# A MINIMALISTIC APPROACH TO
# RUSSIAN-ENGLISH-TURKISH MULTILINGUALISM

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF INFORMATICS

OF

THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

OYA ÖZAĞAÇ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF

MASTER OF SCIENCE

IN

THE DEPARTMENT OF COGNITIVE SCIENCE

APRIL 2002

Approval of the Graduate School of Informatics.

_____
Prof. Dr. Neşe Yalabık
Director

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

_____
Assoc. Prof. Dr. H. Gürkan Tekman
Head of Department

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

| | |
|---|---|
| _____ | _____ |
| Asst. Prof. Dr. Bilge Say | Prof. Dr. Wolf König |
| Co-Supervisor | Supervisor |

Examining Committee Members

| | |
|---|---|
| Prof. Dr. Wolf König | _____ |
| Prof. Dr. Deniz Zeyrek | _____ |
| Assoc. Prof. Dr. H. Gürkan Tekman | _____ |
| Asst. Prof. Dr. Bilge Say | _____ |
| Asst. Prof. Dr. Ayşegül Daloğlu | _____ |

ABSTRACT

A MINIMALISTIC APPROACH TO

RUSSIAN-ENGLISH-TURKISH MULTILINGUALISM

Özağaç, Oya

M.S., Department of Cognitive Science

Supervisor: Prof. Dr. Wolf König

Co-Supervisor: Asst. Prof. Dr. Bilge Say

April 2002, 86 pages

The empirical question which is the focus of present research is: How may the lexicons from different languages interact in the course of one syntactical derivation, resulting in code switching phenomena? We develop the following hypothesis concerning code switching: The units of intrasentential code switching are either heads or functional maximal projections. To get support for this hypothesis, intrasentential code switching instances from Russian-English-Turkish and Dutch-Turkish spoken data are analyzed within the minimalist framework. In the data analysed, it has been observed that the data gathered support this hypothesis and that the Minimalist Program has an explanatory force for bilingual language processing.

Keywords: bilingualism, multilingualism, code switching, minimalism, lexicon

# ÖZ

## MİNİMALİST YAKLAŞIMLA
## RUSÇA-İNGİLİZCE-TÜRKÇE ÇOKDİLLİLİK

Özağaç, Oya

Yüksek Lisans, Bilişsel Bilimler Bölümü

Tez Yöneticisi: Prof. Dr. Wolf König

Ortak Tez Yöneticisi: Y. Doç. Dr. Bilge Say

Nisan 2002, 86 sayfa

Bu araştırmanın odağı olan deneysel soru, farklı dillere ait sözlüklerin, sözdizimsel bir türeme durumunda dil değişimi ile sonuçlanacak şekilde birbirlerini nasıl etkiledikleridir. Bu dil değişimi ile igili olarak şu hipotez geliştirildi: cümle-içi dil değişiminde değişen birimler, ya baş ya da işlevsel büyükçül yansımalardır. Bu hipotezin geçerliliğini gözlemlemek için, Rusça-İngilizce-Türkçe ve Hollandaca-Türkçe konuşma verileri toplanmış ve cümle-içi dil değişimi örnekleri, minimalist çerçevede incelenmiştir. İncelemenin sonunda verilerin bu hipotezi desteklediği ve Minimalist Programın iki dilli dil işlemini açıklayıcı güce sahip olduğu sonucuna varılmıştır.

Anahtar Kelimeler: iki dillilik, çok dillilik, dil değişimi, minimalizm, sözcük

ACKNOWLEDGEMENTS

I would like to thank my husband who has been supportive from the beginning and who has always been my best friend and commentator. I am also indebted for my daughter Hazar's patience. We celebrated her first birthday when this thesis was approved.

Secondly, I would like to thank my supervisors for their invaluable contributions and constant encouragements and the examining committee for their comments.

I would also like to thank Leyla Tağızade and Figen Özdemir for their assistance in the gathering and Evgueniy Stepanov in the transcription of the data.

Lastly, I would like to thank Ad Backus for providing the Dutch-Turkish data, Petek Kurtböke for the written corpus and Michel Paradis for sending me copies of his articles.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

# LIST OF ABBREVIATIONS

1. Abbreviations used throughout the Thesis (in alphabetical order)

| | |
|---|---|
| A | Adjective |
| Adj | Adjunct |
| $C_{HL}$ | Computational System |
| CP | Complementizer Phrase |
| CS | Code Switching |
| CT | Computerised Tomography |
| D | Deep Structure |
| DP | Determiner Phrase |
| EEG | ElectroEncephaloGram |
| EL | Embedded Language |
| EQ | Equivalence Constraint |
| ERP | Event-Related Potential |
| FHC | Functional Head Constraint |
| FMC | Free Morpheme Constraint |
| GB | Government and Binding |
| I-language | Internal Language |
| IP | Inflectional Phrase |
| LF | Logical Form |
| ML | Matrix Language |
| MLF | Matrix Language Framework |
| MP | Minimalist Program |
| MRI | Magnetic Resonance Imaging |
| N | Noun |
| NP | Noun Phrase |
| PET | Positron Emission Tomography |
| PF | Phonetic Form |
| PP | Prepositional Phrase |
| S | Surface Structure |
| SPECT | Single-Photon Emission Computerised Tomography |
| UG | Universal Grammar |
| V | Verb |
| VP | Verb Phrase |

## 2. Abbreviations Used in Transcription

| *Abbreviation* | *Morpheme* |
|---|---|
| 1pl | first person plural |
| 1s | first person singular |
| 2pl | second person plural |
| 2s | second person singular |
| 3pl | third person plural |
| 3s | third person singular |
| pl | plural |
| s | singular |
| f | feminine |
| m | masculine |
| n | neutral |
| ACC | accusative case |
| DAT | dative case |
| GEN | genitive case |
| LOC | locative case |
| ABL | ablative case |
| INF | infinitive case |
| INS | instrumental case |
| PREP | prepositional case |
| Pres | present tense |
| Past | past tense |
| PrProg. | present progressive tense |
| PProg | past progressive tense |
| RepPast | reported past |
| Prog | progressive |
| Aor | aorist |
| Fut | future tense |
| Cond | conditional |
| Neg | negative clitic |
| Abil | abilitative |
| Col. | colloquial affix |
| ObjP. | object participle |
| Fnom | factive nominal |
| Imp | imperative |
| Ref | reflexive |
| Perf | perfective |
| Com.Conj. | comitative conjunction |
| SbjP | subjective participle |
| Madv | adverb of manners |

# CHAPTER I

# INTRODUCTION

## 1.1 General background of the study

Bilingualism has become a worldwide reality since globalisation, worldwide communication and trade. Increasing mobility of mass population has made bilingualism the focus of interest for scientists and linguists (Bloomfield 1927, Weinreich 1953, Clyne 1967, Mackey 1972). According to the 1990 U.S. Census, in the United States, the number of people who speak a language other than English at home is more than 35 million (Paradis, 1996).

The vast amount of bilinguals throughout the world necessitated the examination of bilingualism from various aspects. Among these aspects are neurolinguistic (the study of organisation of two or more languages in the brain, for example, by means of comparison of aphasic patients and healthy people), sociopragmatic (foreign language education, the motivations of bilinguals to use the languages they know in discourse), and linguistic (the analytical examination of the utterances of bilinguals).

In spite of the research done on bilingualism, it would not be wrong to admit that there is still a lot to be uncovered since when the subject is human mind, many hindrances avoid prompt and accurate answers to questions related to the subject of bilingualism. However, there is a positive effect of the hindrances as well. These difficulties lead researchers to look for data in a number of different fields.

When the researcher is obtaining data relevant to language, the data should relate to four kinds of real-world phenomena (Lamb, 1999):

1. A well-developed field of articulatory phonetics which deals with the organs and processes of speech production. Accurate and objective transcription and analysis of human languages are essential for the second kind of data as well.
2. Texts of written or spoken discourse, which are the things people say or write. The analysis of such material is the task of analytical linguistics.
3. Neurocognitive basis of language, which is the human brain.
4. Cognitive processes in speaking, understanding, and learning languages.

The importance of considering bilingualism from various aspects has made us study the mental representation of bilinguals with evidence from all the above-listed bases.

When the spoken discourse of bilinguals is examined, it is possible to observe the phenomenon of alternating between languages. The alternating use of two or more

languages in the same conversational event is called *code switching*[1] (CS henceforth) (cf. Section 1.5). Analysis of instances of CS which occur within the same utterance (Intrasentential CS, cf. Section 1.5) because especially this kind of CS provides us with a lot of information about the mental representation of bilinguals, as will be discussed in Section 2.3.

The third item of the above list involves obtaining neurolinguistic evidence. Neurolinguistics is still a developing and therefore a challenging area since examining the human brain is not possible if it is not pathologically dead. To overcome this serious problem, distinct neuroimaging techniques are used. Some of these techniques record the electrical activity of the brain (i.e. EEG Recording, Event-Related Potential - ERP, Single-Cell Recording). Others scan the brain image using different methods (Computerised Tomography - CT, Positron Emission Tomography - PET, Single-Photon Emission Computerised Tomography - SPECT, Magnetic Resonance Imaging - MRI, Mahnetoencephalogram) (Kolb et al., 1996). Research carried out with these neuroimaging techniques on bilinguals has shown some evidence that different parts of the human brain are active when the subjects are asked to use the different languages they know (Kim et al., 1997; Prince, 1999). In addition, studies on aphasic patients have shown that these patients can recover the languages they know at different levels of proficiency, may forget one totally or mix them unconsciously. The different patterns of

[1] "Code switching" in this study, refers to a speech style in which fluent bilinguals move in and out of two (or conceivably more) languages. A more detailed definition is given in Section 1.5 along with an example. The term is spelled variously in the literature as "code switching" (Gumperz, 1967), "code-switching" (Milroy and Muysken, 1995), and "codeswitching" (Myers-Scotton, 1993). We will use the first of these spellings throughout, and refer to it as CS, except where quoted material differs.

language recovery of these aphasic patients led Paradis (1996, p. 15) to assert that languages are:

> … subserved by different circuits intricately interwoven in the same language areas, so that both are represented in the same area at the gross anatomical level, while still being independently subserved by different neural circuits at the microanatomical level. (Paradis, 1996, p. 15)

To get support for the claim of this study, it is necessary to gather evidence for the fourth basis of the above-mentioned list (p. 2) as well. For this, we start with asking "How is bilingual knowledge mentally represented?" Supported by empirical evidence, we first provide a model of the mental representation of monolingual knowledge of language (cf. Section 2.3, Fig. 4). Then, we make a corresponding design for a bilingual speaker (cf. Section 2.3 Fig. 6). Using the minimalist framework (Chomsky, 1995), we assume that bilingual listeners-speakers have different lexicons for each language they know.

## 1.2 Aims of the study

We are going to support the assumption (cf. Section 1.1) that lexicons of a bilingual speaker are mentally represented separately with evidence we get from the analytical examination of intrasentential CS utterances we gather and with an analysis of these data in a minimalist framework. Based on the Minimalist Program (MP henceforth), this framework holds that all syntactic variation is associated with the lexicon. This view proposes that morphological inflections of words are inherent in the lexical items themselves. And, in the case of CS, the inflected words from the lexicon of either

language are successfully derived by checking their features during syntactic processing, or in another term, computation (cf. Section 2.2). Therefore, in the case of an intrasentential CS, it is expected that words from the different lexicons are switched together with their inflections within the same utterance.

To be able to explain the intrasentential CS phenomenon, we develop the following hypothesis concerning intrasentential CS: The units of intrasentential CS are either functional maximal projections[1] or heads.

In the case of head-switching, we speak of (episodic) borrowing: Words are taken in their bare (not inflected) form from one lexicon and transfered into another lexicon. After this transfer, they are treated as if they were words of the language in which the bare lemma[2] has been transfered, i.e., they get the syntactical-morphological and even phonological properties of the words of the language in which they have been transferred (see Section 2.3, Fig. 7). In bilingual CS, we have the case that borrowing is episodical, i.e., a word has been "borrowed" in the real sense of the word: After the derivation, it is given back to the original lexicon and it is not transfered into the host lexicon. (In cases where such borrowings repeat, the words may become part of the "host"-lexicons of a language community; these words get cases of loan influence).

[1] A functional maximal projection is a category which is the projection of a functional head (INFL, Comp, D, T, ...) and which is not contained within any larger constituent with the same head (e.g., IP, DP, CP, ...) (see p. 30f).
[2] The term is first used by Kempen & Hoenkemp (1987) to denote the nonphonological part of an item's lexical information (Levelt, 1989, p. 6).

In the case of functional maximal projections, initially, we may assume for bilingual CS that the units of CS are always phrases, and that we cannot cut through phrases in processing the units of bilingual speech; what is interchanged are phrases. Furthermore, assuming that lemmata are actualized in the numeration in their inflectional forms, thus being joined with morpho-syntactic information including head, specifier, and complement features, we may assume that the phrasal units of intrasentential CS are the functional extension of the maximal projections of the lexical heads. The morpho-syntactic information that is joined with a lexical head is projected inside these functional units. This leads us to the conclusion that intrasentential CS takes 'functional maximal projections' as their units, in addition to 'heads'.

Thus, another aim of the present thesis is to test the validity of this hypothesis that intrasentential CS is always either between heads or between functional maximal projections with the data gathered.

Our third aim is to show the suitability of applying the MP in the analysis of the CS instances. We believe that MP's derivational way of looking at how sentences are composed and the role MP gives to the lexicon (see Section 2.3.2) support our hypotheses (see Section 1.2), as well.

Finally, this study aims to provide evidence for the organisation of the monolingual brain; furthermore, to propose a similar architecture for the bilingual brain by indicating the units of CS.

## 1.3 Significance of the study

CS studies show under which circumstances bilinguals switch from one language to another. In particular, when the CS cases are analysed, it is possible to see the internal operations of syntactic computation and the morphosyntactic tactics applied at CS. We hope to provide CS data and enhance the CS literature, from this perspective.

## 1.4 Scope and limitations of the study

This thesis is structured in the following manner: in Chapter 2, the literature of intrasentential CS will be reviewed, the key concepts of the MP will be outlined, and a lexicalist view of intrasentential CS will be presented. In Chapter 3, the method of data collection will be specified. Also, the method of analysis of these data will be explained. Chapter 4 discusses the results of the research carried out. Chapter 5 presents a summary of the study, generalizations based on CS examples, and recommendations for future studies on this matter.

In this study, CS is examined from a syntactic point of view only. Thus, the pragmatic issues such as the sociolinguistic factors underlying CS and its place in the discourse, and the topic of bilingual education have not been dealt with. With respect to the modular structure of human cognition, this is a legitimate way of getting access to certain aspects of bilingualism.

## 1.5 Definition of some concepts of bilingualism

One of the founders of bilingual studies, Weinreich (1953) defined the term bilingualism as “the practice of alternately using two languages” (Quoted in Hoffman,

1991, p. 15). While using two (or more) languages in a written or speaking discourse, there is a tendency to move in and out of the languages. This movement is named code switching (MacSwan, 1999), where Skiba (1997) defines the term *code* as "language". There can be two types of CS; intersentential and intrasentential. In the former, each of the sentences that the speaker utters can belong to a different language; whereas, in the latter, the speaker switches to another language within the same sentence or utterance. Bhatt (1997) defines intrasentential CS as the alternation between two linguistic systems within a single clause. Below is an example to an intrasentential CS:

(1) This morning *mi hermano y yo fuimos a comprar* some milk (English/Spanish)
This morning my brother and I went to buy some milk

(Belazi et al., quoted in MacSwan, 1998, p. 49)

Another term that is widely used in the studies of bilingualism is the phenomenon of borrowing. The identification and analysis of the terms borrowing and CS has been a controversial issue. While some researchers argue that these two phenomena should be distinguished (Muysken, 1995; Poplack, 1995), some believe that the two are hard to differentiate (Bentahila and Davies, 1995), and some even claim that making such a distinction would not be fruitful in explaining lexical relations in the data (Kurtböke, 1998).

Muysken (1995) sees borrowing as a kind of intrasentential CS. According to him, CS can be of a phrasal category, which he calls "CS", or of an alien lexical item, which is borrowing. According to him, the only difference between CS and borrowing is

the size and the type of the element switched; i.e. if a noun is switched, this is borrowing and if a noun phrase is switched, this is CS (Muysken, 1995, p. 180). In this thesis, we try to be precise with respect to the size and type of the constituent switched. We adopt Muysken's view of the concept of borrowing as a kind of CS besides phrasal switching. We claim that CS can be in lexical form in cases of borrowing, or of a phrasal category, which we claim are maximal functional projections.

