OVERT VERSUS NULL SUBJECT PRONOUN VARIATION IN THE TURKISH SPOKEN IN TURKEY AND IN NEW YORK CITY

by

DİDEM KOBAN

A dissertation submitted to the Graduate Faculty in Linguistics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2010

This manuscript has been read and accepted for the
Graduate Faculty in Linguistics in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Date	Ricardo Otheguy, Ph.D. Chair of Examining Committee
Date	Robert W. Fiengo, Ph.D. Executive Officer

Gita Martohardjono, Ph.D.

Marcel den Dikken, Ph.D.

Supervision Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

OVERT VERSUS NULL SUBJECT PRONOUN VARIATION IN THE TURKISH SPOKEN IN TURKEY AND IN NEW YORK CITY

by

Didem Koban

Adviser: Professor Ricardo Otheguy

The purpose of this dissertation is to examine the use of subject personal pronouns in the Turkish spoken in Turkey and in New York City from a variationist perspective. Whereas the variable use of subject personal pronouns in Turkish has been extensively analyzed in many studies conducted in Europe, it has received much less attention in the U.S. This study has as one of its aims replicating the study conducted by Otheguy, Zentella and Livert (2007) where the influence of different social and linguistic variables on the expression of Spanish subject pronouns was examined across Latin American and Caribbean immigrant generations in New York. The present study examines several linguistic and social variables that condition the presence and absence of subject personal pronouns in the speech of 20 adult speakers living in Turkey (TT) and 20 living in New York (TNY). The study compares the rate of subject pronoun use in Turkey with that of NYC and whether contact with English has an influence on the overt pronoun rate.

In both the TT and TNY samples, there were an equal number of males and females and an equal number of speakers from working and professional classes. The speakers ranged in age from 20 to 80. Data analysis involved Anovas, correlations, cross-tabulations and multivariate regression analyses of linguistic and social variables. The

linguistic variables, which were also examined in Otheguy et al. (2007) and in other previous studies, are person and number of the pronoun and of the verb, continuity of reference, and TMA of the verb. Social variables that are analyzed are gender, social class, age of the informant, education, age of arrival in NYC, length of residence in NYC and so forth.

The results of the study indicate that TT and TNY resemble each other regarding the linguistic variables that condition the distribution of subject personal pronouns and regarding the order of the variables that account for the most variance in the use of the pronouns. However, the two samples differ from one another with respect to the order and strength of the constraints within the person and number of the verb variable. In addition, we find a significantly higher rate of overt pronoun use for TNY than for TT. These findings are consistent with those obtained in the Spanish study and provide clear support for an English contact hypothesis when the increased use of overt subject pronouns among TNY and differences in constraint hierarchies between TT and TNY are taken into consideration.

ACKNOWLEDGEMENTS

I would like to express my appreciation to everyone who has helped me complete my dissertation. There are many people whose efforts made this work possible. First, I wish to thank my dissertation adviser Prof. Ricardo Otheguy whose continued guidance and constant support made it possible for me to complete my dissertation. He always had time to advise me, helped me improve my work and has given me a great deal of useful advice. I learned so much from him.

I wish to thank each member of my dissertation committee for being in my committee and the knowledge they shared with me. Besides Prof. Ricardo Otheguy, the committee consisted of Prof. Gita Martohardjono and Prof. Marcel den Dikken. They provided their specific expertise and guidance. I am truly grateful to them for their contribution.

I would like to dedicate this doctoral dissertation to my parents, Birsen and Rıdvan Koban and my sister, Sinem. Their love and support have been invaluable. They have stood by me through many difficulties and frustrations. I thank my mother and father who believed in me, supported me and prayed for me every step of the way. This dissertation would not have finished on time without my father's help. I wish to thank him for helping me transcribe the interviews. I am especially grateful to my loving and protective sister for her invaluable support and assistance. I count myself very fortunate to have a sister like her. She was always there for me. I could not have done this without her.

I also want to express my gratitude to the participants who are part of the study. This dissertation draws upon data generated as a result of the 40 interviews I conducted in

Turkey and in New York City. The participants' names have to remain anonymous but I wish to sincerely thank each of them.

I also wish to thank my colleagues and friends who provided general support and encouragement. The data collection would not have been possible without the help of my friends. I want to thank them for introducing me to my participants. Their help meant a lot to me.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1. Description of the study

The purpose of this dissertation is to analyze the alternation between null and overt subject personal pronouns (SPP) in the Turkish spoken in Turkey and in New York City (NYC). Some languages, such as English, require that, with few and well delineated exceptions, each sentence include an overt subject pronoun, whereas in many others, such as Spanish, Italian and Turkish, overt subject pronouns may be omitted. Such languages are referred to as pro-drop or null subject languages. In (1) one can either use an overt pronoun or omit it. In (1) and throughout this dissertation, overt SPPs are presented in bold and null pronouns are represented by Ø. The alternation between overt and null SPPs is shown by /.

(1) **Ben** / **Ø** kitap oku-yor-um.
I book read-IMPF-1SG.
'I'm reading a book.'

In (1) even if the overt SPP *ben* 'I' is missing, it can be easily inferred as the verb *oku-mak* 'to read' is inflected with the first person present progressive tense form. Regardless of the occurrence of the overt or null pronoun, the two structures are equivalent in terms of conveying the same basic referential meaning.

In recent years, several studies have addressed the use of SPPs in the Turkish spoken in Turkey and in Europe, particularly in Germany, France and Netherlands. However, there is little or no research that compares the distribution of overt and null SPPs in the Turkish spoken in the U.S. with the distribution of these forms in the Turkish

spoken in Turkey. In addition, to my knowledge, none of these studies has approached the problem from a variationist perspective. The significance of this dissertation lies in the fact that it examines the use of overt and null SPPs in the Turkish spoken in NYC through a comparison with the Turkish in Turkey from a variationist model based on a corpus of naturally-occurring data.

The dissertation is framed around the replication of a study conducted by Otheguy, Zentella and Livert in 2007, where SPP use is analyzed in several Spanish speaking communities in NYC. Patterning my research on the Spanish study, my aim is to identify the contextual-linguistic and socio-demographic variables that influence speakers' selection of either an overt or a null pronoun. I compare the use of SPPs in Turkish in Turkey with that of Turkish in NYC and determine whether the variables that I have selected have a significant effect on the use of SPPs, that is, whether they are real variables. My goal is also to find out (a) which of the significant variables have greater and lesser effects on the occurrence of overt pronouns, and (b) within each variable, which factors or constraints are ranked higher and lower in terms of this effect.

Further, in keeping up with the Spanish study that this study is replicating, my goal is to establish whether contact with English has an influence over overt subject pronoun use among Turks in NYC. To achieve this, I explore the differences between Turks in Turkey (TT) and Turks in New York (TNY) with respect to their rates of use of overt pronouns, and study the correlations between the overt pronoun rate and several social variables such as place of residence (Turkey vs. NYC), immigrant generation, age of arrival in NYC, years spent in NYC, and the amount of use of English.

The study of language variation is concerned with the social significance of variation through relating linguistic features to social variables such as social class, age,

gender, ethnicity and geographical origin (Labov, 1972; Weinreich, Labov & Herzog, 1968; Fasold, 1990). Thus, this study aims to analyze the relationship between the rate of subject pronoun use and the ordering of predictor variables and of factors within variables, and other non-linguistic variables such as gender, social class, education, and age of the informant.

The results of our study indicate that social variables such as place of residence (Turkey or New York) and, for residents of New York, immigrant generation, age of arrival, age of the informant and the amount of Turkish used with one's father all have a significant impact on the use of overt SPPs. However, a multiple linear regression analysis revealed that place of residence and age of arrival account for the most variance between overt and null subject pronouns. With regard to the effects of linguistic variables, the analysis indicates that TT and TNY are similar with regard to their variable hierarchies; continuity of reference provides the best account of SPP variation in both populations. However, the two groups are different with respect to constraint hierarchies within the variable Person and number ending of the verb.

It is argued in this dissertation that the significant impact on the use of pronouns of place of residence and, for the TNY, of age of arrival as well as the differences in constraint hierarchies between TT and TNY suggest that there is language-contact influence from English on Turkish in New York with regard to the use of overt pronouns.

The dissertation focuses exclusively on nominative personal pronouns acting as subjects. Personal pronouns in the nominative that make reference to inanimates are included, alongside those making animate references, because Turkish verbs whose subjects are inanimate can take either an overt or a null pronoun just as do those whose subjects are animates. This study includes all environments where these pronouns can be

variably used. These environments include some that are familiar from the study of languages such as Italian and Spanish, namely finite verbs, as in (2). Also included are several other environments where these nominative pronouns occur in Turkish, but not in the Western European languages. These less-familiar environments include, for example, non-finite verbs in adverbial subordinate clauses and locative existential expressions, as in (3).

(2) **Ben** / **Ø** kitap oku-yor-um.
I book read-IMPF-1SG.
'I'm reading a book.'

(3) **Ben** / **Ø** orada yok-tu-m.
I there non.existent-P.COP-1SG.
'I wasn't there.'

Also studied in this dissertation are environments where nominative pronouns occur with nouns, adjectives, postpositional phrases, words expressing necessity, and some wh-phrases. In these environments, these forms can function as predicates and SPPs occurring as subjects of these predicates can be variably used. (Full details regarding which items are included and excluded are given in chapter 3). For example, the sentence in (4) consists of a subject pronoun and a noun, functioning as a predicate.

(4) **Sen** / **Ø** satıcı-(y)mış-sın.
You seller-EV.COP-2SG.
'You were a seller.'

2. Organization of the dissertation

The dissertation is organized as follows. The following section provides information on the subject personal pronouns in Turkish. Chapter 2 reviews some of the major works on overt and null pronoun variation in Turkish, along with a discussion of previous variationist research conducted for other languages. Chapter 3 defines the envelope of variation for the present study. Chapter 4 provides the research questions and hypotheses. This is followed by chapter 5 where the informants, data collection procedure and materials are described.

Finally, results for the social variables examined in this study are presented in chapters 6 and 7. Those for the linguistic variables are provided in chapters 8 and 9. The linguistic variables and their constraints are provided in Appendix A and the social variables and their constraints are given in Appendix B. The sociolinguistic survey given to the informants who participated in the study is included in Appendix C.

3. Subject personal pronouns in Turkish

Before turning to an overview of SPPs in Turkish, it would be good to be reminded that Turkish is an agglutinative language, with rich verbal morphology. Verbs are marked for tense, mood, aspect and voice, as well as person and number. In addition, nouns and personal pronouns in Turkish can be inflected for all the six cases: the *nominative, accusative, dative, locative, ablative* and *genitive*.

The following SPPs inflected in the nominative will be analyzed in the study:

Ben	'I'	*Biz*	'we'
Sen	'you' (familiar)	*Siz*	'you' (plural), (formal singular)
O	'he/she/it'	*Onlar*	'they'
İnsan	'one'		
İnsanlar	'people'		

In addition to the personal pronouns, demonstrative pronouns such as *bu* 'this,' *şu* 'that,' *bunlar* 'these,' and *şunlar* 'those' in the nominative case will also be included in the study because in colloquial Turkish, these can be used to make personal references. *Bu* and *şu* can be used to refer to 'he' or 'she' and *bunlar* and *şunlar* can be used to refer to 'they.' For example:

(5) **Bunlar** Ankara-(y)a git-ti-ler.
These Ankara-DAT go-PF-3PL.
'These [They] went to Ankara.'

However, pronominalized definite and indefinite determiners such as *başkası* 'another (one),' *bazısı* 'some (of them),' *birçoğu* 'many (of them),' *biri*(*si*) 'someone,' *birkaçı* 'a few (of them),' and *bütünü* 'all of it' will not be included in the study.

Note that the person and number properties of some pronouns do not always match the person and number inflections on the verbs. For example, when a 3pl overt pronoun occurs in a sentence, predicates do not usually get inflected with the 3pl marker. This depends on whether the subject refers to a human or non-human and specific or non-specific entity (Göksel & Kerslake, 2005). If the overt subject pronoun refers to a human

and is specific, then the plural marking of the verb is optional. Verbs can be marked with either 3pl or 3sg. This optionality can be seen in (6) and (7).

(6) **Onlar** İzmir-e git-ti-ler.
They İzmir-DAT go-PF-3PL.
'They went to İzmir.'

(7) **Onlar** İzmir- e git-ti.
They İzmir-DAT go-PF-3SG.
'They went to İzmir.'

Also, the verb inflection for 2sg (formal) and 2pl pronouns can be the same. Consider (8) where the referent of the subject pronoun *siz* 'you' can either be 2sg (formal) or 2pl. Whether the subject refers to 2sg or 2pl is inferred from the context.

(8) **Siz** İzmir- e git-ti-niz.
You İzmir-DAT go-PF-2SG(FORMAL)/PLURAL.
'You went to İzmir.'

CHAPTER 2

THE REVIEW OF THE LITERATURE

1. Introduction

Overt and null SPP variation has been a widely studied topic in sociolinguistics. Linguists dealing with this topic have conducted studies from the standpoint of both syntax and pragmatics. They have elicited speech as well as written narratives in several languages, from both children and adults, and investigated the effects of linguistic variables such as verb morphology, and continuity of reference, that is, whether the subject of a given clause in a narrative is the same as the subject in the previous clause. They have also investigated the effects of social variables such as age, education, social class and gender. In addition, many linguists have investigated, in a variety of contact languages, including contact Turkish in several European settings, but especially in contact Spanish in the U.S., whether or not influence from a majority language could account for aspects of the use of either an overt or a null SPP.

In the sections that follow, I will review important studies that analyzed SPP variation in the Turkish spoken in Turkey and in Europe. Note that the present study focused on only adults. However, I will also focus on the studies that involved children along with those conducted with adults.

2. Research conducted on the Turkish spoken in Turkey and in Europe

Linguists studying subject pronouns in Turkish have emphasized the importance of discourse-pragmatic conditions, which can account for the use of overt or null SPPs, as well as the importance of the influence of a majority language that can account for the increased use of overt SPPs in the Turkish spoken in Europe. We first discuss studies that

investigated the discourse-pragmatic functions of SPPs. We then turn to the studies that analyzed the influence of a majority language on SPP expression in contact Turkish.

2.1 Topic

One of the discourse functions of overt subject pronouns is to introduce a new topic or change the topic of the conversation (Enç, 1986). Enç refers to topic as what the portion of the discourse is about and argues that introduction of a topic in discourse or a change in conversation results in higher use of overt pronouns because new information is being introduced. Once the topic of the discourse is introduced or changed, it is maintained by null pronouns. Overt SPPs give extra pragmatic information, thus, the use of overt pronouns to maintain the topic of discourse is "pragmatically redundant" (Erguvanlı-Taylan, 1986:196). This is based on Grice's (1975) Maxim of Quantity according to which speakers give as much information as is needed, and no more. Enç's argument is illustrated in (1), where the speaker continues a conversation about being late unless a taxi is taken.

(1) **Ben** banka-(y)a git-me-(y)i unut-tu-m.
I bank-DAT go-INF-ACC forget-PF-1SG.
'I forgot to go to the bank.'

(Enç, 1986:197)

Here, the speaker uses an overt SPP because, according to Enç, he/she wants to switch the topic of the conversation. If the speaker uses a null pronoun instead of an overt, we assume that he/she is commenting on the previously established topic. Consider (2) where the English pronoun in parenthesis corresponds to the null subject, which is shown by Ø.

(2) **Ø** banka-(y)a git-me-(y)i unut-tu-m.
(I) bank-DAT go-INF-ACC forget-PF-1SG.
'I forgot to go to the bank.'

(Enç, 1986:197)

In (2), it can be said that by using a null pronoun, the speaker is making a connection between his/her response and the previously established topic of the conversation. The speaker may imply that he/she does not have any money because he/she forgot to go to the bank and therefore, he/she cannot take a taxi.

Some scholars, however, suggested that topic change need not always be obligatorily signaled by an overt pronoun. According to Turan (1996), subject pronouns do not have to be overtly expressed as the tense of the verb may also indicate that the speaker has switched to another topic. The narrative below elicited from one of the informants in this study illustrates that the informant uses a null SPP instead of an overt even though there is a switch to another topic. The informant is a 41-year-old man who lived in Turkey at the time of the interview.

(3) …**bu** 2, 3 sene İzmir-e gel-miş, yaşa-mış. Ondan sonracığıma,
…[this] he 2, 3 year İzmir-DAT come-PF-3SG live-PF-3SG. Then

ortaokul dönem-in-de iyi-(y)di-m. (8TR)
secondary school term-3SG.POSS-LOC good-PF-1SG.

'This[he] had come and lived in İzmir for 2, 3 years. Then, I was good in secondary school [I was a good student in secondary school].'

In (3), the informant used a null pronoun to switch the topic from his father's living in İzmir to his being a good student in secondary school.

2.2 Contrast

A number of scholars argued that the presence of subject pronouns is considered obligatory for purposes of contrasting referents. Overt SPPs are used to contrast the referent of a pronoun with the referent of another pronoun or an NP (Erguvanlı-Taylan, 1986; Enç, 1986). Consider the following example taken from Erguvanlı-Taylan (1986).

(4) **Ben** iş-e gecik-ti-m ama **sen** henüz gecik-me-di-n.
I work-DAT be late-PF-1SG but you yet be late-NEG-PF-2SG.
'I'm late to work but you're not late to work yet.'

(Erguvanlı-Taylan, 1986:210)

According to the author, the overt SPPs in (4) are necessary to contrast the subject referents of the two clauses as contrast cannot be achieved by only the person suffix on the verb.

Enç (1986) also suggested that overt SPPs are obligatory in yes/no contrast or what Erguvanlı-Taylan (1986) calls new information, that is, the subject of a sentence constitutes new information. One of her examples is repeated here.

(5a) Bu hava-da kimse top oyna-ma-z.
This weather-LOC nobody ball play-NEG-AOR.
'Nobody will play ball in this weather.'

(5b) **Ben** oyna-r-ım.
I play-AOR-1SG.
'I'll play.'

(Enç, 1986:205)

Enç (1986) argues that there is a form of contrast in (5a) and (5b). The sentence in (5b) is a counterexample to the assertion made by the sentence in (5a).

Some Turkish scholars, however, implied that occurrence of overt subjects in contrastive contexts is not necessarily obligatory. According to Kornfilt (1997), when an adverbial such as 'as for' is added in contrastive environments, then an overt SPP becomes optional. An example of this is provided in (6).

(6) Ahmet-e gelince **O / Ø** sen-i çok beğen-ir.
Ahmet-DAT coming he you-ACC very admire-AOR.
'As for Ahmet, he admires you very much.'

2.3 Continuity of reference

Continuity of reference, as mentioned earlier, deals with whether the subject of a given sentence in a narrative is the same as or different from the subject in the previous clause. Of relevance to this variable is the work of Cameron (1995) who referred to this variable as switch reference and defined it as "…two related reference relations that may hold between two NPs. When these two NPs have different referents, they are 'switch' in reference and when they share the same referent, they are 'same' in reference" (Cameron, 1995:4). This can be illustrated as:

NP_1 + Tensed V_1 (X)……(Y) NP_2 + Tensed V_2 (Z)

The first NP is referred to as the trigger and the second one as the target. We look at the target NP with respect to the trigger NP to see whether it is same or different in reference. In other words, the subject of the second clause shows whether it has the same or different referent from the subject of the immediately preceding clause. Consider the following example.

(7) **Ben** orda uyu-du-m. Sabah erken kalk-tı-m.
I there sleep-PF-1SG. Morning early wake up-PF-1SG.
'I slept there. I woke up early in the morning.'

In (7), the subject of 'woke up' is null and its referent is the same as the subject of the finite verb, 'slept' in the first clause. This is an example of 'same reference.' If the subject referent of 'woke up' had been 2sg instead of 1sg, the null subject would constitute an example of 'switch reference', since the subject referent of the verb in the first clause is 1sg.

The influence of reference continuity on SPP expression has been one of the most investigated topics in the Turkish spoken in Turkey and in Europe. Studies have shown that SPP variation is constrained by continuity of reference. Null SPPs are more likely to be used than overt SPPs or NPs when their referents can be predicted or recovered from the discourse context (Kerslake, 1987). The following examples illustrate this.

(8a) Ayşe şarkı söyle-r ve **Ø** piyano çal-ar.
Ayşe song sing-AOR CONJ piano play-AOR.
'Ayşe sings and plays the piano.'

(8b) Ayşe şarkı söyle-r ve Ayşe piyano çal-ar.
Ayşe song sing-AOR CONJ Ayşe piano play-AOR.
'Ayşe sings and plays the piano.'

(Kerslake, 1987:91)

According to Kerslake, (8a) is more acceptable than (8b). The use of the NP in the second clause in (8b) is not ungrammatical, however, "the element of communicative redundancy which (8b) contains makes (8b) infelicitious' (p. 91).

The relationship between continuity of reference and subject pronoun use was examined in a study by Turan (1996). Turan performed a quantitative analysis of subject

pronouns in a corpus of discourse samples culled from published narratives. Her results indicated that speakers used overt pronouns when they needed to give new information about an entity and null subjects for maintenance of discourse coherence.

Ruhi (1992) analyzed the use of null and full NP subjects in written narratives collected from first year university students who were all native Turkish speakers living in Turkey. The students renarrated a story from memory. Ruhi determined which form the students used when introducing a new character or referring to the same character in their narratives. The results indicated that the students used a null pronoun to refer to the main character after it was introduced in the story and a full NP to switch to another character in the story.

In Gűrcanli et al. (2006), the effects of different reference on overt SPPs were explored in the speech of twenty-two children (age ranged from 3 to 4;10) and twenty-four adults (age ranged from 22 to 26) living in Istanbul, Turkey. The researchers showed the informants 17 animated video clips depicting motion events and asked them to talk about those events. The results indicated that the children omitted subjects more than the adults. In addition, discourse pragmatic information had an important effect on the children's null subject use in that when subjects had old referents, the children used null subjects and when subjects introduced new referents, the children used overt subjects.

Continuity of reference and the distribution of overt and null SPPs in two short stories were investigated in Kerslake (1987). The results of the quantitative analyses revealed that null pronouns are used if the subject of a previous sentence has just been introduced into the discourse and if it is immediately accessible. Overt pronouns were used to eliminate ambiguity. The author reported that since the sample in the study was very small, further research is needed to confirm the conclusions.

Research conducted with bilingual children living in Netherlands gave similar results especially with respect to maintenance of reference. In a quantitative study, Aarssen (1996) investigated the distribution of nominal expressions, overt and null SPPs in the speech of monolingual and bilingual Turkish children living in Netherlands. The informants' ages ranged from four to ten. The results showed that the bilingual children preferred nominal forms when there was a switch in reference. As for reference maintenance, the bilingual children in all age groups, except for the four-year-olds, used null pronouns. The four-year-olds used an increasing number of overt pronouns. Aarssen concluded that there may be a transfer from Dutch into Turkish since the bilingual speakers used more pronouns than the Turkish monolinguals.

These findings were supported in a similar study that involved the narrative production of eight-year-old bilingual Turkish-Dutch children (Verhoeven, 1990). With respect to maintenance of reference, most of the bilingual children used null pronouns. Thirteen percent of the children used nominals whereas 11 percent used overt pronouns. As for switch reference, 32 percent of the bilinguals used nominals, 39 percent used overt pronouns, and 24 percent used null pronouns. The results were inconclusive regarding direct influence from Dutch.

In another study conducted in Netherlands, Schaufeli (1994) compared oral and written language samples of twelve-year-old Turkish children in Turkey with those in Netherlands. Her results indicated that the bilingual children used overt SPPs for participant introduction and shift. They used null pronouns for maintenance of reference only.

