

**LINGUISTIC EFFECTS ON THE CONTACT
BETWEEN GREEK AND BULGARIAN LANGUAGES
FOR RECENT BULGARIAN IMMIGRANTS
TO NORTHERN GREECE**

by

Emanuela Yancheva Mileva
B.A., Modern Greek Philology, Sofia University, 1998
M.A., Bulgarian Philology, Sofia University, 2000

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

In the
Department of Linguistics

© Emanuela Yancheva Mileva 2009

SIMON FRASER UNIVERSITY

Fall 2009

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without permission of the author.

APPROVAL

Name: Emanuela Yancheva Mileva
Degree: Master of Arts
Title of Thesis: Linguistic effects of the contact between Greek and Bulgarian languages for recent Bulgarian immigrants to Northern Greece

Examining Committee:

Chair: Dr. Nancy Hedberg
Associate Professor, Department of Linguistics

Dr. Panayiotis Pappas
Senior Supervisor
Assistant Professor, Department of Linguistics

Dr. Juan Sosa
Supervisor
Associate Professor, Department of Linguistics

Dr. Steve Marshall
External Examiner
Assistant Professor, Faculty of Education

Date Defended/Approved:

November 24, 2009



SIMON FRASER UNIVERSITY
LIBRARY

Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website <www.lib.sfu.ca> at: <<http://ir.lib.sfu.ca/handle/1892/112>>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada



SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

STATEMENT OF ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Bennett Library
Simon Fraser University
Burnaby, BC, Canada

ABSTRACT

This study presents the results of an analysis of code-switching behavior observed in bilingual Bulgarian-Greek conversations. On the basis of eight hours of recorded conversations with eight participants, a corpus of 234 instances of intra-sentential code-switching were extracted with the complementizer phrase being a unit of analysis. All these instances are analyzed structurally in accordance with the MLF. Code-switching patterns are also discussed in relation to the shared features within the Balkan Sprachbund to which both Bulgarian and Greek belong. The structural analysis focuses on instances of lexical insertions and more specifically, pays attention to single-switch morpheme such as the definite article. It is found that variation between the Bulgarian and Greek definite article combined with Greek noun is probably determined by the language of the preceding lexical item.

Keywords: code-switching, Bulgarian, Greek, definite article, lexical insertions, Balkan Sprachbund.

ACKNOWLEDGEMENTS

This study was supported by a Joseph Armand Bombardier Canada Graduate Scholarship – Masters from Social Sciences and Humanities Research Council of Canada (Award No. 766-2008-1084).

I would like to thank my senior supervisor Panayiotis Pappas for his constant support in my academic endeavours, as well as Dr. Marshall and Dr. Sosa for their helpful comments.

I would also like to thank all Bulgarian participants in this study living in Greece. They know who they are.

TABLE OF CONTENTS

Approval	ii
Abstract	iii
Acknowledgements	iv
Table of Contents	v
List of Figures	vii
List of Tables	viii
List of Abbreviations	ix
Chapter 1: Introduction	1
Chapter 2: Background	3
2.1. The Balkan Sprachbund	3
2.1.1. Simplified nominal case system	5
2.1.2. Formation of a future tense using reduced form of the verb ‘want’.....	6
2.1.3. Loss of the infinitive.....	6
2.1.4. Evidential	7
2.1.5. Analytic adjectival comparative structures	7
2.1.6. Clitic doubling	8
2.1.7. Postposed definite article.....	8
2.2. Code-switching	11
2.2.1. Code-switching as a term.....	11
2.2.2. Approaches	13
2.2.3. Equivalence-based constraints.....	14
2.2.4. Government-based approaches.....	14
2.2.5. Production-based models	15
2.2.6. Motivations for Code-switching.....	20
2.2.7. Code-switching vs lexical borrowing	21
Chapter 3: Methodology	23
3.1. Participants	23
3.2. Procedure.....	28
3.3. Coding.....	32

Chapter 4: Data description and analysis.....	35
4.1. Lexical Insertions.....	37
4.1.1. Nouns	37
4.1.1.A. Greek nouns with Bulgarian markers.....	38
4.1.1.B. Greek nouns without Bulgarian markers.....	41
4.1.2. Adjectives	43
4.1.3. Other lexical insertions	44
4.1.4. Verbs.....	44
4.2. Bilingual CPs containing Embedded language phrasal elements	48
4.3. Exceptional Tokens.....	50
4.4. Code-switching and the definite article	52
4.4.1. Results and Analysis	54
Chapter 5: Conclusions.....	63
Appendix.....	68
References.....	69

LIST OF FIGURES

Figure 1. Map of distribution of preposed and postposed definite article in the Balkans.....	9
Figure 2. System morphemes.....	17
Figure 3. Language production processor.....	19
Figure 4. Map of Greece.....	25
Figure 5. Map of definiteness.....	58
Figure 6. Sequence of system and content morphemes.....	65

LIST OF TABLES

Table 1. Transformation of demonstrative pronouns into definite articles.....	10
Table 2. Representation of the most significant morphosyntactic Balkan Sprachbund features across languages.....	11
Table 3. Participants information.	27
Table 4. Distribution of the code-switched elements in the data	35
Table 5. General distribution of the code-switched and non-code-switched CPs.....	36
Table 6. Greek lexical items embedded in Bulgarian frames.....	36
Table 7. Distribution of marked and non-marked Greek inserted nouns.....	37
Table 8. Mixed verb constituents.....	45
Table 9. Coding schema for the variants.....	53
Table 10. Distribution of variants across factor groups.	55

LIST OF ABBREVIATIONS

1, 2, 3	1 st , 2 nd , 3 rd person
ACC	accusative
AOR	aurist
BG	Bulgarian
CL	clitic
COMPR	comparative adjective marker
DAT	dative
DEF ART	definite article
EXCL	exclamation
F	feminine
FUT	future particle
FUT NEG	future negative particle
FUT PAST	future in the past particle
G	genitive
GR	Greek
IMP	imperative
IMPRF	imperfect tense
INT PRT	interrogative particle
M	masculine
N	neutral
NEG	negative
NOM	nominative
PL	plural
PP	past participle
PPAA	past participle aorist aspect
PREP	preposition
PRES	present tense
PRES PERFVE	present perfective aspect
PRT	particle
REFL PR	reflexive pronoun
SG	singular
SUB PRT	subordinating particle
SUPRL	superlative adjective marker
VOC	vocative

CHAPTER 1: INTRODUCTION

The majority of Bulgarian and Greek speakers live in the two respective neighboring countries of Bulgaria and Greece. These two language communities belong to a linguistic union – the Balkan Sprachbund, sharing features among themselves as well as with other languages in the area. They are result of a linguistic convergence – an outcome from prolonged and intense contact between the speech communities. The increased contact gives rise to bilingualism where speakers inevitably code-switch. As this code-switching interaction intensifies structural features can pass from one language to another.

Code-switching has been an object of interest for various sub-disciplines within sociology and linguistics. Social factors like age, gender, education, can have influence on the code-switching and can provide significant insights for the better understanding of code-switching behavior. On the other hand, there are also many insights to be gained concerning the structure of language by observing the interaction of the two grammars. For the purpose of this study, I focus on the structural aspects predominantly.

The current study investigates the code-switching patterns among Bulgarian-Greek speakers in order to determine whether existing Balkan Sprachbund features affect code-switching performance. To aid this investigation, I present a description of the structural aspects of code-switching, with a focus on lexical insertions and embedded ‘islands’, and the variation of single-morpheme switches. I also present exceptional tokens that are different from the rest of the data. From a Sprachbund perspective, Bulgarian and Greek differ in two main morphosyntactic features: 1) the positioning of the definite article (preposed in Greek and postposed in Bulgarian), and 2) the presence and wide usage of *evidential* verbal morphology in Bulgarian, compared to its absence in Greek. The analysis of code-switching

patterns regarding these two features reveals some interesting constraints: the choice between a Bulgarian or Greek definite article to accompany a Greek noun was found to be influenced primarily by the type of preceding lexical item. The absence of evidential marking in Greek has influence on the code-switching patterns – Greek aorist stems are supplemented with Bulgarian aspectual morphemes and form mixed constituents in a interesting way.

The data were collected through recorded interviews of naturally-occurring conversations among Bulgarian-Greek bilinguals living in Greece. The participants were eight students, four from each gender, and between the ages of 25 and 30. The data is investigated through the Matrix Language Framework model (MLF) by Myers-Scotton (2002). The current data generally supports the model in a number of ways, especially in terms of system and content morphemes. In the same time it provides examples, which are at odds with the framework. It should be mention that this language pair has yet to be examined in the code-switching literature.

The thesis is structured as follows. Chapter 2 provides the necessary background on the Balkan Sprachbund, describing the morphosyntactic features that are shared among participating languages in the union. I then turn to the code-switching literature, focusing on the MLF and to the debate between code-switching versus borrowing for lexical insertions. Chapter 3 goes over the methodology of the study, and describes the participants, the procedure, the materials, and the coding system. The data are analyzed in detail in Chapter 4, with one section on lexical insertions, another on ‘islands’, and a third on the definite article and definiteness. Exceptional tokens are also discussed here. Chapter 5 outlines the major findings, describes their relation to key issues in the code-switching literature and suggests some directions for future research.