# CHAPTER II

# A REVIEW OF LITERATURE ON BILINGUAL LANGUAGE PRODUCTION

Before investigating the specific research question of this study, several issues should be addressed. In this chapter, we will review the literature on syntactic CS constraint theories firstly, to give the reader an idea of the types of constraints found so far and to give counter-examples for each with CS examples from the literature. Then, we will explain the key concepts of the MP to form a basis to our hypothesis. Lastly, against this backdrop, we will present once again our research question and give arguments for the hypothesis we have formulated.

## 2.1 Literature on code switching

Studies related to CS date back to 1970s. Researchers of CS have focused on different aspects of code switching constraints within current syntactic or psycholinguistic theories and searched for the constraints of code switching on syntactic and pragmatic aspects. The difference between the syntactic and pragmatic approach to the study of code switching is that while pragmatic framework considers the motivation for switching as a stylistic and a discourse phenomenon which cannot be explained in term of the internal structure of sentences, the syntactic approach is concerned with accounting for the linguistic constraints on CS. (Romaine, 1995, p.

121). The focus of syntactic studies was to find the constraints that shape the CS data gathered on specific language pairs such as Spanish-English and Nahuatl-French.

### 2.1.1 Free Morpheme Constraint and Equivalence Constraint

When the history of syntactic analysis of CS is examined, perhaps the most cited constraints belong to Shana Poplack (1980). Her two important constraints are Equivalence Constraint (EQ) and Free Morpheme Constraint (FMC). According to EQ, "switching is free to occur only between sentence elements that are normally ordered in the same way by the monolingual grammars in contact" (Poplack, 1995, p. 209). This means that the word order requirements of the languages involved in CS should match for a CS to occur. For example, a CS involving English and Spanish may occur between determiners and nouns but not between nouns and adjectives because in Spanish, the word order is *NP⟶(det) N (adj),* whereas in English it is *NP⟶(det) (adj) N*. Thus, (2) is prohibited.

(2) *his *favorito lugar*
his favorite spot

(Romaine, 1995, p. 127)

Poplack's FMC (1980) maintains that codes may be switched only when they do not occur at the boundary of a bound morpheme (Cited in Myers-Scotton, 1993). However, Spanish-English examples (3) and (4) are CSs violating FMC:

(3) **estoy* eat-*iendo*
I-am eat-ing
'I am eating'

(Poplack cited in MacSwan, 1999)

(4) *taip*iamos cada dia
'we type every day'

(Pfaff cited in Backus, 1996)

### 2.1.2 Government Constraint

Another syntactic analysis was done by Woolford (1983) who developed a generative model of code switching. During mid-1980s, Government and Binding syntactic theory of Chomsky (1981) was prevalent and this framework also shaped the works of DiSciullo, Muysken and Singh (1986) who argued that switching was not permitted within a maximal projection (Myers-Scotton, 1995). According to the Government Constraint, each governing element (e.g. verb, preposition, auxiliary) creates a matrix structure. If the chain of government were unbroken, the highest element in the tree would determine the language for the whole tree; this would often be the inflection on the finite verb, as in the theory proposed by Klavans (1985) and taken up by Treffers-Daller (1991). In subordinate clauses, this would be the complementiser (Muysken, 1995, p. 182).

The theory was later modified to be more limited by Muysken (1990) but was still inappropriate in explaining the occurrence of Moroccan Dutch-Arabic CS data in which verbal and prepositional object NPs are often in a different language from their governing verb or prepositions, as in (5):

(5) anaka-ndir *intercultureel werk*
'I I-am-doing intercultural work'

(Nortier cited in Muysken, 1995, p. 187)

### 2.1.3 Functional Head Constraint

Yet another CS constraint theory belongs to Belazi, Rubin and Toribio (1994). Taking the X-bar theory as their basic model, they found the Functional Head Constraint (FHC) which says that code switching between a functional head and its

complement is not permitted. However, in the following example from Spanish/Nahuatl, the Nahuatl affix *nik-*, which is a bound morpheme, is attached to the Spanish verb *amar*.

| (6) | Ne | nik*amar*ao | in Maria |
|---|---|---|---|
| | ne | ni-k-amar-oa | in Maria |
| | I | 1-3s-love-suffix | in Maria |
| | 'I | love | Maria' |

(MacSwan, 1999, p. 3)

### 2.1.4 Closed-class Constraint

There has also been a lexicalist analysis of code switching. One of such a study is a constraint on closed class items (Garret, 1975). According to this, items belonging to this class (e.g. determiners, quantifiers, prepositions, possessives, auxiliaries) cannot be switched. A similar constraint was later proposed by Joshi (1985). However, the following Turkish-Russian example, in which a conjunction is switched, presents a counter-example for this constraint:

| (P13) | Tamam hazırım | *no* ayakkabımı | bulamıyorum. |
|---|---|---|---|
| | Tamam hazır-ım | no ayakkabı-m-ı | bul-am-ıyor-um |
| | Okay ready-1s | but shoe-1s-ACC | find-NEG-PrProg.-1s |
| | 'Okay I am ready but I cannot find my shoe.' | | |

### 2.1.5 Matrix Language Framework

One of the major theories of lexical analysis of CS constraints belongs to Myers-Scotton (1997). Her Matrix Language Frame proposal (MLF) suggests that in the languages involved in CS, the principal language is called the matrix language (ML) and the weaker language whose morphemes are inserted into the ML is called

the embedded language (EL). ML sets the morphosyntactic frame in the case of a code switch; thus, the ML morpheme order must not be violated and ML should provide all system morphemes. Her theory proves to be quite strong; however, according to MacSwan the "unfalsifiability" of this theory is "an intractable weakness in … the Matrix Language Frame" (1999, p. 13) since the theory also allows changing of the ML at any time in production, even in the middle of a sentence, as shown in (7). In such utterances where the ML is changed in the middle, it is not clear which language framework would be responsible for the structure of the sentence and which items would be considered EL-islands.

(7) The students had *visto la pelicula italiana*
seen the Italian movie

(MacSwan, 1999)

### 2.1.6 PF Disjunction Theorem

A recent proposal belongs to MacSwan (1997, 1999) whose work is shaped by the MP. According to his model, CS is simply the introduction of features by the lexicons of the languages involved in CS. Lexicons bear the morphological features and thus carry the requirements of the respective systems during convergence (cf. Section 2.2). He asserts that "the grammar used for code switching consists of the union of the two lexicons plus the invariant computational system, with no mediating mechanisms needed" (p. 22). According to his theory, at the point where the words are spelled out (cf. Section 2.2), if the code switched word is treated with the language which enforces its rules to the lexicon, then there are no ill-formed constructions. This means, for a code-switched word to be considered part of the glossary of the host (= matrix) language, it should be adapted by using the

phonological rules of the host language, as in (12). The following examples are given by MacSwan (1999, p. 25):

(8) is an example for an ill-formed CS:

(8) Juan eat-*ará*
'Juan will eat'

However, in (9), the Spanish lexical items are introduced into the Nahuatl lexicon where the rules of word formation add its own inflectional morphology:

| | | | |
|---|---|---|---|
| (9) | Juan está | *parq*ueando | su coche |
| | Juan is | parking | his car |

In this study, we also utilize the minimalist idea that lexicons bear abstract and concrete morphological properties specific to the language. However, the use of the word 'union', as MacSwan did, would not be appropriate when a bilingual lexicon is examined as illustrated in Fig. 5, Section 2.3. We believe that intrasentential CS is the 'interaction' of these lexicons of two (or more) languages during phonological processing (cf. Section 2.3) rather than a 'union' of them.

## 2.2 Key concepts of the Minimalist Program

Being a still-developing research framework, MP is the latest development of an approach to syntax putting an emphasis on and thus taking its name from the idea that a theory of syntax should be simple enough to conform to 'minimalist' concepts. By minimalism, we should understand a linguistic economy in the constructions of structures and in the steps of derivations.

We will introduce the key concepts of MP in juxtaposition to Government and Binding (GB) since MP adopts, reinterprets and elaborates on some of the principles of GB framework (Haegeman, 1997, p. 1). In the GB model, there are four levels of representation Deep (D) Structure, Surface (S) Structure, Phonetic Form (PF), and Logical Form (LF). Items are taken from the lexicon and inserted into D in accord with their thematic relations. As a result of a number of movement transformations, each leaving traces that mark the positions from which the movement took place, S is formed. Transformations of the same character continue the derivation to LF, the conceptual-intentional system of the mind. At the same time, phonological rules continue the transformation to PF, the articulatory-perceptual system. The operations between D structure and S structure are called "overt". The operations between S and LF are called "covert" since these have no phonetic effects.

PF and LF are ineliminable interfaces of human language, and minimalism makes use of these interfaces in the model it develops. In order to minimize the number of transformations, it does not use the D structure. Instead, elements of the lexicon are assumed to be inserted in the course of derivation. Instead of postulating S structure as a level of representation, it is assumed that there is a computational system ($C_{HL}$) which is the fixed and invariant I-language process across languages and only lexicons which are derived ‘on-line’ bear the idiosyncratic properties of individual languages.

According to MP, linguistic variation among various languages is a result of the morphological properties of the lemmata of the languages. The starting point of computational operations is picking words from the lexicon of a language by the

operation s*elect.* The words, then, are inserted into the *numeration*. In the numeration, words are numerated together with their features.

During computation, there are two other operations called *merge,* and *move. Merge* combines a pair of syntactic objects and combines these into a new one by projections of one of the objects. In order for a derivation to converge, all the syntactic objects must be interpretable at one of the interfaces, LF or PF.

*Move* rearranges items in the syntactic tree for several purposes, such as checking its morphological features such as case, gender, number, and person.

If the derivation does not satisfy *Full Interpretation*, which says that a representation must consist of only legitimate objects at the interfaces, then it *crashes*. After having checked the specifier, head and complement features, the derivation arrives at the point of *spell-out*, and it can be converted by linearization to PF.

When the features are *strong*, the movement is called *overt* and the result is seen in the surface structure called PF. When they are *weak*, then the movement is *covert*, and can only be observed at LF (Radford, 1998). These operations can be figured as in Fig. 1.

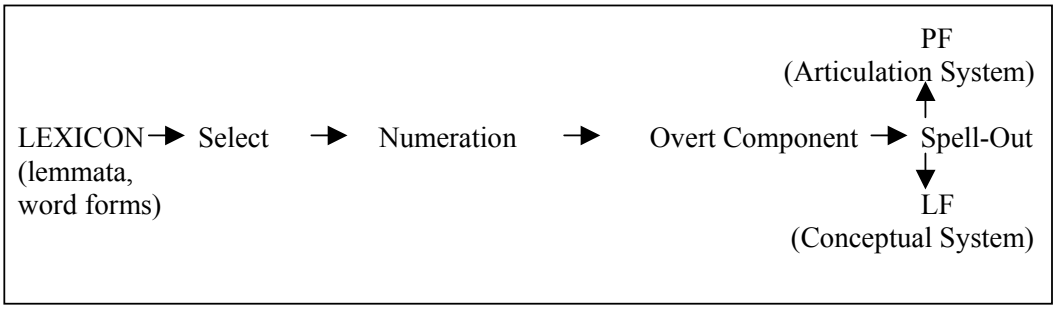



Figure 1: Minimalist Model of Levels of Syntactical Derivation

## 2.3 A minimalist approach to code switching

### 2.3.1 Mental representation of monolingual knowledge of language

To answer the question what we mean by mental representation of 'bilingual/ multilingual knowledge of language', we first consider models of the mental representation of monolingual knowledge of language.

Assuming the minimalist hypothesis, we firstly state that the knowledge of language is derivative, not representative, the latter as assumed inside the framework of GB theory.

At this point, it is necessary to give some explanations on the derivativeness hypothesis of the knowledge of language. We know, in any case, that the human language generator and the human parser ( = language understander, language analyzer) constitute two separate modules of our cognitive equipment with respect to the knowledge of language. From the examples of garden-path sentences[1], we understand that the parser has difficulties in understanding sentences that are easily generated, because the parser works according to some procedural devices that are not part of a generative grammar (minimal attachment, LIFO principle, preferred readings etc.) This shows that we have to assume at least two modules: an independent parser (language understanding system), and the language generator.

Inside the frame of GB, it was presupposed that there is a generator (a phrase structure grammar) that is able to generate more than only the set of the grammatical

[1] Garden-path sentences are sentences which lead the listener or the reader to an incorrect analysis, as in the example "The man who hunts ducks out on weekends." This kind of sentences shows that people mainly use a depth-first strategy. They make an analysis that seems to be working and pursue it as long as possible until it cannot fit into the tree (Pinker, 1994, pp. 212-3).

sentences of a given language. Therefore, besides the generator (and the parser) a further module was presupposed: A system of universal principles and language-specific parameters (or technically: a set of filters) that restricts the possible derivations of the generator such that after filtering all possible derivations of the generator, according to the given principles and parameters, only the set of the grammatical sentences of a given language are left. This third module constitutes the representational knowledge of language.

From a cognitive point of view, there is no evidence any more for an independent module of a representational knowledge of language in addition to the generator that constitutes the derivational knowledge of language. Representational knowledge may completely be formulated as derivational knowledge (see Epstein, et al., 1998). Thus, it can be shown that the knowledge of language as procedural knowledge is at the same time the knowledge of how to generate sentences of a given language, in other words, is at the same time derivational knowledge of language. From an empirical point of view, there is no evidence that we produce ungrammatical structures and sentences, and that we restrict them later by means of the filters of a respresentative module of knowledge of language.

We hypothesize that the knowledge of language is derivative: The knowledge of language means the knowledge of how to generate sentences of a given language L, or, of how to derive the strings that constitute the sentences of L.

The sentence generator is a subpart of an information processor that produces meaningful utterances, and a theory of grammar is a model for the sub-module of that part of our cognitive equipment that is responsible for the production of

meaningful utterances. There must be some interfaces of grammar that are links to the other parts of the information processor. In the minimalist version of generative grammar, we assume two interfaces: the perceptual-articulatory interface and the conceptual-intentional interface. We must see how this minimalist model of a grammar fits into the model of a human as an information processor.

A working hypothesis of a human as an information processor is the model of language production proposed by W. Levelt, which is based on corresponding models developed earlier by Garrett (1971) and Fromkin (1975). This model constitutes the starting point of further investigations on language production, in which conceptual planning as prelinguistic preparation of language production, the process from conceptualisation to syntactic codification, syntactic and conceptual influence on morphosyntactic agreement in language production, meanings of verbs, theta rules and the conceptual realisation of situation, and the timing of the activation of syntactic and phonological features during language production are investigated. The aim of all these investigations is to determine the subparts of this model and to find empirical evidence for these subparts that include the phrasal units of sentences. Methodological focus during present empirical investigations is on psycho-linguistic experiments (Levelt, (n.d.); and the references thereafter).

In any case, the model of Levelt is considered, for the present, as the starting hypothesis on human as an information processor (see Fig. 2).

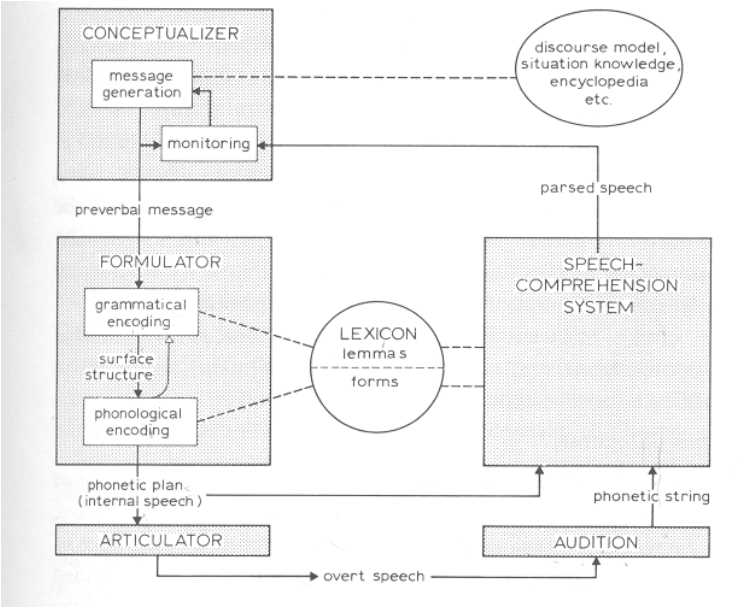



Figure 2: A Blueprint for the Speaker (Levelt, 1989, p. 9).