Here, I would like to point out the difference between topic change and continuity of reference in the literature in terms of SPP use. As we have discussed earlier, scholars

analyzed the two variables separately and argued that both are determining variables in SPP expression. With respect to topic change, the use of an overt pronoun indicates that the topic of discourse has changed or a new topic has been introduced. As for continuity of reference, a switch in reference consists simply of a verb having a different subject from the previous verb. An introduction or change in topic is not required. In other words, continuity of reference does not deal with whether or not an overt pronoun is used to change the topic of the conversation or introduce a new topic. Consider the following example taken from a speaker telling me his life story.

(9) ...1 sene sonra doktora-(y)ı bitir-ecek. **Ben** Rafineri-de-ki
…1 year after doctorate-ACC finish-FUT-3SG. I Refinery-LOC-GEN

çalış-ma-lar-ım-ı 2004 yılı 10 Eylül-ün-de bitir-miş
work-VN-PL-1SG.POSS-ACC 2004 year 10 September-3SG.POSS-LOC finish-PF

ol-du-m. (1TR)
AUX-PF-1SG

'She will finish her doctorate in a year. I retired from the Refinery on September 10, 2004.'

Here, we see a switch in reference. The speaker used an overt pronoun in the second clause to refer to himself. We also see a switch in the topic of the conversation indicated by the overt SPP. Note that in many studies, as well as in the present study, only continuity of reference is examined. Therefore, the relationship between topic change and overt pronoun use is not considered.

2.4 The influence of a majority language on Turkish

This section reviews various studies that investigated the influence of a majority language on SPP expression in Turkish. Linguists analyzed the influence of English and of several languages spoken in Europe on the increased use of overt SPPs in Turkish. Whereas some linguists were able to attribute high rates of overt SPP use in contact Turkish to the influence of English and of the majority languages in Europe, some could not obtain conclusive results due to lack of monolingual data.

In a qualitative study, Backus and Akıncı (2005) analyzed the use of SPPs in the speech of twelve to fourteen-year-old and sixteen to twenty-year-old second generation speakers living in France. The authors found pragmatically redundant uses of overt pronouns in the informants' speeches. For example:

(10) Tamam sen şimdi diyorsun ki **ben** Fransiz kartımı çıkartığım koyduğum zaman karşıdaki insan zaten benim yabancı olduğumu görüyor diyorsun.

'OK, you're saying that when **I** whip out my French ID, the guy across from me will still see that I'm a foreigner.'

(Backus & Akıncı, 2005:158)

The overt pronoun *ben* 'I' in (10) is used grammatically, but according to the authors, the overuse of the overt pronoun indicates that the Turkish spoken in France may be undergoing structural change. However, since the study lacked monolingual data, the authors cannot be sure whether this change is due to the influence of French or general attrition of grammatical knowledge.

Pfaff (1991) examined the use of first and second person pronouns in the speech of several eight-to nine-year-old Turkish-German bilingual children living in Berlin,

Germany. The children are drawn from the EKMAUS study, which investigated the speech of 80 Turkish/German bilingual children between the ages of 5 and 12. Like Backus and Akıncı (2005), Pfaff found discourse-redundant overt subject pronouns in the children's speech. One of her examples is repeated here as (11).

(11) Peter, **sen** ban-a top-u ver-ir mi-sin?
Peter, you I-DAT ball-ACC give-AOR-2SG.
'Peter, will you give me the ball?'

(Pfaff, 1991:120)

The 2sg overt SPP in (11) is considered as redundant. According to Pfaff, utterances with such overt SPPs, which are inappropriate in their discourse context, are still grammatical. The author mentioned the need for a comparison of more discourse contexts for Turkish monolinguals with bilingual children to determine whether or not more frequent use of overt pronouns is due to the transfer of the non-pro-drop parameter setting from German or language attrition.

Pragmatically inappropriate, but grammatically used overt pronouns were also found in a recent quantitative study conducted by Haznedar (2007). Haznedar investigated the use of overt and null SPPs in the speech of a Turkish monolingual (from age 1;7-3;10) and a Turkish-English bilingual child (from age 2;4-4;3) living in Turkey. The results indicated that the monolingual child used subjects around 33% whereas the bilingual child more than 60%, suggesting that the bilingual child used overt subjects in the majority of his sentences. In addition, in environments where the overt SPP use does not serve any pragmatic purpose, such as contrast, topic shift, etc., the bilingual child used overt SPPs much more frequently than the monolingual child, especially when

referring to the 1sg subject, *ben* 'I'. An example taken from the speech of the bilingual child is given below. He uses an overt pronoun to respond to a question addressed by the investigator who arrives just before the child finishes his breakfast.

(12) Investigator: ekmek ye-di-n mi?
'Did you have some bread?'

Ali-John: **ben** ekmek ye-di-m
I bread eat-PAST-1SG
'I ate bread.'

(Haznedar, 2007:130)

While the monolingual child provided overt pronouns for the first and second subjects two percent of the time in this type of environment, this rate was 21 percent for the bilingual child. The author concluded that the high frequency of the use of overt pronouns observed in the bilingual child's speech could be attributed to the cross-linguistic influence of English.

We should bear in mind that most of the studies that examined SPP use in child language acquisition found 1sg and 2sg subject pronouns to be more dominantly used than the other pronouns. It has been argued that the frequent use of 1sg and 2sg pronouns during the very early stages of language production results from the egocentric nature of children's conversations (Halliday & Hasan, 1976; Piaget, 1959).

3. Research conducted in other languages

There have been many variationist studies that examined the use of SPPs in languages spoken especially in Latin America and in the Caribbean as well as in immigrant communities in the U.S. In these studies, linguistic variables such as continuity of reference, person and number of the pronoun and of the verb, verb tense and social

variables such as age of the speaker, gender, social class and education have been the most commonly investigated variables. Studies have also analyzed the hypothesis of the influence of English on the increased use of overt SPPs in contact languages.

With respect to continuity of reference, as is the case in Turkish, speakers use overt pronouns more frequently than null pronouns when there is a change in referent from the immediately preceding clause. The influence of continuity of reference on the choice of an overt or a null pronoun was investigated in the speech of ten speakers from Madrid, Spain and ten from San Juan, Puerto Rico in Cameron (1995). The results of the data analysis indicated that the frequency of overt pronoun usage in both dialects was very similar with respect to the context in which overt pronouns were used. The author reported that the continuity of reference constraint played an important role in these contexts. In addition, when there was a switch in reference, the speakers from San Juan used overt pronouns at a higher rate than those from Madrid.

Another variationist study of the influence of continuity of reference on the use of SPPs is Flores-Ferrán (2004). She analyzed the Spanish spoken by Puerto Ricans living in NYC. Her study involved three groups: recent arrivals, established residents, and native-born NYC Puerto Ricans. She observed that speakers, when considered together, favored null SPPs when there was not a switch in reference. They favored overt pronouns when there was a switch. Also, the NYC group used more overt SPPs than the monolinguals in Puerto Rico when there was no switch. However, this difference was not significant.

Further, Silva-Corvalán (1982) analyzed the presence versus the absence of subjects in their preverbal and postverbal placement in finite clauses in the speech of Mexican-Americans who have lived in California for 10 years or more. Her results indicated that same versus switch reference played an important role on the position of

subjects. Since same reference subjects constituted old information, they took place in the preverbal area and favored null pronouns. Overt pronouns constituted new information, thus they took place post verbally. Overt pronouns were also favored in contrastive contexts and with ambiguous verbal forms.

Continuity of reference and SPP variation has also been one of the central topics in Brazilian Portuguese (Lira, 1982). In a quantitative study, Lira examined the use of SPPs in the speech of thirty people from Rio de Janeiro, Brazil. Lira found that the speakers were more likely to use overt pronouns when there was a switch in reference and when subjects had second person references.

The influence of continuity of reference on SPP variation was also explored in Egyptian Arabic in Eid (1983). Eid analyzed subject pronouns in everyday speech and in a variety of syntactic structures. When the relative and subordinate clauses were considered, it was found that subject pronouns in Arabic were not optionally used; rather, the existence of an overt pronoun indicated that there was a switch in reference.

In Persian, Haeri (1989) investigated the distribution of overt and null subjects in the speech of four adults and six children who were at least six years old. The data consisted of narratives. The children described six pictures that were presented to them. The adults were interviewed about certain topics. On the basis of child and adult data, the authors reported that the distribution of overt and null subjects was influenced by the form used in the preceding clause. In the first sentence, the children used a full NP to refer to the thematic subject, which is the main actor in the narratives. In the second sentence, they used a null pronoun to refer to the same subject. An overt pronoun was used when the subject referent was different from the subject referent of the preceding main clause.

In American Sign Language, Bayley et al. (2001) examined SPP variation in the discourse narratives of nineteen informants who participated from seven different areas of the United States. The results were statistically significant with respect to pronoun usage versus continuity of reference. Overt pronouns occurred more often when there was a switch in reference.

Finally, with respect to Japanese, Clancy (1980) analyzed the reasons for referential choice in a study of English and Japanese discourse narratives. She investigated whether speakers used NPs, overt pronouns or null subjects to refer to a particular entity. Clancy focused only on third person human referents. The study indicated that nominals were used when there was a switch in reference. Null subjects were more frequently used when subject was the same.

Researchers have also attempted to analyze the influence of the person and number of the pronoun on the variable use of SPPs. Silva-Corvalán (1994) analyzed audio-recorded Spanish conversations collected from Mexican immigrants living in the U.S. Of the several independent grammatical variables that Silva-Corvalán (1994) examined, she found that the 1sg *yo* 'I' had a much higher percentage of expression than any other pronoun among Mexican-born speakers who immigrated to the U.S. after the age of eleven. The author attributed this result to the speakers' need to keep themselves overtly present in conversations. Thus she concluded that the expression of *yo* 'I' has a pragmatic function. Similar results were obtained in Flores-Ferrán (2004) where the use of overt SPPs with *yo* was favored more than the other pronouns by the Puerto Rican residents in NYC.

With regard to the person and number of the verb, in a recent large-scale study, Otheguy et al. (2007) analyzed the variable use of SPPs in finite clauses in the Spanish

spoken among Caribbean and Latin American immigrants who have recently arrived in NYC and who were born in NYC. Of the ten linguistic variables that the authors analyzed, the variable Person turned out to be the strongest predictor of the probability of appearance of an overt SPP in the speech of all Caribbeans and Latin Americans.

Another important variable that has engaged the attention of researchers is discourse connectedness, that is, the relationship between clauses or what the Turkish linguists call 'topic of discourse' (Enç, 1986). Linguists suggested that the use of null or overt SPPs depends on how closely connected clauses are. It was found that null pronouns are used in contexts where the topic of the discourse remains the same, even if there is a switch in reference (Paredes Silva, 1993). In a study of written Brazilian Portuguese, Paredes Silva (1993) examined the patterning of null and overt SPPs in the written narratives of 70 people born in Rio de Janeiro, Brazil. The participants ranged in age between 17-24 and 25-35. All of them had graduated from high school. Some had university degrees and some were university students. The author studied the influence of three variables: discourse connectedness, ambiguity, and emphasis. Only subjects with a specific referent were analyzed.

As for discourse connectedness, the results demonstrated that when there was a close relationship between the second clause and the preceding one, overt pronouns tended not to be used. In other words, there was a decrease in null pronominal use as discourse connectedness became weaker. Overt pronouns were mostly used to avoid ambiguity with another referent and clarify the correct referent.

Discourse connectedness was also examined by Bayley and Pease-Álvarez (1997) who analyzed oral and written Spanish narratives collected from Mexican pre-adolescents living in northern California. The results indicated that, as was the case in Paredes Silva

(1993), overt pronouns were not preferred when there was a close relationship between the second clause and the preceding one. Overt pronouns were favored with imperfect, conditional and subjunctive verb forms.

In Chinese linguistics, research into the occurrence of third person null versus overt pronouns in written discourse has demonstrated that null pronouns were so widely used that they have become the most preferred form and that the usage of overt pronouns was highly dependent upon pragmatic knowledge (Li & Thompson, 1979). An important factor that predicted the usage of third person overt pronouns was the degree to which two clauses can be conjoined. Pronouns were more likely to be expressed overtly in the second clause if there was a lack of conjoinability between the second clause and the preceding clause.

Linguists also explored the relationship between verb semantics and SPP variation. Silva-Corvalán (1994) found a positive correlation between subject expression and estimative verbs such as *creer* 'to believe,' and *pensar* 'to think.' Both 1st and 2nd generation speakers living in the U.S. were more likely to use overt subjects with these estimative verbs because these verbs were associated with contrastive contexts. In addition, verbs like *gustar* 'to like' also favored the use of overt subjects.

Finally, some linguists (Hochberg, 1986; Eid, 1983) concentrated on morphological ambiguity and the use of SPPs. From a functional perspective, Hochberg (1986) analyzed SPP usage and /s/ deletion from 2sg verbs in the speech of ten Puerto Rican working-class women who moved to the U.S. when they were teenagers. Her analysis revealed that there was a relationship between the deletion of /s/ from 2sg verb forms and the use of SPPs, that is, SPPs were more likely to be used when the /s/ was deleted from 2sg verb forms to avoid ambiguity because in Spanish, deletion of –s from

verbs results in the loss of semantic information it carries. The speakers in Eid (1983) also showed similar patterns in the use of overt pronouns to avoid ambiguity.

Turning away from linguistic variables, social variables such as contact with English served as an important variable that conditioned the use of an overt or a null SPP. In some studies though, the effect of English did not turn out to be significant. For example, Otheguy and Lapidus (2005) examined the use of the SPP *ellos* in the Spanish spoken in NYC. It was found that instead of the null pronouns, which are used to refer to nonspecific 3pl in non-contact Spanish, overt pronouns are commonly used in the Spanish spoken in NYC. The authors found significant differences between the informants who were born in NYC and two other groups; informants who arrived at age thirteen and those who arrived at seventeen or older. The informants who were born in the City used the non-specific overt *ellos* more frequently than the other two groups. However, since these kinds of usages also exist in the Spanish spoken in Latin America, the authors concluded that English did not have a direct influence on the Spanish spoken in NYC.

Other major studies provided further proof that contact with English did not play an important role on the use of SPPs in the Spanish spoken by different communities in the U.S. (Silva Corvalán, 1994; Flores-Ferrán, 2004; Bayley & Pease-Álvarez, 1997). However, when Otheguy et al. (2007) compared Caribbean and Latin American immigrants who have recently arrived in NYC with those who were born in NYC on the variable use of SPPs in finite clauses in Spanish, they found that contact with English played an important role on the increased usage of overt SPPs in the speech of the Spanish speakers who were born in NYC. There were positive correlations between overt SPP usage and number of years spent in NYC as well as English skills. There were negative correlations between overt pronoun usage and the age at which speakers arrived

in the City as well as the amount of Spanish spoken with interlocutors and in different domains. The authors attributed the increased use of overt SPPs observed in the speech of the New York born speakers to inadequate use of Spanish and constant exposure to English.

Another variable that has been widely studied is social class, which has become a highly important variable in variationist sociolinguistics since Labov's New York and Martha's Vineyard studies. Labov's (1963) study involved the English dialects on the island of Martha's Vineyard, Massachusetts. In his study, Labov observed that the first elements of the dipthongs in words like *right* and *house* are centralized as /əy/ and /əw/ rather than /ay/ and /aw/ particularly in the speech of middle-aged fishermen. He also found that the centralization index for the other groups, especially for the youngest group, the group between the ages 14 and 30 was very low. The significance of this study is that there is a social meaning attached to the phonetic differences observed between the fishermen and the young people. Centralization was found mostly in the speech of the fishermen as a result of as Labov (1963) points out, a reaction to threats to island identity such as the need to leave the island to make a living. Note that linguistic variants spread quickly among speakers in low-status groups because social networks in these communities are stronger than the others (Labov, 1972a).

Another social variable that was found to constrain the use of overt or null SPP is age of the speaker. The relationship between the use of overt SPPs and age was studied in Avila-Jiménez (1996). Her study was based on a corpus of naturally-occurring speech collected from fifty-nine Puerto Rican speakers from different socio-economic backgrounds. The results revealed that age had a significant influence on the use of overt vs. null pronouns. Whereas sixty percent of the speakers in the study preferred to use null

pronouns, speakers who were younger than fifty used overt pronouns at a significantly higher rate. Overt pronouns were dominant in the speech of college level speakers and professionals.

Other social variables, such as gender was also an important social variable affecting the use of overt versus null subjects. In many studies, sociolinguists have found a significant difference between males and females with respect to the use of certain variables. Labov (1972) observed that for some sociolinguistic variables, women use more standard forms than men and that prestige is an important factor that motivates such linguistic behavior.

Based on a study of the influence of gender differences on language variation, Eckert (1989) argued that power differences and social interaction in groups were two important factors affecting the use of new forms by women. In a high school in Detroit, she analyzed the linguistic behavior of girls and boys in two groups: 'Jocks' and 'burnouts'. 'Jocks' who were from middle-class regularly participated in school activities whereas 'burnouts' represented working-class and were against school activities. Eckert reported that Jocks girls used more standard forms than Jocks boys in order to achieve status in the school. As for the burnout group, the girls were more likely to use newer forms than the boys.

With regard to SPP variation, Bayley and Pease-Álvarez (1997) found that females preferred more overt pronouns than males, in their study of Mexican descent Spanish-speaking children between the ages of 8 and 12 years.

CHAPTER 3

THE ENVELOPE OF VARIATION

1. Introduction

This chapter defines the envelope of variation for the present study and lists the items that are included in and excluded from the study. Before defining the envelope of variation, it is important to mention the term, *the sociolinguistic variable,* which was proposed by Labov in the late 1960s and defined as "the option of saying the same thing in several different ways. The alternatives, which are used to say the same thing, are called *variants* (Fasold, 1990).

When the variants, equivalent in meaning and function alternate, researchers determine the possible environments or contexts in discourse in which alternation occurs. These environments are called *the envelope of variability* or what other scholars have called *circumscribing the variable context* (Poplack & Tagliamonte, 1989:60). I will use the term *the envelope of variation* throughout this dissertation. After researchers determine the environments where variants alternate, they measure probabilities for either variant for each environment. They do this by comparing the number of actual occurrences of a particular variant to its potential number of occurrences. For example, initial /h/ in words such as *house* is generally regarded as having two variants [h] and zero. The number of [h] occurrences and the number of zero occurrences are compared with the total number of possible occurrences of /h/ in the data. The categorical environments, where the variants never alternate with one another have to be excluded from the analysis because variation does not exist in these environments. It is important to

explicitly describe the items that are to be included and excluded so that the study can be accurately replicated in future studies.

Many scholars referred to the envelope of variation (Cameron, 1992; Lira, 1982; Silva-Corvalán, 1982; Bayley et al., 2001; Flores-Ferrán, 2004) and described the variants of a linguistic variable in terms of where it is possible or impossible for the variable to occur. In a somewhat different approach, Otheguy et al. (2007) place inside the envelope of variation those environments where it is practical to regard the use of both variants as possible, while placing outside the envelope those environments where variability is so low that it makes practical sense to regard the occurrence of one or the other variant as obligatory.

In the present study, the envelope of variation included all finite verbs; non-finite verbs in adverbial subordinate clauses; nouns, adjectives, words expressing necessity, and some wh-phrases that function as predicates. It also included locative existential expressions that can co-occur with either a null or an overt pronoun in the nominative. In addition to the nominative personal pronouns, demonstrative pronouns such as *bu* 'this,' *şu* 'that,' *bunlar* and 'these,' *şunlar* 'those' that make personal references were also included in the study. Environments in which overt or null pronouns were found with such high frequency that they could reasonably be regarded as being obligatorily expressed or obligatorily absent were excluded.

2. Inclusions

- *Finite verbs*

In (1), the finite verb *alıyorum* 'buying' can occur with both overt and null pronouns. Thus such verbs were included in the study.

(1) **Ben** / **Ø** gazete al-ıyor-um.
I newspaper buy-IMPF-1SG.
'I'm buying a newspaper.'

- *Verbs in causative, passive, reflexive and reciprocal constructions*

The verbs in causative, passive, reflexive and reciprocal constructions were also included in the study. The subjects of these verbs can either be an overt or a null pronoun.

(2) Her gűn öğrenci-ler-i **onlar** / **Ø** çalış-tır-ıyor-lar.
Every day students-PL-ACC they work-CAUS-IMPF-3PL.
'They have the students work every day.'

In addition, verbs formed with more than one voice suffix (causative + passive; reciprocal + causative; reciprocal + causative + passive) can also occur with either overt or null pronouns. Therefore, they were included in the study. However, reciprocal + passive and reflexive + passive constructions are impersonal passive constructions and thus they cannot occur with either overt or null pronouns. These verbs were excluded from the analysis. (Impersonal passive constructions will be discussed in section 3).

- *Verbs whose subject is a nonspecific 2nd sg or 3rd pl*

Verbs marked with 2sg or 3pl making nonspecific reference were also included in the study. The subject of these verbs can either be an overt or a null pronoun.

(3a) **Sen / Ø** bir bak-ıyor-sun film bit-miş.
You once look-IMPF-2SG film end-PF.
'You realize that the movie has ended.'

(3b) **Onlar / Ø** bu bardak-lar-ı kutu-lar-a koy-uyor-lar.
They this glass-PL-ACC box-PL-DAT put-IMPF-3PL.
'They (nonspecific) put these glasses in boxes.'

- *Verbs in conjoined phrases*

When a set of conjoined phrases or clauses carry the same suffixes, these suffixes may be omitted in all of the conjuncts except the last one (Göksel & Kerslake, 2005). In (4), the copular and the person markers are omitted in the first verb, *okur*, but they are marked on the last verb, *tartışırdık*. Since the subject of *okur* can either be a 1pl overt or a null pronoun, this verb was included in the analysis.

(4) **Biz / Ø** gün-de üç dört saat gazete oku-r, çeşitli
We day-LOC three four hour newspaper read-AOR various

konular-da saat-ler-ce arkadaş-lar-la tartış-ır-dı-k. (1TR)
topic-PL-LOC hour-PL friend-PL-CONJ discuss-AOR-P.COP-1PL.

'We would read the newspaper for 3, 4 hours a day and discuss various topics with friends for a few hours.'

- *Verbs in adverbial clauses*

Also included in the analysis were verbs formed with *–DIğI de/zaman/sirada* 'when,' *-DIğI için* 'as,' *–DIğI göre* 'since,' *–DIğI oranda* 'the more… the more,' *–DIğI kadar* 'as

long as,' 'enough for,' *–DIğI beri* 'since,' *–DIğI sűrece* 'throughout the time,' *DIğI gibi* 'the way,' *-(y)AcAğI zaman/sirada* 'when,' *-(y)AcAğI için* 'as,' *-(y)AcAğI göre* 'since,' *-(y)AcAğI oranda* 'the…the,' *-(y)AcAğI kadar* 'enough…to,' *-(y)AcAğIndAn (dolayı/ötűrű)* 'on the account of the fact that'. Non-finite verbs formed with these suffixes can occur with either overt or null subject pronouns. Therefore, they were included in the analysis. In (5), both verbs were included in the study.

(5) **Sen** / **Ø** İstanbul-da yaşa-dığ-ın zaman **ben** / **Ø** sen-i gőr-mey-i
You İstanbul-LOC live-CV-2SG.POSS time I you-ACC see-INF-ACC

iste-miş-tim.
want-PF-PC-1SG

'When you lived in İstanbul, I had wanted to see you.'