CHAPTER 2: BACKGROUND

2.1. THE BALKAN SPRACHBUND

The term *Sprachbund* (Trubetzkoy 1928), or *unité linguistique* (Sandfeld 1930), refers to a situation where linguistic communities, which are not necessarily genetically related, are in geographical proximity and share a number of linguistic features due to contact. The Balkan languages included in the sprachbund are all Indo-European (excluding Balkan Turkic), but they belong to different language branches: Albanian, Modern Greek, the South-Slavic languages Bulgarian, Macedonian and Serbo-Croatian, the Romance languages Daco-Romanian, Aromanian (Vlah) and Judeo-Spanish (Ladino), the Indic language Romani (the language spoken by the Roma on the Balkans), and from the Altaic family (Turkish). (Joseph 2003). Some authors consider as core participants only Albanian, Modern Greek, the South Slavic languages and Romanian, and all the rest as marginally related (Winford 2003, Asenova 2002). It should also be noted, that “Balkan languages” is not equivalent to “languages of the Balkans”. The latter comprises other languages such as Armenian, German, Ukrainian, Yiddish, etc. which are spoken in the Balkan peninsula, but don’t share the structural similarities of the former and thus, are not included in the geolinguistic Balkans.

Linguistic areas emerge as contact between the various speech communities is enhanced by numerous historical events involving periods of conquest, war, colonization, etc. In the case of the Balkan area, these events date back between 800 and 1,700 AD (Winford, 2003). Such dramatic influences aside, language contact and change are enhanced if enough users alter their speech based on peaceful economic forces, such as worker migration to more prosperous neighboring countries. As Joseph (1983b) points out, it is

equally likely that mutual accommodation and shift among the immigrant groups themselves promoted the spread of features.

What makes such linguistic areas complex and intriguing to linguists is the great degree of convergence at several levels of linguistic structure – phonological, morphological, lexical, and syntactic. The intriguing part is that these abstract similarities cannot be explained genetically because the shared features, or arealisms (in this case *balkanisms*), do not represent common inheritances from Proto-Indo-European (Joseph 2003). They are the result of linguistic convergence resulting from prolonged and intense contact between the different speech communities.

Although the features shared by the languages on the Balkans had been observed as early as the 19th century (Kopitar 1829, Schleicher 1852, Miklosich 1861), the first systematic work describing structural parallels is that of Sandfeld (1930). The origin of the Balkan Sprachbund remains controversial, with three main schools of thought. Scholars like Kopitar (1829), Miklosich (1861) and Weigand (1928) defend the idea of *substratum* – the indigenous languages Thracian, Dacian and Illyrian spoken in the Balkans in ancient times having influence on and being responsible for the balkanisms. However, two questions arise immediately – the Balkan Sprachbund originated in the post-Byzantine period and therefore, it would be hard to believe that features from substratum languages which “had been dormant for centuries, become active long after the languages itself had disappeared” (Tomic 2006). Second, the linguistic material if any, is too insufficient to be able to provide reasonable confirmation for whichever Balkan Sprachbund feature. Others argue for the existence of *adstratum* (Sandfeld 1930, Solta 1980), where the influence of Greek or Latin would be the most natural explanation, even sometimes the only possible one. And finally, there are different types of contact models, trying to explain the origins, based on

bilingualism and convergence (Civjan 1965, 1979, Rozencveig 1969, 1976, Lindstedt 2000). As of now there is no complete agreement as to the origin of the Sprachbund. Plenty of facts about the social situation in the medieval period, when the crucial contact between these languages actually occurred, are still unknown. (Joseph 1983b).

The most discussed Balkan features are: stressed schwa, the vowel inventory of [a, o, u, e, i] without phonological contrasts in quantity, openness, or nasalization; shared phraseology, shared loan words; reduction of the nominal case system, merging of dative and genitive cases, formation of future tense using the reduced form of the verb 'want', use of postposed definite article, used of evidential, lost of infinitive, analytical adjectival comparative structures, clitic doubling. My particular concern will be on the morphosyntactic features since they are considered as the most robust with interesting mechanisms of interaction between the languages. They are also relevant in terms of borrowing of morphemes in the past. What follows is a brief description.

2.1.1. Simplified nominal case system

All Balkan languages have undergone reduction of their case systems accompanied by merging of Genitive and dative cases. There is no surface distinction between the form of dative and the form of genitive anymore – the dative just replaced the genitive except in Greek where is the opposite happened. Reduced case systems can be found in Albanian, Romanian, and Greek. In Bulgarian and Macedonian the only traces of case morphology are found in the personal pronouns, and more specifically, in the dative vs accusative forms, in two forms of interrogative pronouns, and a few numerals. As a result the syntactic relations are expressed analytically by prepositions. In this respect the transition from syncretic to the analytic in Bulgarian dates as early as 11th century.

2.1.2. Formation of a future tense using reduced form of the verb 'want'

This feature is found in Greek, Bulgarian, Romanian, Albanian, Serbo-Croatian and Romani. However, in the 'want' + inflected verb construction the choice of the auxiliary verb was not a clear-cut process. There were competing auxiliary verbs like 'have' and 'be' (Latin, Greek, Old Church Slavonic) and it was only in the early modern period when that the situation was settled and 'want' established itself as the predominant auxiliary (Joseph & Pappas, 2002). In Greek the reduced, invariable form of the verb 'want' is represented by *tha* (from *thelo*) and in Bulgarian by *shte* (from *shta*). A grammaticalized vestige of the competing 'have' constructions is the negated future in Bulgarian (also in Macedonian) using a negative possessive: *njama¹ da pisha* ('I will not write').

2.1.3. Loss of the infinitive

Non-finite verbal complementation is replaced gradually by fully finite complement clauses (Joseph 1983a). This was basically done by introducing a subordinating particle, usually of pronominal origin, followed by the inflected verb agreeing with the subject of the main clause. This is widely spread in Greek, Bulgarian, Macedonian, Serbo-Croatian. The replacement began to occur in Greek earlier than in the other Balkan languages which gives some authors the reason to argue that Greek was the source of spreading the feature (Joseph 1983a, Asenova 2002). According to Joseph (1983b) the loss of the infinitive took place in the Medieval period lasting approximately five centuries – from 10th till 14th century. In Modern Greek there is a complete absence of this verb form², while Bulgarian shows traces of it in a limited group of verbs: *moga* ('can'), *smeja* ('dare'), *stiga* ('enough'), *nedej* ('don't'), and a few more. In Greek the particle is *na* or *pos* as in *na grafo*, in Bulgarian is *da* as in *da pisha*

¹ from the verb *imjati* with meaning 'have' + negative particle *ne*

² it can be found only in some established expressions like *apagorevete to kapnizim!* ('smoking is forbidden!')

(‘that I write’). There has been no complete agreement as to how exactly this feature emerged. Whatever the answer, it should take into consideration the language-internal factors, as well as the circumstances of language contact (Joseph 1983b).

2.1.4. Evidential

The use of verb forms representing actions in which the speaker is not a witness, is found in Bulgarian, Macedonian, Albanian, and Turkish. Jacobson (1956) calls this grammatical category *evidentiality*, and in Bulgarian grammars it is described as inferential mood or paradigms describing “unwitnessed events” (BAN 1982-83, Pashov 1990, Bojadžiev et al 1999). According to Friedman (2004) the spread of this feature is strongly influenced by Turkish. Greek, on the other hand, doesn’t have this type of verb morphology. In Bulgarian and Macedonian they are widely used and their existence is in contrast with all the other non-Balkan Slavic languages.

2.1.5. Analytic adjectival comparative structures

The analytic comparative forms of adjectives are available in all Balkan languages. In Bulgarian and the other Slavic languages these are represented analytically by means of the particle *po* as in *po-hubav* (‘nicer’), and in Greek it is *pjo* as in *pjo kalos* (‘nicer’). However, coexisting with these comparative forms, are remnants of synthetic formations in Greek: the comparative form of an adjective can be formed by adding the inflection – *-teros* which is semantically equivalent: *kalos, kaliteros*. With the superlatives the situation is a bit different – the Slavic languages and Turkish use the particle *nai-* as in *nai-hubav* and in Greek these are expressed by the comparative form and a definitive article in front of it: *o kaliteros* or *o pjo kalos*.

2.1.6. Clitic doubling

Direct and indirect objects are cross-referenced by weak (clitic) pronouns. These short pronominal forms agree in gender, number and case with an accusative or dative object. This oft-cited Balkan feature is found in Greek, Bulgarian, Albanian, Romanian, and in Macedonian. Object reduplication, though, is found to be facultative and discourse-conditioned in Greek and Bulgarian, and not so grammaticalized as in Macedonian. It indicates a focused or emphasized element. In Bulgarian an example would be *knigata ia dadoh* ('I gave the book') where the short pronominal form *ia* agrees in gender and number with the object *knigata* ('the book'), and is in accusative case. The same sentence in Greek would be: *to vivlio to edosa* where *to* agrees with the object *to vivlio* ('the book') in gender, number and case, and refers to this object.

2.1.7. Postposed definite article

One of the most discussed topics among Balkan linguists is that of grammaticalized definiteness. Definiteness is a feature of noun phrases distinguishing between entities that are specific and identifiable in a given context and those that are not. One of the ways to express definiteness is to use the definite article. It can be introduced either from a previous discourse context, or accessible through the speech situation through general knowledge, or via a relation that has been established to a separate identifiable referent (a nominal or a relative clause) (Matras 2003).

As for their morphological appearance they can be postposed (in Bulgarian, Macedonian, Albanian, and Romanian), and preposed (in Greek).



Figure 1. Map of distribution of preposed and postposed definite article in the Balkans. Postposed: forward-slanting lines; preposed: backward slanting lines (Perry-Castañeda Library Map Collection, University of Texas; additionally designed).

Historically, the change started with one type native demonstrative pronouns: *ty*³, *ta*, *to* found in Old Bulgarian. These were weak pronouns, i.e. they possess only anaphoric use and had short forms, being encliticized to the pronominals. Through their linking to the immediately preceding noun they started to lose the initial demonstrative meaning and to acquire a meaning and function of a definite article.