It is empirically strongly supported that there are at least three modules of information processing: Firstly, the module of conceptualization, a pre-linguistic domain that we use in the beginning of the planning of an information process. Empirical evidence for such an independent module is, e.g., the 'tip of the tongue' phenomenon; sometimes we cannot find the word for a concept that we have already actualized.

We may assume further that there is a phonetic component for the articulation. This module is independent of the abstract phonological component, which is the result of a linearization process of a hierarchical syntactic structure. Examples for empirical evidence are misspellings like "it certainly run out[s] fast" instead of "it certainly run[z] out fast". This example shows that we misplace abstract phonological units, and that we give the phonetic interpretation after we have misplaced the abstract phonological units. The two modules, conceptualizer and

articulator, constitute the outside parts of the two interfaces (perceptual-articulatory and conceptual-intentional) of a grammar according to the minimalist design.

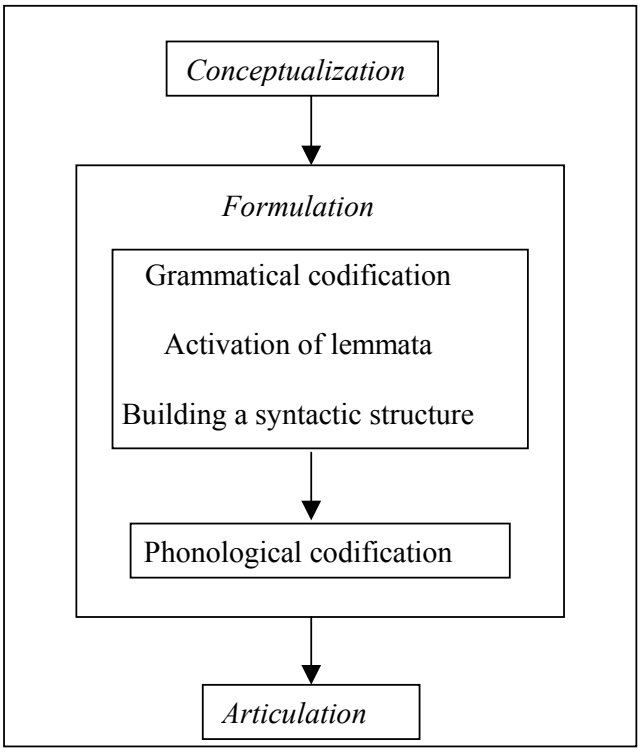



Figure 3: Levelt's Model of a Human as an Information Processor (simplified)

The big box in the middle may be thought of as a grammar. According to the minimalist conception, we have to separate the lexicon from the grammar and introduce the concept of a *numeration* for the activization of lemmata. Under "lexicon" we understand a grammatical lexicon that contains all inflectional forms of words besides their bare forms plus the necessary syntactic information as head, complement, and specifier features. This grammatical lexicon is different from, but in interaction with a mental lexicon[2], which might be thought of as joined to the conceptualizer.

[2] The passive store of information about the words in one's language (Levelt, 1989).

2.3.2 Minimalist model of the knowledge of language

According to the minimalist syntax, we have subparts of the 'computational system' by means of which a grammatical structure is built up in a bottom up manner; this system applies the rules of merging and projecting. At some point of a derivation the stage of *spell out* is reached. At this point, the derived syntactic structure gets linearized and will be assigned a phonological interpretation.

A design of a human as an information processor, modified according to the concepts of minimalist grammar, may look like this:

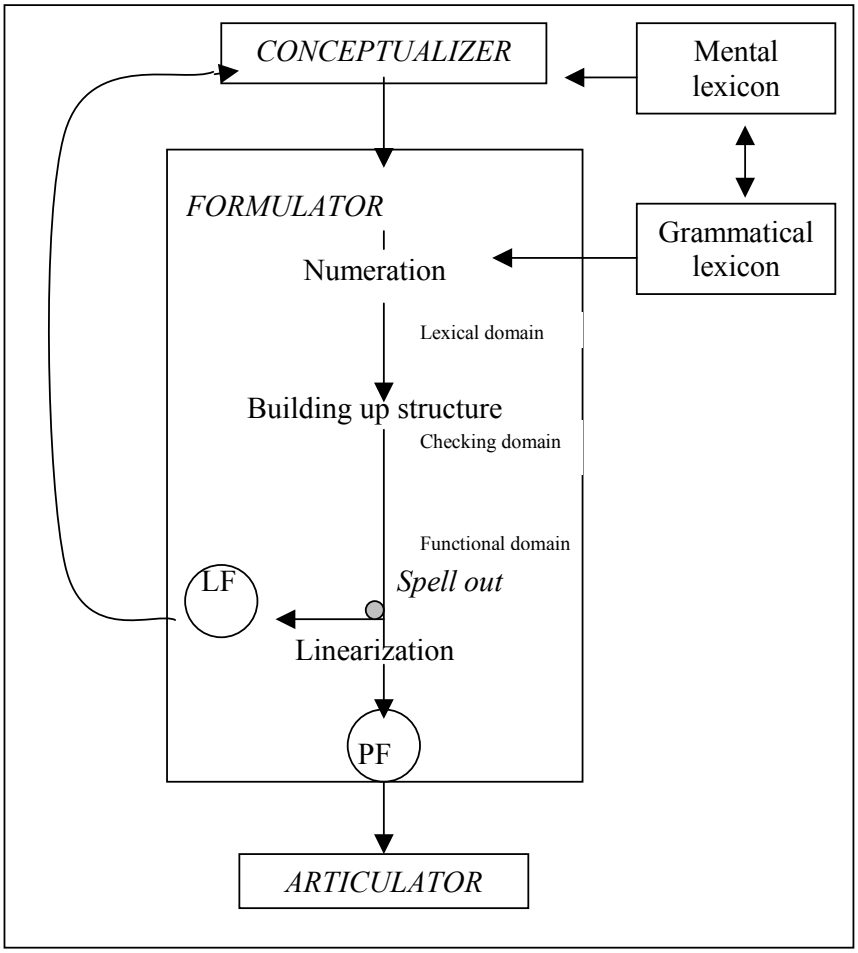



Figure 4: A Minimalist Design of a Human as an Information Processor

2.3.3 Human speech as an incremental language production process

The cognitive process of language production as sketched in Fig. 3 and 4 is an incremental process. This means, we do not start with the conceptualization, and after having finished it we activize the lemmata, and after having activized all the lemmata we are using in a given sentence, we build up the structure etc. Actually, we start with the conceptualization, but short after, we start with the numeration, and build up structure after we have actualized some lemmata, and so on. In the meanwhile, conceptualization is still going on. The process of language production is thus a parallel process, but with retardings with respect to the lower levels. This kind of incremental processing can be shown in the following schema:

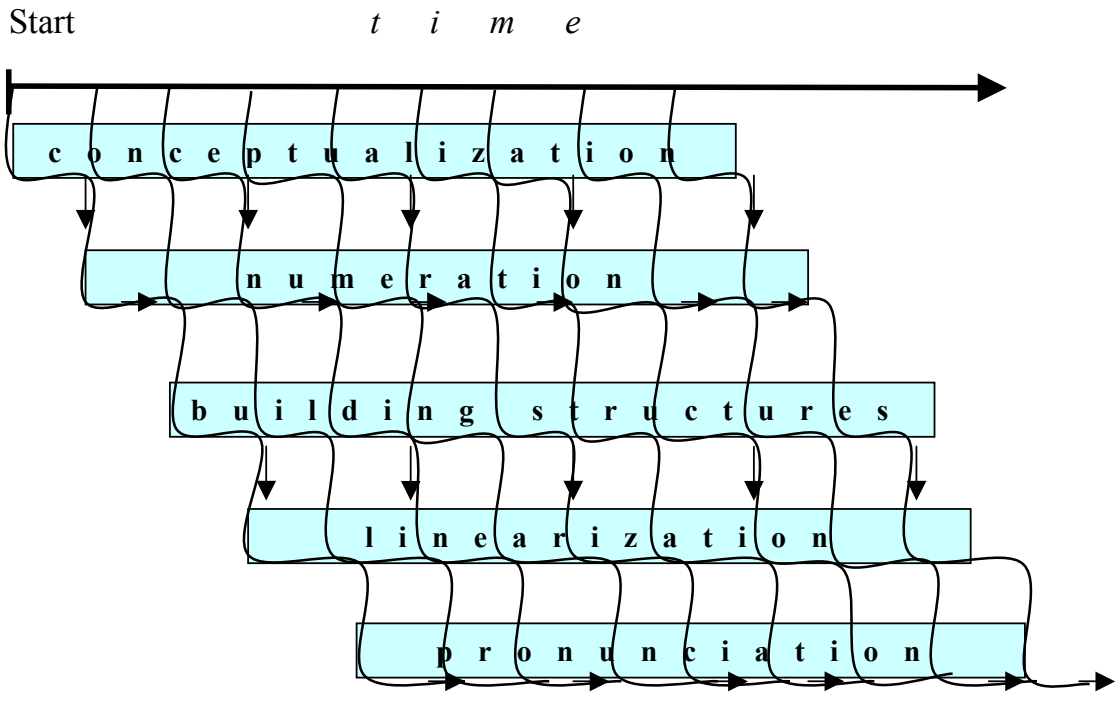



Figure 5: Incremental Processing of Speech (adapted from Levelt's Incremental Production 1989, p. 25)

The length of an arrow inside one box may be considered as an incremental unit. At the end of an incremental unit, we change to a lower level.

The empirical question which is the focus of present research is: What are the incremental units for the case of CS? For the case of building up structure, we may assume that they constitute the maximal projections. In the case of numeration, we may assume that they are the set of lemmata necessary to build up sentences. For conceptualization, we may assume basic event structures, whatever they are in detail. For pronounciation, we may postulate prosodic units that constitute the increments. In any case, we leave open the question of the increments of the conceptualizer or articulator as we are concerned here with the formulator which we assume to be represented by a minimalist version of generative grammar.

### 2.3.4 Bilingual language processing

After we have sketched the design of the cognitive architecture of monolingual language production, we now try to consider the corresponding design for a bilingual speaker. Some of the subparts in the model must occur twice. We first ask if we have two conceptualizers and, as a consequence, two complete different information and language-processors. The answer must obviously be 'no', because we assume the conceptualizer to be a pre-linguistic and a language independent unit. The conceptualizer, being the most abstract component and the starting module, must be at least the component that is common for all languages in bilingual processing.

What about $C_{HL}$ according to the minimalist design for a generative grammar? $C_{HL}$ is supposed to be universal. The derivations are lexicon driven, i.e. they are determined by the idiosyncratic information given together with the lemmata

taken into the numeration. The information may change considerably from language to language, but the rules for processing it, that are given by CS, are the same for any numeration as the input of a derivation. If we accept this as a fact about the cognition of language processing, then we must assume exactly one grammar for the derivations in the different languages of a bilingual user: We have lemmata from different languages, but exactly one system of building up structures upon words in a bottom-up manner and one system of transforming structures into a linearization.

Finally, there is the lexicon as the unit that is language-specific in bilingual language processing. Do we have one lexicon with lemmata from different languages (as MacSwan assumes, see above p. 16), or do we have a separate lexicon for each language? We should assume two different lexicons if we consider that bilingual speaker-listeners are able to discern their two (or more) languages. Otherwise, we must assume that bilinguals cannot discern their languages. Then, how can we explain the CS phenomena where there is the possibility of an interaction between the two (or more) lexicons? From the results of research on bilingualism, we know that there is a scale of phenomena from a strict discretion of the two lexicons at one end, to mixings of languages at the other end of the scale. In a design of bilingual language processes, it must be possible to explain these phenomena. Thus, a model of bilingual language processing may be as in Fig. 6.

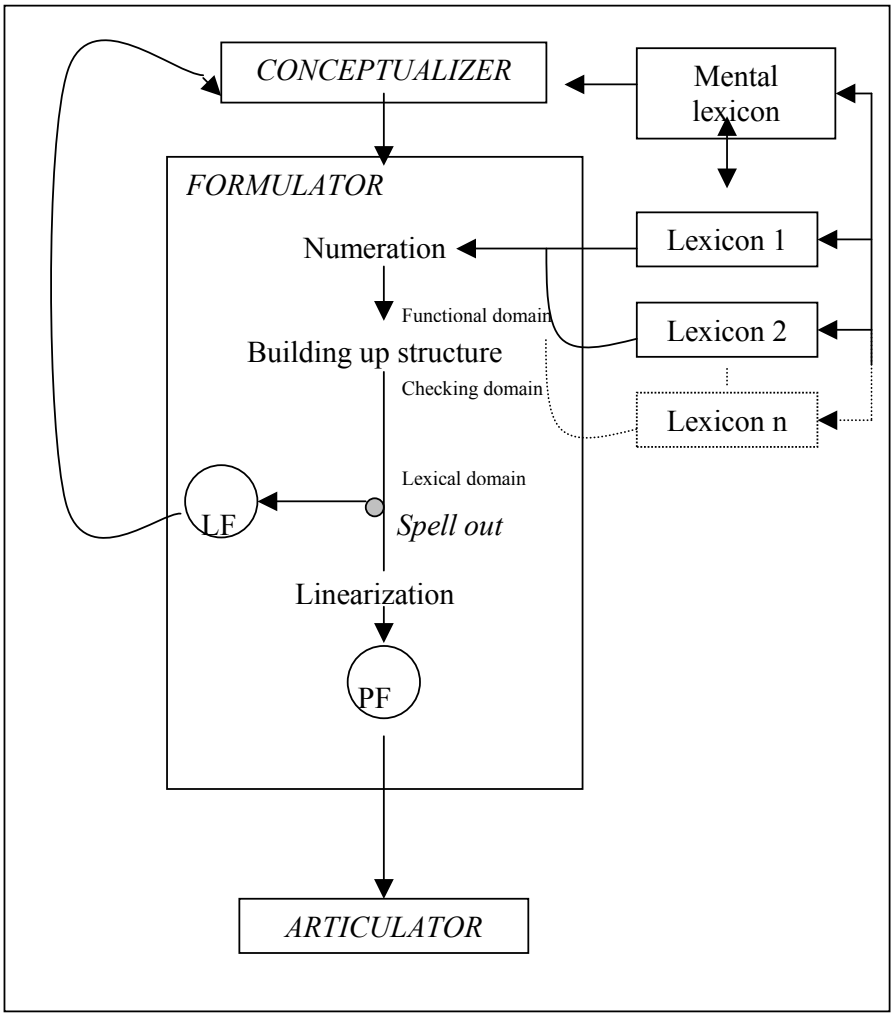



Figure 6: A Model of Bilingual Language Processing

We assume this as the basic model of how multi-/bilingual knowledge is mentally represented. The formulator, the generative grammar, is able to process any input from any lexicon of any language. Learning a language means to fix the parameters given by the morpho-syntactic features and the subcategorial information joined with the lemmata a child acquires as vocabulary learning. In order to be from one language, the morpho-syntactic features of word forms must fit to each other in order be able to be checked. Word forms that fit to each other are collected in one lexicon. In this way, a child can construct several lexicons, by learning which words

fit to each other, so that they may enter into a numeration as input for building up structure and derive without crashing.

## 2.4 The framework and aim of the thesis

According to this model (Fig. 6), the lexicons are mentally represented separately. We ask now, and this is our specific research question, how these two lexicons may interact in the course of one derivation, resulting in CS phenomenon. We develop the following hypothesis concerning CS: The units of CS are either maximal functional projections (cf. p. 30) or heads.

In the case of (episodic) borrowing, a word, in its bare form, is taken from one lexicon and inserted into another lexicon. After this transfer, it is treated as one of the words of the language in which the bare lemma has been transfered, i.e., it gets the syntactical-morphological properties of the words of the language in which it has been transfered (see Fig. 7).