This study examines the alternation between an overt pronoun and what in the generative literature is known as pro or 'little pro.' This pro, whose occurrence is determined by the presence of agreement, occurs as the subject of root sentences, object clauses, and relative clauses. It refers to an antecedent in the higher clause and satisfies the requirements of the condition B of the Binding Theory, which requires a pronoun to be free in its governing category and be able to refer to its antecedent in the higher clause (Őzsoy, 1987). In (5a), pro alternates with an overt pronoun in both clauses. In addition to pro, generative grammars include the construct PRO or 'big PRO' as an additional empty pronominal element. PRO occupies the subject position of the infinitival clause where the infinitival verb does not have any kind of AGR morphology. In these clauses, pronouns are obligatorily deleted. Since there is no surface alternation between an overt pronoun and a null, these elements fall outside the scope of this variationist analysis and were not included in my study.

- *'Var' and 'yok' in locative existential sentences*

In (6a) and (6b), the two expressions, *var* 'there is' and *yok* 'there is not' that occur in locative existential sentences were included in the analysis because overt and null pronouns can be variably used with these expressions.

(6a) Ev-de **ben** / **Ø** yok-tu-m.
Home-LOC I non.existent-P.COP-1SG.
'I wasn't home.'

(6b) **Sen** / **Ø** orada var-dı-n.
You there exist-P.COP-2SG.
'You were there.'

- *Nouns, adjectives and postpositional phrases functioning as predicates*

Adjectives, nouns and postpositional phrases that function as predicates were included in the study because these predicates can occur with either an overt or a null subject pronoun.

(7a) Geçen hafta **ben** / **Ø** hasta-ydı-m.
Last week I sick-P.COP-1SG.
'I was sick last week.'

(7b) **O** / **Ø** mühendis-miş.
He/she engineer-EV.COP-3SG.
'He/she was an engineer.'

(7c) **O** / **Ø** san-a gőre-ymiş.
He/she you-DAT for-EV.COP-3SG.
'He/she was for you.'

- *Wh-phrases: 'kim', 'neci', 'nerede', 'hangisi', 'nasıl'*

Wh- phrases such as *ne* 'what,' when combined with the suffix –ci, *kim* 'who,' *nerede* where,' *hangisi* 'which,' *nasıl* 'how' that function as complements, were included in the study. They can occur with both overt and null pronouns.

(8) **Siz** / **Ø** kim-siniz?
You who-2SG.
'Who are you?'

- *Verbs occurring in finite subordinate and superordinate clauses that are connected via 'diye,' 'ki' and 'DI mI'*

In Turkish, finite subordinate clauses can be connected to superordinate clauses by means of the subordinators *diye, ki* 'that' and *DI mI* 'as soon as.' In (9), the finite verbs in both clauses can occur with either an overt or a null pronoun. Thus both verbs were included in the analysis.

(9) [**O** / **Ø** şarkı söyle-yecek] diye **ben** / **Ø** bil-iyor-du-m.
He/she song sing-FUT-3SG that I know-IMPF-P.COP-1SG.
'I thought that she/he would sing.'

- *Verbs whose subject is animate*

Verbs whose subject is animate were included in the study.

(10) **O** / **Ø** iş-e git-ti.
He work-DAT go-PF-3SG.
'He went to work.'

- *Verbs whose subject is inanimate*

Verbs whose subject is inanimate were also included in the study. Verbs in these contexts can either occur with an overt or a null pronoun.

(11) **O** / **Ø** düş-tü.
It fall-PF-3SG.
'It fell.'

- *Verbs in contrastive contexts*

In contrast to some scholars (Cameron, 1992; Silva- Corvalán, 1982) who did not count subjects in contrastive environments as sites of variation between an overt and null subject, I included in the study, in keeping with Amaral and Schwenter (2005), overt SPPs occurring in contrastive environments. Amaral and Schwenter, in their paper about Spanish, argued that overt SPPs are not obligatory in contrastive contexts and that they do not have a contrastive function. They observed that if, for example, one of the clauses has a topic-introducing adverbial, speaker-oriented adverb or locative adverb, overt and null pronouns can be variably used in that particular clause because these adverbials are compatible with the person and number features of the referent of the subject. Thus the use of an overt or a null subject pronoun is fully grammatical in these types of clauses. This is demonstrated in the following example taken from Amaral and Schwenter (2005). In this example, an overt pronoun does not have to occur since there is a topic-introducing adverbial.

(12a) Bueno, aquí (**Nosotros / Ø**) lo tenemos el viernes.
'OK, in our case we have it on Friday.'

(Amaral & Schwenter, 2005:118)

In Turkish, overt and null pronouns can also be variably used if a clause has a topic-introducing adverbial such as *a gelince* 'as for,' as in (12b). Verbs in these types of clauses were included in the analysis.

(12b) Ahmet-e gelince **O / Ø** sen-i çok beğen-ir.
Ahmet-DAT coming he you-ACC very admire-AOR.
'As for Ahmet, he admires you very much.'

In (12c), the verbs in both clauses were included. Variation between null and overt could not only take place in the first clause but it could also take place in the second clause since there is a topic-introducing adverbial.

(12c) **Ben / Ø** iş-e git-ti-m ama **san-a gelince / Ø / sen** git-me-di-n.
I work-DAT go-PF-1SG but you-DAT coming you go-NEG-PF-2SG.
'I went to work but as for you, you did not go to work.'

In environments where one of the previously mentioned adverbials does not occur, overt SPPs are more frequently used than nulls. Consider (12d):

(12d) **Ben / Ø** çalış-tı-m ama **sen** çalış-ma-dı-n.
I work-PF-1SG but you work-NEG-PF-2SG.
'I worked but you did not work.'

- *Verbs occurring with ile* ‘with’ *in comitative constructions*

In Turkish, the clitic *–(y)la / ile* ‘with’ can be used with the comitative meanings. Verbs occurring with *ile* that has this meaning were included in the study. SPPs, which can occur as either the first or second item, can be variably used in these contexts. This can be seen in (13a) where the SPP *ben* occurs as the first item.

(13a) **Ben** / **Ø** Ayşe ile ev-e gid-iyor-um.
I Ayşe CONJ home-DAT go-IMPF-1SG.

[I with Ayşe going home.]
‘I am going home with Ayşe.’

In contexts where SPPs occur as the second item, verbs have plural marking. This can be seen in (13b) where the verb is marked with 1pl. In this case, deletion of the overt SPP may result in ambiguity because the null subject referent may either be 1sg or 1pl. Whether the referent is 1sg or 1pl is inferred from the context.

(13b) Ayşe ile **ben** / **Ø** ev-e gid-iyor-uz.
Ayşe CONJ I home-DAT go-IMPF-1PL.

‘*I* am going home with Ayşe.’
‘*We* are going home with Ayşe.’

- *Verbs occurring with ve ‘and,’ veya / ya da ‘or’ in clausal coordination*

Ve ‘and,’ and *veya / ya da* ‘or’ conjoin all types of phrases and clauses. When they are used to conjoin clauses, overt SPPs can be variably used. In (14), *ve* ‘and’ conjoins two verbs, *gitti* and *seyretti*. Since the subject of the verb in the first clause is *babam* ‘my dad,’ the verb was excluded from the study. The verb in the second clause was included in the study although it is more likely to occur with the null pronoun.

(14) Baba-m sinema-(y)a git-ti ve **O** / **Ø** gűzel bir film
Dad-1SG.POSS movie theatre-DAT go-PF-3SG CONJ he nice a film

seyret-ti.
watch-PF-3SG.

'My dad went to the movie theatre and watched a nice film.'

Note that although these conjunctions allow for both nulls and overts in clausal coordination, they only allow overts in NP coordination. In other words, when they join noun phrases, SPPs have to be expressed overtly and thus they are obligatory. Verbs in such contexts were excluded from the study. This will be discussed in section 3.

- *Verbs whose subject occurs with focus particles in clausal coordination*

Focus particles such as *da….. da, hem….hem, gerek… (ve) gerek (se de)* 'both… and;' *ister…ister, ha…ha, ama… ama* 'whether… or' can not only conjoin phrases, but they can also join clauses. When they join clauses, overt SPPs and nulls can alternate with each other. This is shown in (15) where the verbs *gittim* in both clauses were included in the study. However, it will be shown later in section 3 that these particles, like the coordinating conjunctions, can also be used to conjoin noun phrases. In those cases, overt SPPs are obligatory and thus they were excluded from the analysis.

(15) Hem **ben** / **Ø** iş-e git-ti-m, hem (de) **ben** / **Ø** pazar-a git-ti-m.
Both I work-DAT go-PF-1SG as well as (and) I bazaar-DAT go-PF-1SG.
'I went to work and to the bazaar.'

- *Verbs in small clauses*

Verbs in small clauses were also included in the study because their subjects can be either an overt or a null pronoun. In (16), both verbs were included in the analysis because the subjects of both the main and the small clause can either be an overt or a null pronoun.

(16) **Biz** / **Ø** [**sen** / **Ø** ye-din] zannet-ti-k.
We you eat-PF-2SG think-PF-1PL.
'We thought you had eaten.'

- *Verbs whose subject occurs in contexts where there is a switch from one discourse topic to another*

As mentioned in section 2.1 of chapter 2, overt subject pronouns are used to establish or change the topic of the discourse (Enç, 1986). Enç refers to 'topic' as what the discourse is about. Consider (17b):

(17a) Ali yarın Ankara'ya gid-iyor.
Ali tomorrow Ankara-DAT go-IMPF.
'Ali is going to Ankara tomorrow.'

(17b) **O** bugűn-ler-de çok dalgın.
He this day-PL-LOC very absent-minded.
'He's been very absent-minded lately.'

(Enç, 1986:199)

In (17b), the overt SPP *O* 'he' is used to change the topic from Ali's trip to Ankara to his recent absent-minded behavior. Also consider (17c), where the speaker uses a null pronoun to answer a question such as 'why are you putting your coat on?'

(17c) **Ø** çarşı-(y)a gid-iyor-um.
(I) market-DAT go-IMPF-1SG.
'I am going to the market.'

(Enç, 1986:196)

Here, the use of a null pronoun indicates that the speaker is either commenting on a topic or answering a question about a topic that has been previously mentioned in the discourse. As we have also discussed in chapter 2, the use of an overt pronoun is not the only way to introduce another topic or change the topic of the conversation. In the corpus of the present Turkish study, there were cases of null pronoun use when the speakers had the choice of using an overt SPP to introduce a new topic or change the topic of the discourse. The example presented in section 2.1 of chapter 2 is also repeated here.

(17d) …**bu** 2, 3 sene İzmir-e gel-miş, yaşa-mış. Ondan sonracığıma,
…[this] he 2, 3 year İzmir-DAT come-PF-3SG live-PF-3SG. Then

ortaokul dönem-in-de iyi-(y)di-m. (8TR)
secondary school term-3SG POSS-LOC good-PF-1SG.

'This[he] had come and lived in İzmir for 2, 3 years. Then, I was good in secondary school [I was a good student in secondary school].'

In (17d), the informant used a null pronoun to switch the topic from his father's living in İzmir to his being a good student in secondary school. The informant could have also used an overt pronoun instead of a null. Therefore, the adjectival predicate was included in the study.

- *Verbs whose subject occurs in construction with a noun*

Verbs whose subject occurs in construction with a noun to refer to a group of people were included in the analysis. In these contexts, overt SPPs can be variably used.

(18) **Biz** / **Ø** Türk-ler futbol-u çok sev-er-iz.
We Turk-PL soccer-ACC very love-AOR-1PL.
'We Turks love soccer.'

- *Verbs whose subject co-occurs with kendi 'self'*

Verbs whose subject co-occurs with *kendi* 'self' were included in the study. *Kendi* agrees with the person and number of the pronouns that it co-occurs with. These emphatic forms are possessively inflected.

(19a) **Ben** / **Ø** kendi-m git-ti-m.
I self-1SG.POSS go-PF-1SG.
'I myself went.'

In (19b) and (19c), *kendim* 'myself' functions like the possessed head noun of phrases such as:

(19b) Ben-im kalem-im dűş-tű.
I -GEN pencil-1SG.POSS fall-PF-3SG.
'My pencil fell.'

(19c) Sen-in kalem-in dűş-tű.
You-GEN kalem-2SG.POSS fall-PF-3SG.
'Your pencil fell.'

In the examples above, the possessor is marked with the genitive case suffix, while the possessed head noun of the phrase has agreement suffixes that alternate for the person and number features of the possessor. *Kalemim,* which is inflected for the person of the possessor, is third person and the verb *dűş-tű* is inflected for the 3sg because *kalem* is third person. So it does not matter if *kalem* is possessed by a first person, *benim kalemim* is still a third person noun phrase. The same case can be made for *Kendi* in that it is also a noun phrase in the third person and the fact that the verb in (19a) takes first person inflection indicates that *kendim* is not itself a subject because the first person morphology

that might appear to be getting this agreement in fact is the result of agreeing with the SPP *ben*. Therefore, it can be concluded that *kendim* functions as a modifier of either an overt or a null subject. In (19a), in both cases *kendim* modifies both the overt and the null pronoun on the basis of the fact that the first person inflection is not coming from *kendim*, it is coming from *ben*.

3. Exclusions

There are certain environments, however, within which either overt or null pronouns are best regarded as obligatorily. Since there was very little or no variation in these environments, the verbs were excluded from the analysis.

- *Verbs used in meteorological and seasonal expressions*

Verbs used in meteorological expressions were excluded from the analysis because the subject of these verbs cannot be an overt subject pronoun.

(1) **Yağmur** / **Ø** yağ-ıyor.
Rain rain-IMPF.
'It is raining.'

- *Verbs in impersonal and double passive constructions*

Impersonal and double passive constructions were not included in the study. Neither overt nor null pronouns can occur with these constructions.

(2) İzmir-e araba-(y)la gid-il-ecek.
İzmir-DAT car-CONJ go-PASS-FUT.
'[People] will go to İzmir by car.'

- *Subject pronouns that do not occur with any verb*

In some cases, pronouns in the nominative do not occur with any verb. The pronouns that are not in construction with a verb were excluded from the study.

(3) Bir gűn **ben**… annem bir market-te çalış-ıyor-du. **Ben** ora-(y)a
One day I… my mother a market-LOC work-IMPF-P.COP. I there-DAT
git-ti-m.
go-PF-1SG.
'One day, I… My mother worked at a supermarket. I went there.'

In (3), the first SPP *ben* 'I' was excluded from the study because it is not in construction with any verb.

- *Verbs whose subject is a pronominal quantifier*

Verbs whose subjects are pronominal quantifiers were excluded from the study. Some of the pronominal quantifiers are *herkes* 'everyone,' *her şey* 'everything,' *bir şey* 'anything,' *kimse* 'no one,' *hiçbiri*(*si*) 'none of (them),' and *hiçbir şey* 'nothing.'

(4) **Herkes** dondurma sev-er.
Everyone ice cream love-AOR.
'Everyone loves ice cream.'

- *Verbs that are non-finite: infinitives, participles and gerunds*

Verbs whose subject is an infinitive, gerund, sentence or an overt lexical NP were excluded from the study.

(5a) Sokağ-a çık-mak gűzel-dir.
Street-DAT go out-INF nice-GM.
'Going out is nice.'

(5b) Deniz ve Ahmet Londra-(y)a git-ti-ler.
Deniz CONJ Ahmet London-DAT go-PF-3PL.
'Deniz and Ozan went to London.'

- *Verbs whose subject is a demonstrative, locative, and directional non-personal pronoun*

This study also excluded verbs whose subject is a demonstrative, locative, and directional non-personal pronoun. However, as mentioned in section 2, verbs whose subjects are demonstrative pronouns making personal references were included in the analysis.

(6) En gűzel elbise bu.
Most nice dress this.
'This is the nicest dress.'

- *'Var' and 'yok' in possessive existential sentences*

The two expressions, *var* 'there is' and *yok* 'there is not' that occur in possessive existential sentences were excluded from the analysis because the subject pronouns of these verbs have to receive genitive case.

(7) Ben-im **yok**.
I-GEN non.existent.
'I don't have [it].'

- *Verbs in relative clauses*

Verbs formed with participle suffixes in relative clauses can occur either with an overt or a null pronoun, however, overt pronouns have to be marked with the genitive case. Therefore, such verbs were excluded from the analysis.

(8) **Ben**-im / **Ø** dűn konuş-tuğ-um adam.
I-GEN yesterday speak-PART-1SG man.
'The man whom I talked to yesterday.'

- *Verbs in headless relative clauses*

Verbs with participle suffixes in headless relative clauses where the head noun is omitted from the sentence were excluded from the analysis. These verbs cannot occur with overt and null pronouns.

(9) Műzik dinle-me-(y)en-ler.
Music listen-NEG-PART-PL.
'Those who don't listen to music.'

- *Verbs whose first or second person subject is one of a set of people*

Verbs whose first or second person subject is one of a set of people involved in some action were excluded from the analysis. These are contexts where each subject has to be addressed separately. In (10), the verb was excluded.

(10) **Sen**, Ahmet ve Meral yemek yap-acak-sınız.
You, Ahmet and Meral food make-FUT-2PL.
'You, Ahmet and Meral will cook.'

Realize that these kinds of contexts also allow for two nouns to be in an appositional relationship. The sentence in (10) can have an appositional reading without the 'coma', which occurs after the SPP *sen*. When *sen* occurs in construction with the noun *Ahmet*, *sen* can be variably used. This being the case, the verb *yapacaksınız* could be included.

- *Verbs whose subjects occur with ve 'and' and veya / ya da 'or' in NP coordination*

As mentioned in section 2, *ve*, 'and,' and *veya / ya da* 'or' can not only join clauses but they can also join noun phrases. When they join clauses, overt pronouns can be variably used. Verbs occurring in these clauses were included in the study. However, when these conjunctions join noun phrases, overt pronouns are obligatory and thus they were excluded from the analysis. For example, *ve* in (11) joins two SPPs that have to be expressed overtly.

(11) **Sen** ve **ben** Londra-(y)a git-ti-k.
You CONJ I London-DAT go-PF-1PL.
'You and I went to London.'

- *Verbs whose subject occurs with focus particles*

Verbs whose subjects occur with focus particles such as *da… da, hem…hem, gerek… (ve) gerek (se de)* 'both… and;' *ister…ister, olsun…olsun, ha…ha, ama… ama* 'whether… or' were excluded from the analysis. When NPs occur with these focus particles, they have to be overtly expressed. In (12a), the SPPs that occur with the focus particle *hem…hem* are obligatory and thus the verb *aradik* was excluded from the analysis.

(12a) Hem **ben** hem **O** Deniz-i ara-dı-k.
Both I and he Deniz-ACC call-PF-1PL.
'Both he and I called Deniz.'

- *Verbs whose subject occurs with bile* 'even'

The particle *bile* 'even' can be placed in different positions in a sentence. It not only occurs with an object, as in (13a) but it can also occur in the post-verbal position, as in

(13b). In these contexts, the subjects of verbs can either be an overt or a null pronoun. Therefore, such verbs were included in the analysis.

(13a) **Ben / Ø** kitap bile al-dı-m.
I book even buy-PF-1SG.
'I even bought a book.'

(13b) **Ben / Ø** kitap al-dı-m bile.
I book buy-PF-1SG already.
'I've already bought a book.'

However, when this particle is associated with a subject, then the subject has to be expressed overtly, otherwise, the sentence is ungrammatical. Therefore, verbs whose subject is construction with *bile* were excluded from the analysis.

(13c) **Ben** bile al-dı-m kitap.
I even buy-PF-1SG book.
'Even I bought a book.'

- *Verbs whose subject occurs in response to 'who?'*

Verbs whose subject occurs in response to the wh- question, *kim* 'who' were excluded from the study. Overt pronouns are obligatory in these contexts.

(14) Bu mektup-u kim oku-du.
This letter-ACC who write-PF.
'Who read this letter?'

Ben oku-du-m.
I read-PF-1SG.
'I read [it].'

- *Verbs whose subject occurs with mI*

Yes/no questions are formed by inserting *mI* after a phrase or a sentence. Verbs whose subject occurs with *mI* were excluded from the study because overt pronouns are obligatory in these environments.

(15) **Siz** mi git-ti-niz?
You INT go-PF-2PL.
'Did you go?'

- *Verbs whose subject occurs with dA with topic-shifting and additive functions*

In Turkish, the clitic *dA* is a discourse connective which has a variety of functions i.e. topic-shifting. Verbs whose subject occurs with *dA* with topic-shifting and additive functions were excluded from the study because subjects occurring with *dA* with these functions must be expressed overtly. Null pronouns cannot be used in such environments.

(16) Bu pasta-yı da **ben** al-dı-m.
This cake-ACC as for I buy-PF-1SG.
'As for this cake, I bought it.'

- *The verb de-/söyle- 'say' in direct speech*

The subjects of verbs such as *de-/söyle-* 'say' were excluded from the analysis. In these environments it is grammatical to use one of these verbs after every word in Turkish. However, when repeated, speakers rarely use the overt pronoun with these verbs. Null pronouns are almost obligatory. The variability between the overts and the nulls is so low that these environments were excluded and treated as if they were categorical. In the analysis, the first environment was treated as a variable environment and the subsequent ones categorical.

(17) **O** / **Ø** “Ahmet” de-di “otur-du” de-di.
She Ahmet sit-PF sit-PF say-PF.
‘She said “Ahmet” she said “sat” she said.’

- *Verbs formed with the doubled use of the –sA form in conditional clauses*

Verbs formed with the doubled use of the *–sA* form in conditional clauses were excluded from the analysis. These verbs can only be used with an overt pronoun, which receives focus. In (18), *söylersin* could not be included in the analysis because it cannot occur with a null pronoun.

(18) Bu-nu söyle-se-n söyle-se-n **sen** söyle-r-sin.
This-ACC say-COND-2SG say-COND-2SG you say-AOR-2SG.
‘If anyone can say this, it’ll be you.’

CHAPTER 4

RESEARCH QUESTIONS AND HYPOTHESES

1. Introduction

This study investigates the effects of linguistic variables such as *person and number of the pronoun, person and number ending of the verb, continuity of reference tense/mood/aspect of finite and non-finite verbs*, as well as the effects of social variables such as *place of residence, immigrant generation, gender, age, education* and *social class* on the use of an overt or a null pronoun in the Turkish spoken in Turkey and in NYC (see Appendix A and B for a complete list of linguistic and social variables and their factors examined in this study).

Expressed in statistical terms, the dependent variable here is pronoun use; this dependent variable has two levels, use of a null and use of an overt. The independent variables are the grammatical and socio-demographic variables that are associated with probabilities of occurrence of an overt or a null pronoun. In the variationist literature, the levels of an independent variable are called *factors* or *constraints*. For example, verb tense is an independent variable, while present tense and past tense are constraints within that variable.

2. Research Questions

My goal in this study is to have a better understanding of the alternation between overt and null SPPs in Turkish. The analyses reported and discussed in the following chapters aims to answer the following research questions:

1. Taking the sample as a whole, what are the grammatical variables from among the ones we are studying that constrain the use of overt and null pronouns? What variables turn out not to be constraints?

2. What are some of the socio-demographic variables, among the ones that we have chosen, that favor the use of overt and null pronouns? Which of our socio-demographic variables turn out not to be constraints? That is, what differences are there between informants who live in different places, arrived in NYC at different ages, use more or less Turkish, etc.?

3. Dividing the sample between Turks and New Yorkers, what are the differences between TT and TNY with respect to (a) their rates of use of SPPs and (b) the variables, and constraints within variables that influence or fail to influence this use?

4. Which of the variables and constraints that turn out to be operative, among TT, TNY, or both, exerts a greater influence on the use of SPPs? That is, what are the hierarchies of variables and hierarchies of constraints that exist in the corpus as a whole, or among TT only, or among TNY only?

