Yang, C. (2002). *Knowledge and learning in natural language*. Oxford: Oxford University Press.

Yavuz, K. (Ed.). (2000). *Garib-name* (Vols. 1/1,1/2). İstanbul: Türk Dil Kurumu Yayınları.

Yenice, A.B. (2024). *Locative case in Turkish temporal adverbials.* (Unpublished MA thesis). Boğaziçi University, İstanbul, Turkey.

Yılmaz, E., & Demir, N. (2007). Uygur Edebiyatı (VIII-XIV. yüzyıl): Nesir. *Türk Edebiyatı Tarihi, 1*, 154-176.

Yılmaz, M. (2020). *İkramiyeli gazete transkripsiyon ve değerlendirmesi* (MA thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=Eb5EkakJlp3olBdo_wNEGSs_BG00uf_lEaTi_8SwQ7g-_pear1HdetbD3P01iNwg