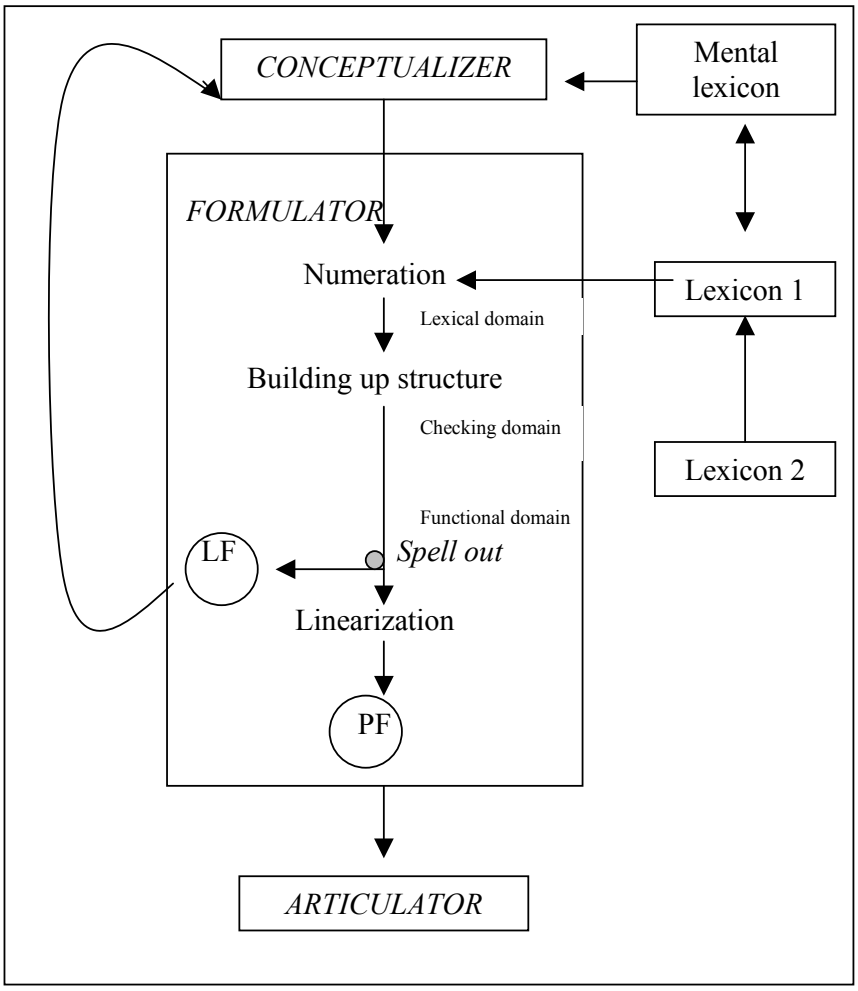



Figure 7: A Model of Borrowing

The more complex and critical issue is that intrasentential CS which is not lexical switching. Bilingual intrasentential CS cannot happen without regularities; at least we should pressupose that the principle of dependency on structure for grammatical rules is valid in bilingual speech, too.

Then, we may think that the units of CS are always phrases in bilingual speech. We take for granted that lemmata are actualized in the numeration in their inflectional forms, thus are being joined with morpho-syntactic information including head, specifier, and complement features. Therefore, the phrasal units of

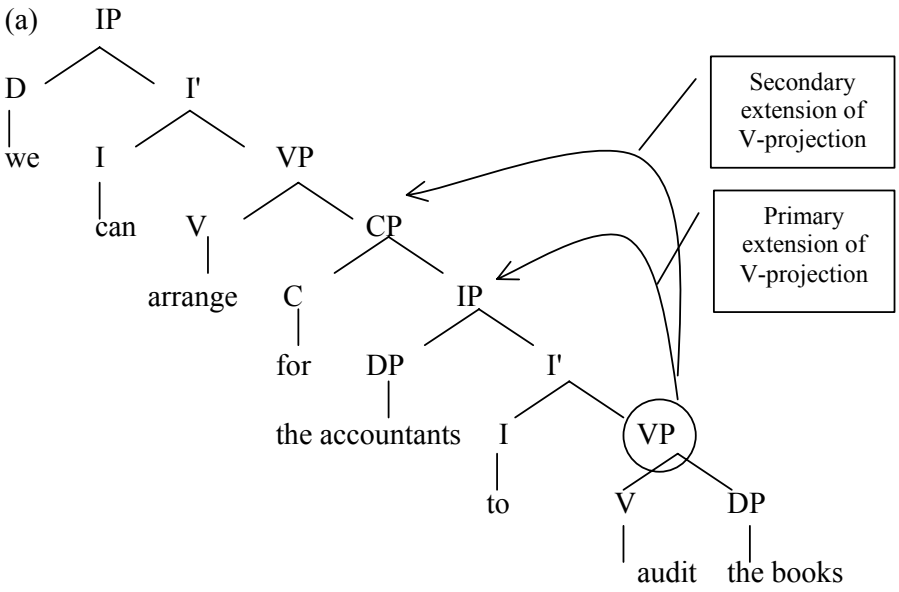
intrasentential CS are the functional extension of the maximal projections of the lexical heads.

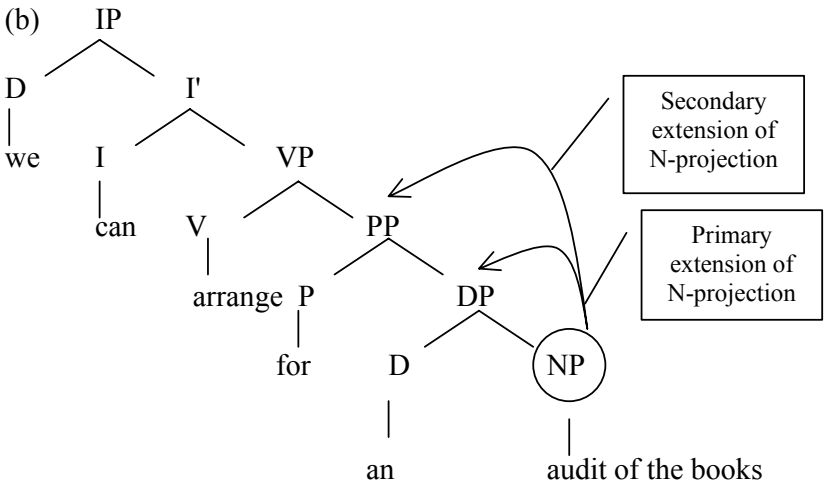
The terms "primary" and "secondary" maximal projections can be illustrated with the help of the following examples:

(a) We can arrange for the accounts to audit the books.

(b) We can arrange for an audit of the books.

These example sentences have the following trees:



In (a), the VP *audit the books* has an extended projection into the IP *the accountant*, which we call a "primary extension", and a further extended projection into the CP *for the accountants to audit the books*, which we call "secondary extension".

In the same way, in (b), the NP has a primary extension into the DP and a secondary extension to the PP.

The morpho-syntactic information that is joined with a lexical head is projected inside these functional units. Thus, we conclude that intrasentential CS takes heads or functional maximal projections as their units. This is a strong hypothesis with a high amount of explanatory force, and it supports the cognitive relevance of generative grammar in its minimalist version. Because of its explanatory force, this hypotesis is easy to be falsified, just by one example that does not fit to the hypothesis (or cannot be explained in another way as exception).

In this chapter, we have reviewed the CS literature. We have explained the key concepts of the MP. We have also examined CS from a lexicalist view. Finally, we have provided our specific research question, which is whether the data gathered support the hypothesis that intrasentential CS is always either between heads or between maximal functional projections.

# CHAPTER III

# METHOD OF DATA COLLECTION AND ANALYSIS

In this chapter, the data collection methodology and the selection criteria for the informants contributed to this study are introduced. At the end of this chapter, the reader should have a fairly good idea about what kind of data will be encountered in the rest of this study. It should also be clear why the data were collected in the way they were.

## 3.1 Data collection

Though in studies related to bilingual speech data collection methods vary extensively in relation to research fields and objectives of these studies, it is possible to find the three main approaches that contribute to code switching data collection (Dabène and Moore cited in Kurtböke, 1998):

(i) Elicitation of speech samples using standardized experimental procedures in which informants are chosen according to strict selection criteria[1].

(ii) Recording informal interviews with bilinguals who are again recruited according to a strict selection criteria.

[1] The criteria selection varies depending on the purposes of the study; this is why the selection criteria are not mentioned specifically. The criteria for this study are listed in section 3.1.2.2.

(iii) Collection of authentic natural language in interaction situations.

The discussions and interviews with bilinguals have shown that CS is a common way of speaking and a natural process as long as the setting is informal (Backus, 1992) and the speakers know the languages involved in CS more or less at the same proficiency. This makes it necessary for the researcher to carry out pilot studies to be able to create an environment informal enough to hear day-to-day conversation.

### 3.1.1 The pilot study

In this thesis, we first carried out a pilot study to find the data collection method that suits best to the kind of data relevant to the hypothesis of this study. This pilot study included spoken and written data collection procedures.

#### 3.1.1.1 The spoken data

The spoken data collection was done in accordance with (ii) explained above. Speech samples were gathered in an informal setting. The conversation group consisted of one middle-aged lecturer from Azerbaijan, one PhD student, and the researcher herself. The topic of conversation was not restricted. The group members knew each other at least by name, if not more familiar. The selection criterion for these informants was that the informants needed to have a certain level of spoken proficiency in the three languages, namely Turkish, English and Russian, to be able to switch into them. The researcher also contributed to the conversation, however, she neither informed the informants about in what terms the data would be analysed nor used intrasentential CS during conversation. The total amount of recording was approximately two hours. The recording was later transcribed and the utterances that

included intrasentential CS were coded as explained in Section 3.2.1.3. A list of these utterances was put in the Appendix C.

#### 3.1.1.2 The written data

Collection of written data has been quite rare in the history of CS research since CS is a basic phenomenon of spoken data, not written (Halliday, 1987; cited in Kurtböke, 1998). Believing that any theory of knowledge, and of language, must encompass both written and spoken language, as these are complementary, first, we carried out a pilot study for the collection of written data.

##### 3.1.1.2.1 Data collection

The informants were shown on an individual basis, a videotape including several commercials that are silent so that the informants were not affected by any language. Each commercial was shown twice so that the informant could remember them easily. After each commercial, the informants were asked to jot down on a piece of paper what they had seen. The researcher had no interference with the informants both during the show and in the writing process. There was no time limit for the writing section; it generally took approximately 10 minutes for each commercial. The selection criterion for the collection of the written data was exactly the same with the spoken data; the informants needed to have knowledge of Turkish, English, and Russian enough to be able to switch into one another. A total of 6 bilinguals were employed for the study and each was shown 4 commercials. However, there was only one written CS instance at the end of the process. This result was in accordance with the general conviction that collection of written data for CS studies did not contribute much; thus the researcher decided to enrich the study with an alternative source of written data, which is corpus analysis.

3.1.2.2 Corpus analysis

In accordance with the belief that a study on language would not be sufficient without written data, Kurtböke (1998) used an authentic written source of data for her studies, which are the ethnic newspapers of Turkish immigrants living in Australia. However, when we examined these data, we found that the instances of language switches were all head-switching cases and there were no phrasal switches. Moreover, as we have no information about the changes the language that Turkish immigrants use in Australia has undergone, we would not be able to comment on whether these head switches were episodic borrowing cases or not. Thus, we have decided not to use the written corpus.

Since the attempts to use written data were unsuccesful, we decided to concentrate on spoken language and carried out a primary data collection process only for the spoken data. However, we will also analyse the CS instances gathered in the pilot study together with the primary ones.

3.1.2 The Primary Data

3.1.2.1 Data Collection

The primary data to be analyzed are gathered in the following ways:

(i) Natural conversations are taped and transcribed word for word after tape-recording,

(ii) An already transcribed corpus of spoken data gathered among a group of Turkish immigrants in Holland was analysed.

These methods were preferred as a result of the pilot study. The reason for employing more than one method is firstly the need to gather as much utterance

containing CS as possible and secondly the fact that all these methods have both disadvantages as well as advantages.

The first method of data collection has been widely used in CS research (Poplack, 1980; Backus, 1992; Myers-Scotton, 1995; MacSwan, 1999). The main advantage of such a method is that the researcher is able to observe utterances in a natural setting. However, since the data are collected in authentic or almost authentic circumstances, this method may also have a disadvantage such that the kind of utterances may not add much in quantity to the data to be analyzed for a specific purpose, which is intrasentential CS in this thesis. This means, the researcher may have to spend more time with the informants but may not be able to get much intrasentential CS samples.

The second method of data collection is also favored in CS studies because a ready corpus may be analysed from different points of view. For this study, the data collected by Backus (1992) were used. His data were also used by other researchers (Bhatt, 1997; Myers-Scotton *et al.*, 2001). The informants were from the first and second generation of Turkish immigrants in Holland. They were not specially picked on an individual basis. The researcher was not present during the recording to avoid any interference (for detailed information on the informants and the data collection procedure see Backus, 1992) and the discretion when to record was left to the informants.

3.1.2.2 Informant Selection

For the first method that will be used in this thesis, the informants were selected according to particular criteria:

(i) The informant should be from one of the former Soviet Union countries, such as Azerbaijan, Kazakstan, Kirgizistan, etc.

(ii) The informant should have a high verbal fluency in Turkish and English (though he has to have a higher ability in Russian) to be able to switch back and forth among these languages.

The informants were not selected on an individual basis, but as members of a social network, which is the small community of foreign students in the Middle East Technical University that come from the former Soviet Union countries who conform to the above criteria. The reason for choosing a particular social group was that if all the informants present at a recording session knew each other well, they would be more at ease and thus more likely to speak their vernacular. In addition, they would know how to speak to one another, as they did so in their daily lives (Backus, 1992).

To be able to have an idea about the kind of data to be encountered, it is necessary to have an overall information about this community. There are communities of foreign students in the Middle East Technical University, each forming a semi-closed unit to support each other in their daily lives and courses. The students from the former Soviet Union form one of the larger groups among such communities. After the revolution, though some countries changed the language of education from Russian to their own and some even accepted the Latin alphabet (e.g. Azerbaijan), the imposing strength of Russian is still observed in countries such as Kazakstan, Tadjikistan, Turkmenistan, Uzbekistan, and Kirghizia. This language policy throughout the history of Russia is the reason for the dominance of Russian

today in the community of students from the former Soviet Union. As for Turkish, the native languages of some of these students resemble Turkish phonologically, so they are familiar with most of the words. For the other students (e.g. Russian), Turkish is the language of the country they live in so they are expected to have some knowledge of this language. Thirdly, English is a compulsory course for these students therefore they are expected to be able to speak and read in this language, too. As a result, we expected that the spoken data gathered by method (i) would consist of the languages Russian-Turkish-English.

#### 3.1.2.3 Data recording

The group gathered various times in informal settings, generally at the student canteens in various buildings in the university campus so that the informants felt at ease. The time and place of the gathering, and the length of the conversation were chosen by the informants. In each session, the number of the informants varied from 3 to 6. The speech was recorded by an assistant to avoid any dismay or introversion that might be caused by the presence of the researcher. The assistant himself was one of the informants; however, he did not know what kind of data was needed.

## 3.2 Method of analysis

For the first method of data collection, assistance of a native bilingual was employed. Born in Azerbaijan, the assistant is able to speak both Russian and Turkish with native-like fluency. She is also giving Russian courses to undergraduate students in METU. The researcher transcribed and coded the data to indicate the morphological, syntactic, and lexical information, relevant to CS in cooperation with her.

The transcription of Russian was done according to Turkish phonology. The reason behind this is that Turkish and Russian morphology are more alike than Russian and English since both Turkish and Russian have the letters 'ш, х, ж, ч, ы' which correspond to the letters in Turkish 'ş, h, j, ç, ı' respectively. For example, the Russian verb 'читает' is transcribed in Turkish 'çitayet'. The transcription rules are listed in Appendix A.

In Chapter IV, the utterances analysed are presented in the following format to enable the reader to see clearly the code-switched unit:

P16 Dün bütün gün **[IP** *ya çitala knigu***]**
Yesterday all day I read a book.

Moreover, all the sentences that are analyzed analytically in this thesis are coded and translated in the following format and put in Appendices B and C:

| | | | | |
|---|---|---|---|---|
| 2 | Dün | bütün gün | *ya çita-l-a* | *knigu*. |
| | Dün | bütün gün | ya çita-l-a | knig-u |
| | Yesterday | all day | I read-PAST-f/3s | book-ACC |
| | Yesterday | all day | I read | a book. |

The first line of (2) is the datum; the second line is a morphological parse of the datum; the third line is a morpheme-by-morpheme gloss, and the fourth line is an approximate translation into English. If the datum occupies more than a line, the rest continues as a fourth line and the translation is given at the end.