3. Hypotheses

The hypotheses of the present study relate to the expectation that Turkish in NY is undergoing a process of language contact wherein English pronominal usage patterns are influencing the usage patterns of Turkish pronouns. This dissertation attempts to replicate, for contact Turkish, the study of contact Spanish presented in Otheguy et al. (2007). It investigates the different ways in which pronominal usage differs according to place of residence (Turkey or NY) and, for NY residents, between different immigrant generations (first, Turkish-born vs second, NY-born). In keeping with Otheguy et al.

(2007), the terms first and second generation refer to immigrant generations. First generation speakers are the ones who were born in Turkey and came to the U.S. after the age of three and second generation speakers are those who were either born in the U.S. or Turkey, but came to the U.S. before the age of 3 and were raised in the U.S. Thus the second generation group not only includes speakers who were born in the U.S. but it also includes those who were born in Turkey but arrived in the U.S. before the age of 3.

The dissertation also investigates, for those who live in NY, the effects of years spent in the City and age of arrival in the City on pronominal usage. The assumption is made throughout, to be tested in the dissertation, that bilingual speakers of Turkish perceive a parallel between English and Turkish pronouns, and that the usage of the socially dominant language, English, is influencing the socially less dominant language, Turkish. In addition, we make the obvious assumption that none of this is happening among monolinguals in Turkey. The specific hypotheses are the following:

1. Turkish speakers who live in NYC will show a higher pronoun rate than those who live in Turkey. This expectation is based on the assumption that the English usage pattern is having an effect on the usage pattern of Turkish in NY, and on the experience of other immigrant communities whose language includes variable use of subject pronouns. (Otheguy et al., 2007).

2. Speakers who have spent more years in NYC will use overt pronouns at a higher rate than those who have spent fewer years in NYC. This expectation is, again, based on the assumption that (a) the pattern of usage of English is having an effect on the pattern of usage of Turkish, assuming that exposure to English increases as Turks spend more time in NY, and (b) on the assumption that Turkish in NY may be undergoing a language

contact process similar to that reported for languages such as Spanish (Otheguy et al., 2007).

3. Speakers who arrived in NYC at a younger age will use more overt pronouns than those who arrived in the City at an older age.

4. Speakers living in NYC with strong English skills will use more overt pronouns in Turkish than those who had weaker English skills. Similarly, speakers with weak Turkish skills will use more overt pronouns in Turkish than those with stronger Turkish skills.

5. TNY who used Turkish to a great extent with different kinds of interlocutors will use fewer overt pronouns than those who used Turkish less. Similarly, TNY who spoke Turkish to a great extent in different domains such as home, work place and school will use fewer overt pronouns than those who used Turkish less.

6. With respect to the independent grammatical variables and constraints within those variables, I predict, based on the Spanish study, that the variable hierarchies will not differ between the different residential groups, but that the constraint hierarchies will differ. For example, I expect that the variable, Continuity of Reference, to be on top of the hierarchy and it will be ranked the same among all the groups.

However, I expect the constraints inside this variable to be different. In the Turkish spoken in Turkey the preferred form for reference maintenance is null pronouns. A verb with the same reference does not get many overt pronouns, but a verb with switch reference may get many overt pronouns. Therefore, same reference constraint may be significant in Turkey, however, because of the influence of English on the Turkish spoken in New York, I expect TNY to use more overt pronouns than the TT in the same reference environment.

In order to address the research questions, the following linguistic variables were analyzed (see Appendix A).

- Person and number of the null or overt SPP
- Person and number ending of the verb
- TMA I in finite verbs
- TMA II in finite verbs
- TMA III in finite verbs
- TMA IV in finite verbs
- TMA I in non-finite verbs
- TMA II in non-finite verbs
- TMA III in non-finite verbs
- TMA I in non-verbal predicates
- Continuity of reference

In addition to the grammatical variables, the following social variables were analyzed. The variables were based on a sociolinguistic survey (see Appendix C) obtained from each informant.

- Age
- Sex
- Social class
- Education

As mentioned earlier, this study also determined whether contact with English has an influence over SPP use among Turks in NYC. To determine this, several other social variables were also analyzed.

For informants living in NYC:

- Age of arrival in NYC
- Length of residence in NYC
- Level of Turkish skills
- Level of English skills
- Amount of use of Turkish with different interlocutors
- Amount of use of Turkish in different domains

CHAPTER 5

METHODOLOGY

1. Introduction

This study was approved by the Graduate Center Institutional Review Board (IRB) - the Committee for the Protection of Human Subjects. The data were collected via face-to-face interviews from a total of forty speakers; twenty Turkish speakers who were currently living in Turkey and twenty Turkish speakers who were living in New York (details about the informants are given in the following section). The interviews were not less than thirty minutes. The informants were told that they were participating in a linguistic study but were not told that this study investigates the use of SPPs in Turkish. The informants chose among several different topics to talk about, such as a positive event that had an important influence on their lives, their life history, movies, marriage, books, their children, and jobs etc. The interviews were collected by me.

2. Informants

Scholars believed that analyzing a speech community gives us a better understanding of the linguistic behavior of an individual (Labov, 1989). Thus, sociolinguistic studies are based on the speech of many people who represent a speech community in which people are chosen according to stratification criteria such as age, educational level and social class. In order to obtain a representative sample of the population, I also chose the informants based on stratification criteria such as gender, age, age of arrival in NYC, education and social class. The distribution of informants by these criteria is shown in the following sections. Informants are divided into two groups: Turks living in Turkey and Turks living in NYC.

2.1 Turks living in Turkey

This group consisted of twenty Turkish adults who were currently living around the west part of Turkey and had minimal or no knowledge of the English language. They did not move into another language environment. The distribution of speakers by age and gender for informants residing in Turkey is presented below.

Table 5.1: Distribution of informants living in Turkey by age and gender

Age	**Male**	**Female**	**Total**
20-34	2	3	5
35-55	2	4	6
56-80	6	3	9
Total	10	10	20

Informants in this group ranged in age from twenty to eighty years and consisted of ten males and ten females. The following table presents the distribution of speakers by social class.

Table 5.2: Distribution of informants living in Turkey by social class

Social class	**Male**	**Female**	**Total**
Professionals	6	4	10
Workers	4	6	10
Total	10	10	20

As shown in the table above, the informants are classified into two categories: professionals and workers. In keeping with Labov (1966), I categorized the informants in

both subsamples as members of the working and professional classes based on the type of work they performed, and the extent of formal education.

The Turkey sample included a secretary, a student, three engineers, two housewives, two bank directors, a pharmacist, two soldiers, a math professor, two cleaning ladies, a teacher, a machinist, an editor, and a doorman. One informant had his own business. Informants working as secretaries, low-ranking soldiers, cleaning ladies, drivers, waiters, technicians, housewives and doormen who have less than a college education or who have not finished a two-year or a four-year college were categorized as members of the working class. Engineers, bankers, pharmacists, professors, teachers, analysts, attorneys and editors were categorized as professionals. These informants graduated from either college or graduate school. Informants who were retired at the time of the interview were categorized under their profession prior to retirement.

We now turn to the level of formal education obtained by the informants living in Turkey. Note that before 1997, education in Turkey was divided into 3 levels: Primary, secondary and high school. Primary school, which starts at the age of 7, was compulsory for 5 years. This was followed by 3 years of secondary school education and 3 years of high school education. After 1997, however, compulsory basic education was extended from 5 to 8 years. This is followed by high school, which still normally provides 3 years of education. Since all the informants living in Turkey had gone to school before the new system replaced the old one, I included secondary school as a factor in the study. In the following table, I present a distribution of speakers by level of formal education.

Table 5.3: Distribution of informants living in Turkey by level of formal education

Level of education	**Male**	**Female**	**Total**
Attended/s elementary	1	1	2
Attended/s secondary	1	2	3
Attended/s high school	2	2	4
Attended/s college	5	5	10
Attended/s graduate school	1	0	1
Total	10	10	20

2.2 Turks living in NYC

This sample consisted of twenty Turkish speakers who were currently living in NYC at the time of the interview. The informants lived in different parts of the city. Nine informants lived in Manhattan; four around Murray Hill/Midtown, one around the West Village, one in the East Village, two around the Upper West Side, and one in Tribeca. Four informants lived in Brooklyn; two in Brooklyn Heights, and two around Ocean Parkway. Two informants lived in Queens; one lived around Bayside and the other in Sunnyside. One lived in Long Island City. I interviewed two informants who have recently moved to Washington, D.C. and to Ohio but they had lived and worked in NYC all their lives before they moved. Two informants lived in New Jersey but they worked in Manhattan. Almost all informants who came to the City after the age of three were from İstanbul, Turkey. One informant was from Ankara. Turks in NYC is divided into two groups:

A. First generation Turks (Turks who came to the U.S. after the age of three)

This group included ten speakers who came to the U.S. after the age of three and were exposed to English for the first time in the U.S. They were currently living in NYC. The following tables present the distribution of age, gender, age of arrival, social class, level of education and length of residence for all informants living in NYC.

B. Second generation Turks (Turks who were born in the U.S. or brought to the U.S. at or before the age of three)

In keeping with prior research (Otheguy et al., 2007), this group included ten informants who were born in the U.S. or were brought to the U.S. at or before the age of three. The reason why 3 was the cut-off point is that the use of pronouns among children who are younger than 3 is not fully developed. Literature on the acquisition of pronouns demonstrates that although children acquire the grammatical knowledge at an early age, they learn the pragmatic factors much later (Wexler & Chien, 1985).

Table 5.4: Distribution of informants living in NYC by age and gender

Age	**Male**	**Female**	**Total**
20-34	3	6	9
35-55	7	3	10
56-57	0	1	1
Total	10	10	20

The informants ranged in age from twenty to fifty-seven years. This sample was also balanced with respect to gender.

Table 5.5: Distribution of informants living in NYC by age of arrival

Age	**Male**	**Female**	**Total**
0-3	5	5	10
3-12	4	3	7
13-19	1	2	3
20+	0	0	0
Total	10	10	20

Table 5.6: Distribution of informants living in NYC by social class

Social class	**Male**	**Female**	**Total**
Professionals	5	5	10
Workers	5	5	10
Total	10	10	20

This group included an analyst, an accountant, a limousine driver, a physics professor, a housewife, an event planner, a flight attendant, two attorneys, a jeweler, an investment banker, a student, a pianist, a lab assistant, an adjunct instructor, a waiter, a doctor, a technician, and a consultant. One of the informants had his own business.

Table 5.7: Distribution of informants living in NYC by level of formal education

Level of education	**Male**	**Female**	**Total**
Attended/s elementary	0	0	0
Attended/s secondary	0	0	0
Attended/s high school	3	1	4
Attended/s college	3	5	8
Attended/s graduate school	4	4	8
Total	10	10	20

Table 5.8: Distribution of informants living in NYC by length of residence

Length of residence	**Male**	**Female**	**Total**
3-15	1	3	4
16+	3	2	5
Native	6	5	11
Total	10	10	20

3. Data collection procedure

The interviews for this study were conducted both in Turkey and in New York between the months of January and June in 2008. I interviewed twenty Turkish speakers in İzmir, the third largest city located on the west coast of Turkey. As of December 2007, the population of İzmir was 3,739,353. In both İzmir and New York, I used the social network approach (Milroy, 1980) when collecting data. The advantage of such an approach, according to Milroy, is that it is easier to enter a speech community and meet more speakers within contacts' social networks. So I contacted the speakers through my

parents, friends and relatives. Since most of the speakers knew my parents, and friends, they were willing to participate in the study. Some of the interviews were recorded in my house and the informants' houses. Some were conducted in their offices.

After returning to New York City from İzmir, I started looking for potential informants for my second group, Turks in NYC. Finding speakers that represented this group was more difficult than finding speakers in Turkey since I knew only few potential speakers. Most of the people that I knew came to the U.S. after their twenties. Once again, I made contact with most of the informants through my friends via e-mail and phone to find out if they would be interested in participating in my study. The informants I interviewed introduced me to their spouses and friends. Since the speakers in NYC were very busy, some of the interviews were conducted on the phone and some of them in the informants' offices.

Before the interviews, I had informed the speakers that everything, including their names, would be confidential. Some informants living in NYC were more interested than the informants in Turkey about the nature of my study. When they asked me what my research was about, I tried not to give a lot of information to the informants about the purpose of my study since it could have affected their speech. According to Labov (1972b), if speakers become aware of the fact that their speech will be used for a study, they are likely to use a formal speech. This can lead to collecting data that is not representative of the speakers' natural speech.

Labov also found that speakers pay less attention to their speech when they talk about emotional topics. Following Labov, I asked my informants to talk about a positive event that had an important effect on their lives. Nevertheless, some informants were more willing to tell me their life histories. Some other topics that the informants talked

about were movies, TV series, their children, jobs and favorite vacations. I asked the speakers to talk about one of the topics giving as many details as they could and told them that I would not interrupt them throughout the interview. When the speakers finished talking about a topic, I asked them to talk about another topic in detail.

The interviews in both İzmir and New York lasted no less than 30 minutes. After the interviews, the informants filled out a sociolinguistic survey (see Appendix C) where they stated their ages, professions, gender, education level, and English and Turkish skills. Following the interview sessions, the informants' speeches were transcribed.

4. Materials

The interviews were recorded using a Sony ICD P210 digital voice recorder. After the interviews, the data regarding the sociolinguistic variables were obtained from all informants by means of a survey. The informants living in Turkey filled out a questionnaire regarding their ages, professions, and education. The informants living in NYC filled out a more detailed survey in which they explicitly responded to questions such as how long they have been living in the U.S., how frequently they use English and Turkish on a daily basis. The corpus I collected for this research was then analyzed using the Statistical Package for the Social Sciences, SPSS.

CHAPTER 6

RESULTS: SELECTING SOCIAL VARIABLES FOR A MULTIVARIATE ANALYSIS

1. Introduction

The report of my results for the social and linguistic variables is divided into four chapters. The purpose of this chapter is to select the social variables to be included in the multivariate regression analyses. I do this here by determining which of the variables, when studied in a bivariate analysis, have on their own a significant effect on the occurrence of overt SPPs in Turkish. The more definitive regression analyses of these significant variables will be provided in chapter 7. Chapters 8 and 9 will present the results for the linguistic variables. Recall that this study included social variables such as place of residence, age of arrival in NYC, years in NYC, level of English and Turkish skills among NYC residents, gender, social class, education, age of informants, amount of use of Turkish with different kinds of interlocutors and amount of use of Turkish in different domains. To see the relationship between overt pronoun rate and each of these social variables, I ran correlations, Anovas and multiple linear regressions. The analyses found the following social variables significant:

- Place of residence

  This variable was significant for the entire sample (N = 40).

- Age of arrival in NYC

  This variable was significant for NYC speakers as a whole (N = 20).

- Amount of Turkish spoken with father

  This variable was also significant for NYC speakers as a whole (N = 20).

- Age of informants

  This variable was significant for the entire sample (N = 40).

2. Residence

The first analysis to be considered is the overt pronoun rate by (place of) residence. The table below shows the overt pronoun rates by the two groups: Turks in Turkey (TT) and Turks in NYC (TNY).

Table 6.1: One-way ANOVA: Overt pronoun rate by residence

Groups	**N speakers**	**% Overt**	**St. Deviation**
TT	20	.11	(.05)
TNY	20	.20	(.05)
F = 35.33 p < 0.01			

For this analysis, a one-way analysis of variance was conducted to evaluate the relationship between the two informant groups and the use of an overt or a null pronoun. The dependent variable here was qualitative, the use of an overt or a null pronoun. The results of the ANOVA were, as I will discuss presently, as expected. They showed that the average rate of overt pronouns for Turks living in NYC (*M*=.20, *SD*= .05) was higher than Turks living in Turkey (*M*=.11, *SD*= .05). The results are statistically significant overall and thus this variable will be included in the multiple regression. The

interpretation of these results, as of those of the subsequent tables to be presented, will be postponed until chapter 7.

3. Age of arrival

For the continuous variable that I have called Age of arrival, I present results in terms of a Pearson correlation between the speakers' pronoun rate and the ages at which they arrived in NYC.

Table 6.2: Pearson correlations: Overt pronoun rate and age of arrival for NYC speakers

	N speakers	**r**
Pronoun rate * Age of arrival	20	-.69**
** = $p < 0.01$		

As the table reveals, there was a significant negative correlation between the dependent and the independent variables. The table shows that, as expected, the younger a Turkish speaker arrived in NYC, the more overt subject pronouns were found in his or her speech. Since this variable is also bivariately significant, it will go into the multiple regression.

4. The use of Turkish with interlocutors among NYC speakers: Turkish with father

I probed the connection between the pronoun rate and the amount of Turkish used with different kinds of interlocutors by speakers in NYC. The results were all in the expected direction, but the only statistically significant one was the negative correlation between overt pronoun rate and the amount of Turkish spoken with father.

Table 6.3: Pearson correlations: Overt pronoun rate and amount of Turkish used with father for NYC speakers

	N speakers	**r**
Pronoun rate * Turkish with father	20	-.58*
*= $p < 0.05$		

The table shows that, when all NYC speakers are taken into account, the higher the reports are of speaking in Turkish with one's father, the lower is the use of overt subject pronouns in Turkish. This variable will also be included in the multiple regression analysis.

Further correlational analyses between overt pronoun rate and the amount of Turkish used in different domains were not statistically significant for the entire sample. Therefore, these variables will not be included in the regression analyses.

5. Age of informants

Finally, I discovered, to my surprise, that age of informants had a significant effect on the rate of overt pronouns when the entire sample was taken into account. The results are shown in the table below.

Table 6.4: Pearson correlations: Overt pronoun rate and age of informants for the entire sample

	N speakers	**r**
Pronoun rate * Age	40	-.56**
** = $p < 0.01$		

The table shows an overall significant negative correlation between age of consultants and overt pronoun rate. In other words, the younger the informants were, the more overt pronouns they used in Turkish. I further tested this connection in a separate one-way analysis of variance to find out whether or not discretely different age groups differed significantly from each other. Recall that for the purpose of this study, the informants were divided into three different age groups: 20-34; 34-55; 56-80. The relationship between age categories and overt pronoun rate was evaluated. Post hoc comparisons with the use of a Tukey test were conducted to evaluate pair-wise differences among the means. The table below presents the results.

Table 6.5: One-way ANOVA: Overt pronoun rate by three different age groups in the entire sample

Age Category	**N speakers**	**% Overt**	**St. Deviation**
20-34	14	.18	(.05)
35-55	16	.16	(.06)
56-80	10	.10	(.04)
F = 5.74 $p < 0.05$			

The results of the ANOVA indicated that there was a significant difference between the speakers in the youngest age group (*M*=.18, *SD*= .05) and speakers who were older than 56 (*M*=.10, *SD*= .04). There was also a significant difference between middle-aged speakers (*M*=.16, *SD*= .06) and speakers who were 56 and older (*M*=.10, *SD*= .04). No significant difference was found between the youngest and the middle-aged group with respect to overt pronoun rate. The table below repeats this result, presenting a clearer

comparison between the two significantly different groups, namely those who are between the ages of 20-55 and those between the ages of 56-80.

Table 6.6: One-way ANOVA: Overt pronoun rate by two different age groups in the entire sample

Age Category	**N speakers**	**% Overt**	**St. Deviation**
20-55	30	.17	(.06)
56-80	10	.10	(.04)
$F = 11.324$ $p < 0.05$			

The table shows that, in my sample of Turkish speakers, the younger consultants used overt pronouns at a higher rate than the older ones.

The unexpected discovery that there was a statistically significant negative correlation between age and pronoun rate meant that I had to probe whether my other findings were vitiated by the age factor. That is, I needed to find out whether the findings in terms of residence, age of arrival in NYC, and amount of Turkish used with father were simply caused by the fact that the members of these different categories were of different ages, and that it was their age, and not their place of residence, their age of arrival etc. that was truly associated with the pronoun rate.

For this purpose, I first studied the connection between age and the first variable that I found to be predictive of the pronoun rate, namely residence. At the beginning of this chapter, it was shown that TNY used overt pronouns at a higher rate than TNY, and that this difference between the two residential groups was statistically significant. In order to find out whether this result might be partially compromised by the age factor, we

tested whether there is a correlation between the different groups and the age of the consultants.

Table 6.7: Pearson correlations: Age of informants by residence

	N speakers	**r**
Age * Residence	40	-.52
$p < 0.01$		

As can be seen in the table, there is a significant negative correlation between age of informants and residence. This suggests that younger informants lived in NYC. In addition, I ran an Anova that compared the ages of consultants residing in Turkey and in NYC. The results are in Table 6.8.

Table 6.8: One-way ANOVA: Age of informants by residence

Place of residence	**N speakers**	**Mean Age**	**St. Deviation**
Turkey	20	50	(14)
New York	20	35	(10)
$F = 14.802$ $p < 0.05$			

The table shows that there was a significant difference between the ages of the consultants in Turkey ($M = 50$, $SD = 14$) and those in NYC ($M = 35$, $SD = 10$). In retrospect, it might have been wise to control for my informants' age. But I did not anticipate this problem when gathering my data, so that it has turned out that the Turkish subsample is somewhat older, in a statistically significant way, than the NYC subsample.

It remains an outstanding problem, to be addressed further below, to determine whether my previously reported finding with respect to the effect of residence on the use of pronouns (TNY use more overt pronouns than TT) may be simply an artifact of age differences.

Correlations were also performed on age of arrival and age of informants. Recall that age of arrival also had an influence on overt pronoun rate, with younger arrivers using more overt pronouns than older arrivers. We wanted to find out whether this finding would be compromised by age differences between the two groups.

Table 6.9: Pearson correlations: Age of informants by age of arrival

	N speakers	**r**
Age * Age of arrival	20	.39
p = .084		

The table shows that the correlation between age of informants and age of arrival was also non-significant, suggesting that our findings with regard to age of arrival still stand.

Finally, we also tested whether or not amount of Turkish used with father was correlated with age of informants.

Table 6.10: Pearson correlations: Age of informants by amount of Turkish spoken with father for NYC speakers

	N speakers	**r**
Age * Turkish with father	16	.13
p = .62		

Again here, the table shows, the correlation between age of informants and the amount of Turkish spoken with father was also not significant. This suggests, in this case too, that the findings with regard to the amount of Turkish used with father and the pronoun rate still stand.

In sum, my purpose in running these correlations was to find out whether age of the consultants and not residence, generation, age of arrival and amount of Turkish spoken with father caused the increased percentages in the use of overt pronouns by NYC speakers. The only significant correlation was between residence and age of the consultants. The other correlations gave non-significant results, suggesting that my previous findings with respect to the overt pronoun rate and its correlation with age of arrival and amount of Turkish spoken with father were still valid and that those findings were not compromised by the age factor. Thus the higher overt pronoun rate observed among NYC speakers cannot be explained away by appealing simply to age differences.

Other social variables such as speaker gender, education and social class were found not to be significant. These variables did not have a significant effect on the use of an overt pronoun when the whole sample or the subsamples were taken into consideration. Same findings were also reported in Otheguy et al. (2007) and Silva-Corvalán (2001) who found that social variables of this type tend to have little or no effect on syntactic variation such as subject pronoun variation. Therefore, I will not discuss these variables.

6. Summary

This part of the study has shown that SPP variation in the Turkish spoken in NYC is significantly influenced by a number of social variables. Residence is one of them.

Turks who live in NYC showed a higher overt pronoun rate than those who live in Turkey. Overt pronoun rate is influenced by whether a speaker lives in Turkey or NYC.

Further, I had predicted that the younger a Turkish speaker arrived in NYC, the more overt subject pronouns would be found in his speech. This was also based on Otheguy et al. (2007) who found a significant negative correlation between overt pronoun rate and age of arrival. My prediction was also confirmed by the data. As the results showed, within the NYC sample, the speakers who arrived in the City at a younger age used more overt pronouns than those who came to the City at an older age.