³ With *y* in *ty* is represented a shwa sound

Old Bg (Sg)	Modern Bg (Sg)	Meaning
<i>rodь ty</i> → <i>rodoty</i> kin this kin the	<i>rodьt</i>	the kin
<i>peshtera ta</i> → <i>peshterata</i> cave this cave the	<i>peshterata</i>	the cave
<i>miasto to</i> → <i>miastoto</i> place this place the	<i>miastoto</i>	the place

Table 1. Transformation of demonstrative pronouns into definite articles. Examples taken from Codex Suprasliensis, 11th century (Mirchev 1978:200).

Subsequently, the respective definite articles in Modern Bulgarian are *-yt/-iat* for masculine, *-ta* for feminine, and *-to* for neuter singular.

From all the demonstrative pronouns in Old Bulgarian, it was *-ty*, indicating the speaker's generic indication of subjects and objects without specifying their remoteness that became established as the definite article.

In Greek, on the other hand, the definite article is preposed, i.e. placed before the nominals: *o anthropos* ('man') and was based, too, on native material: the pronoun that became an article was mostly demonstrative in the earliest periods of the history of the language (Mirchev 1978).

These are the most noted shared morphosyntactic features in the Balkan Sprachbund. It should be noted that some authors also include *Future in past as conditional* (Friedman 2006, Golab 1964, Tomic 2006), *Perfect in 'have'* (Solta 1980, Campbell 1998, Thomason 2001), the *Formation of numerals* and *Deictic pronouns* (Friedman 2006). The pattern of sharing is summarized in Table 2.

Feature	Gr	Bg	Mac	SC	Ro	Alb
1. Simplified nominal case system	+	+	+	+	+	+
2. Formation of a future tense using reduced form of the verb 'want'	+	+	+	+	+	+
3. Loss of the infinitive	+	+	+	(-)	+	+
4. Evidential	-	+	+	(-)	+	+
5. Analytic adjectival comparative structures	+	+	+	-	+	+
6. Clitic doubling	+	+	+	(+)	+	+
7. Postposed definite article	-	+	+	(-)	+	+

Table 2. Representation of the most significant morphosyntactic Balkan Sprachbund features across languages⁴

As the table shows, Greek and Bulgarian differ only in two respects – Bulgarian has evidential as a grammatical category and postposed definite article and Greek lacks of them.

2.2. CODE-SWITCHING

2.2.1. Code-switching as a term

As Myers-Scotton (2002) notes, code-switching is a common phenomenon in linguistic areas such as the Balkans. Code-switching occurs among multilingual speakers, and broadly refers to the mixing of two or more languages in discourse. Increased contact leads to multilingualism, multilinguals code-switch and as code-switching becomes more intense

⁴ the brackets indicate that the feature is partially represented, i.e. in some dialects

structural features can pass from one language to another. Vogt (1954) suggested that language contact, including language alternation, is an important element of language change. Toribio (2004) concludes that the simultaneous presence of languages in code-switching further favors the search for parallels between them and enhances convergence. For Gardner-Chloros (2004) code-switching is a major vehicle for convergence and language change.

Code-switching is a frequently observed phenomenon in the speech of bilingual communities around the world (Backus 2003, Auer 1995, Milroy & Muysken 1995). Code-switching allows for alternating between complete utterances from the two languages, or between structures related to the sentence or clause, and insertion of lexical items from one language to another. These mixtures are considered a result of diverse socio-historical forces (Winford 2003). CS is sometimes inevitable when speakers master both languages. Consider the communities that live around boundary regions, where close interaction is encouraged by a variety of commercial, cultural and political reasons. The geographically determined contact between Greeks and Bulgarians is an excellent case in point.

A broad spectrum of code-switching definitions can be found in the literature from the 19th century to date. They range from perspectives of anthropology to sociolinguistics, and in pure structural linguistic terms there are points of conflict between different models. Auer (1995) provides a good classification of the patterns of code-switching, which according to him are four. Pattern I demonstrates switches from code A to code B or switches within a single speaker's turn. This type of switching often indicates a topic shift. Pattern II involves a negotiation of a language of interaction. Pattern III shows switching between languages in a single turn without being possible to identify a base language. And

finally, pattern IV code-switching behaviour, where momentary switches occur without changing the language of interaction.

This intra-clause code-switching creates utterances with lexicon and morphosyntactic structure from one language and insertion of a single word or phrases from the other. As a result three kinds of constituents can be produced: mixed constituents consist of materials from both languages; embedded language islands are phrases incorporated from the Embedded language; and Matrix language islands coming entirely from the matrix language.

2.2.2. Approaches

The recent developments in the field of intra-sentential CS are grounded in research on different language pairs. Early studies were conducted mostly on Spanish and English (Lance 1975, Pfaf 1979, Poplack 1980, Klavans 1983). Later, other CS research includes Turkish-Dutch (Backus 2003), French-Dutch (Treffers-Daller 1991), Greek-English (Gardner-Chloros 1995), as well as pairs including more typologically different languages: English-Maori (Elliasson 1989), Dutch-Moroccan Arabic (Nortier 1990), Arabic-French (Sankoff, Poplack & Vanniarajan 1990), Swahili-English (Myers-Scotton 1993), Spanish-Hebrew (Berk-Seligson 1986) among others.

Many initial attempts at understanding code-switching were more concerned with the sociolinguistic aspects of the phenomenon, focusing on factors such as age, gender, education, and attitude. Syntactic and psycholinguistic studies on CS appeared later on, examining grammatical constraints and parallelism to speech production. All of the frameworks in this area seek to establish the universal principles that determine the nature of intra-sentential code-switching patterns. Admittedly, however, there exist counterexamples to all of the principles and constraints formulated so far. Recent research on intra-sentential

CS investigates the ways that the integration of the EL units is accomplished morphosyntactically. Investigative approaches fall into one of several major streams of thought.

2.2.3. Equivalence-based constraints

Pfaff (1979) first suggested that surface structures common to both languages are favored for switches. The Equivalence constraint introduced by Poplack (1980) aims to predict where switches are most likely to occur. If two languages have similar (that is, equivalent) syntactic structures, a switch can take place since the change of language does not violate the structure of either language. Muiyeken (1997) refers to this type of CS as a “congruent lexicalization”. In other words, speakers with two languages that have similar grammatical structures can provide lexical elements from either language. For Poplack, CS can happen only between those sentence elements that are normally ordered in the same way by the monolingual grammars.

2.2.4. Government-based approaches

These are generative grammar approaches mainly interested in the relationship between the head of a construction and its complement. DiSciullo et al. (1986) proposed that “when a government relation holds between elements, there can be no mixing; when that relation is absent, mixing is possible” (ibid.4). For example the complement of a preposition in a PP should be drawn from the same lexicon as that of the preposition (p.9). The theory predicts that in the sentence like: ‘I went to Rome’ the preposition ‘to’ and ‘Rome’ will be drawn from the same language in a code-switched data. They form a prepositional phrase where the preposition ‘to’ is a head and respectively governs its complement “Rome”.

2.2.5. Production-based models

Joshi (1985) expresses the idea that certain functional items come from the matrix language, which leads to certain constraints on code-switching. He also notes that the switching is asymmetric and one-directional – from the matrix to the embedded language. His Closed Class Items Constraint posits that determiners, quantifiers, prepositions, possessives, auxiliaries, tense helping verbs, etc. can not be switched; only major class items as nouns, verbs etc. can. In this respect his view is in line with the Functional Head Constraint Model of Belazi et al (1994). Both constraints would predict that switching between functional elements, closed items in Joshi's description, like COMP, INFL, DET and their complements, is not possible. However, counterexamples for this statement were found in Joshi's Marathi/English data (Joshi 1985) where a Marathi postposition was following an English noun.

Following the Joshi (1985) Closed class Items Constraint, Myers-Scotton (1993) suggests a model for looking at the CS data based on a language production process. Her main rationale is based on previous psycholinguistics studies (Grosjean 1988) providing empirical evidence that one language is dominant in bilingual processing. Code-switching, Myers-Scotton states, can be properly examined only if we look first on how "language is accessed and retrieved before it takes its final form" on the sentence surface. (1993:45).

The underlying idea of the MLF is that in code-switching interaction, one of the languages has a dominant role defined as a Matrix Language, while the other is less active – the Embedded Language (EL). In terms of defining the matrix language, there is no agreement among the scholars in this area. Klavans (1983) as well as Treffers-Daller (1991) suggest that ML is determined by the language of the main verb. Yet, many counterexamples indicate this claim to be inaccurate. The best way to define ML according to

Nortier (1990) is to cover a set of criteria, including the language of the first words of the utterance, the syntax of the sentence, and the frequency of constituency of each language. Myers-Scotton (Myers-Scotton 1993:68) suggests this is “the language of more morphemes in interaction types including intra-sentential code switching”. She also suggests that frequency counts must be based not only on single sentences, but on a discourse sample.

Two major principles are operating in the code-switching processes, the morpheme order principle and the system morpheme principle. Both of these come from the same language which should be determined as the ML. In other words, the ML gives the morphosyntactic frame of the code-switched sentences by providing the surface morpheme order and productive system morphemes (productive inflections and function words).

Another cornerstone in the MLF model is the distinction between system and content morphemes. Psycholinguistically, this demarcation is based on their different behavior in the mental lexicon of the speakers. They are activated on different levels of a language production process. The criteria for distinguishing them are [+/- Quantification], [+/- Thematic Role-Assigner] and [+/- Thematic Role-Receiver]. A system morpheme is “any lexical item or affix that is a member of a syntactic category which involves quantification across variables and thus shows [+Quantification] property,” while content morphemes are “any categories which show the [-Quantification] property but have either the [+Thematic Role-Assigner] or [+Thematic Role-Receiver] property” (Myers-Scotton 1993:100).