In the datum line, a code switch is indicated in *italics* merely to show a change in the language, not with the specific purpose of identifying the 'matrix' or 'embedded' language (see Myers-Scotton, 1993).

In the gloss line, bound morphemes are separated by a hyphen (-) and free morphemes by a space. If a morpheme has more than one function, the other functions are indicated by a slash (/). The meanings of the inflectional morphemes are glossed in CAPITALS. Abbreviations used in the glosses are listed in the List of Abbreviations section (see p. xii).

# CHAPTER IV

## ANALYSIS OF THE DATA AND DISCUSSION OF THE RESULTS

The type of data gathered and the method of analysis were described in Chapter 3. In this chapter, the spoken data were analysed on the basis of the framework and aims presented in 2.4. Based on the model of mental lexical representations of bilinguals presented in Section 2.3, Fig. 5, we asserted that CS was the interaction of two lexicons from either language in the course of one derivation. According to our hypothesis, intrasentential CS can be in lexical or in phrasal form. If CS is in lexical form, then we propose that the switched element is a head. In this case, we call it borrowing. The words are borrowed in their bare forms and if they have any inflections, they are inflected in the language in which they are transferred. If CS is in phrasal form, we assume that the switched element is a maximal functional projection of a head. The research question was whether the spoken data we have collected falsified this assumption or not.

We have organised this chapter in the following manner: Section 4.1 will be about our data on other theories we have mentioned in Section 2.1. In Section 4.2, we will divide and analyse our data in two main groups: Lexical and phrasal CS.

4.1 Analysis of the data for the reviewed code switching constraints

Poplack's (1980, 1981) FMC approach (cf. Section 2.1.1) predicts that a CS will not occur at the boundary of a bound morpheme. Although it is sometimes difficult to know whether a morpheme is bound or free (Muysken, 1995), the following examples appear to indicate that this assumption is false.

P8 Mıy idyom na **[**$_{N}^{o}$*pazar***]** pakupat **[**$_{DP}$ **[**$_{N}^{o}$*havuç***]**ku**]**, kartoşku, pomidori, **[**$_{N}^{o}$*semizotu***]**
'We are going to the bazaar to buy carrots, potatoes, tomatoes, and purslane'

P24 Rusça'da benim bildiğim, **[**$_{DP}$ **[**$_{N}^{o}$ *subject***]**i**]** söylemeyebilirsiniz, **[**$_{DP}$**[**$_{N}^{o}$ özne**]**yi**]**.
'To my knowledge, in Russian you do not have to say the subject.'

In (P8), the speaker added a russian suffix (*ku*) at the end of (*havuç*), having a parallelism with (*kartoşku*). In the second example, the switched word (*subject*) takes Turkish accusative case. Thus, Poplack's FMC cannot be the operative principle which governs CS.

Poplack's EC theory (cf. Section 2.1.1), the idea that a CS is allowed between constituents only if the relevant word order requirements of both languages are met at S-structure appears to be empirically incorrect as well, as illustrated in the following example. In (P16), the word order of the switched element (SVO) does not match with the Turkish word order requirement (SOV).

P16 Dün bütün gün **[**$_{IP}$ *ya çitala knigu***]**
Yesterday all day I read a book.

Di Sciullo, Muysken and Singh (1986) proposed an anti-government condition on CS, claiming that a CS cannot occur where a government relation holds (cf. section 2.1.2). In GB Theory, subjects are assumed to receive nominative case from the Infl node under government. Thus, any sentence in which a subject DP occurs with a verb from another language will serve as a counter-example to this claim. Example (13), therefore, is a counter-example to this approach, showing that Di Sciullo, Muysken and Singh's system does not capture the operative principles which govern CS.

13 **[**$_{N^0}$ *Negr***]**, **[**$_{DP}$ *krutoi negr***]** rus mafyalarıyla savaşıyor.
'Nigger, cool nigger, fighting with the Russian mafia.'

Belazi, Rubin and Toribio (1994) claim that a CS may not occur between a functional head and its complement (cf. section 2.1.3). (P7) is an example of a case in which a switch occurs between an adjective and a noun, presenting a counter-example to the FHC.

P7 **[**$_{IP}$ **...** **[**$_{DP}$ üç **[**$_{N^0}$ *faza***]]** var**]**
'There are three phases.'

Joshi's (1985) account predicts that a switch into the "embedded language" may not occur at the boundary of a closed-class item (cf. section 2.1.4). Both the example (P13) and the Dutch-Turkish example (92) below form a counter example for this constraint. In (P13) there is a conjunction switch (*no*), and in (92) *waarom* is a closed class item. Joshi's constraint could not therefore be the operative principle governing CS.

P13 Tamam hazırım **[**$_{C^0}$ *no***]** **[**$_{IP}$ ayakkabımı bulamıyorum**]**
'Okay I am ready but I cannot find my shoe.'

92 **[**$_{C}$ *Waarom*] sana **[**$_{A}$ *dof*] geliyor?
'Why does it come out all dull on you?'

Another theory, Matrix Language Framework (MLF) (Myers-Scotton, 1993) claims that it is the matrix language that defines the position of content words and functional elements (cf. Section 2.1.5). When we consider the relatively free word order requirements of the languages (Russian and Turkish) analysed in this thesis, we find it a little bit hard to find counter examples to this theory. In (22), the speaker corrects himself by repeating the IP in the matrix languages, as well. The word orders of the two IPs used in the same sentence are different (head initial for English and head final for Turkish). If there were no correction, then, this could be a counter-example. (144) involves Dutch and Turkish; however, the word order of the switched element agrees with neither Turkish nor Dutch, which might be considered as a counter-example. In that case, if noncanonical word orders are allowed, then MLF must be revised accordingly.

22 Perşembeden önce **[**$_{IP}$ *no chance*], **[**$_{IP}$ şansımız yok**]**.
'We have no chance before Thursday.'

P9 Pazara gidiyorum. Gelir misin **[**$_{PP}$ *s mnoi*]?
'I am going to the bazaar. Will you come with me?'

144 Die deel dat ze **[**$_{IP}$ *hani şeye gitmişlerdi ya dükkana hepsi*]
'that episode in which they all went to the thing, the shop'

A weakness of MLF is that it is difficult in Myers-Scotton's system to know which language counts as the matrix language and which as the embedded language since she adopts "a frequency based criterion" to differentiate these, the matrix language being the one which contributes "more morphemes" to the expressions (1993, p. 68).

According to frequency criterion, it would be hard to guess the ML in the following examples in which three languages are involved in an utterance.

16b **[PP** For example**]**, **[V** bak**]**: "**[CP** Kimler bizi kurtarır**]**" **[Adj** *şto li***]**?
'For example, look: isn't it "Who will save us"?'

25a/b/c **[N°** *Sapojnik***]**, **[Adj** şey yani**]**, **[N°** *shoemaker***]**
'I mean shoemaker.'

Finally, according to MacSwan's proposal (1997, 1999), the words of a language enter into the word formation system of another language, taking the necessary inflections from the system. This process is called the "union" of two lexicons.

A counter-example can be found for this theory in our data. In (P3), the word is switched however, it does not take the correct inflection from the system ( = matrix) language. According to Russian grammatical rules, the verb (*jdat*) takes a genitive case. Then, the correct inflection would be a genitive case, and the correct CS element according to this theory would be "*dolmuş*a".

P3 Ya dolga jdala **[N°** *dolmuş***]** iz za etava ya apazdala
'I waited for *dolmus* for a long time, because of this I was late'

4.2 An analysis of the data

Given the theoretical framework in Sections 2.2, 2.3 and the research question in Section 2.4, our strategy in the present section will be analysing the data according to the type of switches; firstly lexical and secondly phrasal CS.

Below is Table 1, illustrating the general distribution of head and phrase switches in our data:

| | | Type of Switch | | | |
|---|---|---|---|---|---|
| | | Head Switch | Phrase Switch | **Total** | **%** |
| **Languages Involved** | Russian/Turkish | 36 | 32 | **68** | **77** |
| | Russian/English | 6 | 6 | **12** | **14** |
| | Turkish/English | 2 | 6 | **8** | **9** |
| | **Total** | **44** | **44** | **88** | |
| | **%** | **50** | **50** | | |

Table 1: General Distribution of Switched Heads and Phrases

4.2.1 Lexical CS

We have observed that in 88 CS instances recorded, 44 of the cases are lexical CS cases. The distribution of the instances according to the type of heads switched are as follows:

| Languages Involved | | Type of Head Switched | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | $N^o$ | V | A | C | Lex. Adj. | **Total** | **%** |
| | Russian/Turkish | 27 | 2 | 2 | 2 | 4 | **36** | **82** |
| | Russian/English | 6 | 0 | 0 | 0 | 0 | **6** | **13** |
| | Turkish/English | 3 | 0 | 0 | 0 | 0 | **2** | **4** |
| | **Total** | **35** | **2** | **1** | **2** | **4** | **44** | |
| | **%** | **79** | **6** | **2** | **4** | **9** | | |

Table 2: Distribution of Head Switches in Russian/Turkish/English Data

4.2.1.1 N - Switch

In the Russian-Turkish-English lexical CS cases, the majority of the switched heads are nouns (35 cases out of 44 lexical CS cases).

P1 Stambul prekrasnıy gorad s ogromnim ojivleniyem, şto sozdayot **[$_N{}^o$** *kalabalık***]**
'İstanbul is a beautiful city with very much liveliness that creates crowd'

P3 Ya dolga jdala **[$_N{}^o$** *dolmuş***]** iz za etava ya apazdala
'I waited for dolmus for a long time, because of this I was late'

P12 **[$_N{}^o$** *Transport***]** çok pahalı
'Transportation is very expensive'

In the utterances where $N^o$ is switched, the cases are not processed. For example, (P4) would require an accusative case both in Turkish and in Russian.

P4 İstediğim **[**$_{N}^{o}$ *pasta***]** bulamadım
'I could not find the toothpaste I wanted'

If the case is nominative and the word is in bare form, it is difficult to decide if we have borrowing or phrasal CS. Nominative case is an inflection and an inflected N is a phrase. Then, the example P4 would be in accordance with our hypothesis. If we consider (*pasta*) an episodical borrowing, then, we would expect it to take the correct inflection from Turkish; however, it did not take an accusative case.

In some cases, the speaker arranges the case of the switched noun according to the requirements of the matrix language. For example, in (P7), "faza" is in singular form; thus, is grammatical according to the requirements of Turkish. However, in Russian, nouns have to agree with the adjective. One explanation may be that, being a Russian, the speaker paid attention to the agreement rules in Turkish.

P7 üç **[**$_{N}^{o}$ *faza***]** var
'There are three phases.'

We believe the above are all episodic borrowing cases. This means, the words are borrowed for one time usage and then given back to their original languages, as explained in Chapter 2.4.

However, the following examples are instances where, we believe, are not episodical borrowing. On the contrary, the switched words are widely used as part of the language.

3a U niyo **[**$_{N}^{o}$ *nick***]** Gizka bıla, karoçe, Nargizka.
'Her nickname was Gizka, short form of Nargizka.'

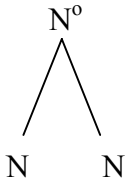
3b Saifa gavarit tipa svai **[**$_{N}^{o}$ **[**$_{N}$ *mail***]** **[**$_{N}$ *address***]]** dai.
'Saifa says: "give me your own mail address".'

3c Ana gavarit, vot u Mirzı spraci **[**$_{N}^{o}$ **[**$_{N}$ *mail***]** **[**$_{N}$ *address***]]** u padrugi yest.
'She tells him to ask Mirza for the mail address of her girlfriend.'

In (3a), (3b), and (3c), the switched lexicons are words that are internet terms. The same comment can be made for *Pepsi Generation Next* (7a) since this is an advertorial motto and is used worldwide.

7a Ladna, **[**$_{N}^{o}$ **[**$_{NP}$ *Pepsi Generation Next***]]**, gavari dalşe.
'Okay, Pepsi, Generation Next, ask more.'

Examples (3b) and (3c) above and (P2), (P6) and (P12) below are interesting cases. Our classifying these utterances as borrowing examples may seem as a controversy with Chapter 2 in which we claimed that only lexical intrasentential CSs are borrowing cases. Here, we would like to point out that we take compound nouns as a head, not a projection of a noun. We do this by basing our claim that compound nouns are the result of an incorporation rule deriving the following structure:




P2 Maya **[**$_{N}^{o}$ *çamaşır makinesi***]** vçera vışla iz stroya
'Yesterday my washing machine broke down'

P6 Na uroke **[**$_{N}^{o}$ *Hayat Bilgisi***]** mıy sevodniya çitali pro **[**$_{N}^{o}$ *Atatürk'ün Samsun'a Girişi***]**
'Today we read about *Atatürk'ün Samsun'a Girişi* in *Hayat Bilgisi* lesson'

12 Yest je film "**[**$_{N}^{o}$ Çarlinin Melekleri**]**".
'There is a film called "Charlie's Angels".'

4.2.1.2 V Switch

We have two examples of V-switch in our data. In both cases, there is a subject-verb agreement and the sentences are grammatical.

Balşoi futbol ni dlya minya, **[**$_{V}^{o}$ *de***]**, ya ni budu igrat.
21 'Say that big football is not for you and you won't play.'

5a/b Nu karoçe, ya c etoi c Nargizoi, **[**$_{PP}$ *by chat***]**, **[**$_{V}^{o}$ *chat***]**uem, ana i gavarit: "Tıy minya ni pugai"
'In short, I was chatting with that, with Nargiza by chat, and she said: "Don't scare me".'

4.2.1.3 C Switch

C switch is a debated area as discussed in Section 2.1.4. As was pointed out there, it was proposed that an element of a closed class could not be switched. We have one instance (P13) where only C is switched. However, we have taken examples of other instances from the Dutch-Turkish example where not only the conjunction but also the clause coming after was switched as well. These examples, (101), (144), (109) are listed in section 4.3.2.

P13 Tamam hazırım **[**$_{C}^{o}$ *no***]** **[**$_{IP}$ ayakkabımı bulamıyorum**]**
'Okay I am ready but I cannot find my shoe.'

#### 4.2.1.4 Lexical Adjunct Switch

The following examples are classified as adjuncts. Adjuncts are also maximal projections, thus adjunct switching as such is compatible with our hypothesis. Here, the speakers are from different former-Soviet Union countries. In lexical adjuncts, their language preference is Russian.

20b **[$_{Adj}$** *A*], dün bayramdı ya, olabilir ondan.
'Yesterday was a holiday, might be because of that.'

23 **[$_{Adj}$** *Davai*] Atatürk'ü çaldırayım.
'Let me call (by phone) Ataturk (dormitory).'

18a **[$_{Adj}$** *Da*] araştırma, ben araştırma görevlisiyim ya.
'A research, you know, I am a research assistant.'

### 4.2.2 Phrasal CS

The phrasal switches can be summarised in the folowing table:

| | **Type of Phrase Switched** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Languages nvolved** | | IP | DP | PP | VP | CP | Phr. Adj. | **Total** | **%** |
| | Russian/Turkish | 10 | 17 | 3 | 0 | 0 | 2 | **32** | **72** |
| | Russian/English | 0 | 1 | 1 | 1 | 2 | 1 | **6** | **14** |
| | Turkish/English | 3 | 1 | 2 | 0 | 0 | 0 | **6** | **14** |
| | **Total** | **13** | **19** | **6** | **1** | **2** | **3** | **44** | |
| | **%** | **29** | **43** | **14** | **2** | **5** | **7** | | |

Table 3: Distribution of Phrase Switches in Russian/Turkish/English Data

The total amount of phrasal CS cases is 44 out of a total of 88.

4.2.2.1 IP Switch

11 out of 43 phrasal CS cases belong to IP switches. Some examples are given below.

P19 Beli noçi **[**$_{N}^{o}$ iyun**]**, **[**$_{IP}$ **[**$_{N}^{o}$ *Haziran***]** *'da oluyor***]**
'White Nights are in June'

P16 Dün bütün gün **[**$_{IP}$ *ya çitala knigu***]**
'Yesterday all day I read a book.'

P15 Türkçe konuştuğumda Türkçe düşünmek gerekiyor; fakat hâlâ **[**$_{IP}$ *ya dumayu pa ruski***]**
'Speaking in Turkish requires thinking in Turkish. However, I still think in Russian.'