Turning to the amount of Turkish used with different kinds of interlocutors by speakers in NYC, I had hypothesized that speakers who used Turkish to a great extent would use fewer overt pronouns in Turkish than those who used Turkish less. Otheguy et al. (2007) found significant negative correlations between overt pronoun rate and the amount of Spanish used with many types of interlocutors. In my study, the only significant negative correlation was between overt pronoun rate and the amount of Turkish spoken with father, suggesting that the less a speaker spoke Turkish with his father, the more overt pronouns he used in Turkish.

Another result was a significant negative correlation between age of consultants and overt pronoun rate by all speakers in the study. I did not expect to find this result for the reason that most of the previous studies did not find any effects of this variable on subject pronoun variation. Only few studies (Avila-Jiménez, 1996; Flores-Ferrán, 2004) found that informants who are younger than 50 are more likely to use overt SPPs than those who are older than 50. Flores-Ferrán reported that informants between the ages of 20 and 39 used more overt SPPs than informants who were older than 50. It should also be noted that the NYC and Turkey samples were not equal to each other with respect to

the age of informants. Therefore, it is necessary to conduct a further study with samples, which include an equal number of speakers in each age group so that we can see whether or not age will have a significant influence on overt pronoun rates.

Finally, contrary to my predictions, correlations between overt pronoun rate and years in the U.S. and overt pronoun rate and the amount of Turkish used in different domains did not give significant results. These findings differ from those for Spanish in Otheguy et al. (2007). In that study, these social variables had a significant effect on the use of an overt pronoun. The authors found a significant positive correlation between overt pronoun rate and years in the U.S. and significant negative correlations between overt pronoun rate and the amount of Spanish used in different domains.

As mentioned in section 1, the purpose of this chapter was to select the significant social variables that will go into the regression analyses. Since residence, age of arrival, the amount of use of Turkish with father, and age of the informant were found to be significantly associated with overt pronoun rate, they will be included in the multiple regression analyses.

CHAPTER 7

RESULTS: MULTIVARIATE ANALYSIS OF SOCIAL VARIABLES

1. Introduction

In the previous section, I considered the variables separately, in several bivariate analyses and selected the variables that had a significant effect on the use of an overt subject pronoun. In this section, I further analyze via multiple regression those social variables which have been found to have significant effects on overt pronoun rate to see which of these variables has the greatest impact on the use of an overt subject pronoun.

2. Multivariate analyses of social variables

The first analysis to be considered involves residence and age of consultants. Recall that, to my surprise, there was a significant negative correlation between age of informants and overt pronoun use, suggesting that the younger a speaker was, the more overt pronouns were found in his or her speech. Further analysis showed that the speakers in NYC were, in fact, much younger than those in Turkey. This led us to worry that it could be the age of the NYC informants, and not the fact of living in NYC, that was causing the higher rate of overt subject pronouns in Turkish. Therefore, in a multiple linear regression analysis, I further probed the possibility that it is age that is the predictor of overt pronoun rate, and not residence. The predictors were residence and age of consultants, while the dependent variable was overt pronoun rate. The table below presents the findings of this analysis.

Table 7.1: Multiple linear regression: Residence and age of informants by overt pronoun rate for the entire sample

Predictors	**R^2 change**
Residence	0.64
Age of informants	0.02
	Beta
Residence	0.70**
Age of informants	-0.17
$R^2 = 0.659$ ** = p < .01	

The results indicate that approximately 66 percent of the variance between overt and null pronoun in the sample can be accounted for by the linear combination of the two independent variables in the regression. But whereas residence accounted for 64 percent of the variance, age of informants accounted for only an additional two percent of this variance. The table also shows the relative strength of the individual predictors. Only residence ($\beta = .70$, p=.000) demonstrated significant effects on overt pronoun rate. Age of informants ($\beta = -.17$, p=1.29) did not have any significant influence. Therefore, we can conclude that, out of these two, residence is the only predictor of overt pronoun rate. Even though age of the informant was found to be significantly correlated with overt pronoun rate in the bivariate analysis, and even though the NYC sample shows a lower average age than the Turkey sample, the regression in Table 7.1 shows that age is not a significant predictor of overt pronoun rate. The age factor can now be discarded.

Recall that as shown in section 5 of chapter 6, Residence and Age of the informant were significantly correlated with one another. In the literature, it has been reported that

the presence of two or more correlated predictor variables in a regression may cause a problem, referred to by the term *collinearity* (Allison, 1999). According to Allison, multiple regression is designed for separating the effects of two or more independent variables on a dependent variable, when the independent variables are either not correlated or show low to moderate levels of correlation. Two variables that are highly correlated would not be appropriate input for a regression. For example, if two independent variables showed a level of correlation in the order of .75 or higher, then the results of a regression analysis would be exceedingly difficult or impossible to interpret.

In our case, the correlation between Residence and Age ($r = -.52$) can be described as moderate, creating therefore what we can in turn call moderate collinearity. This moderate collinearity does not preclude us from running a multivariate regression comparing the effect of Residence and Age on pronoun overtness, nor does it preclude us from interpreting the results. The regression suggests very strongly that Residence is the significant predictor, and that Age is not. To be sure, the conclusion needs to be modulated by the awareness of the small amount of collinearity that characterizes the two variables, but it remains nevertheless a strong and valid conclusion.

I ran another multiple linear regression analysis that involved age of arrival in NYC and amount of Turkish spoken with father. In sections 3 and 4 of chapter 6, I showed that there was a significant negative correlation between overt pronoun use by NYC speakers and age of arrival and again a significant negative correlation between overt pronoun use by NYC speakers and the amount of Turkish spoken with father. My aim now is to find out which of these two social variables is the best predictor of overt pronoun rate. The age factor was not included in this analysis as we established through the first regression analysis that age of consultants was not truly associated with the

pronoun rate. The variable residence was also not included in this analysis although we found significant correlational effects of this variable on the pronoun rate. The reason is that the residence variable involves all 40 speakers in the sample, whereas age of arrival and Turkish with father are relevant for only NYC speakers (N = 20). I now present the results of the regression analysis.

Table 7.2: Multiple linear regression: Age of arrival and Turkish with father by overt pronoun rate for TNY

Predictors	**R^2 change**
Age of arrival	0.50
Turkish with father	0.05
	Beta
Age of arrival	-0.55*
Turkish with father	-0.27
$R^2 = 0.552$ * = p < .05	

The table shows age of arrival as the best predictor. According to the results, age of arrival and amount of Turkish spoken with father accounted for 55 percent of the variance between overt and null pronoun in the corpus. Of this variance, age of arrival accounted for 50 percent and was the only variable that demonstrated significant effects. Based on these results, age of arrival appears to be a better predictor of overt pronoun rate than amount of Turkish spoken with one's father.

3. Discussion

In this section, I will interpret the findings of the multivariate analyses and discuss whether the findings provide support for the language contact hypothesis. First, when residence and age was submitted to a regression analysis, it was found that residence, and not the age of informants was the actual predictor of overt pronoun rate for all speakers in the study. In other words, overt pronoun rate was most influenced by whether speakers lived in Turkey or NYC. Age turned out to be not significant in this analysis. Therefore, we can conclude that residence has significant effects on overt pronoun rate independently of whether or not the speakers in NYC are younger than those living in Turkey.

This result provides evidence for language-contact influence from English on Turkish with regard to the use of overt pronouns. Since the Turks living in NYC have been using English and exposed to it in many different domains, they are more likely to be influenced by English than the Turks in Turkey who have limited opportunities to use English. This finding is also in line with that of Veltman (2000) who examined the process in which immigrants living in the U.S. came to adopt English. Veltman found that immigrants learn English very rapidly and adopt English as their primary language. In fact, in the long run, the high rates of language shift to English results in the abandonment of minority languages. Veltman shows this to be true for all minority language groups in the U.S.

Further multiple linear regression analysis was performed on the other social variables that had proved to be significant in the separate bivariate analyses, namely age of arrival and amount of Turkish used with father, to find out the best predictor of overt pronoun rate for the portion of our sample consisting of NYC residents. According to the

results, age of arrival was the only significant variable in this analysis. Once subjected to a multivariate analysis, Turkish with father turned out to be not significant.

This finding also favors the hypothesis that there is English influence on Turkish. Veltman (2000) found that the use of English by immigrants declines progressively with increasing age at the time of arrival. In other words, the older the immigrants at time of arrival, the less likely they are to speak English predominantly. His results showed that immigrants between the ages of 5 and 9 at time of arrival in the U.S. adopted English at a rate of more than 65 percent. This rate dropped to 40 percent for immigrants who came to the U.S. between the ages of 10 and 14 and to 25 percent for those who were aged 15-19 when they arrived in the U.S. The reason for this decline, according to Veltman, is that immigrants who arrive in the U.S. at an early age have more opportunities to use English on a daily basis. For example, younger immigrants receive most of their schooling in the U.S. In addition, younger arrivers are more receptive to the influence of English as their first language skills are not fully developed by the time they arrive in the U.S.

Note that in addition to residence and age of arrival, the findings regarding a few other social variables also provide further evidence for the English contact hypothesis. A social variable that was found to be significantly related to the selection of overt or null subject pronouns was immigrant generation. Otheguy et al. (2007) showed for Spanish in NYC that there was a significant difference between Newcomers (speakers who arrived in NYC at age 17 or older and have lived in the City for no more than five years) and NYBRs (speakers who were either born in the U.S. or came to the U.S. at or before the age of 3) regarding the use of overt pronouns. In the Spanish study, NYBRs used SPPs at a significantly higher rate than the Newcomers. In the present Turkish study, I have found

a significant positive correlation between overt pronoun rate and immigrant generation. The results are shown in the table below.

Table 7.3: Pearson correlations: Overt pronoun rate and immigrant generation for TNY

	N speakers	**r**
Pronoun rate * generation	20	.74**
** = $p < 0.01$		

This suggests that overt pronoun use is influenced by whether a speaker is a first or a second generation Turk.

This result can also be attributed to the influence of English. It can be said that English has an influence on the SPP use by second generation Turks because since they have been speaking English all their lives in many domains such as schools, they are more likely to use overt pronouns than first generation speakers do. Second generation speakers usually adopt English as their primary language and abandon the daily use of their first language (Veltman, 2000).

Also relevant to the issue of language contact is the different levels of language skills among my NYC speakers. In the Spanish study, Otheguy et al. (2007) found significant positive correlations between the level of English skills and overt pronoun use and significant negative correlations between the level of Spanish skills and overt pronoun use. On the basis of these findings, I also explored the relationship between overt pronoun rate and the level of English and Turkish skills for TNY. The following table presents the results.

Table 7.4: Pearson correlations: Overt pronoun rate and level of English and Turkish skills for NYC speakers

	N speakers	**r**	**p**
Pronoun rate * English skills	20	.25	.27
Pronoun rate * Turkish skills	20	-.25	.26

As the tables show, the results of the correlations are, as we will discuss presently, in the expected direction, but they are not statistically significant. There is a positive correlation between overt pronoun use by NYC speakers and their level of English skills, and a negative correlation between overt pronoun use by these speakers and their level of Turkish skills. This means that the more the speakers report their English skills to be good, the more overt pronouns they use in Turkish. In contrast, the more they report their Turkish skills to be good, the fewer overt pronouns they use in Turkish. So even though the findings are not significant, they go in the direction predicted by a language contact hypothesis.

These results, though not significant, still provide evidence for language-contact influence from English on Turkish. The use of English on a regular basis by the Turks living in NYC has led to the worsening of their Turkish language skills. As a result, their Turkish has lost its resistance to the influence of English. However, speakers who have been speaking Turkish frequently ended up maintaining good Turkish skills and therefore, the Turkish of these speakers have become more resistant to the influence of English.

These findings provide support for the hypothesis that under conditions of language contact, the languages of bilinguals can influence one another, with the dominant language influencing the grammar of the usage of the less dominant one. As is

well known, language contact happens when languages spoken by bilingual or multilingual speakers interact, causing innovations in the grammar of a language and usually resulting in language change over time (Thomason, 2001; Sankoff, 2001; Hamers & Blanc, 1989; Appel & Muysken, 1987). The fact that the bilinguals in NYC tend to use more overt SPPs than the Turks in Turkey suggests very strongly that there is English influence on Turkish with regard to the use of pronouns.

Previous studies in other languages have provided mixed support for a language contact hypothesis with respect to the use of pronouns. In the most widely studied case, that of Spanish in the U.S., while Silva Corvalán (1994), Flores-Ferrán (2004), and Bayley and Pease-Álvarez (1997) showed that contact with English did not play an important role on the use of SPPs in the Spanish spoken by different communities in the U.S., Morales (1986) and Klein-Andreu (1985) have presented evidence showing that contact with English served as an important factor that conditioned the use of an overt or a null subject pronoun. The more recent study (Otheguy et al., 2007) showed that contact with English had an important effect on the frequent usage of overt pronouns in the speech of the bilingual Spanish speakers who were born in NYC.

The findings of the multivariate analysis with regard to residence and age of arrival as well as the significant negative correlation between overt pronoun rate and generation provide evidence for language contact hypothesis. Note that the conclusion about language contact is not due to a single social variable, rather, it comes from the findings of all social variables, namely residence, age of arrival, and generation. It can be said that that the use of overt pronouns in the Turkish spoken in NYC is becoming similar to the use of overt pronouns in English, and that the Turkish spoken in NYC may be undergoing a language contact process to a certain extent similar to that reported for

Spanish in Otheguy et al. (2007). However, note that the grammatical structure of the Turkish spoken in NYC is not affected. What appears to be happening is that as a result of the frequent usage of SPPs in English, SPPs have been used more frequently in the Turkish spoken in NYC than they would be used in the Turkish spoken in Turkey. The frequent usage of SPPs in the Turkish spoken in NYC does not necessarily affect the grammar of Turkish but certainly leads to changes in the grammar of the usage of Turkish in NYC.

CHAPTER 8

RESULTS: BIVARIATE ANALYSIS OF LINGUISTIC VARIABLES

1. Introduction

This chapter presents an analysis of the linguistic variables that play a role in favoring or disfavoring the occurrence of overt SPPs in Turkish. I discuss here results regarding variables such as person and number of the SPP, person and number ending of the verb, continuity of reference, and the tense, modality and aspect (TMA) suffixes that appear with finite verbs. The TMA suffixes that appear with non-finite verbs and with non-verbal predicates were also investigated, but it turns out that there are no statistically significant relations between them and the use of pronouns, and so these suffixes will be set aside and not discussed here.

We will use simple abbreviations to refer to these variables. Continuity of reference will henceforth be called simply Continuity. The variables that take into account the different persons of the pronoun and of the verb will be referred to as Person-Pronoun and Person-Verb respectively. The variables that make reference to tense and modality suffixes occurring with finite verbs will be named, making reference to the morphological slot that they occupy in Turkish verbs, as TMA II, TMA III, and TMA IV. (The reason that there is no TMA I is that the relationship between SPPs and the suffixes of this slot turned out to be non-significant for the entire sample and the subsamples). This chapter presents only bivariate results, using cross-tabulations and percentages. Multivariate results are in the next chapter.

To explore the bivariate relationship between these linguistic variables and the use of overt pronouns, I ran a cross-tabulation between each variable and overt pronouns for

the entire sample (N = 40) and for the two subsamples: Turks in Turkey (TT) (N = 20) and Turks in NYC as a whole (TNY) (N = 20). The results show that the variables that influence the alternation of overt and null pronouns are:

- Continuity

  This variable was significant for the entire sample, for the TT, and for the TNY.

- TMA II

  This variable was significant for the entire sample and for the TNY.

- TMA IV

  This variable was significant for the entire sample and for the TNY.

- TMA III

  This variable was significant for the TT and for the TNY.

- Person-Pronoun

  This variable was significant for the entire sample, for the TT, and for the TNY.

- Person-Verb

  This variable was significant for the entire sample, for the TT, and for the TNY.

I begin by presenting the variables Continuity, TMA II and TMA IV for the entire sample and the subsamples. The order of factors within these variables was the same for the entire sample as for the subsamples. The order of factors within other variables differed for between the entire sample and the subsamples.

As mentioned in chapter 4, the terms *factors* or *constraints* in the variationist literature refer to the levels of an independent variable. For example, Person-Pronoun is

an independent variable whereas ‘first person singular pronoun’ and ‘second person singular pronoun’ are factors or constraints within this variable.

2. Continuity of Reference

I begin with an analysis of Continuity for the entire sample.

Table 8.1: Crosstabs: Overt pronoun rate by Continuity

Whole sample (N speakers = 40)

Continuity	**% Overt**	**N verbs**
Different reference	20	7412
Same reference	11	6288
Total		13700
Chi-square: 212.131 $p < 0.01$		

The variable Continuity has two factors: different-reference and same-reference. The difference between the two factors is statistically significant. The table shows that different-reference environments favor overt pronouns much more than same-reference environments. The entire group of speakers, when considered together, used overt pronouns at a substantially and significantly higher rate in the different-reference environment than in the same-reference environment. We now proceed to compare TT and TNY as a whole with respect to the overt pronoun rate by the same variable.

Table 8.2: Crosstabs: Overt pronoun rate by Continuity

TT (N speakers = 20), and TNY (N speakers = 20)

Continuity	**TT**		**TNY**	
	% Overt	**N verbs**	**% Overt**	**N verbs**
Different reference	16	3647	24	3765
Same reference	6	2978	15	3310
Chi-square	156.367		84.230	
Cont. coefficient	.152		.108	

The table shows that the greater favoring of overt pronouns in the different-reference environment continues among TNY but that the difference between the two types of linguistic environments has diminished among NYC speakers. Whereas the percentage point spread is 10 points in Turkey, it is slightly less, nine percentage points**,** in NYC. Moreover, the contingency coefficient, which is a measure of association between two variables (in this case, overt pronoun and residence) has also diminished in NYC when compared to Turkey, from 0.15 to 0.10. These results, as well as those of the subsequent tables to be presented, will be interpreted in section 8 of this chapter.

3. TMA II

We now turn to the next variable, which we label TMA II, and which refers to the relationship between the particular suffix chosen and the occurrence of an overt pronoun in the second morphological slot associated with Turkish verbs. The factors of this variable are the different TMA suffixes that can occur in slot II. Slot II can also be found empty, that is, without any suffix. Absence of any suffix is also considered a factor within

this and all other TMA variables. In this analysis, the TMA II variable has only two factors: One is absent suffix and the other is what Turkish grammarians call the possibility suffix, *-(y)Abil.* The other suffixes of this slot are excluded from the TMA II variable because of a criterion under which if the number of overt pronouns that occurred with any suffix in my corpus was less than 5, then these suffixes were excluded from the analysis.

The results for TMA II indicated that the association between the choice of suffix and the occurrence of an overt subject pronoun was significant for the entire sample and for the TNY, but not for the TT. Since the results are the same for these two samples (that is, the order of factors that favor overt pronouns is the same for the whole sample and for the TNY), I report only the results for the sample as a whole. (Note that in Turkish, most suffixes are variable in form according to the rules of vowel harmony and consonant alternation (Göksel & Kerslake, 2005). Therefore, in this dissertation, parts of some suffixes that vary are shown in capital letters.

Table 8.3: Crosstabs: Overt pronoun rate by TMA II

Whole sample (N speakers = 40)

TMA II	**% Overt**	**N verbs**
Suffix absent	15	11083
-(y)Abil (Possibility)	9	175
Total	11258	
-(y)Iver, -(y)Agel, -(y)Ayaz, -(y)Akal, and *-(y)Adur* not shown Chi-square: 5.085 $p < 0.05$		

As the table illustrates, when the whole sample is taken into account, overt pronouns are favored when the second slot is empty, that is, when the TMA suffixes are absent from the second slot of a finite verb. Overt pronoun rates were also the highest for the absence of suffix factor for the TNY.

4. TMA IV

I now present the results for TMA IV. The relationship between the two factors that qualify to be included in this variable was statistically significant for the entire sample and for the TNY. The order of factors that favor overt pronouns is the same for the whole sample and for the TNY. Therefore, I report only the results for the sample as a whole.

Table 8.4: Crosstabs: Overt pronoun rate by TMA IV

Whole sample (N speakers = 40)

TMA IV	**% Overt**	**N verbs**
-(y)DI (Past copula)	18	826
Suffix absent	15	10356
-sA (Conditional copula)	5	167
Total		11370
-mIş (Evidential Copula) not shown Chi-square: 19.877 p < 0.01		

As the table shows, overt pronouns were most favored with the Past copula, *-(y)DI.*

5. TMA III

The factors of this variable are, again, the different tense, aspect and modality suffixes, but now as they are found in morphological slot III associated with Turkish verbs. These suffixes may also be absent from this slot, that is, this slot too may be empty. The results of the cross-tabulations were significant for the subsamples. The following table presents overt pronoun rate by these suffixes for the TT and TNY.

Table 8.5: Crosstabs: Overt pronoun rate by TMA III

TT (N speakers = 20), and TNY (N speakers = 20)

TT			**TNY**		
TMA III	**% Overt**	**N verbs**	**TMA III**	**% Overt**	**N verbs**
-(y)A (Optative)	12	222	*-AcAK* (Future)	24	212
-DI (Perfective)	12	2628	*-DI* (Perfective)	21	1978
-sA (Conditional)	12	68	*-mIş* (Evidential/Perfective)	20	228
-(A/I)r/-z (Aorist)	11	573	*-(I)yor* (Imperfective)	18	2287
-AcAK (Future)	9	307	*-(A/I)r/-z* (Aorist)	15	602
-(I)yor (Imperfective)	9	1622	*-(y)A* (Optative)	12	284
mIş (Evidential/Perfective)	7	295	*-sA* (Conditional)	12	78
Total		5715			5669
Absence of any suffix, *-mAlI,* and *-mAktA* not shown Chi-square: 13.473 P < 0.05			Absence of any suffix, *-mAlI,* and *-mAktA* not shown Chi-square: 25.423 P < 0.01		

The TT and the TNY differed from each other with regard to the order of factors. Whereas in Turkey, the order of the three top suffixes that most favored overts was

Optative, Perfective, and Conditional, in New York, the order was Future, Perfective and Evidential/Perfective.

In the sort of bivariate analysis presented here, one cannot be sure that the order of factors that favor overts in a variable, and the group differences with respect to order that we have just seen, are to be trusted completely; they may be due to imbalances in the data. The order of factors, and differences between groups with regard to factor orders, are much more reliable when produced by a multivariate analysis. These results are thus offered here for the sake of completion. They will not be discussed, the matter being postponed until the next chapter where multivariate results are offered.

6. Person-Pronoun

The relationship between Person-Pronoun and the occurrence of an overt subject pronoun was significant for the entire sample and for the two subsamples. The table below presents the results for the entire sample.

Table 8.6: Crosstabs: Overt pronoun rate by Person-Pronoun

Whole sample (N speakers = 40)

Person-Pronoun	**% Overt**	**N verbs**
3SG & 3PL PRO (ONE, PEOPLE)	91	23
2PL PRO	31	36
1SG PRO	20	5141
3PL PRO	14	1301
1PL PRO	13	2828
3SG PRO	13	3445
2SG (FORMAL) PRO	12	89
2SG (FAMILIAR) PRO	8	864
Total		13727
Chi-square: 254.475 P < 0.01		

As shown in the table, there was a big difference between the overt pronoun rate for the factors 3SG, *insan* 'one' and its plural form 3PL *insanlar* 'people' (these pronouns were coded as one factor) and the overt pronoun rates for the other factors. This led to a concern that it could be this difference that caused the results to be significant. Therefore, I ran another cross-tabulation without this factor to see whether or not the relationships continued to be significant. (This factor was also excluded from the crosstabs regarding the subsamples for the same reason). The results indicated that when *insan* and *insanlar* were excluded, the cross-tabulation remained significant. The following table presents the overt pronoun rate by Person-Pronoun for the entire sample but with *insan* and *insanlar* not counted.