Stemming from this classification quantifiers, possessives, determiners, dummy pronouns, tense/aspect, complementizers, agreement markers, copula, ‘do’ verb, possessive ‘of’, degree adverbs are system morphemes. The occurrence of morphemes in ML + EL

constituents is determined by their status. Under the MLF model all the system morphemes should come from the ML but content morphemes can be from both participating languages.

Myers-Scotton (1999) develops a submodel for the MLF – the 4-M Model. The model gives additional sub-classification of the system morphemes, which is:

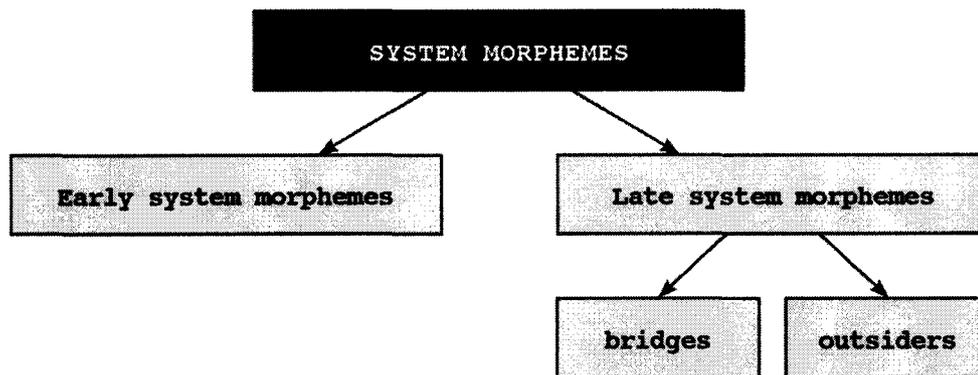


Figure 2. System morphemes (Myers-Scotton 1993)

The reason behind this addition is that the early system morphemes according to Myers-Scotton (2002:75) are not going outside the maximal projection of the content morphemes that select them. That means that “they appear in the same surface-level maximal projections as their heads, and they depends on their heads for information about their forms”.

Another important issue in this framework, and in CS literature in general, is the domain of intra-sentential code-switching. For the purposes of this study, we follow the Myers-Scotton view that the best unit of analysis for examining code-switched data (and in

general for any contact phenomena) is the CP or the maximal projection of complementizer (Myers-Scotton 2002:53).

What favours the CP to be the best choice? Myers-Scotton gives several arguments. First, the CP can be determined clearly in terms of phrase structure as a complementizer or an element in Specifier (Spec) position followed by an IP. Second, it is a category recognized by most syntactic frameworks. Third, since CP can have null elements in Comp (complementizer) position or elsewhere, having CP as a unit of analysis resolves the problems with the status of constituents with nulls.

The model also states two hypotheses, one of which is related to the notion of congruence, a topic greatly discussed in the CS literature. The *Blocking hypothesis* states that a blocking filter impedes any EL content morphemes which are not congruent with the ML with respect to three levels of abstraction regarding subcategorization. (Myers-Scotton 1993:120). That is to say, if EL content morphemes are not in agreement with ML morphemes regarding the thematic-role assignment, such morphemes wouldn't occur in ML + EL constituents.

Another level of congruency is related to categorical congruency between system and content morphemes. If an item is a system morpheme in one language (one syntactic category) but a content morpheme in the other (different syntactic category) then the EL content morpheme cannot occur in ML+EL constituents. Illustrative here would be the following case: if pronominal forms are clitics in ML (therefore, system morpheme) but they are free-form pronouns in EL (content morpheme) then ML clitics occur in ML + EL constituents but not the EL's free-form pronouns. Finally, the way she views the mechanisms resulting in code-switching behavior is given in Figure 3, which represents a language production processor.

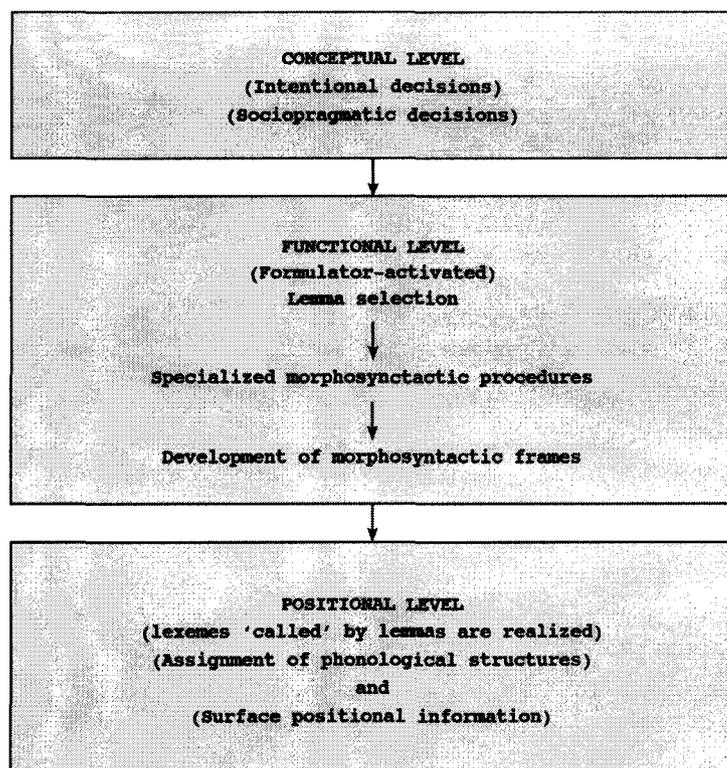


Figure 3. Language production processor (Myers-Scotton 1993)

The crucial mechanisms occur in the Formulator and they are actually described by the two principles – the Morpheme-Order Principle and the System Morpheme Principle. Once the frame is set by the two principles, content morphemes can be “invited” (by either ML or EL) but they both will appear in “slots prepared by ML lemmas”. (Myers-Scotton 1993:118).

The MLF was chosen as the theoretical background to this study for several reasons. Without doubt, this proposal has had significant impact on the code-switching literature, and even critics of Myers-Scotton's hypotheses agree with some of her generalizations. In addition, as this framework allows one to look at lexical insertions as code-switching it is more congruent with the research objectives of this study. Of particular

interest is Myers-Scotton's distinction between system and content morphemes as we will see in the discussion of the use of Bulgarian definite articles with Greek nominal elements.

2.2.6. Motivations for Code-switching

There have been many attempts to give reasons, motivations, and potential answers to the question of why code-switching occurs. Taking the sociolinguistic path we come across studies like those of Blom and Gumperz (1972), Gal (1979) and Gumperz (1982). These researchers are of the opinion that code-switching has intentional and communicative purposes. That the code-switching is a symbol of a speaker's intention is also the idea of Myers-Scotton's Markedness model (1993). Poplack (1978) and Zentella (1982) suggest that among the factors influencing code-switching are the level of language proficiency, social identity and language preference.

Myers-Scotton lists five factors as causes for code-switching behavior. First is the lack of concepts from the ML lexicon. Second is that the inserted item might be a better choice for a certain register, or style of speech. The third one is that an embedded element can narrow down the meaning of an ML element and this way can be more compatible with the speaker's intention. The fourth one is that the embedded item has connotations that are available within ML. The fifth factor is that the EL element may draw the listener's attention.

Backus (2003:4) expresses the idea that the motivation behind the use of large amount of EL lexical units (whether simplex or complex) is semantic specificity. In cases when the right equivalence doesn't exist in the ML it would be cognitively easier for the speaker to insert an EL unit. It might also be psycholinguistically optimal since it "requires less production effort". (Myers-Scotton 1997).

2.2.7. Code-switching vs lexical borrowing

In the CS literature there is still no consensus as to the boundary between the code-switched and borrowed lexical items lies. The term ‘borrowing’ refers to a single lexeme form being phonologically, morphologically and syntactically integrated into the recipient language. (Poplack 1980). The notion of nonce loans is first expressed in Haugen (1950) and later fully developed by Sankoff et al (1986). Poplack and her associates see borrowing as involving different mechanisms than code-switching. Poplack, Sankoff and Miller (1988) identify as different these forms ‘that occur only one in our corpus’ (‘nonce borrowings’) and those used by many speakers (‘widespread loans’). The idea of single-word insertions as fully borrowed items is shared also by Shaffer (1978), Sankoff et al. (1986), Macswan (1999), Muysken (2000). In conclusion, these authors reserve the term code-switching to be exclusively valid for “multiword sentences, which remain lexically, syntactically, and morphologically unadapted to recipient language patterns” (Poplack 1988:97).

On the other side of the spectrum are scholars who consider singly occurring switched forms as instances of code-switching. Among these are Myers-Scotton (1993), Treffers-Daller (1991), Gardner-Chloros (1995), Backus (2003). They all look at single-word switching and borrowing as being governed by similar mechanisms and actually falling into one dynamic continuum, based on the degree of integration and assimilation (Winford 2003). Crucial role here plays the term ‘frequency’ which, according to Myers-Scotton, determines when a word becomes borrowed. She is also of the opinion that code-switching is a major way for borrowing to occur and that single word switches have the status of a borrowed form only through increase frequency of use and adoption by monolingual speakers of the recipient language (1993:182). Backus (2003:124) based on his

Turkish/Dutch data claims also that single-word insertions are a limiting case of the general principle, i.e. a conventional unit in the EL. The present study treats single word insertions as instances of code-switching, because wherever it is possible to determine, switched items do not show phonological integration. For example the pronunciation of the word *katēgōda* ‘storm’ preserves the Greek voiced dental fricative sound [ð], which Bulgarian lacks and it is not pronounced with the alveolar [d] as it would be the case if it was phonologically integrated. The same is with the velar fricative [ɣ] and voiceless dental fricative [θ] in words like *eryasia* ‘paper’, *γλωσολογία* ‘linguistics’, *eθnikotita* ‘nationality’, *eksakolonθitikes* ‘durative’. Their closer equivalent in Bulgarian would be the velar plosive [g] and the alveolar plosive [t]. Everywhere where the Greek words possess these sounds they are preserved in the pronunciation of the participants.