8a/b Da, **[**$_{IP}$ eta normalna**]**, **[**$_{IP}$ *olan birşeydir***]** , **[**$_{IP}$ *normaldir***]**.
'Yes it is normal, it happens, it is normal.'

19c ... kasiyer hala gazete okuyor, 'Ya kardeşim, duymuyor musun, sıraya geçsene'. **[**$_{N}$ *On***]** **[**$_{Adj}$ *opyat***]** oraya buraya bakıyor, **[**$_{IP}$ *nikavo nyetu***]**....
'Cashier still reads a newspaper,"Hey, brother, don't you hear, wait for your turn". He again looks around, there was not anybody else...'

21a Karoçe, ... **[**$_{IP}$ *Arkadaşlar, ben yokum***]**, bez minya budite igrat.
'Shortly,... Friends, do not count me in, the play will be without me.'

7g Dalşe davai. **[**$_{IP}$ *Huylanma oğlum***]**.
'Let's proceed. Don't fret, man. '

14b Ni znayu, **[**$_{IP}$ *got no information***]**.
'I don't know, I got no information.'

W1 **[**$_{DP}$ *The thing they wanted to infer***]** **[**$_{IP}$ Coca Cola aldığımızda her yerde ve her zaman şampiyon oluruz**]**
'The thing they wanted us to infer is that when we buy Coca Cola we become the champions every where and every time'

22 Perşembeden önce **[**$_{IP}$ *no chance***]**, **[**$_{IP}$ şansımız yok**]**.
'We have no chance before Thursday.'

14c Şto, **[**$_{CP}$ şto on tam pateryal**]**, **[**$_{CP}$ *ne kaybetmiş***]**?
'What he did he lose there?'

16c *For example*, **[**$_{V}^{o}$ bak**]**: "**[**$_{IP}$ Kimler bizi kurtarır**]**" **[**$_{Adj}$ *şto li***]**?
'For example, look: isn't it "Who will save us"?'

17 Orada tek başına karda **[**$_{IP}$ *pro plavayet***]**.
'And there alone he swims in snow.'

The reason for the frequent use of IP switch can be explained by the relative ease of switching. The switched IP does not need to agree with anything. In the above examples, they act more like an adjunct, an autonomous sentence. When we deduct the switched IP from these utterances, the sentence would still be meaningful. The same can be said for the adjuncts as well (see Section 4.2.2.5).

#### 4.2.2.2 DP Switch

The most frequent CS is DP switches. According to MP, if a noun is not in its bare form, it is a DP[1]. Therefore, like $N^{o}$-switches, DP is the most common in the phrasal CS cases. Some examples are given below.

13 **[**$_{N}^{o}$ *Negr***]**, **[**$_{DP}$ *krutoi negr***]** rus mafyalarıyla savaşıyor.
'Nigger, cool nigger, fighting with the Russian mafia.'

18b Büyük insan, **[**$_{DP}$ *krutoi çelavek***]**.
'You are a big man, cool guy.'

1 Karoçe, **[**$_{DP}$ *ilahiyatdan***]** sadnoy paznakomilciya tı pomniş, tı je gavaril: "Vo prikolnaya", tipa takova, da.
'Do you remember, you were saying: "that cool", something like that.'

2 Pomniş, mıy **[**$_{DP}$ *Kızılay'dan***]** ili je?
'Do you remember we were walking through Kizilay?'

[1] According to Radford (1995) and Fromkin (2000), NP is a complementizer of a DP. Because of lack of source that implements MP to Turkish and Russian languages, we have followed this rule in the analysis of Russian and Turkish CS instances.

4.2.2.3 PP Switch

In the examples where a Russian preposition is switched, the preposition (*s*) has the majority. However, (P22) is an interesting case. This PP is colloquial and is only used by Russian speakers who also know Turkish well enough. (*sbaşki)* has the same meaning with *(kafama göre*). So, (*baş*) has been taken directly from Turkish and cliticised with the preposition (*s*) to mean *as I want*.

P9 Pazara gidiyorum. Gelir misin **[**$_{PP}$ *samnoi***]**?
'I am going to the bazaar. Will you come with me?'

P22 **[**$_{PP}$ kafama göre**]**, **[**$_{PP}$ *sbaşki***]**, zor olur
'It would be hard to do it as I want.'

Examples (P20b) and (P22) are also discussed in section 4.3 as there is repetition.

P20a /b **[**$_{C}^{o}$ *işte***]** dvaretz **[**$_{PP}$ *taa şeyden***]** **[**$_{PP}$ z daleka**]** vidna bolşoy kaskad.
'From a distance you could see the palace, the big fountain.'

(5a) is also worth mentioning here. We have already stated that *chat* is not considered an episodical borrowing. This is because there is no equivalence of this word in Russian. We may even say that (5b) is not CS either.

5a Nu karoçe, ya c etoi c Nargizoi, **[**$_{PP}$ *by chat***]**, **[**$_{V}^{o}$ *chat***]**uem, ana i gavarit: "Tıy minya ni pugai"
'In short, I was chating with that, with Nargiza by chat, and she said: "Don't scare me".'

16a **[**$_{PP}$ *For example*] desin.
'Let him give an example.'

4.2.2.4 CP Switch

The situation with CPs resemble that of IPs in that they are quite independent of the rest of the sentence. They act like an adjunct and they do not need to agree with any subject or object. However, we have not encountered much examples of CP switch in our data. In fact, we have only one English CP in a Russian sentence (7c).

7c Verite li vıy v lubov, **[**$_{CP}$ *if yes, give example*].
'Do you believe in love, if yes, give an example.'

4.2.2.5 Phrasal Adjunct Switch

It is observed that adjunct as a phrase is switched rarely. We have one instance of adjunct switching in each language pair.

15b A bolşe nikavo tam ni bıla **[**$_{Adj}$ *şansına senin*]?
'And to your luck wasn't there anybody?'

16c *For example*, **[**$_{V^0}$ bak**]**: "**[**$_{CP}$ Kimler bizi kurtarır**]**" **[**$_{Adj}$ *şto li*]?
'For example, look: isn't it "Who will save us"?'

7b Nu ya yişo ni videl, **[**$_{Adj}$ *maybe later*].
'I haven't seen anything yet, maybe later.'

4.3 Correction/Re-Analysis/Repetition

In the data gathered, we have come across instances where either a lexical item, or a phrase have been repeated in different languages. Of all the CS instances (a total of 88 utterances), 12 of them bear this phenomenon. Therefore, we believe this phenomenon is worth mentioning.

We have found no information about the repetition and correction phenomena in the literature, though we found instances of repetition in the examples in the literature, such as (5):

(5) anaka-ndir *intercultureel werk*
'I I-am-doing intercultural work'
(Nortier cited in Muysken, 1995, p. 187)

In our data, we have realised a pattern in these instances. In some cases, the speaker is aware that she has made a CS, and she tries to revert to the main language. Examples (P18), (P24), and (P25a/b/c) are such cases.

P18 *kanikula,* şey, tatilden yemeklere geçtik.
'We changed the subject from vacation to food.'

P24 Rusça'da benim bildiğim, *subject*i söylemeyebilirsiniz, özneyi.
'To my knowledge, in Russian you do not have to say the subject.'

P25a /b/c *Sapojnik*, şey yani, *shoemaker.*
'I mean shoemaker.'

In some cases, there is a mere repetition. This repetition may be done to give emphasis to the repeated idea.

P21 yeterince arkadaşı, *dastatıçnıy druzey* vardı.
'She had enough friends.'

P22 kafama göre, *sbaşki*, zor olur
'It would be hard to do it as I want.'

8a/b Da, eta normalna, *olan birşeydir*, *normaldir*.
'Yes it is normal, it happens, it is normal.'

14b Ni znayu, *got no information.*

'I don't know, I got no information.'

14c Şto, şto on tam pateryal, *ne kaybetmiş*?
'What he did he lose there?'

18b Büyük insan, *krutoi çelavek.*
'You are a big man, cool guy.'

22 perşembeden önce *no chance*, şansımız yok.
'We have no chance before Thursday.'

In another case, the repetition is done automatically and the speaker might not be aware that he has repeated the same morpheme, as in (6).

6 V etam *dönemde* u tibya kagda ekzamenıy zakançivayutsiya?
'When do your exams finish this term?'

The reason behind this repetition could be the fact that locative case is a bound morpheme in Turkish, however, a free morpheme in Russian.

## 4.4 Dutch-Turkish Data

We have decided to use a corpus compiled by Ad Backus (1996) in order to test the validity of our hypothesis with other languages and more data. We again analyse the instances of CS in two groups; lexical and phrasal CS.

We have summarised the CS examples we have selected from Dutch/Turkish data in the following tables:

| **Type of Head Switched** | | | | | |
|---|---|---|---|---|---|
| N$^{o}$ | V | A | C | Lex. Adj. | **Total** |
| 1 | 1 | 2 | 1 | 0 | **5** |

Table 4: Distribution of Switched Heads in the Dutch/Turkish Data

| Type of Phrase Switched | | | | | |
|---|---|---|---|---|---|
| IP | DP | PP | CP | Phr. Adj. | **Total** |
| 1 | 1 | 2 | 1 | 0 | **5** |

Table 5: Distribution of Switched Phrases in the Dutch/Turkish Data

### 4.4.1 Lexical CS

There were no examples of an A switch from the naturally occuring Russian-English-Turkish data. Thus, analysing other data also enabled us to see examples of different switched elements as well, i.e., A switch.

84 Ya, belli **[**$_{A}$ *donker***]** olmuş.
'Yeah, it definitely got darker'

We have found another example for a C switch, as below:

92 **[**$_{D}$ *Waarom***]** sana **[**$_{A}$ *dof***]** geliyor?
'Why does it come out all dull on you?'

We have also taken an $N^{o}$ incorporation example from Dutch-Turkish data:

250 heeft Gül ook, en die gaat Gül effe dragen onder mijn **[**$_{N}^{o}$ *kına gecesi***]** rok.
'Gül has those too, and she is going to wear them under my kina night skirt.'

### 4.4.2 Phrasal CS

In the examples we have taken, the majority belongs to CP switches.

101 Çok koymuyorum biliyor musun, **[**$_{CP}$ *en laatst had ik het hardst warm he***]**.
'I do not put on a lot you know, and recently I had it going on the hottest, warm, right.'

144 hani çarşıya çıktılar ya **[**$_{CP}$ *en toen daarna de volgende deel***]**.
'well, they went out to town, right, and then after that the next episode'

109 Müjde kadar ufaktı **[**$_{CP}$ *en die zei alles ech waar heel mooi, dat joch***]**
'About as little as Müjde, and he said everything in really such a nice way, that little guy.'

We also have another adjunct phrase switched, (174), just as the examples in Russian-English-Turkish data.

174 Ik wilde Turkse nou niet beledigen, maar **[**$_{Adj}$ *yok bir düşüneyim***]**,
'I didn't want to insult Turks or something but let me think first'

# CHAPTER V

# CONCLUSION AND FURTHER PROSPECTS

## 5.1 Summary of the study

This study has demonstrated how intrasentential CS data provide some empirical evidence for certain specifications about the nature of how different lexicons interact in the bilingual speech production process.

In Chapter 2, we showed that CS could be in two forms. In Fig. 7, we illustrated CS in lexical form, which is also called borrowing. In this kind of CS, the switched elements are always heads. In the second kind of CS, the switched element is in phrasal form. Our hypothesis was that this switched element is maximal functional projections.

Our specific research question was whether the data we have collected supported this assumption or not.

We used two kinds of data for the thesis. The first kind consisted of the data we gathered from Russian-English-Turkish multilinguals, who were students of the Middle East Technical University. For the second kind of data, a ready corpus was used. This corpus had been prepared in Holland, among first and second generation

of Turkish immigrants (Backus, 1996). We transcribed and analysed each utterance with CS phenomenon, as explained in Chapter 3.

In Chapter 4, we grouped the data according to size of the switched element, i.e. lexical CS and phrasal CS. Then, we formed subgroups according to the themes of the switched elements, e.g. N-switch, DP-switch. In the same chapter, we also mentioned re-analysis phenomenon, which we encountered in the naturally occurring data we gathered. The last thing we did in Chapter 4 was to find support for our hypothesis with Dutch-Turkish spoken data examples.

In this chapter, we are going to present our deductions as a result of our analysis of the data. Based on our experience, we are also going to make remarks for future studies in this field.

## 5.2 Generalisations

The examples of CS both in the Russian-English-Turkish and the Dutch-Turkish data presented so far are compatible with our specific hypotheses. This supported our hypotheses that:

1. In lexical CS, the switched item is always a lexical head.

2. In phrasal CS, the switched element is a functional maximal projection of a head.

Our starting point, which was the claim that bilinguals have discrete and separate lexicons for the specific languages they know (Fig. 6) and that speaking is

an incremental process, is in accordance with the CS examples gathered. Then, the following hypotheses can be established:

3. At the conceptual level, the bilingual makes choices about the semantic/pragmatic information that he wishes to convey.

4. The speaker has discrete and separate lemmas for the different languages he knows.

5. The cognitive process of language production is an incremental process.

6. It is possible for the speaker to have access to these lemmas during conversation, which enables CS.

We will also question whether taking a minimalist view in analysing bilingual CS was appropriate or not. Firstly, in MP, elements of the lexicon are assumed to be inserted in the course of derivation. Secondly, MP assumes that there is a computational system ($C_{HL}$) which is fixed and invariant across languages and only lexicons, which are derived 'on-line' bear the idiosyncratic properties of individual languages. In our data, we have lexical heads and functional maximal projections, which are derived from discrete and separate lexicons. These assumptions are in line with the information-processor model of Levelt, and thus our model of bilingual language processing (Fig. 6).

7. The Minimalist Approach has an explanatory force for bilingual language processing.

Lastly, we have found out that the most common lexical switching is the switch of N (35 out of 44, see Table 2). This phenomenon is also supported by the current research (Hock, 1991; Sobin, 1982; van Hout, Roeland and Muysken, 1994) which suggest that nouns are crosslinguistically more borrowed than verbs (Hock, 1991). As a result of our findings, we can also generalise as follows:

8. In lexical CS, the most common type of switching is seen in N.

## 5.3 Further implications and prospects

For further studies, it is recommended to gather larger corpora to be able to test the hypothesis to increase its validity. Since we claim universality of the hypothesis, the testing should also be done for other language pairs as well.

An implication of such a study would be investigating CS phenomenon as an aspect of bilingual development process. How does the bilingual lexical organization of the brain affect the language learning process of bilinguals? How would it be possible to modify the educational system? The educational aspect was beyond the scope of this study as mentioned in Section 1.4; however, understanding the CS phenomenon may help to improve the bilingual development process.

Another implication would be analysing the CS instances from a psycholinguistic perspective. This is a major area of investigation and it covers wider phenomena (Gumperz, 1964; Hymes, 1967; Labov, 1972; Clark, 1996, Zentella,

1997). Pragmatic approaches to CS deal with the relation between structure and function in everyday speech exchanges.

As a further prospect, the hypotheses in this study can be viewed in conjunction with the Activation Threshold Hypothesis (Paradis, 1996). Paradis explains this hypothesis as follows:

> The strength of a representation or memory trace (i.e., its propensity to being activated, its threshold level) is a function of both frequency of its activation and the recency of its last activation. Recall or retrieval of a linguistic item requires a lower threshold of activation than recognition (i.e., activation by an impinging outside stimulus) (Paradis, 1996, p. 17).

According to this explanation, a bilingual may consciously or subconsciously switch language during speech because that phrase, word, sentence, or language has been used either more frequently or more recently than another language. In the present thesis, we have encountered examples of such a phenomenon such as (24), below.

| | |
|---|---|
| Kagda u vas | *finaller* naçinayutsiya? |
| Kagda u vas | final-ler naç-ina-yut-siya? |
| When Pos 2pl/Pos pronoun | final-pl begin-Prog-Pres/3pl-Ref? |
| 'When do your finals begin?' | |

The student who is talking is from a former Soviet Union countries. He is also a student at Middle East Technical University, a Turkish university with an English medium. The reason for his preference of the word "finaller" can be

explained by the Activation Threshold Hypothesis as a result of the high frequency of the use of this word among his friends.

Whatever is the reason for CS, CS seems to have a function of supporting communication. It remains for us to see how we might approach such facilitation in the bilingual classroom.

Bilingualism is a subject yet to be discovered both to understand the underlying principles that apply to all human languages and to investigate the relationship of each language with each other. The benefits of the applications of such studies are numerous. This fact makes the CS phenomenon worth examining.