Table 8.7: Crosstabs: Overt pronoun rate by Person-Pronoun

Whole sample (N speakers = 40)

Person-Pronoun	**% Overt**	**N verbs**
2PL PRO	31	36
1SG PRO	20	5141
3PL PRO	14	1301
1PL PRO	13	2828
3SG PRO	13	3445
2SG (FORMAL) PRO	12	89
2SG (FAMILIAR) PRO	8	864
Total	13704	
Chi-square: 155.086 P < 0.01		

The table shows that the overt pronoun rate was the highest for the 2pl. This is followed by 1sg and 3pl pronouns. 2sg (familiar) pronoun occurred at the lowest rate.

We now turn to an analysis of overt pronoun rate by the same variable for TT and TNY.

Table 8.8: Crosstabs: Overt pronoun rate by Person-Pronoun

TT (N speakers = 20), and TNY (N speakers = 20)

TT			**TNY**		
Person-Pronoun	**% Overt**	**N verbs**	**Person-Pronoun**	**% Overt**	**N verbs**
2PL PRO	30	27	2PL PRO	33	9
1SG PRO	17	2038	1SG PRO	22	3103
2SG (FORMAL)	13	64	3SG PRO	20	1739
1 PL PRO	11	1732	3PL PRO	20	591
2SG (FAMILIAR)	9	353	1PL PRO	16	1096
3PL PRO	9	710	2SG (FORMAL)	12	25
3SG PRO	6	1708	2SG (FAMILIAR) PRO	7	511
Total		6630	Total		7074
Chi-square: 137.677 P < 0.01			Chi-square: 70.651 p < 0.01		

Overt pronouns occurred at the highest rate with the 2PL, followed by the 1SG for both the TT and for the TNY. As for the plural pronouns, the overt pronoun rate was higher for the 1PL than for the 3PL for the TT. This order was the reverse for the TNY. The two groups also differed with respect to the ranking of the 2SG (familiar) pronoun. Again here, the comparison between factor orders in the two residential groups is presented only for the sake of completeness. No discussion of this will appear in this chapter; the matter will be taken up when a comparative multivariate analysis is given in the next chapter.

7. Person-Verb

In the case of Person-Verb, the relationship between the choice of person inflection and the occurrence of an overt subject pronoun was significant for the entire sample, and for the two subsamples. Here, it is important to note that the ranking of certain subject pronouns that was shown in the previous section does not match the person and number inflection on finite verbs. For example, if the 3SG pronoun ranked as the second factor for a certain sample, we would also expect the 3sg inflection to be ranked as the second factor for that sample. However, the person and number properties of some pronouns do not always match the person and number inflection on verbs in Turkish.

The reason for this is that the person inflection for some singular and plural pronouns can be the same. For example, the person inflection for the 3PL (when they refer to a human and are specific) and for the 3SG pronouns can be the same. Likewise, the person inflection for the 2SG (formal) and 2PL pronouns is the same. The problem is that when coding for these verb inflections, I did not make a distinction as to whether or not the subjects were 3SG or PL and 2SG (formal) or PL pronouns.

We begin by presenting the results for the entire sample.

Table 8.9: Crosstabs: Overt pronoun rate by Person-Verb

Whole sample (N speakers = 40)

Person-Verb	**% Overt**	**N verbs**
1sg	20	4929
2pl or formal	17	124
3sg	15	3480
1pl	13	2768
3pl	10	1204
2sg (familiar)	8	857
Total		13362
Chi-square: 157.277 $p < 0.01$		

The table shows that when the whole sample is taken into account, the overt pronoun rate was the highest for 1sg inflection, followed by 2pl or formal and 3sg inflections. The overt pronoun rate was the lowest for 2sg (familiar) inflection.

We now turn to the overt pronoun rate by the same variable for TT and TNY, bearing in mind that, again here, this is an informational table only; order differences between the groups will be postponed until we can run the comparison on a multivariate analysis.

Table 8.10: Crosstabs: Overt pronoun rate by Person-Verb

TT (N speakers = 20), and TNY (N speakers = 20)

TT			**TNY**		
Person-Verb	**% Overt**	**N verbs**	**Person-Verb**	**% Overt**	**N verbs**
1sg	17	1970	1sg	22	2959
2pl or formal	17	90	3sg	22	1751
1pl	11	1698	2pl or formal	18	34
2sg (familiar)	9	351	1pl	16	1070
3sg	7	1729	3pl	16	550
3pl	6	654	2sg (familiar)	8	506
Total		6492	Total		6870
Chi-square: 123.843 $p < 0.01$			Chi-square: 79.312 $p < 0.01$		

As the table reveals, the rate at which overt pronouns were expressed was the highest for 1sg inflection for both groups. The order of the other factors was different for both groups.

8. Discussion

In this section, I interpret the findings regarding the linguistic variables and how they relate to overt subject pronoun use in Turkish. Further, I address the issue of whether these findings provide any evidence for the language contact hypothesis. Among the linguistic variables that I examined in bivariate analyses, statistically significant effects on the use of an overt subject pronoun were found for Continuity, Person-Pronoun, Person-Verb, TMA II, TMA III, and TMA IV. But as mentioned above, we will only

discuss here those results that are most revealing and easiest to interpret, namely those associated with the variable Continuity. The others will be discussed in the next chapter.

A comparison of TT with TNY with respect to the overt pronoun rates showed that there was an increase in the use of overt pronouns in both the same- and different-reference environments among the TNY. Recall that in the different-reference environment, TNY showed a 24 percent rate, an increase of eight percentage points from the results for TT. They also showed a 15 percent rate, with an increase of nine percentage points, in the same-reference environment. The higher rate of overt pronoun use in the same-reference environment by the TNY strongly supports the hypothesis regarding the influence of English on Turkish as it is spoken in New York. In addition, the reduced contingency coefficient value for TNY suggests that the control over the difference between overt and null subject pronouns in same- and switch-reference environment has diminished among the TNY.

The finding that Continuity is a weaker predictor of pronoun use among TNY than among TT constitutes strong support for a contact hypothesis. We can presume that Turks in New York are more bilingual than those in Turkey. The strategy of favoring overts for switched referents and null for same referents is only to a very limited extent part of the usage repertoire of English speakers, and this disregard for considerations of switch reference in the use of overt pronouns appears to be coloring the usage of Turkish speakers as they become increasingly bilingual.

A number of other interesting points can be drawn from the results with regard to Continuity. In tables 8.1 and 8.2, we find that the use of overt pronouns in different-reference environments was 20 percent for the entire sample, 16 percent for TT and 24 percent for TNY. A comparison of results from other studies of Spanish subject pronoun

variation (Cameron, 1992; Flores-Ferrán, 2002; Ávila-Jiménez, 1996) reveals that the rates obtained in the present study are comparatively very low. For example, in Cameron (1992), overt pronoun use in the switch-reference environment among Puerto Ricans living in San Juan was 57 percent. In Flores-Ferrán (2002), it was 64 percent for Puerto Rican speakers born in NYC. In the study by Ávila-Jiménez (1996), of Puerto Rican speakers living in NYC it was 54 percent. These results, where Spanish rates of overtness in the switch-reference environment are much higher than the ones I found for Turkish, are consonant with the general observation that in Romance languages the rates of overt pronouns increases in the switch-reference environment much more than is the case in Turkish.

The comparatively very low rates of overt subject pronouns found in different-reference environments in Turkish raise the important question of why this is so. I believe the answer is in discourse-pragmatic factors such as *topic change* that has been found in previous qualitative studies to be related to the presence or absence of overt subject pronouns in Turkish. As we discussed in section 2.1 of chapter 2, Enç (1986) argued that one of the functions of overt subject pronouns in Turkish is to either establish the topic of the discourse or change it. Thus the more speakers change the topic of the conversation, the more overt pronouns they use.

In the corpus used for the present Turkish study, contexts where overt subject pronouns are used to signal topic change are also evident. An example of this is provided in (1).

(1) ...1 sene sonra doktora-(y)ı bitir-ecek. **Ben** Rafineri-de-ki
…1 year after doctorate-ACC finish-FUT-3SG. I Refinery-LOC-GEN

çalış-ma-lar-ım-ı 2004 yılı 10 Eylül-ün-de bitir-miş
work-VN-PL-1SG.POSS-ACC 2004 year 10 September-3SG.POSS-LOC finish-PF

ol-du-m. (1TR)
AUX-PF-1SG

'She will finish her doctorate in a year. I retired from the Refinery on September 10, 2004.'

In (1), the informant was talking about her daughter who was going to finish her Ph.D. in a year. He then switched the topic to mention the fact that he retired from the Refinery he was working for. The speaker used an overt pronoun to switch the topic.

During the interviews, most speakers especially those living in Turkey talked about only one topic. If the speakers had frequently changed the topic of the conversation or had introduced new topics, the use of overt pronouns in different-reference environments would have been probably higher. It can be said that the fewer use of overt pronouns for switch in reference subjects is related to the same topic being maintained in the discourse. According to Givon (1983), overt pronoun use increases in interviews if speakers themselves have the control of changing the topics of the conversations, but most of the time the interviewer is the one who decides on the topics to be discussed. This was observed in Ávila-Jiménez (1996) where the relationship between topic change and the use of overt pronouns was studied in Puerto Rican Spanish. In her corpus, the total number of environments in which topics changed was 219, which was also very low. The author attributed this to the fact that the topics to be discussed in the interviews in her study were determined mostly by the interviewer and not by the participants.

This study addressed the relationship between overt pronoun use and same or switch reference, not between overt pronoun use and topic change. That is, for us a switch consisted simply of a verb having a different subject from the previous verb, with no requirement that a new topic be introduced. Consider, again, the example in (1) above, repeated here as (2).

(2) ...1 sene sonra doktora-(y)ı bitir-ecek. **Ben** Rafineri-de-ki
…1 year after doctorate-ACC finish-FUT-3SG. I Refinery-LOC-GEN

çalış-ma-lar-ım-ı 2004 yılı 10 Eylül-ün-de bitir-miş
work-VN-PL-1SG.POSS-ACC 2004 year 10 September-3SG.POSS-LOC finish-PF

ol-du-m. (1TR)
AUX-PF-1SG

'She will finish her doctorate in a year. I retired from the Refinery on September 10, 2004.'

Here, the speaker used an overt pronoun to switch the topic of the conversation as well as the referent. Thus it might be the case that topic change triggered the use of the overt pronoun. Unfortunately, by studying switch reference alone in such contexts where there is a switch in referent as well as a topic change, and vice versa, the relationship between the use of overt pronouns and topic change was neglected. Further study should analyze this relationship and determine whether or not topic change has a larger quantitative effect on the use of overt pronouns than does simple switch reference.

When analyzing this relationship, we should also consider contexts where topic change and switch reference are disconnected, that is topic change occurs without a switch in referent and vice versa. An example of an instance in which topic change occurs without a switch in referent is:

(3) Bu gün-ler-de çok yorul-uyor-um. **Ben** bu sene dışarı-da ol-duğ-um için
This day-PL-LOC very tired-IMPF-1SG. I this year outside-LOC AUX-PART-1SG for

şehir-e hiç gid-e-me-di-m. (2TR)
City-DAT never go-PSB-NEG-PF-1SG.

'I get tired nowadays. Since I was outside [the City] this year, I could not go to the City.'

In (3), the informant first mentioned that she was tired. She then talked about the fact that she could not go to the City this year because she was outside the City. Note that the subject referent remained the same in both clauses.

Conversely, there can be a switch in referent without there being a topic change. This type of disconnectedness is provided in the following example taken from another informant living in Turkey.

(4)… daha ayrılalı çok ol-ma-dı filan de-di-ler. **Biz** hemen
… yet leave very AUX-NEG-PF etc. say-PF-3PL. We immediately

bir taksi-(y)e bin-di-k. (5TR)
a taxi-DAT get on-PF-1PL.

'They said that it has just left. We immediately got on a cab.'

In this conversation, the speaker was telling me about how he and his friend missed the bus while they were traveling from one city to another. They decide to take a taxi after they are told by other taxi drivers to do so. Note that there is a switch in referent in the second clause while the topic of the discourse remains the same.

I would also like to comment on the fact that the percentage point difference between different-reference and same-reference environments for the entire sample, for the TT, and for the TNY is statistically significant (and for the whole sample it is quite

large) but that it is not very large when compared to the studies on Spanish and other Romance languages. It may be the case, then, that the switch-reference versus same-reference distinction, which is a major piece of evidence supporting the English contact hypothesis, is not in and of itself a major consideration in the decision to use a null or an overt in Turkish, or at least not as major as it is in other languages. Still, for our purposes here, which have to do with establishing English contact influences, this remains an important piece of evidence. Also, the results of the multivariate analysis, which will be shown in the next chapter, indicate that Continuity was the top variable in the variable hierarchies of all the groups in the study, suggesting that Continuity accounted for the most variance between nulls and overts for all the groups. This variable is playing a relatively important role in the choice between null and overt pronouns whether or not the spread between the same- and different-environments is not very high.

Further, the gap between the different- and same-reference environments for the entire sample and the subsamples remains roughly the same. At this stage, it is not possible to provide a clear explanation for the fact that the gap between same- and different-reference environments remained more or less the same for the subsamples in this study. However, it can be said that if the use of overt pronouns in different-reference environments had been higher, the percentage point gap between the same- and different-reference environments might have been much bigger.

The general finding regarding (a) increased overall rates from Turkey to New York and (b) the specific finding regarding the diminished role of switched reference in making decisions regarding the choice between nulls and overts among all New Yorkers, constitute the strongest support that can be drawn from these results for the proposal that Turkish pronominal usage in New York is clearly under the influence of English.

Continuity of reference, as discussed in chapter 2, refers to the reference relations between two NPs. When the NP of the second clause has the same referent as the NP of the immediately preceding clause, these two NPs are ‘same’ in reference, and when these two NPs have different referents, they are ‘switch’ in reference (Cameron, 1995: 4).

CHAPTER 9

RESULTS: MULTIVARIATE ANALYSIS OF LINGUISTIC VARIABLES

1. Introduction

In the previous chapter, I have presented in separate bivariate analyses the results of variables that have a statistically significant impact on the occurrence of overt SPPs in the Turkish spoken by all informants (N = 40), Turks in Turkey (TT, N = 20) and Turks in NYC as a whole (TNY, N = 20). Of further interest is to find out which of these significant variables have greater and lesser impacts on the appearance of overt pronouns, and within each variable, which factors are ranked higher and lower in terms of this impact.

The multivariate results presented in this chapter are, like this whole study, intended to replicate the analysis of Spanish pronouns in Otheguy et al. (2007). The results in this chapter were obtained by using binary logistic regression. The regressions simultaneously included the variables that we have called Continuity, Person-Verb, TMA II, TMA III and TMA IV. I included in the regressions all the variables except for the following. I excluded TMA I in finite verbs and all the variables dealing with TMA suffixes with non-finite verbs and with non-verbal predicates, because there were no statistically significant relations between these variables and the use of pronouns for the entire sample and the subsamples. In addition, Person-Pronoun was not included in the regression for the reason that it was strongly correlated with Person-Verb, and that such highly correlated variables tend to distort the result of regression runs.

As with other statistical analyses in the present study, the regressions in this chapter were run using the Statistical Package for the Social Science (SPSS). As Bayley

(2002b), Otheguy et al. (2007), and other researchers have shown, SPSS results are fully equivalent to those obtained using the other program found in the literature, namely VARBRUL. This chapter presents both hierarchies of variables and hierarchies of constraints within certain variables. I present the variable hierarchies first.

2. Variable hierarchy for the entire sample

A variable hierarchy, in quantitative sociolinguistic research, ranks all variables according to how great an influence they exert on the dependent variable, in this case, the use of an overt pronoun; or, to put it more precisely, according to how much variance they account for, in this case, how much variance between nulls and overts they account for. In Otheguy et al. (2007), the variable hierarchies were shown to be mostly the same in the different Spanish region-of-origin and generational groups. As part of my replication of the Spanish study, I also wanted to explore whether my different residential groups, TT and TNY, share the same variable hierarchies. To create a variable hierarchy, I adopted binary logistic regression, which allows us to predict an outcome from a set of independent variables and to rank their relative strength with respect to the dependent variable.

I begin by presenting the variable hierarchy for the entire sample. The ranking of variables, as produced when the data is subjected to regression analysis in SPSS, is according to the Wald coefficient, which gives a value, relative to all the other values, of the amount of variance that is accounted for by the variable. Note, then, that whereas the values in constraint hierarchies (which we will see below) are expressed in terms of Exp(B), which gives the odds of occurrence of the dependent variable when the

constraint is present, the values in variable hierarchies are expressed in terms of Walds, which rank how much variance the entire variable accounts for.

Table 9.1: Logistic regression: Variable hierarchy regarding overt pronoun use. Whole sample (N speakers = 40)

Whole sample		
Rank	**Variable**	**Wald**
1st	Continuity	189.559**
2nd	Person-Verb	169.071**
3rd	TMA IV	18.239**
4th	TMA III	15.755*
5th	TMA II	4.254*
N verbs = 13727		
** = p <0.01 * = p < 0.05		

The results shown in the table indicate that Continuity accounts for the most variance between nulls and overts, that is, Continuity is the best predictor of the appearance of an overt pronoun for the entire sample. Continuity is followed by Person-Verb, TMA IV, TMA III and TMA II. Recall that the TMA slots are labeled, from left to right, TMA I, TMA II, etc., so that TMA I is the slot next to the stem and TMA IV the one farthest from the verb stem. These results show an 'onion effect,' that is, the choice of suffix in the right-most (farthest from the verb stem) slot has the strongest impact on the choice of null or overt pronoun.

In the literature it has been reported that morpheme ordering within verbs arise from semantics (Bybee, 1985). That is, morphemes whose meaning is more relevant to

the basic meaning of the verb occur closer to the verb. Bybee proposed a relevance scale where the categories of voice, aspect, tense, mood and agreement are ranked for relevance to verbs in that order. For example, passives or causatives tend to appear more closely connected to verb lexical semantics than factive mood marking. If this is true for Turkish, then the reason why the suffixes in Turkish in the order that they are, is because TMA II is more closely connected to a verb in some sense than, say, TMA III and TMA IV. Thus, we would expect the results of the multivariate analysis to be in that direction. However, they showed that the best predictor of the probability of an occurence of an overt pronoun for the entire sample is just the opposite, TMA IV which, under Bybee's theory, has the least semantic relevance to the verb. Even though I am not able at the moment to provide a clear explanation for this result, this lack of congruence between semantic affinity with the stem and predictive power for the pronoun is worthy of note.

3. Variable hierarchies for TT and TNY

Next, I show variable hierarchies separately for TT and TNY, comparing the two. The results are presented in the following table.

Table 9.2: Logistic regression: Variable hierarchies regarding overt pronoun use

TT (N speakers = 20), and TNY (N speakers = 20)

TT			**TNY**		
Rank	**Variable**	**Wald**	**Rank**	**Variable**	**Wald**
1st	Continuity	112.813**	1st	Continuity	86.571**
2nd	Person-Verb	105.006**	2nd	Person-Verb	76.711**
3rd	TMA IV	9.168**	3rd	TMA III	25.861**
4th	TMA III	7.998	4th	TMA IV	4.937*
5th	TMA II	2.771	5th	TMA II	2.259
N verbs = 6637			N verbs = 7090		
** = p < 0.01 * = p < 0.05					

We now see from the results of this table that the two Turkish groups have essentially the same variable hierarchies, as do the groups in Otheguy et al. (2007). As reflected on the table, Continuity followed by Person-Verb, is the best predictor for both TT and TNY. In addition, Continuity, Person-Verb, and TMA IV significantly account for variance among both TT and TNY. And in both populations TMA II is not significant. The only difference between the two populations is in TMA III and TMA IV, which exchange places among the TNY. Also, TMA III, which was significant among TNY, was not relevant to the behavior of TT.

4. Constraint hierarchy for the Person-Verb variable for the entire sample

The Spanish pronoun study being replicated not only compared the subsamples with regard to variable hierarchies, but also with regard to constraint hierarchies, where the variable that we call Person-Verb proved to be the strongest variable for both newcomers and second generation speakers of Spanish in the U.S. When these two

Spanish-speaking groups were compared with respect to this variable, it was found that the order of the factors (constraints) within Person-Verb was different across the two generations. In the present Turkish study, this variable was also found to be significantly associated with variation in the use of pronouns. Therefore, I compare here the subsamples of Turkish speaker with respect to Person-Verb, to see whether or not there are differences between them.

Otheguy et al. (2007) defined a constraint hierarchy as "a listing of the change in likelihood of appearance of the overt pronoun associated with each constraint of an independent variable, expressed as an Exp(B) estimate, in order of descending strength" (2007:38). A constraint hierarchy ranks all constraints according to how great an influence they have on the use of an overt pronoun. Such a hierarchy is obtained from the results of a binary logistic regression run that allows us to rank the strength of the factors within a variable with regard to the dependent variable when all the factors of all the variables included in the regression are considered at the same time.

We begin by looking at the constraint hierarchy for the variable Person-Verb for the whole sample. The persons of the verb are indicated on the table as '1sg' for first-person singular inflections, '2sg' for second-person singular inflections, etc. The second-person plural, which is also second-person formal, whether singular or plural, is however simply indicated as '2pl.' The reader should thus bear in mind that '2pl' means *second-person formal whether singular or plural*. Similarly, '2sg' means *2sg familiar*. Two asterisks next to the Exp(B) value for each constraint indicate that the value is significant at the level of $p < 0.01$; one asterisk indicates that the significance value is at $p < 0.05$.

Table 9.3: Logistic regression: Hierarchy of constraints within Person-Verb. Whole sample (N speakers = 40)

Whole sample		
Person-Verb	**N Verbs**	**Exp (B)**
1sg	4240	1.789**
2pl	109	1.523*
3sg	2267	1.075
1pl	2503	.935
3pl	1082	.715**
2sg	699	.511**
* = p < 0.05 ** = p < 0.01		

The Exp(B) value shows the change in the probability of an overt pronoun occurring when the particular constraint or factor is present. The different person inflections are ranked in descending order based on their Exp(B) values. Exp(B) values above 1.00 mean that the factor favors overts; Exp(B) values below 1.00 mean that the constraint disfavors overts (i.e., it favors nulls). The table shows that 1sg inflection is the strongest predictor of the appearance of an overt pronoun for the entire sample. This is followed by the 2pl inflection. Verb endings for the 3sg and 1pl yielded statistical values for Exp(B) that are not significant, which means that these person inflections neither favor nor disfavor the appearance of overts. Verbs with 3pl and 2sg inflections significantly disfavor overts, that is, they favor nulls.

5. Constraint hierarchies for the Person -Verb variable for TT and TNY

We now turn to a comparison of the Person-Verb constraint hiearchies of TT and TNY. The results in Table 9.4 are the same as in Table 9.3, but now broken down by

residential group. The results for each group come from separate regression runs containing the same predictor variables.