CHAPTER 3: METHODOLOGY

In this chapter, I go over some pertinent information about the participants and the construction of the database.

3.1. PARTICIPANTS

The Balkans are characterized by high frequency of migrations through the millennia of its history. A major crossing area, it has attracted numerous and multidirectional migrations as a result of the historical and political changes in Europe and the Middle East. Among such dramatic events with socio-historical, economic and political consequences are the reunification of Germany, the falling of the Iron Curtain, and the fragmentation of the former Soviet Union in the last century. The latter led to the collapse of all the socialist countries among which was Bulgaria, and became a push-factor for people from those countries (Albania, Bulgaria, Romania, Poland, former Yugoslavia, Georgia, Ukraine) to search for better place of work or living during the first half of 1990s. One of their destinations has been Greece and as a result they form the largest group of immigrants there.

According to the last official Census conducted in 2001 in Greece, the immigrants number 762,191. (Baldwin-Edwards 2004). Albanians represent 57.5% (438,036) of those immigrants and are the largest immigrant group. The second place is occupied by Bulgarians but their numbers are much lower in comparison – 4.6% (35,104), followed by Georgians, Romanians, etc.

The reasons to choose Greece as a point of immigration vary according to the nationality but the main reason is finding a job, often a seasonal one. Greece provides those opportunities in the domain of agriculture, industry and tourist services. The majority of

Bulgarians make the decision to work in Greece for at least 6 months a year and the geographical proximity facilitates this choice in terms of time and money. Not much sociological research has been done on Bulgarian immigrant community in Greece but from what is available (Baldwin-Edwards 2004, Cavounidis 2004, Cholezas & Tsakloglou 2008, Markova 2003) it is clear that working in the abovementioned sectors, combined with higher standard of living in the host country, is the main reason to immigrate. Other reasons are family reunion (4,189), and study (473) (Baldwin-Edwards 2004). From those 473 Bulgarians who have decided to pursue their higher education in Greece, most of them are located in the two major cities in Greece – the capital Athens in the south, and Thessaloniki, the biggest city in Northern Greece. (See Figure 4). It was the Bulgarian student community and more specifically, its representatives in Thessaloniki, that was chosen as a community under investigation in the present study, What makes this group suitable for research is the fact that they apply code-switching on a daily basis at university and home, and some of them are even living with Greeks. The other reasonable factor was that this particular community was more accessible to me.



Figure 4. Map of Greece (Perry-Castañeda Library Map Collection, University of Texas).

As can be seen from the map, Thessaloniki is the nearest biggest Greek city to Bulgaria and the distance is no more than four hours by car. In Greco-Bulgarian relations in the last 50 years there have been numerous bilateral agreements regarding student exchanges. However, the prevalent case has been Greek students going to Bulgaria to pursue their higher education, predominantly in the area of dentistry, surgery or medical education in general. In the last 20 years programmes like Socrates-Erasmus has made the student exchange possible within the European union. It gives opportunity to young people to visit neighbouring and other countries and to pursue university degrees there.

This is also one of the reasons for the Bulgarian participants in the present study to choose Greece as a place of immigration. A similar opportunity for having a higher degree in Greece for Bulgarians is provided by the Greek Ministry of Foreign Affairs through different type of fellowships, based on academic and personal achievements. Six of our participants listed this reason for being in a Greek University. An additional factor was that some of them (those three participants who actually are of Greek origin) had their families in Greece.

As it was mentioned above, the geographical proximity plays a beneficial role in terms of traveling. All of the participants are able to go to their home country at least twice a year. All of them prefer to spend holidays like Christmas and Easter with their relatives and friends in Bulgaria, but except for that, they rarely travel to Bulgaria, excluding short periods of time in mid summer. All of them express positive attitude towards Greece as a country, climate, nature, hospitality. Although Bulgarians and Greeks share more or less the same cultural background found in the Balkans, there are still differences on a micro level, and these were mentioned by the participants in their interviews. All of them have a large circle of friends consisting of both Greeks and Bulgarians, predominantly students, and they have established strong bonds with Greeks over time – classmates, academic contacts, personal relationships. Three of my female participants have Greek partners and one of the males is married to a Greek woman. I was exposed to many gatherings where the predominant language of communication was Bulgarian till the moment when a Greek person came to the party. That made the Bulgarians switch and converse in Greek only in order to accommodate that person. In their free time they hang out with both Greeks and Bulgarians, going to the beaches, taverns or having fun in Greek traditional places like *bouzoukia* (night clubs). Their friendly relationships with Greeks are not limited to leisure only but also working together on school projects and activities. Many of them had invited Greek friends

to visit their Bulgarian homes for a week or two. Most of them share the desire to stay in Greece after graduating, but will look for a job in both Greece and Bulgaria.

The study was conducted in Greece and involved eight adult participants (four males and four females) between the ages of 25 and 30. Their native language is Bulgarian and they have been living in Greece for more than three years, some of them even more. They also possess a high level of bilingualism, and live in a social environment that requires continuous shifts from one language to another. These subjects were recruited by way of a ‘friend of a friend’ approach (Tagliamonte, 2006). To ensure their anonymity, their names have been changed to Subject 1, Subject 2, etc.

Six of them have been learning Greek intensively in their native country, Bulgaria, since age of 18, and one female and one male since their early childhood.

All eight participants knew each other through common friends, six of them doing the same major at the university. Information about the participants is listed below in Table 3.

Subjects	Gender	Age	Occupation	Years in Greece
Subject 1	F	30	PhD student	5
Subject 2	F	30	MA student	8
Subject 3	F	30	MA student	8
Subject 4	F	25	Undergrad student	3
Subject 5	M	30	MA student	12
Subject 6	M	28	Undergrad student	9
Subject 7	M	25	Undergrad student	3
Subject 8	M	26	Undergrad student	3

Table 3. Participants information.

3.2. PROCEDURE

The participants were involved in digitally-recorded (Marantz recorder, model No. PMD 660) one-hour interviews in an informal setting without being aware of the study's ultimate objectives. The interviews took place either at the researcher's hotel or the participants' homes. Data collection, especially in face-to-face community-based fieldwork can be challenging in terms of many factors. I briefly discuss some of the difficulties I faced and tried to overcome.

One of the challenges I faced was that most of the participants were conscious about their language performance in Bulgarian at the beginning of the interviews. They were assuming my data collection had focus on Bulgarian, more precisely, something, related with the grammaticality of their utterances or phonetic aspects. An effort for correctly used grammar and pronunciation was noted at the beginning but as the interview progresses this concern seems to diminish. This self-consciousness is caused by the presence of the interviewer and the recorder itself. When people are aware that their speech or behavior is observed/recorded, they tend to switch their style of speech to a more formal one (The Observer Paradox, Labov 1984). In order to avoid that I tried to make the recorder less conspicuous. I also tried to group the participants in pairs – people who knew each other or were close friends – as that would provide for a more comfortable and casual atmosphere and would allow for more intimacy.

In terms of type of data I needed I faced another challenge. As some of the participants mentioned later, even though they knew I spoke both languages and was able to understand them in both languages, they found it difficult to use both simultaneously. To those of them who met me for the first time, my presence was new to their usual circle of code-switching friends. As Poplack (1980) states, belonging to the same ethnic community is

an important factor for gathering information in sociolinguistic research. The fact that I belong to the same ethnic and age group helped them feel more comfortable to a certain degree. I first introduced myself and, in order to further bridge the gap between us, I talked about my own immigration and education experience, emphasizing the common difficulties one faced in a foreign country. At the same time I could feel that the fact that I normally resided in an English-speaking country accounted for a subtle sense of distance – the participants seemed to be unable to readily “position” me. That is why I tried to code-switch as much as possible before the recorder had been set in order to give them a chance to assess my level of proficiency of both languages as well to give information how and when I started to learn and speak Greek.

Ideally, highly representative data can be obtained if the researcher can afford to devote time, money and effort to record naturally occurring speech for a relatively long period of time. However, the object of this study is to get data that will illuminate structural aspects of code-switching and is not so much focused on the social aspect of code-switching. As Saville-Troike (1997:133) states “The appropriate procedures depend the relationship of the ethnographer and the speech community, the type of data being collected, and the particular situation in which fieldwork is being conducted.” In the remainder of this section I explicate the situation in which my fieldwork took place.

A commonly used method of research is participant-observation. Being a member of the same community I could be easily in the position of observer of my participants. The problem however, is that this method requires a substantial period of time, which I did not have. That is why I used a questionnaire of some warm-up questions, in order to establish a comparable initial context for all subjects; it also helped facilitate dialog with those subjects, whom I had just met for the purpose of this interview. In these cases of less commonality

between researcher and interviewee, there was more of a need for my prompting in order to elicit good speech flow.

The elicitation type of questions I asked were related to their studies in Greece, their way of living and the type of difficulties they had to overcome in order to adapt to their new environment. Other questions inquired whether they considered staying in Greece after graduation, the cultural differences and similarities between the two countries, and how they made friends there.

Apart from these common topics, the talks were shaped according to each interviewee's interests and background. Often, once I got an idea about the participants' interests and backgrounds, I asked more community-specific questions (Tagliamonte 2006), or more tailor-made ones, related to their personal experiences like: "Tell me something more about your research work last summer on the islands" or "What happened when you found out your exam had been marked with unreasonably low grade?"

In conclusion, the questions were drawn from my questionnaire prepared beforehand, and additionally, I had clarifying, spontaneous questions depending on the interaction dynamics of the discourse.