## REFERENCES


Backus, A. (1996). *Two in one. Bilingual speech of Turkisch immigrants in the Netherlands.* Tilburg: Tilburg University Press.

Bentahila, A. and E. E. Davies. (1995). Patterns of code-switching and patterns of language contact". *Lingua, 96*, 75-93.

Bhatt, M. (1997). Codeswitching, constraints, and optimal grammars". *Lingua*, *102*, 223-251.

Calvin, H. M. and G. A. Ojemann. (1994). *Conversations with Neil's brain: The neural nature of thought and language*. New York: Addison-Wesley.

Chomsky, N. (1995). *The minimalist program*. Cambridge: MIT Press.

Cook, V. J. (1988). *Chomsky's universal grammar*. Oxford: Basil Blackwell.

Crystal, D. (1986). Bilingualism. In D. Crystal (Ed.). *The Cambridge Encyclopedia of Language*, 360-367. Cambridge: Cambridge University Press.

Dabene, L. and P. Muysken. (1995). Bilingual Speech of Migrant People. In L. Milroy and P. Muysken (Eds.). *One speaker, two languages* (pp. 17-44). Cambridge: Cambridge University Press.

Di Sciullo, A., P. Muysken and R. Singh. (1986). Government and code-switching. *Journal of Linguistics, 22*, 1-24.

Duran, L. (1994). Toward a better understanding of code switching and interlanguage in bilinguality: Implications for bilingual instruction". *The Journal of Educational Issues of Language Minority Students, 14*, 69-88.

Edwards, J. (1995). *Multilingualism*. London: Penguin Books.

Epstein, S., E. Groat, R. Kawashima & H. Kitahara. (1998). *A derivational approach to syntactic relations*. Oxford: Oxford University Press.

Fromkin V. (2000). *Linguistics : an introduction to linguistic theory.* Blackwell, Massachusetts.

Garrett, M. (1975). The analysis of sentence production. In G. Bower (Ed.) *Psychology of Learning and Motivation.9*. New York: Academic Press.

Gerda, M. (1993). *Multilingualism and nation building*. Bristol: Longdunn Press.

Grosjean, F. (1982). *Life with two languages*. Cambridge: Harvard University Press.

Gumperz, J. and A. J. Toribio. (n.d.). Code switching. Retrieved <http://cognet.mit.edu.tr/MITECS/Articles/gumperz.html>

Hock, H. (1991). Principles of historical linguistics. (2nd ed.). Hamburg: Mouton de Gruyter.

Hoffman, C. (1991). *An introduction to bilingualism*. New York: Longman Inc.

Kim, K. H. S., N. R. Relkin, K. Lee and J. Hirsch. (1997). Distinct cortical areas associated with native and second languages". *Nature, 388*, 171-4.

Kolb, B. and I. Q. Whishaw. (1996). *Fundamentals of human neuropsychology*. New York: W. H. Freeman and Company.

Kurtböke, P. N. (1998). *A corpus-driven study of Turkish-English language contact in Australia*. Unpublished doctoral dissertation, Monash University, Melbourne.

Lamb, S. (1999). *Pathways of the brain: The neurocognitive basis of language*. Amsterdam and Philadelphia: John Benjamins Publishing Co.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, Mass.: MIT Press.

Levelt, W. J. M. A theory of lexical access in speech production. Retrieved <http://www.cogsci.soton.ac.uk/bbs/Archive/bbs.levelt.html>

Longxing, W. (2001). The bilingual mental lexicon and speech production process. Doi:10.1006/brln.2001.2557. Retrieved <http://www.ideallibrary.com>

MacSwan, J. (2000). The threshold hypothesis. *Hispanic Journal of Behavioral Sciences, 22(1)*, 3-45. Retrieved <http://www.ed.asu.edu/coe/faculty/macswan.htm>

MacSwan, J. (1999). A minimalist approach to intrasentential code switching. Retrieved < http://www.ed.asu.edu/coe/faculty/macswan.htm>

MacSwan, J. (1999). The architecture of the bilingual language faculty: Evidence from intrasentential code switching". *Bilingualism: Language and Cognition, 3(1)*, 37-54. Retrieved <http://www.ed.asu.edu/coe/faculty/macswan.htm>

MacSwan, J. (1997). The status of NPs in Southeast Puebla Nahuatl: Comments on the polysynthesis parameter." Forthcoming paper to be presented at the Annual Meeting of the Linguistic Association of the Southwest (LASSO XXVI), October 3-5, Los Angeles.

Meisel, J. M. (2000). Revisiting universal grammar. *Working Papers in Multilingualism*. (Issue Brief No. 1B).

Muysken, P. (1995). Code-switching and grammatical theory. In L. Milroy and P. Muysken (Eds.). *One speaker, two languages* (pp. 177-198). Cambridge: Cambridge University Press.

Myers-Scotton. (1993). *Duelling languages: grammatical structure in code switching*. Oxford: Clarendon Press.

Myers-Scotton, C. (1995). A lexically based model of code switching. In L. Milroy and P. Muysken (Eds.). *One speaker, two languages* (pp. 223-256). Cambridge: Cambridge University Press.

Nishimura, M. (1995). Varietal conditioning in Japanese/English code-switching. *Language Sciences, 17( 2)*, 123-145.

Paradis, M. (1996). Bilingualism. In J. G. Beaumont, P. M. Kenealy and M. J. C. Rogers (Eds.). *Blackwell dictionary of neuropsychology* (pp. 148-154). Cambridge: Blackwell Publishers.

Paradis, M. (2000). Generalizable outcomes of bilingual aphasia research. *Folia Phoniatrica et Logopaedica, 51(1-3)*, 54-64.

Pinker, S. (1994). *The language instinct*. New York: W. Morrow and Co.

Poplack, S. and M. Meechan. (1995). Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilinguals. In L. Milroy and P. Muysken (Eds.). *One speaker, two languages* (pp. 199-232). Cambridge: Cambridge University Press.

Radford, A. (1998). *Syntax – a minimalist introduction*. New York: Cambridge University Press.

Ramat, A. G. (1995). Code-switching in the context of dialect/standard language relations. In L. Milroy and P. Muysken (Eds.). *One speaker, two languages* (pp. 45-67). Cambridge: Cambridge University Press.

Romaine, S. (1995). *Bilingualism*. Second Edition. Oxford: Blackwell Publishers Ltd.

Skiba, R. (1997). Code switching as a countenance of language interference. Retrieved <http://www.aitech.ac.jp/~iteslj/Articles/Skiba-CodeSwitching.html>

Sobin, N. (1982). Texas Spanish and lexical borrowing. In J. Amastae & L. Elias-Oliveras, (Eds). Spanish in the United Sattes: Sociolinguistic aspects (pp. 166-181). Cambridge: Cambridge University Press.

Springer S. P. and G. Deutsch. (1998). *Left brain, right brain*. New York: W. H. Freeman and Company.

Stepanova E. M., Z. N. Ievleva, L. B. Trushina, R. L. Baker. (1994). *Russian for everybody* (V. Kostomarov, Ed.). (7th ed.). Moscow: Russky Yazyk Publishers.

Van Hout, Roeland and Pieter Muysken. (1994). Modelling lexical borrowability. *Language variation and change, 6*, 39-62.

Zentella, A. C. (1997). *Growing up bilingual*. Oxford: Blackwell Publishers.

## APPENDIX A

## TRANSLITERATION RULES

The transliteration rules for codification of Russian utterances are as follows:

| Russian - Turkish | Russian - Turkish |
|---|---|
| а - a, | п - p, |
| б - b, | р - r, |
| в - v, | с - s, |
| г - g, | т - t, |
| д - d, | у - u, |
| е - ye, | ф - f, |
| ё - yo, | х - h, |
| ж - j, | ц - c, |
| з - z, | ч - ç, |
| и - i, | ш - ş, |
| й – ıy, | щ - şş, |
| к - k, | ы - ı, |
| л - l, | э - e, |
| м - m, | ю - yu, |
| н - n, | я - ya. |
| о - o, | |

## APPENDIX B

## DATA FROM PRELIMINARY STUDY

First Recording

(P1)

Stambul prekrasnıy gorıd c ogromnim ojivleniyem,
Stambul prekrasnıy gorıd c ogromni-m ojivleniye-m,
İstanbul beautiful city with enormous-GEN animation-GEN

şto sozdayët *kalabalık*
şto sozda-yët kalabalık
that create-3s crowd

'İstanbul is a beautiful city with very much liveliness that creates crowd'

(P2)

| | | | |
|---|---|---|---|
| Maya *çamaşır makinesi* | vçera | vışla | iz stroya |
| Ma-ya *çamaşır makine-si* | vçera | vış-l-a | iz stroya |
| My-f washing machine-3s | yesterday | go out-Past-f | from road |

'Yesterday my washing machine broke down'

(P3)

| | | | |
|---|---|---|---|
| Ya dolga jdala | *dolmuş* iz | za etava ya apazdala | |
| Ya dolga jda-l-a | dolmuş iz | za eta-va ya apazda-l-a | |
| I long wait-Past-1s-f | dolmus | for this I late-Past-1s-f | |

'I waited for *dolmus* for a long time, because of this I was late'

(P4)

İstediğim *pasta* (diş macunu) bulamadım

İste-diğ-im *pasta* bul-ama-dı-m
Want-ObjP-1s toothpaste find-Neg-Past-1s
'I could not find the toothpaste I wanted'

(P5)

Dışarıya çıkıyorsan *şapku* giy
Dışarı-ya çık-ıyor-sa-n şapk-u giy
Out-DAT go out-PrProg-COND-2s hat-ACC wear
'If you are going out, wear your hat'

(P6)

Na uroke *Hayat Bilgisi* mıy sevodniya çitali pro
Na urok-e *Hayat Bilgisi* mıy sevodniya çita-l-i pro
in lesson-PREP *Hayat Bilgisi* we today read-Past about

*Atatürk'ün Samsun'a Girişi*
*Atatürk-ün Samsun-a Giriş-i*
*Atatürk-GEN Samsun-DAT Giriş-ACC*

'Today we read about *Atatürk'ün Samsun'a Girişi* in *Hayat Bilgisi* lesson'

(P7)

üç *faza* var
üç faza var
three phase be
'There are three phases.'

(P8)

P8a/b Mıy idyom na *pazar* pakupat *havuç*ku, kartoşku, pomidori, *semizotu*
Mıy id-yom na *pazar* pakupat *havuç*ku, kartoşku, pomidor-i, *semizotu*
We go-1p/Pres to bazaar to buy carrot, potatoe, tomatoe-pl, purslane
'We are going to bazaar to buy carrots, potatoes, tomatoes, purslane

(P9)

Pazara gidiyorum. Gelir misin *s* *mnoi*?
Pazar-a gid-iyor-um gel-ir mi-sin s mnoi
Bazaar-DAT go-PrProg.-1s come-AOR Q-3s with me
'I am going to the bazaar. Will you come with me?'

(P10)

I will buy apples *i banan.*
I will buy apple-s i banan

I buy-FUT apple-pl and banana-pl
'I will buy apples and bananas.'

(P11)

P11a /b Yürüyüşe giderken *sobaki oçin* korkuyorum
Yürüyüş-e gid-er-ken sobak-i oçin kork-uyor-um
Walking-DAT go-AOR-WHEN dog-pl very scared-PrProg.-1s
'When going for a walk, I am scared of dogs.'

(P12)

*Transport* çok pahalı
*Transport* çok pahalı
Transportation very expensive
'Local transportation is very expensive'

(P13)

Tamam hazırım *no* ayakkabımı bulamıyorum.
Tamam hazır-ım no ayakkabı-m-ı bul-am-ıyor-um
Okay ready-1s but shoe-1s-ACC find-Neg-PrProg.-1s
'Okay I am ready but I cannot find my shoe.'

(P14)

P14a /b/c/ d İyi İngilizce bilen bir Türk'le konuşurken *industrializatsiya,*
İyi İngilizce bil-en bir Türk'-le konuş-ur-ken *industrializatsiya,*
Good English know-X one Turk-WITH talk-Aor-WHILE industrialization

*revolutsiya, sinonim* ve *lokalizatsiya* gibi kelimeleri Rusça söylüyorum.
*revolutsiya, sinonim* ve *lokalizatsiya* gibi kelime-ler-i Rusça söyl-üyor-um.
Revolution, synonym and localization like word-pl-ACC Russian say-PrProg-1s

'When speaking to a Turkish, I say words like industrialisation, revolution, synonym, and localisation in Russian.'

(P15)

Türkçe konuştuğumda Türkçe düşünmek gerekiyor; fakat hâlâ
Türkçe konuş-tuğ-um-da Türkçe düşün-mek gerek-iyor; fakat hâlâ
Turkish speak-ObjP.-1s-WHEN Turkish think-INF require; however still

*ya dumayu pa ruski*
ya duma-yu pa ruski
I think-1s/Pres in Russian

'Speaking in Turkish requires thinking in Turkish. However, I still think in Russian.'

(P16)

| | | |
|---|---|---|
| Dün bütün gün | *ya çita-l-a* | *knigu*. |
| Dün bütün gün | ya çita-l-a | knig-u |
| Yesterday all day | I read-Past-f | book-ACC |

Yesterday all day I read a book.

Second Recording

(P17)

| | | |
|---|---|---|
| Bizde | *weekend* anlayışı | yok |
| Biz-de | weekend anlayış-ı | yok |
| We-LOC | weekend concept-ACC | none |

'We do not have the concept of weekend'

(P18)

| | | | |
|---|---|---|---|
| *kanikula,* şey, tatilden | | yemeklere | geçtik. |
| Kanikula, şey, tatil-den | | yemek-ler-e | geç-ti-k |
| Vacation well vacation-ABL | | food-pl-DAT | pass-Past-1p |

'We changed the subject from vacation to food.'

(P19)

| | |
|---|---|
| Beli noçi iyun, | *Haziran'da oluyor*. |
| Beli noç-i iyun, | Haziran-da ol-uyor |
| White night-pl June | June-LOC be-PrProg. |

'White Nights are in June.'

(P20)

| | | | | | | |
|---|---|---|---|---|---|---|
| P20a | *işte* dvaretz *taa* | *şeyden* | z daleka | vidna bolşoy | kaskad. |
| /b | İşte dvaretz taa | şey-den | z daleka | vidna bolşoy | kaskad |
| | There palace there | well-ABL | from distance | view big | fountain |

'From a distance you could see the palace, the big fountain.'

(P21)

| | | | |
|---|---|---|---|
| yeterince arkadaşı, *dastatıçnıy druzey* | | | vardı. |
| Yeterince arkadaş-ı, dastatıçnıy | | druz-ey | var-dı |
| Enough friend-pl enough | | friend-pl | have-Past |

'She had enough friends.'

(P22)

| | | | |
|---|---|---|---|
| kafama | göre, | *sbaşki*, | zor olur |
| kafa-m-a | göre | s-baş-ki | zor ol-ur |

head-GEN-DAT according to with-head-Col hard be-AOR

‘It would be hard to do it as I want.’

(P23)

| | |
|---|---|
| Dubai yedut | dlya *alışveriş* |
| Dubai yedut | dlya *alışveriş* |
| Dubai go (by car)-3p | for shopping |

'They go Dubai for shopping.'

(P24)

| | | | |
|---|---|---|---|
| Rusça’da | benim | bildiğim, | *subject*i |
| Rusça-da | ben-im | bil-diğ-im, | subject-i |
| Russian-LOC | I-GEN | know-FNom.-1s | subject-ACC |

söylemeyebilirsiniz, özneyi.
söyle-me-yebil-ir-siniz, özneyi
say-Neg-ABIL-AOR-2p subject-ACC

‘To my knowledge, in Russian you do not have to say the subject.’

(P25)

P25a /b/c

| | |
|---|---|
| *Sapojnik*, | şey yani, *shoemaker* |
| Sapojnik, | şey yani, shoemaker |
| Shoemaker | well I mean shoemaker |

‘I mean shoemaker.’