Table 9.4: Logistic regression: Hierarchies of constraints within Person-Verb

TT (N speakers = 20), and TNY (N speakers = 20)

TT			**TNY**		
Person-Verb	**N verbs**	**Exp (B)**	**Person-Verb**	**N verbs**	**Exp (B)**
2pl	80	2.070**	1sg	2527	1.616**
1sg	1713	1.924**	3sg	1029	1.504**
1pl	1536	1.061	2pl	29	1.195
2sg	262	.707	1pl	967	.930
3sg	1238	.634**	3pl	484	.926
3pl	598	.528**	2sg	437	.400**
* = $p < 0.05$ ** = $p < 0.01$					

As shown in the table, the constraint hierarchies are different in TT and TNY. A constraint by constraint analysis shows the following:

- **3sg**, in perhaps the biggest change, favors nulls in TT but favors overts in TNY.
- **2pl** is the first-ranked constraint in TT but has dropped to third place and become non-significant in TNY.
- **1sg** is the second-ranked constraint in TT but has moved to first place in TNY.
- **2sg** is not significant in TT but favors nulls in TNY.
- **3pl** favors nulls in TT but is not significant in TNY.
- **1pl** is not significant in either TT or TNY.

6. Discussion of hierarchies and supplementary analyses

In this section, I discuss the implications of the findings for the questions of whether Turkish in NY has undergone changes with respect to Turkish in Turkey, of whether these changes are to be attributed to contact influence from English, and of whether, and in what way, Turkish pronouns in New York are comparable to the pronouns of Spanish in the study being replicated. It should be noted that some of the ideas that I discuss in Section 6.2 below, when an explanation for findings regarding constraint hierarchies is offered, come from discoveries about TNY that I made during the course of this study and therefore I do not have in all cases the further data that would be needed to secure an explanation, so my explanations have to be tentative. Still, the initial findings here prompted me to perform further detailed analyses. These results make it possible for a clear picture to start to emerge regarding not only the role of English in general but the specific ways in which Turkish usage appears to be Anglicized in New York.

6.1 Variable hierarchies for TT and TNY

Following the approach of Otheguy et al. (2007), the first analysis in this chapter involved comparing the variable hierarchies of TT and TNY. We showed that TT and TNY do not have identical hierarchies, but that they are very similar. This finding is essentially the same as that of the Spanish study, where the variable hierarchies for Hispanic newcomers and for New York born-or-raised Hispanics (NYBR) were nearly identical (the variables of the Spanish study were similar to the ones analyzed here).

Within this general framework of similarities between the two studies, it is of great interest to determine in which language we find the greater or lesser similarities between the two groups being compared, namely newcomers vs. second generation speakers in the Spanish case, and TT vs. TNY in the present study. Therefore, a detailed comparison of the variable hierarchies in the two studies is needed.

In the Spanish study, speakers were classified into two groups: Caribbeans, which included Puerto Ricans, Dominicans and Cubans and Mainlanders, which included Columbians, Mexicans and Ecuadorians. An analysis of the variable hierarchy of Caribbean newcomers and of their NYBRs revealed that the two groups shared similarities with regard to the order and strength of the variables that are called here Person-Verb, Continuity and Tense. Person was the best predictor of the occurrence of overt pronouns in Spanish, followed by Connect (Continuity) and Tense. The two groups, however, differed from one another with regard to the variables that followed Tense. Almost all of these variables were placed in different positions, although some of them were in the same descending order such as Clause, Lexical, Genre and Specific. In addition, the variables Set Phrase and Reflexive, which were significant for newcomers, lost their significance for NYBRs.

When Mainland newcomers and their NYBRs were considered, the top two variables for both groups were, again, Person and Connect. However, all the other variables that followed Connect were in different positions and had different strengths. For example, Specific, which was not relevant to the behavior of the newcomers, was significant for NYBRs. Also, Lexical was significant for newcomers, was not significant for NYBRs.

Turning to the variable hierarchies of TT and TNY, we find that Continuity and Person-Verb are also the top two variables. The order of the two variables is not the same as that of the Spanish study, but both variables are significant for both TT and TNY. This indicates that for both Spanish and Turkish, the occurrence of an overt or a null pronoun is most influenced by (a) the person of the verb and (b) whether the subject referent of a verb is the same as or different from the subject referent of the previous verb. Thus, the Spanish spoken by newcomers and second generations and the Turkish spoken by TT and TNY resemble each other considerably in the use of subject pronouns with regard to the two variables. Regarding the differences between TT and TNY, the only difference is in TMA III, which ranked fourth for TT but was placed in the third position for TNY and whereas it was non-significant for TT, it was found to be significant for TNY. It thus clear that that the differences between TT and TNY are lesser than the differences between Caribbean newcomers and their NYBRs and between Mainland newcomers and NYBRs.

The reason that the Latin American newcomers differ from NYBRs somewhat more than TNY have diverged from TT may lie in the fact that the Spanish study analyzed ten variables whereas the Turkish study analyzed only five. We might have found more differences between TT and TNY if more variables had been involved in the study. Added to this is the different type of groups involved, that is, the Spanish study involved generational groups in NY whereas the present study focused on country of residence groups. In Otheguy et al. (2007) there was no one from the home countries truly equivalent to the TT in the present Turkish study. The Hispanic newcomers who arrived in NYC at age 17 or older, and have lived in the City for five years or less, are

neither equivalent to the TT, who spoke little or no English and never lived in the U.S., nor equivalent to the first generation TNY1, many of whom have lived in the U.S. for up to 40 years.

The only similarity between the two studies is that both studies involved second generation speakers who were born in NY or came to NY at or before the age of 3. However, the two generational samples differed from one another with regard to their size. Whereas there were 28 NYBR speakers in the Spanish study, there were only 10 in the Turkish study. Moreover, the second generation sample in the Turkish study, unlike the NYBRs in the Spanish case, could not be analyzed as a separate group due to its small sample size. If there had been as many second generation speakers as those in the Spanish study, we might have observed more differences between TT and TNY with regard to their variable hierarchies. We now proceed to discuss the constraint hierarchies for the Person-Verb variable.

6.2 Constraint hierarchies for TT and TNY (Person-Verb variable)

Turning now from discussion of variable hierarchies to constraint hierarchies, the second analysis in this chapter involved comparing the constraint hierarchies of TT and TNY with regard to Person-Verb. We discovered that, as was the case in the Spanish study, the two country of residence groups in the Turkish study were different with respect to the ranking of the constraining factors that articulate this variable. The results show that the use of Turkish in NY is not the same as in Turkey, as the factors that guide the use of pronouns in NY are ranked differently than in Turkey and, in some cases, factors that played no role in Turkey do so now in NY.

Given that TT and TNY have different constraint hierarchies, it can be said that TT and TNY have membership in two communities (Tagliamonte & Poplack, 1999; Poplack & Sankoff, 1987). In Guy's (2005) terms, this may also mean that these two groups have two different grammars for their variable component. According to Guy, speakers have knowledge of the fact that they are using a certain structure in a certain context and at a certain rate and they know which hierarchies of constraints have an impact on the use of a certain structure. Thus, constraint hierarchies are part of speakers' grammatical knowledge and differences in constraint hierarchies between different speech communities are grammatical differences. From Guy's point of view, I have shown that in Turkish, as well as in Spanish, the type of grammatical change that has to do with the variable component of languages can take place very rapidly. It is reasonable to assume that my TNY had, prior to immigration, a grammar like the one of my TT, and that this grammar has changed in the short span of time involved in the immigration experience.

Here, it is also important to mention Newmeyer (2003) who argues for a sharp distinction between our knowledge of language and how we use it. He states that "speakers mentally represent full grammatical structure, however fragmentary their utterances might be" (p. 682), suggesting that the fact that speakers use structures differently when they speak does not mean that the mental representation of structures is changing. Speakers make use of structures without changing the mental representation of grammar. With respect to sociolinguistics, he adds that probabilities associated with constraints are outside the grammar and differences in hierarchies between speech communities are just differences in usage and not differences in grammar. From

Newmeyer's point of view, I have demonstrated that usage involving grammatical forms like pronouns changes under the conditions of immigration, but I have not shown any changes in grammar.

With respect to the question regarding English influence, the most important result is that, within the Person-Verb variable, the significance and the Exp (B) values of the 3sg constraint of the TT turned out to be different for TNY. Whereas the 3sg inflection significantly favors nulls in TT, it significantly favors overts in TNY. It is quite possible that this finding may be due to the use by TNY of the 3sg inflection with overts for inanimate references. During the interviews, I realized that TNY were using 3sg with overts for inanimates more than did TT. Below are examples taken from the speech of a first generation speaker living in NYC. As before, overt 3sg pronouns are in bold.

(1a) **O** yani öyle. (6NY)
It I mean so.
'It is so.'

(2a) **O** o zaman biraz stresli ol-uyor. (6NY)
It then a little bit stressful AUX- IMPF-3SG.
'Then it is a little bit stressful.'

(3a) **O** tatil gibi zaten. (6NY)
It vacation like anyway.
'It is like vacation, anyway.'

Examples of these were also evident in the speech of second generation speakers living in NY. From informant number 25 we have:

(4a) Onun için hemen Türkçe konuş-uyor-lar ben-im-le. **O**
Therefore immediately Turkish speak-IMPF-3PL I-GEN-INS. It

birazcik zor ol-uyor bazen. (25NY)
a little bit hard AUX- IMPF-3SG sometimes

'Therefore, they immediately start speaking Turkish with me. Sometimes, it is a little bit hard.'

(5a) **O** o-nun Türkçe-si-ni tam doğru dürüst Türkçe
It he-GEN Turkish-3SG.POSS-ACC exactly properly Turkish

konuş-ma-dığ-ı için. (25NY)
speak-NEG-CV-3SG.POSS for

'It is because he did not speak [his] Turkish properly.'

In each of these examples from TNY, the equivalent sentences from TT would more likely have been uttered with a null pronoun.

(4b) Onun için hemen Türkçe konuş-uyor-lar ben-im-le. **Ø**
Therefore immediately Turkish speak-IMPF-3PL I-GEN-INS. (It)

birazcik zor ol-uyor bazen.
a little bit hard AUX-IMPF-3SG sometimes

'Therefore, they immediately start speaking Turkish with me. Sometimes, it is a little bit hard.'

(5b) **Ø** o-nun Türkçe-si-ni tam doğru dürüst Türkçe
(It) he-GEN Turkish-3SG.POSS-ACC exactly properly Turkish

konuş-ma-dığ-ı için.
speak-NEG-CV-3SG.POSS for

'It is because he did not speak [his] Turkish properly.'

As we also showed in section 1.1 of chapter 3, it is grammatical to use overts for inanimates in Turkish. That is, the Turkish pronoun *O* is properly equivalent not only to

English *he* and *she*, but also to English *it*. But whereas the general tendency in Turkey is to refer to inanimates with nulls, it has become common to do it with overts in NY for the reason that the use of 3sg with overts for inanimates is categorical in English, using *it*. In order to establish definitively the effect of the use of overts for inanimates on the changes in constraints that we have discovered for TNY, I compared the proportion of overt pronouns used for inanimate referents in TT and TNY. The results are in Table 8.5.

Table 9.5: Crosstabs: Overt pronouns for inanimate reference in TT and TNY

TT (N speakers = 20), and TNY (N speakers = 20)

Residence	**% Overt**	**% Null**	**N verbs**
TT	6	94	332
TNY	29	71	586
Total		918	
Chi-square: 68.222 P < 0.01			

As the table illustrates, the rate of overts with inanimates is higher for TNY than the TT. In Turkey, inanimates prompt the use of 3sg overts but rarely (in 6 percent of cases) whereas in New York, inanimates induce the use of 3sg overts in nearly one third of the cases (29 percent). It appears, then, that the much stronger influence that 3sg inflections exert over the occurrence of an overt in New York than in Turkey (in Table 9.2, 3sg inflections in general favor null in TT but overts in TNY) can reasonably be attributed, to a considerable extent, to the handling of inanimate referents. The model of overt English

it has taken a grammatical but rarely used option in Turkish and expanded it to become normal usage in New York.

The 3pl constraint was also different between the TT and the TNY. This factor significantly disfavored overts (that is, it significantly favored nulls) for TT but had increased its favoring of overts in TNY, moving up in the scale in such a way that now, among the TNY, it no longer significantly disfavors overts but has become non-significant. The explanation here, parallel to the explanation having to do with inanimates for 3sg, may have to do with non-specific. It might be the case that the TNY have increased their use of 3pl with overts for non-specific reference, and decreased the use of nulls, which is the more common way to refer to non-specific 3pl in the Turkish spoken in Turkey. This may be attributed to the fact that English usage allows for non-specific references with overts. Once again, it seemed to me while I was conducting the interviews that TNY were using 3pl with more overts for non-specific reference than TT did. The following example taken from the speech of a second generation speaker living in NY illustrates such a case.

(6a) …yoksa konuş-ma-n lazım şirket-ler-le ilgili falan,
… otherwise talk-VN-2SG have to company-PL-INS about etc.

yatırım-lar-la ilgili ve orda hala takıl-ıyor-um.
investment-PL-INS about CONJ there still get stuck-IMPF-1SG

Ama ben İngilizce konuş-uyor-um **onlar** Türkçe konuş-uyor-lar. (25NY)
but I English speak-IMPF-1SG they Turkish speak-IMPF-3PL

'…otherwise you have to talk about companies and investments etc. and I still get stuck there. But I speak English; they speak Turkish.'

(7a) … bir yer-e gid-iyor-sun, insan-lar-la konuş-ma-n
…. One place-DAT go-IMP-2SG people-PL-INS talk-VN-2SG

gerek-ebil-iyor ve **onlar** anla-mı-yor-lar ben-i. (25NY)
have to-PSB-IMPF-3SG CONJ they understand-NEG-IMPF-3PL I-ACC

'…you go to a place where you may have to talk to people and they do not understand me.'

These sentences in NY are grammatical, but they are unlikely to have had overts in TT, as is the case with the use of 3sg with nulls for inanimates. If these sentences had been uttered in Turkey, they would have been more likely uttered with a null instead of an overt pronoun.

(7b) … bir yer-e gid-iyor-sun, insan-lar-la konuş-ma-n
….One place-DAT go-IMP-2SG people-PL-CONJ talk-VN-2SG

gerek-ebil-iyor ve **Ø** anla-mı-yor-lar ben-i.
have to-PSB-IMPF-3SG CONJ (they) understand-NEG-IMPF-3PL I-ACC

'…you go to a place where you may have to talk to people and they do not understand me.'

Again here, in order to provide definitive explanations, I compared the use of 3pl overts for non-specifics by TT and TNY. The results are in the table below.

Table 9.6: Crosstabs: Overt pronouns for non-specific reference in TT and TNY

TT (N speakers = 20), and TNY (N speakers = 20)

Residence	**% Overt**	**% Null**	**N verbs**
TT	2	98	257
TNY	10	90	180
Total		437	
Chi-square: 13.773 P < 0.01			

The table shows that TNY as a whole used 3pl with overts that make non-specific reference at a higher rate than TT. As was the case with the use of the 3sg for inanimate reference, the use of overt 3pl for non-specific reference is very rare among TT, but has become common among TNY. It can be reasonably concluded, then, that the increased use of overt 3pl pronouns among TNY is partly due to the different strategies used to deal with non-specific referents.

Similar findings were also reported in a Spanish study conducted by Otheguy and Lapidus (2005). The authors compared speakers who were born in NY with newcomers in NY with regard to the use of the overt 3pl pronoun *ellos* for non-specific reference, which is grammatical, but rarely used in non-contact Spanish. The results indicated that NY born-or-raised speakers used overt 3pl *ellos* for non-specifics much more frequently than among newcomers.

Another difference between Turks in Turkey and Turks in NY was the place in the hierarchy of 2sg (familiar), which was nonsignificant among TTs but significantly

disfavored overts (that is, it favored nulls) among TNYs. Unlike the previous two cases, an explanation cannot be provided for this surprising result.

The general finding regarding the constraint hierarchy for the whole sample is also worth mentioning. We discovered that 1sg significantly predicts overt pronoun use for the whole sample. The Exp(B) value of 1.78 indicates that the odds of occurrence of an overt SPP with verbs inflected for 1sg is greater than the odds of occurrence of an overt with other type of verbal inflection. This is an unexpected result considering the fact that 1sg pronouns are easier to recover than the other pronouns.

A possible explanation that has been offered as to why one would expect overt 1sg pronouns to be favored more than the other pronouns is that speech is egocentric and that when people talk about themselves, they emphasize themselves by virtue of using the overt more than when they talk about others. In other words, people focus more on the subject when the subject is them. Considering the fact that speakers find themselves more interesting than anything else in the world and thus talk about themselves (Dahl, 1997), egocentric expressions such as 1sg pronouns are more likely to be used in conversations than any other pronouns.

In a sociolinguistic study where prosody and pauses were excluded, Dahl (1997) investigated 30-minute recordings of conversations obtained from Swedish speakers. A smaller corpus of English conversation chosen from the published version of the London-Lund corpus served as a baseline for the study. The results indicated that 1sg pronouns occurred ten times more in the conversational corpus than in the newspaper corpus. In Biber (1988), the use of overt 1sg pronouns in conversations was even higher. Dahl (1997) concluded that an important discourse-grammatical difference between

conversations and written language is egocentric expressions. Silva-Corvalán (1994) reported similar results for the expression of overt 1sg pronouns in Spanish. Among the Mexican-born speakers who immigrated to the U.S. after the age of eleven, the 1sg pronoun *yo* 'I' was more frequently used than the other pronouns. The author concluded that *yo* has a pragmatic function because the speakers used it to refer to themselves overtly in conversations.

7. Summary and conclusions

In this chapter, we performed regression analyses on the linguistic variables that were found to have significant effects on the variable use of subject pronouns in the separate bivariate analyses presented in the previous chapter. These linguistic variables are Continuity, Person-Pronoun, Person-Verb, TMA II, TMA III and TMA IV in finite verbs.

A comparison of the variable hierarchies of newcomers and NYBRs in the Spanish study and of TT and TNY in the present Turkish study reflected that there were more similarities between Turkish TT and TNY than between Spanish newcomers and NYBRs. This result might be due to the fact that the variables analyzed in the Turkish study were fewer than those examined in the Spanish study. In addition, the informant groups involved in both studies were not exactly equivalent to each other, since none of the Spanish groups lived in the home countries as do our TTs.

Turning to the study of constraint hierarchies, in Otheguy et al. (2007), the variable Person-Verb ranked at the top of all the variables studied and accounted for the most variance in both newcomers and second generation speakers of Spanish in the U.S.

Within this variable, the order of the constraints was different between these groups. Parallel to the finding in Spanish, I have also found a difference between TT and TNY with regard to their constraint hierarchies within the Person-Verb variable. That is, the order and strength of factors within Person-Verb turned out to be different for the two groups. This finding supports two different theoretical interpretations. According to Guy (2005), this indicates that when this variable is considered, TT and TNY have two different grammars and therefore, they are two different groups. To Tagliamonte and Poplack (1999) and Poplack and Sankoff (1987), it means that TT and the TNY have two memberships in two different communities.

Some of the findings regarding the constraint hierarchy for the TNY point towards a contact hypothesis. The fact that the 3sg inflection significantly favored nulls in TT whereas it significantly favored overts in TNY and the 3pl constraint significantly disfavored overts for TT but had increased its favoring of overts in TNY appears due in great measure to the categorical use of 3sg inflection with overts for inanimates and of 3pl inflection with overts for non-specific references in English.

Further, the study of the constraint hierarchy for the whole sample revealed that 1sg inflection significantly predicts overt pronoun use and is ranked first. Further, the Exp(B) value of this constraint indicates that overt pronouns were most favored by the 1sg inflection. This was unexpected since 1sg pronoun is the most easily recoverable pronoun. This finding could be attributed to the fact that people use the 1sg pronoun more than the others when the topic of the conversation is about themselves (Dahl, 1997).

To sum up then, it can be said that the results regarding the constraint and variable hierarchies in the present chapter tend to replicate, with some qualifications, those

obtained from the Spanish study (in much the way that the bivariate results of the previous chapter had shown that in Turkish in New York, as in Spanish in New York, there are large increases in the use of overt pronouns, largely due to the role of English in both communities). In both studies the variable hierarchies show many fewer differences than the constraint hierarchies; in both there are important generational or residential group differences in the Person-Verb hiearchies; and in both these differences point to areas, such as the treatment of indefinites, where English appears to be playing a role. We have gone well beyond the Spanish study, however, in showing the importance for an English contact hypothesis of the treatment of inanimates. And we have shown that the differences between the study fit well with the differences in the informant groups.

APPENDIX A

LINGUISTIC VARIABLES AND THEIR CONSTRAINTS

1. Introduction

I analyzed the influence of twelve linguistic variables and the factors within these variables on the use of an overt or a null SPP. As mentioned in section 1 of chapter 4, the term *factor* in variationist sociolinguistics refers to the levels of a dependent or an independent variable. Specifically, each clause was coded for the following linguistic variables and their factors.

1.1 Person and number of the null or overt SPP

The factors of this variable are the following:

- First person singular pronoun

(1) **Ben** şimdi alış-tı-m. (22NY)
I now get used to-PF-1SG.
'I have now got used to [it].'

- Second person singular pronoun (familiar)

(2) Herkesi **Ø** davet et-me-z-se-n… (29NY)
Everybody (you) invite AUX-NEG-AOR-COND.C.-2SG
'If you don't invite everybody…'

- Third person singular pronoun

(3) **O** büyü-dü. (22NY)
She grow-PF.
'She has grown up.'

- First person plural pronoun

(4) **Biz** beraber-iz. (22NY)
We together-1PL.
'We are together.'

- Second person singular pronoun (formal)

(5) **Siz** bil-ir-siniz. (21NY)
You know-AOR-2SG.
'You know.'

- Second person plural pronoun

The same person marker is used for 2sg (formal) and 2 pl.

(6) **Siz** bil-ir-siniz.
You know-AOR-2SG/2PL
'You know.'

- Third person plural pronoun

(7) **Onlar** yan-ım-a gel-ebil-ecek-ler. (13TR)
They side-1SG.POSS-DAT come-PSB-FUT-3PL.
'They will be able to come to me.'

- 'One' and 'people'

(8) İnsan-lar anla-mı-yor-lar.
People-PL understand-NEG-IMPF-3PL
'People do not understand.'

1.2 Person and number ending of the verb

- First person singular
- Second person singular (familiar)
- Third person singular
- First person plural
- Second person plural or formal
- Third person plural

1.3 TMA I in finite verbs

In Turkish, the different slots of a finite verb are occupied by a number of inflectional affixes that express tense, mood and aspect (Göksel & Kerslake, 2005; Sezer, 2001; Cinque, 2001). Some of the suffixes indicate only modality whereas others may express both tense and aspect. For example, the morpheme, *–DI* that appears in the third slot of a finite verb indicates both past tense and perfective aspect.

In the present study, the different slots of a finite verb, namely TMA I, TMA II, TMA III, and TMA IV, are the variables and the inflectional suffixes that appear in these slots are the constraints. Some slots are occupied by more than one inflectional suffix. The first slot of a finite verb is TMA I and the factor of this variable is the modality suffix, *-(y)A.*

- *-(y)A*

This slot takes only the modality suffix, *-(y)A* that expresses ability and occurs only in negative verb forms (Göksel & Kerslake, 2005).

(9) … **Ø** yetiş-e-me-y-iz. (4TR)
… (we) reach-PSB-NEG-1PL.
'We cannot catch [it].'

1.4 TMA II in finite verbs

- -(y)Abil
- -(y)Iver
- -(y)Agel
- -(y)Ayaz
- -(y)Akal
- -(y)Adur

The modality suffixes that occur in the second slot of a finite verb consist of one of the following verb roots: *bil-* 'know,' *ver-* 'give,' *yaz-* 'make a mistake,' *dur-* 'stay,' and *kal-* 'remain.' The suffix, *-(y)Abil,* expresses possibility; *-(y)Iver* expresses swiftness and suddenness of an action; *-(y) Agel* expresses habitual actions; *-(y)Adur* expresses continuous actions; *-(y)Akal* is usually combined with verbs such as *bak-* 'look,' *don-* 'freeze,' and *şaş-* 'be surprised.' When combined, these verbs have the meaning of 'be amazed,' 'be stunned' and 'be dumbfounded.' The possibility suffix, *-(y)Abil* is more commonly used than the other suffixes (Göksel & Kerslake, 2005). An example of *-(y)Abil* is provided in (10).