The conversation often took different directions from the ones I wanted to follow but that actually was beneficial because it made the whole experience much more spontaneous, without me interrupting the interviewee to purposefully shift the direction. My role here was more as an observer, than interviewer because I was no longer asking, but just passively participating in the discussions and letting them unfold. Turn-taking between different participants made the conversation lively. The more relevant and "controversial" the topic was to them, the more they conversed and enriched the discourse. Nevertheless, I found they were somewhat cautious and reserved when it came to their political views or

their opinion on some of the university professors. As the different participants come from different parts of Bulgaria representing different vernaculars, I was trying to approximate the vernacular of the respective interviewee (in terms of phonology only). This way I sounded more familiar and informal to them without being “stigmatized” as a foreigner or simply as an academic person using the formal grammar or pronunciation. This helped me record more natural speech without self-corrections.

In evaluating my data, I would say that the participants code-switched more when the recorder was off. This fact comes to confirm that the “observer paradox” plays crucial role in gathering data in a sociolinguistic interview. It was maybe the main obstacle that limited the amount of code-switched items in my data compared to the other challenges I faced.

With some of the participants my presence was factor as well. One of them said that if I was not there they would code-switch more with her roommate, as they usually do. To the question “what is the reason?” she answered that for them I was coming from Bulgaria and was not residing in Greece – temporary or permanent for whatever reason – education or work. That however, was not the case with other participants, who were my friends and knew me closely.

The topic was a factor also that influenced the degree of code-switching. Overall, it was observed that code-switching occurred mostly and with all participants when they talked about their way of living in Greece and their problems related to that. This is the case also with topics related to their student life, both about the university and leisure. All of them showed code-switching when telling funny stories or jokes, retelling movie plots, or something they had recently watched on television. Not so much code-switching occurred when they talked about friends, places or events that had happened in Bulgaria.

Overall, the interviews can be characterized as casual in terms of style and despite the concerns listed above they can be defined as naturally-occurring conversations. The purpose of this study was to obtain enough data for analyzing structural aspects of code-switching and not for social ones. In this respect, the collected data were appropriate and suitable.

3.3. CODING

The speech of the participants was transcribed using Praat software, and then transferred to Word, Excel and Goldvarb software during the data analysis stage. From the eight and a half hours of transcribed speech only the clauses including code-switching served as the base for the analysis dataset. Approximately 1% of these were indistinct tokens due to overlapped speech and were therefore excluded.

The proportion of code-switched speech wasn't equal for all participants. Some of them didn't code-switch so much as they did when the recorder was off. That comes to confirm that the Observer paradox played an extremely important role in gathering data in a sociolinguistic interview.

Overall, it was observed that code-switching occurred mostly and with all participants when they talked about their way of living in Greece and their problems related to that. This is the case also with topics related to their student life, both about the university and leisure. All of them showed code-switching when telling funny stories or jokes, retelling movie plots, or something they had recently watched on television. Not so much code-switching occurred when they talked about friends, places or events that had happened in Bulgaria.

As mentioned in Chapter 2, this study focuses on intra-sentential code-switching that creates utterances with lexical and morphosyntactic structure from one language, referred to

4) Za **arsi** **varon** deto im bialia dali doping.
 About lifting weight that them be.2PL A give.2PL PPAA dope
 It was the weightlifters who were doped.

5) A, **den** **exei,** vsichko e prazno
 PRT not has everything is empty
 "There's nothing here; everything is empty."

On this ground the available data was analyzed.

CHAPTER 4: DATA DESCRIPTION AND ANALYSIS

In this chapter the data for single-word as well as phrasal insertions will be presented followed by a separate section on the definite article and definiteness. As it was mentioned earlier, the code-switched single words are considered as instances of code-switching and not borrowing. The latter might be realized when “single words first occur as code-switches and gradually, over time may become fully accepted borrowings” (Naseh 2002).

For the explanation of the data in this study I follow Myers-Scotton’s (1993) MLF of code-switching. It states that the Matrix Language is “the language of more morphemes in interaction types including intra-sentential code switching” (Myers-Scotton 1993:68). She also suggests that frequency counts must be based not only on single sentences, but on a discourse sample. Later, she adds that the ML is the one that shows the system morphemes in CPs.

Having said that, on an intra-sentential level we first identified inserted Greek lexical items into Bulgarian frames, second, inserted Greek phrases or clauses (“islands”), and on an extra-sentential level – whole Greek sentences embedded in the discourse, either within a single turn or as a part of different ones. Table 4 below represents the distribution of the code-switched items in the data.

Type of code-switched item	Number	Percentage
Inserted Greek lexical items	172	54%
Greek phrases and clauses	62	19.4%
Greek sentences	85	26.6%
Total	319	100%

Table 4. Distribution of the code-switched elements in the data

The amount of inserted Greek lexical items comprises about half of the total number of the code-switched elements. The other half is divided between whole switched sentences from one hand and phrases and clauses on the other.

Table 5 shows the proportion between the total amount of monolingual and bilingual CPs and what percentage of them is taken from CPs with code-switched elements in the entire set of data:

Type	Number
Monolingual and bilingual CPs	1888
CPs with Greek elements (lexical items + phrases)	234 (12.3% of Total CPs)

Table 5. General distribution of the code-switched and non-code-switched CPs

Inserted Greek lexical items can be further divided according to their lexical category, which is as follows:

Lexical category	Number	Percentage
Nouns	127	73.8%
Verbs	22	12.9%
Adjectives	17	9.9%
Adverbs	3	1.7%
Others	3	1.7%
Total	172	100%

Table 6. Greek lexical items embedded in Bulgarian frames

As one can see, nouns represent 3/4 of the total amount of embedded Greek elements, while verbs are a distant second with only 22 instances. This matches the results of

previous studies such as Poplack et al (1990), Gardner-Chloros (1995), Myers-Scotton (1993), Köppe and Meisel (1995).

4.1. LEXICAL INSERTIONS

4.1.1. Nouns

Inserted nouns occur in two different forms: suffixed by Bulgarian markers for definiteness or number, unmarked (some of them with different gender assigned). As can be seen in Table 7, unmarked nouns are almost twice as many as the marked ones. However, the environments where both groups occur will be discussed further. The examples that follow illustrate each group presented in Table 7:

Lexical category	With Bg markers	Without Bg markers	Double marking
Nouns	46 (33%)	81 (66.4%)	1 (0.6%)

Table 7. Distribution of marked and non-marked Greek inserted nouns

Greek nouns with Bulgarian marker fall into two groups:

a) marked for definiteness:

- 6) Hodih dneska do **komotirio-to**
 went 1SG today to hairdresser+BG DEF.ART.N.SG
 “I went to the hairdresser saloon today”

In 6) we see the Greek noun *komotirio* (‘hairdresser’) with a postposed Bulgarian definite article *-ta*.

b) marked for number:

- 7) Dobre be, kyde ima
 okay PRT where have.3SG

falen-i v tui more be?
 whales in this sea PRT
 “Where can you find whales in this sea?”

In 7) we have the Greek noun *falena* ('whale') but instead of the expected Greek ending for case and number (-es), we find the corresponding Bulgarian ending for Pl, feminine: -i. Greek noun without Bulgarian marker is shown in 8) where the Greek noun *komboloi* ('worry-beads') appears with no marking:

- 8) Vinagi idvashe s **komboloi**
 always came.3SG with worry-beads
 "He would always come with worry-beads"

4.1.1.A. Greek nouns with Bulgarian markers

The two markers occurring in the data are for number and definiteness. There are two nouns receiving Bulgarian inflectional morpheme -i for plural feminine: *falemi*, *erevni*. These nouns are feminine in Greek, too.

- 9) Biah minaloto liato na Samotraki,
 Was 1SG last summer on Samothraki

 da pravim tia **erevn-i**
 SUB PRT make 1PL these research+BG F.PL
 "I was last summer on (the island of) Samothraki to do this research"

Out of the 46 morphologically marked nouns, 44 are marked with a morpheme for definiteness. The two languages, however, differ in the position of the definite article. In Bulgarian we find it postposed – a bound morpheme added to the stem – a feature, shared by all language-members of the Balkan Sprachbund. This way, the definite article morpheme “exhibits a clitic-like behaviour” (Avgustinova 1994). In Greek, on the other hand, we find it preposed — making the language an exception to the linguistic union. Both the Greek and the Bulgarian definite articles have gender and number features. This is illustrated in 10):

- 10) to **maθima** vs **ðilima** **-ta**
 GR DEF.ART.N.SG GR NOUN GR NOUN BG DEF.ART.N.SG
 “lesson” “dilemma”

Ten of the Greek nouns with Bulgarian definite articles have masculine gender, 17 feminine and 17 neuter. Formal factors such as morphological ending seem to play role in assigning different gender when a Bulgarian determiner is attached. All code-switched neuter Greek nouns ending in *-ma* receive the definite article morpheme for feminine gender *-ta*. The *a* from the Greek ending overlaps with the ending *-a* corresponding to feminine nouns in Bulgarian:

- 11) Tochno tia beshe s **ðilima-ta.**
 exactly she was.3SG with dilemma+BG DEF.ART.F.SG
 “It was exactly her, with the dilemma”

Unsurprisingly, all the Greek feminine nouns ending in *-a* preserve their gender when the determiner is joined:

- 12) Toi prosto otishyl zaradi **empiria-ta**
 he simply go.3SG.M.PPAA because of experience+BG DEF.ART.F.SG

 da vidi kakvo e po-goliama zhena
 SUB PRT see.3SG.PRES PERFVE what is COMPR big woman
 “He went to see just for the experience of how it was like with an older woman”

As for the nouns in masculine, their endings in Greek are *-as*, *-as*, and *-is*, but similarly to the feminine case above, their *s* overlaps with the ending of the masculine nouns in Bulgarian, which is always a consonant (with a few exceptions). Thus, the masculine assignment of the determiner is preserved everywhere.