Written Data

W1

| | |
|---|---|
| *The thing they wanted to infer* | Coca Cola aldığımızda |
| the thing they want-ed to infer | Coca Cola al-dığ-ımız-da |
| the thing they want-Past infer-INF | CocaCola buy-ObjP.-1pl-When |

| | |
|---|---|
| her yerde | ve her zaman şampiyon oluruz |
| her yer-de | ve her zaman şampiyon ol-ur-uz |
| every place-LOC | and every time champion be-AOR-1pl |

‘The thing they wanted us to infer is that when we buy Coca Cola we become the champions every where and every time’

## APPENDIX C

## MAIN SPOKEN DATA

First Recording

(1)

| | | | |
|---|---|---|---|
| Karoçe, | *ilahiyatdan* | sadnoy | paznakomilsiya |
| Karoçe, | ilahiyat-dan | s-adn-oy | pa-znakom-il-siya |
| In short, | theology-ABL | with-one-f/3s/INS | Perf-meet-Past/3s/m-Ref |

'I have met someone from faculty of theology'

| | | |
|---|---|---|
| Tı pomniş, | tı je gavaril: | "Vo prikolnaya", |
| Tı pomniş, | tı je gavar-il: | "Vo pri-kol-n-aya", |
| You remember-2s/Pres | you col say-s/PProg/m | That Perf-cool-f/3s/NOM, |

Tipa takova, da.
Tipa tak-ova, da.
like that-GEN, yes.

'Do you remember, you were saying: "that cool", something like that.'

(2)

| | | | | |
|---|---|---|---|---|
| Pomniş, | mıy | *Kızılaydan* | ili | je? |
| Pomn-iş, | mıy | Kızılay-dan | il-i | je? |
| Remember-2s, we | | Kizilay-ABL | walk-pl/PProg | Col? |

'Do you remember we were walking through Kizilay?'

(3)

| | | |
|---|---|---|
| U niyo *nick* Gizka | bıla, | karoçe, Nargizka. |
| U niyo nick Gizka | bı-l-a, | karoçe, Nargizka. |
| Her nick Gizka | be-Past-f/3s, | short/Col NarGizka. |

'Her nickname was Gizka, short of Nargizka.'

3a

Saifa gavarit tipa svai *mail address* dai.
Saifa gavar-it tipa svai mail address da-i.
Saifa say-3s/Pres Col own mail address give-Imp.
'Saifa says: "give me your own mail address".'

3b

Ana gavarit, vot y Mirzı spraci *mail address* u padugi yest.
Ana gavar-it, vot y Mirz-ı sprac-i mail address u padug-i yest.
She say-Pres/3s that Mirza-Gen ask-Imp/s, mail address girlfriend-GEN be/Pres

3c 'She tells him to ask Mirza for the mail address of her girlfriend.'

(4)
Dialog:
A vıy gde sideli to?
A vıy gde sid-eli to
You where sit-PProg/pl col?
'And where were you sitting?'

Mıy v etam v *endüstri*
Mıy v etam v endüstri
We in this-m/PREP, in Industry
'We were sitting in this, in Industry'

4a

Mog dagatatsiya i zabejat tuda
Mog-[] da-gad-at-siya i za-bej-at tuda
Could-m/sg/Past guess-INF-Ref and in-run-INF to there.
'Could guess and run in there.'

Znayiş *harç* skolka budit?
Zna-yiş *harç* skolka bud-it?
Know-2s/Pres fee how much be-Fut/3s?
'Do you know how much will the fee be?'

4b

Skolka?
Skolka?
How much?
'How much?'

182 milyona
182 milyon-a
182 million-pl ....
'182 millions ....'

Znayiş paçemy 182 milyona budit?
Zna-yiş paçemy 182 milyon-a bud-it?
Know-2s/Pres why 182 million-GEN be-3s/Fut?
'Do you know why it will be 182 million?'

Paçemu?
Paçemu?
Why?
'Why?'

3 liyama s etava, s tavo *dönem*'a.
3 liyam-a s eta-va, s ta-vo dönem-a
3 million-pl from this-m/3s/GEN, from that-m/3s/GEN term-m/3s/GEN
'3 millions from this, that term.'

4c
3 liyama?
3 liyam-a?
3 million-pl?
'Are you sure it is 3 millions?'

3 liyama na etat *dönem*,
3 liyam-a na et-at dönem,
3 million-pl on this-m/3s/ACC term-ACC,

4d *faiz,* karoçe budit.
faiz, karoçe bud-it.
interest, in short, be-3s/Fut

4e 'In short, there will be a 3 millions interest on this term (fee).'

(5)

Nu karoçe, ya c etoi c Nargizoi,
Nu karoçe, ya c et-oi c Nargiz-oi,
'In short, I with that-f/3s/Ins with Nargiza-f/3s/INS, '

*by chat*, *chat*uem, ana i gavarit:
by chat, chat-u-em, ana i gavar-it:
by chat, chat-Prog-1pl/Pres, she and say-3s/Pres

5a/b "Tıy minya ni pugai"
"Tıy minya ni pug-ai"
"You me not scare-Imp".

'In short, I was chatting with that, with Nargiza by chat, and she said: "Don't scare me".'

(6)

V etam *dönemde* u tibya kagda ekzamenıy
V et-am dönem-de u tibya kagda ekzamen-ıy
In this-m/3s/PREP term-LOC Poss [2s/Gen pronoun] when exam-pl/NOM

zakançivayutsiya?
Zakanç-iva-yut-siya?
заканч-ива-ют-ся?
Finish-PROG-pl/Pres-Ref?

'When do your exams finish this term?'

(7)
(Questionnaire)

A Ladna, *Pepsi Generation Next*, gavari dalşe.
Ladna, Pepsi Generation Next, gavar-i dalşe.
Okey, Pepsi, Generation Next, say-Imp/s further.
'Okey, Pepsi, Generation Next, ask more.'

7a

B Şto tibya vo mnye ni nravitsiya?
Şto tibya vo mnye ni nrav-it-siya?
What you/DAT in me not like-Pres-Ref?
'What is it that you don't like with me?'

A Nu ya yişo ni videl, *maybe later.*
Nu ya yişo ni vid-el, maybe later.
But I still yet not see-Past/1s/m, maybe later.
'I haven't seen anything yet, maybe later.'

7b

B Verite li vıi b lubov, *if yes, give example*.
Verit-e li vıi b lubov, if yes, give example.
Believe-2pl whether you in love/ACC, if yes, give example.
'Do you believe in love, if yes, give an example.'

7c

A *Romeo and Julieta.*
'Romeo and Juliet.'

B Nyet, eta banalnoye slişkom. *'Farhat ve Şirin'*, karoçe.
Nyet, eta banaln-oye slişkom. 'Farhat ve Şirin', karoçe.
No, this banal-n/3s/NOM too. 'Farhat ve Şirin', shortly/col.
'No, this is too banal. "Ferhat and Şirin".'

7d

A *Şirin*, şto takoye?
Şirin, şto tak-oye?
Şirin, what that-n/s?
'What is that, Şirin?'

7e

B Nyet, *'Farhat ve Şirin'* u uzbekof eta ...
Nyet, 'Farhat ve Şirin' u uzbek-of eta ...
No, Ferhat and Şirin Poss Ozbek-ACC that ...

A "Tahir ve Zehra"?

'"Tahir and Zehra"?'

B Nu, tipa "Tahir ve Zehra"...
Nu, tipa "Tahir ve Zehra"
Col. like "Tahir ve Zehra"
'Yes, like "Tahir and Zehra". '

7f

...

B Dalşe davai. *Huylanma* *oğlum*.
Dalşe davai. Huy-lan-ma oğl-um.
Further let. Irritate-Imp-2s-Neg son-GEN/1s.
'Let's proceed. Don't fret, man. '

7g

Yest li u vas sikretiy?
Yest li u vas sikret-iy?
Be/Inf whether you/... secret-pl/NOM?
'Have you got any secret?'

A Da, no, *top secret*.
Da, no, top secret.
Yes, but top secret.
'Yes, but top secret.'

7h

(8)

Da, eta normalna, *olan birşeydir* , *normaldir*.
Da, eta normalna, ol-an bir-şey-dir , normal-dir.
Yes this normal, be-SbjP one-thing-Aor, normal-Aor
'Yes it is normal, it happens, it is normal.'

8a/b

(9)

Çay pastav, çay haçu.
Çay pa-stav, çay haç-u.
Tea Perf-put/Imp, tea want-1s/Pres
'Put some tea, I want some tea.'

Çay *Iğdırckii,* şto li?
Çay Iğdır-ckii, şto li?
Tea Iğdır-GEN/m/s what whether?
'The tea is from Igdir, isn't it?'

(10)

Prikol tselava barana varit *kazanda*.
Prikol tsel-ava baran-a var-it kazan-da.
Fun whole-m/3s/ACC/ ram-m/3s/ACC cook-INF bowl-LOC.
'It's fun to cook whole ram in a bowl.'

(11)

*Fener* viyigrali, znayiş, da?
*Fener* viy-igra-li, zna-yiş, da?
Fener win-3pl/Past, know-2s/Pres yes?
'Do you know they defeated Fener?'

(12)

Yest je film "*Çarlinin Melekleri*".
Yest je film "Çarli-nin Melek-ler-i".
-Be/INF col film/NOM "Charlie-GEN Angel-pl-ACC"
-There is a film called "Charlie's Angels".'

Second Recording

(13)

A Kto snimaetciya?
Kto snim-a-et-ciya?
Who/NOM play-Prog-3s/Pres-Ref?
'Who is playing?'

B Denzel Washington.

A Kim o?
Kim o?
Who him?
'Who is he?'

B *Negr, krutoi negr* rus
Negr, krutoi negr rus
Nigger/m/3s/NOM, cool-m/NOM nigger/NOM russian

mafyalarıyla savaşıyor.
Mafya-lar-ı-yla savaş-ıyor.
Mafia-pl-ComConj fight-PrProg.

'Nigger, cool nigger, fighting with the Russian mafia.'

(14)

A *Antalya'da kışın* on şto zabıl?
Antalya-da kış-ın on şto zabıl?
Antalya-LOC winter-SbjP he what/ACC forget/m/3s/Past?
'What did he forget in Antalia during winter?'

14a

B Ni znayu, *got no information.*
Ni zna-yu, got no information.
Not know-1s/Pres, got no information.
'I don't know, I got no information.'

14b

| | | | |
|---|---|---|---|
| | Şto, şto | on | tam pateryal, |
| | Şto, şto | on | tam pa-ter-yal, |
| | What, what | he/3s/m/NOM | there Perf-lose-Past/m/3s, |

*ne kaybetmiş*?
ne kaybet-miş?
what lose-RepPast

14c 'What he did he lose there?'

(15)

| | | | |
|---|---|---|---|
| Ve *fsyeh* | *tvaih* | dedin. | |
| Ve fsye-h | tva-ih | de-di-n. | |
| And all-Gen | your-Gen | say-Past-2s. | |

'And you said ... all yours.'

15a

| | | | | |
|---|---|---|---|---|
| A bolşe | nikavo | tam ni | bıla | *şansına senin*? |
| A bolşe | ni-kavo | tam ni | bı-l-a | şans-ın-a sen-in? |
| But more | nobody-Gen | there not | be/3s/Past | luck-Gen/2s-Dat you-Gen/2s? |

15b 'And to your luck wasn't there anybody?'

(16)

*For example* desin.
For example de-sin.
For example say-Imp/2s.
'Let him give an example.'

16a

| | | |
|---|---|---|
| *For example*, bak: | | "Kimler bizi kurtarır" *şto li*? |
| *For example*, bak: | | "Kim-ler biz-i kurtar-ır" şto li? |
| For example, look-Imp: | | "Who-pl/Nom us-Acc save-Aor" what whether. |

'For example, look: isn't it "Who will save us"?'

16b/c/d

(17)

| | | | | |
|---|---|---|---|---|
| A | O şeye | gitti, | Konya'ya ... | |
| | O şey-e | git-ti, | Konya-ya ... | |
| | He thing-Dat | go-Past/3s, | Konya-Dat... | |

'He went there, to Konya...'

| | | | |
|---|---|---|---|
| Orada tek | başına | karda | *plavayet.* |
| Orada tek | baş-ın-a | kar-da | plav-a-yet. |
| There alone | head-2s/Gen-Dat | snow-Loc | swim-Prog-Pres/m/3s. |

'And there alone he swims in snow.'

B And then what?

(18)

A Domaşka var mı?
Domaşka var mı?
Homework be/Aor Q?
'Is there homework?'

B Şto za domaşka?
Şto za domaşka?
What kind homework?
'What kind of homework?'

A *Da* araştırma, ben araştırma görevlisiyim ya.
Da araştırma, ben araştırma görevlisi-yim ya.
Yes research/Nom, I research assistant/Acc-1s col
'A research, you know, I am a research assistant.'

18a B Büyük insan, *krutoi çelavek*.
Büyük insan, krut-oi çelavek.
Big man, cool-m/s/Nom man.
'You are a big man, cool guy.'

18b (19)

(from an anecdote)

... kasiyer hala gazete okuyor, 'Ya kardeşim,
...kasiyer hala gazete oku-yor, 'Ya kardeş-im,
...Cashier still newspaper read-PrProg/3s, 'col. brother-Gen/1s,

duymuyor musun, sıraya geçsene'.
duy-mu-yor mu-sun sıra-ya geç-se-n-e'.
hear-neg-PrProg Q-3s, queue-Dat pass-Cond-2s-A'.

*On opyat* oraya buraya bakıyor, *nikavo nyetu*. ....

On opyat ora-ya bura-ya bak-ıyor, nikavo nyetu...
He again there-Dat here-Dat look-PrProg/3s, nobody Neg/be...

19a/b/c 'Cashier still reads a newspaper,"Hey, brother, don't you hear, wait for your turn". He again looks around, there was not anybody else...'

(20)

A Vçera zvanil, yivo nye bıla.
Vçera zvan-il, yivo nye bıl-a.
Yesterday call-1s/Past/m, he/Acc not be/Past-m/3s
'I called yesterday, he wasn't there.'

B Nereye gitmiş peki?
Nereye git-miş peki?

Where/Dat go-RepPast/3s Okey?
'Where has he gone then?'

A Kajıtsya *yurtta* vçera nikavo nye bıla
Kajıt-sya yurt-ta vçera nikavo nye bıl-a
Seem/3s-Ref dormitory-LOC yesterday nobody not be/Past-3s.
'It seems that there was nobody in the dormitory yesterday. '

B *A*, dün bayramdı ya, olabilir ondan.
20a A, dün bayram-dı ya, ola-bil-ir o-ndan.
But, yesterday bayram-Past Col, be-Abil-Aor that-ABL.
'Yesterday was a holiday, might be because of that.'

(21)

20b
A Organizatır - kapitan komandıy Davut.
Organizatır - kapitan komand-ıy Davut.
Organiser - captain/m/Nom team-Gen/f Davut.
'Organiser - captain of the team, Davut.'

B Karoçe, ... *Arkadaşlar ben yokum*, bez minya budite igrat.
Karoçe, ... Arkadaş-lar ben yok-um, bez minya bud-ite igr-at.
Shortly,... Friend-pl i be/Neg/1s, without me be/Fut-3pl play-Inf.
'Shortly,... My friends I will be absent, so you will play without me.'

21a A Balşoi futbol ni dlya minya, *de*,
Balşoi futbol ni dlya minya, *de*,
Big/Nom/m football/Nom/m not for me, say/Imp,

ya ni budu igrat.
21b ya ni bud-u igr-at.
I not be/Fut-1s play-Inf.

'Say that big football is not for you and you won't play.'

(22)

perşembeden önce *no chance*, şansımız yok.
perşembe-den önce no chance, şans-ımız yok.
Thursday-ABL before no chance, chance-1pl/Gen none.
'We have no chance before Thursday.'

(23)

A *Davai* Atatürk'ü çaldırayım.
Dav-ai Atatürk'ü çal-dır-a-yım.
Give-2s/PrProg Ataturk-Acc call-Imp-Madv-1s.
'Let me call (by phone) Ataturk (dormitory).'

B S kem budiş razgavarivat?

S kem bud-iş raz-gavar-iva-t?
with who-sg/Ins be-2s/Fut ref-speak-Prog-Inf?
'With whom will you speak?'

(24)

A Kagda u vas *finaller* naçinayutsiya?
Kagda u vas final-ler naç-ina-yut-siya?
When Pos 2pl/Pos pronoun final-pl begin-Prog-Pres/3pl-Ref?
'When do your finals begin?'

B Finalıy u nac naçinayutsiya
Final-ıy u nac naç-ina-yut-siya
Final-pl/Nom Pos. 1pl/Pos pronoun begin-Prog-Pres/3pl-Ref

7-ova yanvarya.
7-ova yanvar-ya.
7-Seq January-GEN

'Our finals begin on 7th of January.'