(10) **Onlar** yan-ım-a gel-ebil-ecek-ler. (13TR)
They side-1SG.POSS-DAT come-PSB-FUT-3PL.
'They will be able to come to me.'

1.5 TMA III in finite verbs

The factors of this variable are the following suffixes:

- -DI (perfective)
- -mIş (perfective/evidential)
- -sA (conditional)
- -(A/I)r/-z (aorist)
- -(y)AcAK (future)
- -(I)yor (imperfective)
- -mAlI (obligative)
- -mAktA (imperfective)
- -(y)A (optative)

An example of -DI (perfective) is:

(11) **Ø** üç sene orda oku-du-m. (7TR)
(Ben) three years there study-PF-1SG.
'I studied there for 3 years.'

1.6 TMA IV in finite verbs

This slot is occupied by different copular markers.

- -(y)DI (past copula)
- -(y)mIş (evidential copula)
- -(y)sA (conditional copula)

An example of past copular is given below.

(12) **Ø** 180 lira-ya mı ne al-mış-tı-m yani o ceket-i. (4TR)
(Ben) 180 lira-DAT INT what buy-PF-P.COP-1SG I mean that jacket-ACC.
'I mean I had bought that jacket for 180 Turkish liras.'

1.7 TMA V in finite verbs

- -DIr (generalizing modality marker)

The generalizing modality marker, *–DIr,* which is the last marker in a finite verb, is most commonly used in formal writing and official announcements. Since this suffix was rarely found in the informants' transcripts, I excluded it from the analysis.

1.8 TMA I in non-finite verbs

Like the finite verbs, the first slot of a non-finite verb is occupied by the ability affix, *-(y)A.* It occurs only in negative non-finite verb forms (Göksel & Kerslake, 2005).

- -(y)A

(13) İlkinci dönem-e **Ø** final-ler-e yetiş-e-me-diğ-i için… (7TR)
Second semester-DAT (O) final-PL-DAT reach-PSB-NEG-VN-3SG.POSS for.
'Since he could not make it to the second semester, to the finals…'

1.9 TMA II in non-finite verbs

The modality suffixes that occur in the second slot of finite verbs also occur in the same slot of non-finite verbs.

- -(y)Abil
- -(y)Iver
- -(y)Agel
- -(y)Ayaz
- -(y)Akal
- -(y)Adur

An example of the possibility suffix, *-(y)Abil* is given below.

(14) … **Ø** oraya geri dön-ebil-diğ-imiz-de…
…(Biz) there back go back-PSB-CV-1PL.POSS-time.
'…when (we) were able to go back there…'

1.10 TMA III in non-finite verbs

- - DIK
- -(y)AcAK
- -(y)IncA
- -(A/I)r…-mAz
- -mAdAn
- -(y)ken

- -(y)Ip...-(y)ken

The suffixes that appear in this slot are the different subordinating suffixes, which express present, past or future time. The suffixes *–DIK, -(y)IncA, -(A/I)r…-mAz, –mAdAn, -(y)ken, -mAsına rağmen,* and *-(y)Ip...-(y)ken* express present and past time while *-(y)AcAK* expresses future.

(15) …**Ø** hazırlığ-ı oku-ma-dığ-ı için… (10TR)
…(O) prep. school-ACC read-NEG-VN-3SG.POSS for.
'…since he did not study at prep. school…'

1.11 TMA I in non-verbal predicates

- -(y)DI (past copula)
- -(y)mIş (evidential copula)
- -(y)sA (conditional copula)
- -(y)ken

An example of past copula is:

(16) … **Ø** ertesi gün öğlen-de Gaziantep-te-ydi-k. (5TR)
…(Biz) next day noon-LOC Gaziantep-LOC-P.COP-1PL.
'…We were in Gaziantep the next day, at noon.'

1.12 Continuity of reference

This variable examined whether a continuity or shift of reference influenced the use of overt versus null SPPs. In the present Turkish study, we not only look at NPs that are subjects of finite verbs, but also those that are subjects of non-finite verbs in adverbial subordinate clauses, locative existential expressions, nouns, adjectives, postpositional phrases, words expressing necessity, and some wh-phrases. These forms can function as predicates and SPPs occurring as subjects of these predicates can be variably used. All

referring expressions were coded as overt or null pronouns for same reference and different reference. This variable has three factors:

- There is no trigger/beginning of discourse

If a subject of a verb occurring at the beginning of discourse did not have a trigger, it was coded as 0. In the example, the subject of verb (1) does not have a trigger, so it was coded as 0.

(17) **Ø** İstanbul-da (1) doğ-muş-tu-m. (40NY)
(I) İstanbul-LOC born-PF-P.COP-1SG.
'I was born in İstanbul.'

- The target NP is the same as the trigger NP

If the target NP was the same as the trigger NP, the target NP was coded as 1. In (18), the subject of predicate (5) was coded as 1 because it is the same as the trigger NP, which is null 1sg *ben* 'I.'

(18) …**Ø** orda (4) oku-du-m. **Ben** (5) küçük-ken… (34NY)
….(I) there study-PF-1SG. I little-COP
'… I studied there. When I was little…'

Same-reference relation may also hold between trigger and target NPs that are demonstrative pronouns, but used as personal pronouns. As mentioned in chapter 3, this study included demonstrative pronouns such as *bu* 'this' and *şu* 'that,' which can be used to make personal references in colloquial Turkish. *Bu* and *şu* can be used to refer to 'he' or 'she.' In (19), the demonstrative pronoun *bu* 'this' was used to refer to the personal pronoun *O* 'he.' The subject of verb (166) was coded as 1 because it is the same as the subject of verb (165).

(19) Profesörlük-ten **bu** şey-e (165) gir-di. **O** şimdi (166) Arizona-da. (19NY).
Professorship-ABL this thing-DAT enter-PF. He now Arizona-LOC.
'From being a professor, he entered [thing]. He is in Arizona now…'

- The target NP is different from the trigger NP

If the target NP was different from the trigger NP, it was coded as 2. The subject of verb (2) was coded as 2 because the null 1pl *biz* 'we,' is not the same as the trigger NP, which is null 1sg *ben* 'I.'

(20) **Ø** İstanbul-da (1) doğ-muş-tu-m**.** Ondan sonra iki yaş-ın-da
(I) İstanbul-LOC born-PF-P.COP-1SG. Then two year-POSS-LOC

aile-m-le birlikte, anne-m, baba-m ve
family-1SG.POSS-INS together mother-1SGPOSS father-1SG.POSS CONJ

abla-m-la birlikte **Ø** bura-ya (2) taşın-dı-k. (40NY)
sister-1SG.POSS-INS together (we) here DAT move-PF-1PL

'…I was born in Istanbul. Then when I was two years old, my parents, my sister and we moved here.'

Different-reference relation may hold between trigger and target NPs, which have the same forms but refer to different referents. From informant 39, we have:

(21) …Asım Can Gündüz, **O** ora-ya (275) taşın-mış-tı. **Ø** tamamen
… Asım Can Gündüz, he there-DAT move-PF-P.COP. (It) complete

(276) tesadüf…(39NY)
coincidence

'…Asım Can Gündüz, he had moved there. It is complete coincidence…'

In (21), the subject of (275) is an overt NP, *O* 'he,' and the subject of verb (276) is a null NP, *O* 'it.' Although these NPs have the same forms, they refer to different referents. Therefore, the subject of verb (276) was coded as 2.

APPENDIX B

SOCIAL VARIABLES AND THEIR CONSTRAINTS

1. Introduction

This dissertation analyzed the influence of twenty-nine social variables and the factors within these variables on the use of an overt or a null pronoun. The variables and the factors to be analyzed were taken from the Spanish study that this study is replicating. Information regarding the social variables was obtained through a sociolinguistic questionnaire from all informants (see Appendix C). The social variables and their factors are shown below.

1.1 Place of residence

The informants were coded depending on where they lived.

- Turkey
- New York

1.2 Sex

The informants were coded depending on their gender.

- Male
- Female

1.3 Age, expressed in categories

- 20-34
- 35-55
- 56-80

1.4 Age expressed in figures

1.5 Class

- Professionals
- Workers

1.6 Education

- Elementary
- Secondary
- High school
- College
- Graduate

1.7 Age of arrival in categories

- Native = born in the US or came at or before age 3
- Child = came to US between the ages of 3 and 12
- Teenager = came to US between 13 and 19
- Adult = came to US age 20 or older.

1.8 Age of arrival in figures

Born in US have age of arrival 0

1.9 Years in the US in categories

- Long 3-15 years
- Established 16+ years
- Native had any 5 of first 8 years of education in the US.

1.10 Years in the US in figures

1.11 English skills

- Poor
- Passable

- Fair
- Good
- Very good

1.12 Turkish skills

- Poor
- Passable
- Fair
- Good
- Very good

1.13 The amount of Turkish spoken with interlocutors

1.13a Turkish with father

- never
- rarely
- usually
- often
- always

The factors of the following variables are the same as those of variable 13a.

1.13b Turkish with mother

1.13c Turkish with siblings

1.13d Turkish with children

1.13e. Turkish with friends

1.13f. Turkish with spouse/boyfriend/girlfriend

1.13g. Turkish with boss

1.13h. Turkish with classmates/workmates

1.14. The amount of Turkish spoken in domains

1.14a. Turkish at home

- never
- rarely
- usually
- often
- always

The factors of the following variables are the same as those of the above variable.

1.14b. Turkish in school

1.14c. Turkish in social activities

1.14d. Turkish in reading

1.14e. Turkish in listening to radio

1.14f. Turkish in watching TV

1.14g. Turkish in internet

1.14h. Turkish in store/shopping

1.14i. Turkish in place of worship/mosque etc.

APPENDIX C

SOCIOLINGUISTIC SURVEY

All personal information you will provide is confidential. Please answer the questions as thoroughly and specifically as possible.

Basic information

1. Age: _____
2. Gender: M __ F __
3. Occupation: _____
4. What is your country of birth? _____
5. In which city were you born in Turkey? _____
6. In which city did you grow up in Turkey? _____
7. What is your native/first language? _____
8. How old were you when you arrived in the US? _____
9. How many years have you been in the US? _____
10. Briefly explain why you came to the US.

_____

_____

_____

11. Education background (check all that apply):

Elementary school	□ in Turkish	□ in English	□ in another language
High-school	□ in Turkish	□ in English	□ in another language
College/university	□ in Turkish	□ in English	□ in another language
Graduate school	□ in Turkish	□ in English	□ in another language

Location (check all that apply)

Did you attend elementary school? If yes, where? □ in the US □ in Turkey □ elsewhere

Did you attend high-school? If yes, where? □ in the US □ in Turkey □ elsewhere

Did you graduate from high school? Yes ____ No _____
If no, how many years did you attend high school? _____

Did you go to college/university? If yes, where? □ in the US □ in Turkey □ elsewhere

Did you graduate from college? Yes ____ No _____

If no, how many years did you attend college? _____

Did you go to graduate school? If yes, where? □ in the US □ in Turkey □ elsewhere

Did you graduate from graduate school? Yes ____ No _____

If no, how many years did you attend graduate school? _____

Turkish language

1. How often do you use Turkish in the following situations?

Place a check.

	1 Never	2 Rarely	3 Usually	4 Often	5 Always
Father					
Mother					
Siblings					
Children					
Friends					
Spouse/boyfriend/ girlfriend					
Boss					
Classmates/ Workmates					

	1 Never	2 Rarely	3 Usually	4 Often	5 Always
Home					
School					
Social activities					
Reading					
Listening to radio					
Watching TV					
Internet					
Store/Shopping					
Place of Worship Mosque etc.					

2. How well do you speak Turkish?

Please use the chart below to indicate how well you speak, read, write, and understand Turkish. Please check the box under the number that best represents your proficiency from 1 (poor) to 5 (very good).

	1 Poor	2 Passable	3 Fair	4 Good	5 Very Good
Speaking					
Reading					
Writing					
Understanding					

English language

1. How old were you when you began to learn English? _____

2. How many years did you study it? (please choose a range from below)

1-3 years 4-6 years 7-9 years 10-12 years 13+ years

3. How well do you speak English?

Please use the chart below to indicate how well you speak, read, write, and understand English. Please check the box under the number that best represents your proficiency from 1 (poor) to 5 (very good).

	1 Poor	2 Passable	3 Fair	4 Good	5 Very Good
Speaking					
Reading					
Writing					
Understanding					

4. How often do you use English in the following situations?

Place a check.

	1 Never	2 Rarely	3 Usually	4 Often	5 Always
Father					
Mother					
Siblings					
Children					
Friends					
Spouse/boyfriend/ girlfriend					
Boss					
Classmates/ Workmates					

	1 Never	2 Rarely	3 Usually	4 Often	5 Always
Home					
School					
Social activities					
Reading					
Listening to radio					
Watching TV					
Internet					
Store/Shopping					
Place of Worship Mosque etc.					

BIBLIOGRAHY

Aarssen, J. (1996). *Relating events in two languages: Acquisition of cohesive devices by Turkish-Dutch bilingual children at school age.* Tilburg, Netherlands: Tilburg University Press.

Allison, P. D. (1999). *Multiple regression: A primer*. California: Pine Forge Press.

Amaral, P. M., & Schwenter, S. C. (2005). Contrast and the (non-) occurrence of subject pronouns. In D. Eddington (ed.), *Selected Proceedings of the 7th Hispanic Linguistics Symposium* (pp. 116-27). Somerville, MA: Cascadilla Proceedings Project.

Appel, R. & Muysken, P. (1987). *Language contact and bilingualism*. London: Edward Arnold.

Avila-Jiménez, B. (1996). *Subject pronoun expression in Puerto Rican Spanish: A sociolinguistic, morphological, and discourse analysis* (Doctoral Dissertation, Cornell University, 1996).

Backus, A. & Akinci, M. (2005). The structure and the role of code-switching in Turkish-French conversations. *Copenhagen Studies in Bilingualism, 36*,156-166.

Bayley, R. (2002b, October). The study of variable subject personal pronouns: methodological and analytical issues. Paper presented at NWAV 31, Stanford University.

Bayley, R. & Pease-Álvarez, L. (1997). Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change, 9*, 349-371.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

Cameron, R. (1992). *Pronominal and null subject variation in spanish: constraints, dialects, and functional compensation* (Doctoral dissertation, University of Pennsylvania, 1992).

Cameron, R. (1995). The scope and limits of switch reference as a constraint on pronominal subject expression. *Hispanic Linguistics, 6/7*, 1-27.

Cinque, G. (2001). A note on mood, modality, tense and aspect affixes in Turkish. In E.E. Taylan (ed.), *The verb in Turkish* (pp. 47-61). Amsterdam: John Benjamin Publishing Co.

Clancy, P. M. (1980). Referential choice in English and Japanese narrative discourse. In

Wallace L. C. (ed.), *The pear stories: Cognitive, cultural and linguistic aspects of narrative production* (pp. 127-202). Norwood, NJ: Ablex Publishing Corporation.

Dahl, O. (1997). Egocentricity in discourse and syntax. Retrieved from http://www.ling.su.se/staff/oesten/egocentric

Eckert, P. (1989). The Whole Woman: Sex and Gender Differences in Variation. *Language Variation and Change, 1*, 245-67.

Eid, M. (1983). On the communicative function of subject pronouns in Arabic. *Journal of Linguistics 19*(2), 287-303.

Erguvanli-Taylan E. (1986). Pronominal versus zero representation of anaphora in Turkish. In D. Slobin & K. Zimmer (eds.), *Studies in Turkish Linguistics* (pp. 209-231). Amsterdam: John Benjamins.

Enç, M. (1986). Topic Switching and pronominal subjects in Turkish. In D. Slobin & K. Zimmer (eds.), *Studies in Turkish Linguistics* (pp. 195-207). Amsterdam: John Benjamins.

Fasold, R. (1990). *The sociolinguistics of language.* Oxford: Blackwell Publishers.

Flores-Ferrán, N. (2002). *A sociolinguistic perspective on the use of subject personal pronouns in Spanish narratives of Puerto Ricans in New York City*. Munich: Lincom-Europa.

Flores-Ferrán, N. (2004). Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English contact? *Language Variation and Change, 16*, 49-73.

Givon, T. (1983). Topic continuity in discourse an introduction. In G. Talmy (ed.), *Topic continuity in discourse: A quantitative cross-language study* (pp.4-41). Amsterdam: John Benjamins Publishing Co.

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

Grice, H. Paul (1975). Logic and conversation. In Cole P. and Morgan, J. L. (eds.), *Syntax and semantics III: Speech acts* (pp.41-58). New York: Academic Press.

Gürcanli, Ő., Nakipoglu, M. & Özyürek, A. (2006). Shared information and argument omission in Turkish In H. Caunt-Nulton, S. Kulatilake & I. Woo (eds.), *Proceedings of the 31st annual Boston University conference on language development* (pp. 262-273).

Guy, G. (1980). Variation in the group and the individual: The case of final stop deletion. In W. Labov (ed.), *Locating language in time and space* (pp. 1-36). New York: Academic Press.

Guy, G. (2005). Grammar and usage: A variationist response. *Language, 81*(3), 561-563.

Haeri, N. (1989). Overt and non-overt subjects in Persian. *PrA Papers in Pragmatics, 3* (1), 155-166.

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

Hamers, J. F. & Blanc, M. H. A. (1989) Implications of language contact phenomena for linguistic theory. Chapter 10.4. *Bilinguality and bilingualism*. Cambridge: Cambridge University Press.

Haznedar, B. (2007). Crosslinguistic influence in Turkish-English bilingual first language acquisition: The overuse of subjects in Turkish. In Alyona B., Luisa M. & Mari U. (eds.). *Proceedings of the 2nd conference on generative approaches to language acquisition North America (GALANA)* (pp. 124-134). Somerville, MA: Cascadilla Proceedings Project.

Hochberg, J. (1986). Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language, 62*(3), 609-621.

Kerslake, C. (1987). Noun phrase deletion and pronominalization in Turkish. In Hendrik E.B. & Ludo V. (eds.), *Studies on modern Turkish* (pp.91-104). Tilburg University Press.

Klein-Andreu, F. (1985). La cuestión del anglicismo: apriorismos y métodos. *Thesaurus, 40*, 533-48.

Kornfilt, J. (1981). *Case marking, agreement and empty categories in Turkish* (Doctoral Dissertation, Harvard University, 1981).

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

Labov, W. (1963). The social motivation of a sound change. *Word, 19*, 273-309.

Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, W. (1972). *Sociolinguistic patterns.* Philadelphia, PA: University of Pennsylvania Press.

Labov, W. (1972b). Some principles of linguistic methodology. *Language in society, 1* (1), 97-120.

Labov, W. (1989). Exact description of the speech community: Short a in Philadelphia. In Ralph F. & Deborah S. (eds.), *Language change and variation* (pp. 1-59). Amsterdam: John Benjamins Publishing Co.

Lapidus, N. & Otheguy, R. (2005a). Overt nonspecific *ellos* in the Spanish of New York. *Spanish in Context, 2*, 157-176.

Li, C. N & Thompson, S. (1979). Third-person Pronouns and Zero-anaphora in Chinese Discourse. In Talmy G. (ed.), *Syntax and semantics: Vol.12. Discourse and syntax* (pp.311-335). New York: Academic Press.

Lira, S. (1982). *Nominal, pronominal and zero subject in Brazilian Portuguese* (Doctoral Dissertation, University of Pennsylvania, 1982).

Lucas, C., Bayley, R., Valli, C., & Rose, M. (2001). *Sociolinguistic variation in American Sign Language.* Washington, D.C.: Gallaudet University Press.

Milroy, L. (1980). *Language and social networks*. Baltimore, MD: University Park Press.

Morales, A. (1986). *Gramáticas en contacto: Análisis sintácticos sobre el Español de Puerto Rico.* Puerto Rico, Madrid: Editorial Playor.

Newmeyer, F. J. (2003). Grammar is grammar and usage is usage. *Language, 79*, 682-707.

Otheguy, R., Zentella, A. C. & Livert, D. (2007). Language and dialect contact in Spanish in New York: Towards the formation of a speech community. *Language, 83*, 770-802.

Özsoy, S. (1987). Null Subject Parameter in Turkish. In H.E. Boeschoten and L. Th. Verhoeven (Eds.), *Studies on modern Turkish: Proceedings of the third conference on Turkish linguistics* (pp. 82-91). Tilburg: Tilburg University Press.

Paredes Silva, V. (1993). Subject omission and functional compensation. *Language variation and change, 5,* 35–50.

Piaget, J. (1959). *The language and thought of the child*. Canada, U.S.: Routledge.

Pfaff, C. W. (1991). Turkish in contact with German: language maintenance and loss among immigrant children in West Berlin. *International Journal of the Sociology of language, 90,* 97-129.

Poplack, S & Sankoff, D. (1987). The Philadelphia story in the Spanish Caribbean. *American Speech, 62*, 291–314.

Poplack, S & Tagliamonte, S. (1989). There's no tense like the present: verbal –s

inflection in early Black English. *Language Variation and Change, 1*, 47-84.

Poplack, S. & Tagliamonte, S. (1999). The grammaticalization of going to in (African American) English. *Language Variation and Change, 11,* 315-42

Ruhi, S. (1992, August). Zero NP anaphora versus full NP anaphora in Turkish paper presented at the 6th International Turkish Linguistics Conference, Anadolu University, Eskisehir.

Sankoff, G. (2001). Linguistic outcomes of language contact. In P. Trudgill, J. Chambers &N. Schilling-Estes (eds.). *Handbook of Sociolinguistics*. Oxford: Basil Blackwell.

Schaufeli, A. (1994). First language text cohesion in a Turkish-Dutch bilingual setting. In G. Extra & Ludo V. (eds.), *The cross-linguistic study of bilingual development* (pp. 199-218). New York.

Sezer, E. (2001). Finite inflection in Turkish. In E.E. Taylan (ed.), *The verb in Turkish* (pp. 1-47). Amsterdam: John Benjamin Publishing Co.

Silva-Corvalán, C. (1982). Subject expression and placement in Mexican-American Spanish. In J. Amaste & E. Olivares (eds.), *Spanish in the United States: Sociolinguistic aspects* (pp. 93-120). Cambridge University Press.

Silva-Corvalán, C. (1994). *Language contact and change*. Oxford: Clarendon Press.

Silva-Corvalán, C. (2001). *Sociolingüística y pragmática del español*. Washington, DC: Georgetown University Press.

Thomason, S. G. (2001). *Language contact.* Washington, D.C.: Georgetown University Press.

Turan, U.D. (1996). Null vs. overt subjects in Turkish discourse: A centering analysis. (Doctoral dissertation. University of Pennsylvania, 1996).

Veltman, C. (2000). The American linguistic mosaic: understanding language shift in the United States. In S. McKay and S. C. Wong (eds.), *New immigrants in the United States* (pp. 58-95). Cambridge: Cambridge University Press.

Verhoeven, L. (1990). Acquisition of discourse cohesion in Turkish. In S. Koç (ed.), *Studies on Turkish linguistics. Proceedings of the fourth international conference on Turkish linguistics* (pp. 437-452). Ankara, Middle East Technical University.

Weinreich, U., Labov, W. & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. Lehmann & Y. Malkiel (eds.), *Directions for historical linguistics* (pp. 95-188). Austin: University of Texas Press.

Wexler, K. & Chien, Y. (1985). The development of lexical anaphors and pronouns. *Papers and reports on child language development, 24,* 138-49.