The preceding lexical item in 27 of the cases is a Bulgarian preposition. The most frequent one is *na*, meaning ‘on’, ‘at’ or ‘of’ with 10 instances, then *za* (‘for’, ‘about’) with 6 instances, *v/vav* (‘in’ or ‘inside’) with 4 instances, and *do* (‘to’) with 6.

- 13) Maikoo, minalata godina na paralia-ta
 mother F.VOC last year on beach+DEF.ART.F.SG
 “Oh my god, last year on the beach”

Verbs are preceding elements in 4 examples, 3 of them in subjunctive function.

- 14) da, da, pak za da napraviat **γamos-a,**
 yes yes again for SUB PRT make.3PL.PRES PERFVE marriage+BG
 DEF.ART.M.SG
 da se stigne do proslovutia **γamos-a**
 SUB REFL reach.3SG.PRES to famous marriage+BG DEF.ART.M.SG
 PRT PR PERFVE
 “Yes, it is about that marriage, so that they can end up in a marriage”

In 19 of the cases the Greek noun was found in the beginning of the sentence:

- 15) **Priki-a-ta** e vazhna.
 dowry+BG DEF.ART.F.SG is important
 “The dowry is important”

The rest of the preceding items include verbs (5) noun (1), demonstrative pronoun (1), conjugation (3).

Greek nouns with Bulgarian articles occur far more frequently as objects (31) in the clause than as subjects (13) as the preferences are given to direct object position. The issue of definiteness will be discussed in more detail in the last section of this chapter.

4.1.1.B. Greek nouns without Bulgarian markers

These constitute 2/3 of the Greek noun insertions, or 81 out of 127. 18 of them participate in mixed noun phrases where the first nominal constituent is a Bulgarian lexical item. In Bulgarian the modifiers must agree in gender and number with the head of the noun phrase. Through them we can see that 5 of the Greek embedded nouns have different gender assignment. Different word classes such as demonstrative pronouns, adjectives, possessive pronouns, numerals, and indefinite pronouns all serve as modifiers in these mixed noun phrases. All of them have different forms for gender and number. In three of the cases we find neuter gender assigned to a Greek noun that is actually feminine. In 16) the Greek feminine noun *vaftisi* is transformed to a neuter gender through agreement with the neuter numeral before it:

- 16) Edno **vaftisi** kolko struva?
one N.SG baptism NOM F.SG how much cost.3SG
“How much is a baptism ceremony? “

In 17) we have a mis-match from neuter to masculine where *to peripteron* (“kiosk”) having neuter gender in Greek is used in masculine in Bulgarian, shown by the agreement with the demonstrative pronoun (*toz*) before it.

- 17) Ha, toz **peripter** kyde tam na moreto imame?
EXCL this kiosk N.SG where there on sea have.1PL
“Where do we have that kiosk on the seaside?”

Finally, we have a case where the gender is preserved but the number is changed. In the example below the Greek word for department *tomeas* is used in its plural form *tomis* and

it is the head of a noun phrase where the first constituent is inflected with a singular morpheme for masculine, namely *-ia*.

18)	Te	sa	go	dali	na
	they	are.3PL	3SG.N.ACC.CLI	give.3PL.PPAA	of
	niakoi	nashi	kolegi	ot	drugia
	some	our	colleagues	from	other
	tomis	da	go	testvat.	
	department M.SG.NOM	SUB PRT	3SG.M.ACC.CLI	test.3PL	
	“They gave it to some other students from the other department to test it”				

Since Modern Greek has morphological case marking, the noun should be in the case determined by its syntactic function in the sentence. In 19) the Greek noun *isologizmos* plays the role of direct object and, accordingly, should be in accusative case. According to the declension its form is *isologizmo*, without the final *s*. This way, the Greek noun ends in *o*, which overlaps with one of the endings for neuter nouns in Bulgarian. And not surprisingly, we find this noun in gender agreement with the preceding adjective *falshivo* (“falsified”), which is in neuter and ending in *-o* as well.

19)	Dade	im	falshivo	isologyzmo,
	gave.3SG	3PL.DAT.CLI	falsified	balance sheet
	deto	za	balansa	ikonomicheskia.
	that	for	balance	economic
	“She gave them a falsified economic balance sheet”			

Overall, the same formal factors as in the case with the definite article bound morpheme can be observed here. If the Greek nouns’ ending coincides with an ending for nouns in feminine or neuter, then naturally the Greek noun receives one of them. The rest of the unmarked nouns occur in EL islands.

In 90% of the cases the preceding lexical item is a Bulgarian word. 41 from these are verbs, 18 prepositions, and the other 14 include nouns, pronouns, numerals, particles. From

the 41 verbs 14 are forms of the verb *sam*, ('to be'), and these are overwhelmingly (13) in the present tense.

- 20) Mi da, tva si e **yonidio** tuka.
 PRT yes this REFL PR is gene N.SG.NOM here
 "This is a gene here"

After the verb 'to be' the next most frequently occurring verb is *imam*, ('to have') with 7 instances mostly in the present tense.

- 21) To i oktopoda ima **melani,** ama...
 so and octopus have.3SG ink but
 "The octopus also has ink, but..."

In regards to the temporal use the preferences are for present tense (25 verb forms), past tense (7 verb forms) and just one in future. There are 8 verb forms having a subjunctive-like interpretation:

- 22) Ako beshe Elza shteshe da
 if be.3SG FUT PAST SUB PRT
 kazhe imam da pisha **eryasia**
 say.3SG.PRES PERFVE have.1SG SUB PRT write.1SG paper
 "If Elza was here, she would say "I have to write a paper""

4.1.2. Adjectives

The embedding of adjectives is not so active as the embedding of nouns. There are just 7 of them, incorporated into bilingual Bulgarian/Greek CPs. In both monolingual Greek and Bulgarian grammars adjectives precede the noun; therefore, when inserted, they do not violate the syntactic structure of the ML.

- 23) Frapeto kak go praviat – **metrio** ili **yliko?**
 frappé how 3SG.N.ACC.CLI make.3PL moderate or sweet SG.N
 "How do they prepare the frappé – with a little or with plenty of sugar?"

Only one comparative form was found:

Table 8 shows the pattern of morphological mixing for these verbs:

Verb	Greek stem in aorist	Bulgarian suffix (3Sg past tense)	Bulgarian suffix (3Sg present tense)	Bulgarian suffix (past active aorist participle)
paratisa	paratis	a		
paratival	paratis	a		l
ekopsal	ekops	a		l
lipothemisal	lipothemis	a		l
lisptesali	listeps	a		li (Pl)
pliroforisa	pliroforis	a		
pirovolisa	pirovolis	a		
ektiposva	ektipos		(v)a	
apotikevsva	apotikevs		(v)a	
diagrapsi	diagrap		i	

Table 8. Mixed verb constituents

Note: (v)+a gives infinite form of a verb

It should be noted that the ending *-a* in the third column is also a Greek suffix, corresponding to 1st Sg past tense but here looks more Bulgarian one. Four of the mixed verb constituents are in inferential mood (when the speaker is not witness to the event). The unit of perfective Greek stem and the Bulgarian ending *-a* receives an additional element – the suffix *-l* (for masculine Sg or *-li* for plural). The new form is past active aorist participle.

The rest of the inserted verb items are non-mixed and they all are Greek verbs in singular, present tense, indicative or imperative forms and just one in plural. The majority of the mixed verbs discussed above are transitive. They take a complement, which in 90% is an object clitic. The majority of the pure Greek inserted verbs are di-transitive as the preference to transitivity and non-transitivity is equally distributed. Only two of the verbs behaving as transitive take object clitic (ex. 27 and 28):

31) Pырvo vsichkite snimki gi
 At first all photos 3PL.ACC.CLI
apotikesv-a tam na komputera.
 save.3SG there on computer
 “First he saves all the photos on the computer”

32) Dokato nas ni niama, prozoretsa stoi
 while 1PL.ACC.PR 1PL.DAT.CLI have.3SG.PRES.NEG window stay.3SG
 zatvoren da ne mu **listepsa-li** komputera.
 closed SUB PRT NEG 3SG.M.DAT.CLI steal.3PL.PPAA Computer
 “While we’re away, the window stays closed so that his computer doesn’t get stolen”

In this respect the two grammars are not in conflict, since both require proclitics to be placed before finite verbs. No instances of Bulgarian enclitics were found accompanying Greek verbs.

The great flexibility, which both Greek and Bulgarian possess in terms of word order, makes the placement of the verbs adequate to the ML frame. All the inserted verbs are in consistency with grammatical structure in Bulgarian and occur in the same spot where their Bulgarian equivalent would be.

4.2. BILINGUAL CPs CONTAINING EMBEDDED LANGUAGE PHRASAL ELEMENTS

Following the definition of Myers-Scotton (2002) that Embedded Language islands are “full constituents consisting *only* of Embedded Language morphemes occurring in a *bilingual* CP that is otherwise framed by the Matrix Language” we found 62 instances qualifying as EL islands under the MLF model in our data. The EL islands should show structural dependency and act as constituents and that’s why examples such as 33) and 34) are excluded from the analysis. The reasons behind this are: in 33) the two EL morphemes *mini* (‘little’) *megalo* (‘big’) do not form a constituent and do not show any structural dependency; in 34) – although it is considered bilingual sentence it is formed of a monolingual main clause (*To orao itan*, ‘the good thing was’) and an embedded, also monolingual clause, but in a second language (*che nie vliazohme*, ‘that we got in’) and this way the two parts are not really in contact. An example like in 34) was cited by Myers-Scotton (2002:55) from English/French data: *I think qu’il pleuvra après midi* (‘I think it will rain this afternoon’). The structural configurations in the two examples show no interaction between the two languages, although they are complex sentences comprised of main and embedded clauses.

33) Oh, kak pochnat s tia **mini**, **meyalo!**...
 EXCL how begin.3PL.PRES.PERFVE with these little big
 “Oh, when they start with those “little, big”!”

34) **To** **oreo** **itan**, che nie vliazohme.
 GR DEF.ART.N.SG.NOM good was.3SG that we got in.1PL
 “The good thing was that we got in”

Very much like the nouns discussed earlier in the section 4.1.1, here, too, the noun phrases are predominant. 65% of the EL islands in the data are noun phrases. Illustrations are the next two examples consist of two content morphemes (a noun and modifier):

- 35) Kakto sym hvanal da cheta za **anonimi** **eteria**,
as am hold.1SG.M. SUB PART read.1SG about anonymous company
PPAA
vikam dai da vidia za Bulgaria v edin forum,
say let SUB PART see.1SG.PRES about Bulgaria in one forum
PERFVE
“Yesterday, since I’d already been reading about limited companies,
I said to myself, let me check out Bulgaria in that forum”

- 36) Tolkova e kysmetliika na tia **tihera** **pexniðia**
so is lucky on these lucky games
“She’s so lucky in that lotto”

There are just two instances of an EL island as verb phrase:

- 37) **ða** **pezi** farsi obache mine se ne
FUT play.3SG perfectly however pass.3SG.PRES REFL PR NEG
PERFVE
mine obazhda se da me pita.
pass.3SG.PRES PERFVE call.3SG REFL PR SUB PRT 1SG.DAT.CLI ask.3SG
“He speaks them perfectly but every now and then he’d call to ask me things”

As Myers-Scotton (2002) points out many EL islands play the role of adjuncts. The prepositional phrase in 38) is a such adjunct. It is an adverbial phrase of place:

- 38) Az sym roden v Plovdiv, shest godini,
I am born in six years
posle sedem godini v Sofia, sti mesi kapou
then seven years in PREP + DEF.ART middle somewhat
“I was born in Plovdiv – that’s six years (I spent there) – and then seven years in Sofia,
so it’s somewhat in the middle”

Interesting bilingual CP was the following example where we have double expression of a personal pronoun. Once we have it in a prepositional phrase being an EL island - the

Greek personal pronoun in its emphatic form *emena* (accusative, 1st Sg), and second time a Bulgarian personal pronoun (dative, 1st Sg) but in its short form *mi*.

39) **γία** **mena** arabskite tanci ne mi haresvat.
 for PR.SG.DAT Arab dances NEG 1SG.DAT.CL like.3PL
 “I personally don't like Arab dances”

4.3. EXCEPTIONAL TOKENS

I present here those tokens of code-switching that are exceptional with respect to the rest of the data. Two of them are double marked, i.e. they show Bulgarian and Greek markers for definiteness simultaneously. In example 40 double marking is demonstrated by the presence of the Greek definite article *to* for Sg neuter and Bulgarian definite article *-to* for Sg neuter but added to the noun as a bound morpheme. In 41) we have the same Bulgarian and Greek definite article for neuter, Sg, this time applied on a demonstrative pronoun *to tetio* (‘such’). Demonstrative pronouns are not marked for definiteness in Bulgarian. Adding a definite article to them makes the word more noun-like. Similar examples of double morphology across language pairs are found predominantly with plural markers, not so much with definite articles (Crawhall 1990 on Shona-English, Kamwangamalu 1989 on Lingala-French, Myers-Scotton 1993 on Swaili-English). Myers-Scotton explains their existence within the 4-M model (Myers-Scotton & Jake 2001) discussed earlier and states that plurals and definite articles belong to the so-called “early” system morphemes category and are “accessed simultaneously with their Embedded Language noun during language production” (Winford 2003).

One way to look at this CP is the IP *kati pou iperveni ta epitrepomena oria tis malakias* found within the CP headed by the subordinating conjunction *che* ('that') in complementizer position. However, it is odd to have a complementizer as a head and under its umbrella to find another complementizer. The other reading is to see *kati* ('something') as a head of the phrase but again the result is a structurally ungrammatical sentence. It is clear that one of the complementizers is not needed and it is more likely that to be the Bulgarian one.

4.4. CODE-SWITCHING AND THE DEFINITE ARTICLE

In this section I pay closer attention to the choice of definite article for inserted Greek nominals and I discuss the concept of definiteness in terms of its formal expression in Bulgarian and Greek and its semantic content. As was mentioned early on, Bulgarian and Greek differ with respect to the placement of the definite article; in the former it is a post-position while in the latter it precedes the noun phrase (noun and preceding adjectives). In the code-switching data of this project, this difference shows up as a choice between two possible variants: using a Greek (preposed) definite article with a Greek noun versus using a Bulgarian (postposed) definite article with a Greek noun:

to	maθima	vs	ðilima	-ta
GR DEF.ART.N.SG	GR NOUN		GR NOUN	BG DEF.ART.N.SG
"lesson"			"dilemma"	

In other words, the choice between the two types of construction can be treated as a linguistic variable.

In order to explore the variation, the following coding schema with four factors groups was designed:

Factor Group	Factors
Variable: definite article	1 = Bg definite article 0 = Gr definite article
FG2 Greek noun ending	[a], [i], [o], C
FG3 Preceding word	B = Bg lexical item G = Gr lexical item
FG4 Subject - Object	S = subject O = object

Table 9. Coding schema for the variants

Why are these groups considered to have influence? As designed, the coding schema lists four independent variables. FG2 is observed to determine the choice of the Bulgarian definite article related to the endings of the Greek noun. Endings like these are found in Bulgarian nouns and they may have influence on the gender assignment of the Bulgarian definite article in one of the variants. Nouns that are masculine and end in a consonant (C) use *a/yt*⁶ in Sg as a definite article. Nouns that are feminine and end in [a] use *ta* in Sg. Nouns that end in [o] use *to* in Sg. Choosing this factor group I was trying to test if the speakers are influenced by the Greek noun ending, and assign definite article from the grammar of the ML. FG3 is chosen in relation to the matrix language on a sentence level. Hence my hypothesis that the matrix language of the sentence immediately registered before the token governs the choice or maybe activate the Triggering Hypothesis.

The definiteness marking was seen by Spyropoulos (2004) to interact in a complex way with the encoding of the grammatical relations of subject and object in Cappadocian

⁶ The choice between the two forms is based on an artificially created rule in Standard Bulgarian according to which *-a* is reserved for direct/indirect object and *iat/yt* for subject.

Greek dialects. FG4 is an attempt to test whether the choice of the definite article is related to the syntactic categories subject and object.

The extracted tokens were coded according to the Greek noun ending, the nature of the preceding lexical item and according to its syntactic function in the clause (subject or object). The initial coding schema included four more factors groups: *Greek Noun Gender*, *Greek Noun Number*, *Following Lexical Item*, and *Clause Type*. The first run of the data in GoldVarb gave three knockouts, which represented either a very low-frequency items in the data, or a 0 presence. They were omitted lately due to no influence on the variation. However, the main problem was the fact that in the step-down and step-up analysis, no match between the best runs occurred. This indicated that there are too many factor groups for this small number of tokens, and required decreasing the numbers of the factor groups. Therefore, the above-listed factors were excluded.

4.4.1. Results and Analysis

The final run of the data showed FG3 to be the most significant factor with influence on the variants. However, the number of tokens was low so that a logistical regression could not be run but interesting observations can be made from the overall distribution. Table 10 presents the distribution of the two variants across factor groups.

FACTOR GROUP (N=60)	VARIANT				TOTAL	
	BG definite article with Greek noun		GR definite article with Greek noun		Tokens	%
	Tokens	%	Tokens	%		
Greek noun ending						
[a]	22	73.3	8	26.7	30	50.0
C (consonant)	10	71.4	4	28.6	14	23.3
[o]	7	100.0			7	11.7
[j]	5	55.5	4	44.5	9	15.0
Preceding lexical item						
Bulgarian	37	84.1	7	15.9	44	73.3
Greek			10	100	10	16.7
no presence	4	66.6	2	33.4	6	10.0
Syntactic function of the Greek noun						
Subject	13	72.2	5	27.8	18	30.0
Object	31	73.8	11	26.2	42	70.0

Table 10. Distribution of variants across factor groups.

The overall proportion in percentage between the two variants is 73% (Bulgarian definite article) vs. 27% (Greek definite article). Only one error was detected, but given the insufficient data set, this is no large surprise. The step-down and step-up analysis selected the *Preceding Lexical Item* as the most significant factor group compared to the others. Bulgarian definite article is used in 37 tokens (vs. 7 for the Greek), where the preceding element is a Bulgarian word where half of these is a preposition.

Looking at the ML of the sentence where these tokens are found, it can be easily seen that this is Bulgarian, and the Greek noun is considered as EL. Greek word as a preceding element doesn't occur at all. The Greek definite article is used with just 7 tokens when the preceding element is a Bulgarian word, and in 10, when it is a Greek word. The factor group *Greek Noun Ending* was not selected by the program as a significant factor. Yet it is interesting to see that from a total 30 Greek nouns in the data ending in -a to 22 of them, were assigned the Bulgarian definite article for feminine singular or plural.

My hypothesis stated that the similarity in the endings between Greek and Bulgarian nouns facilitates the code-switching to occur between stems and bound morphemes. Even though this was a small dataset, it provides evidence that this is possible, violating the Free morpheme constraint (Poplack 1980), not just with single instances. Moreover, this micro-level code-switching can be seen not only with nouns, but between verbs and inflections taken from the two languages respectively. Binomial one-step analysis showed 15 tokens (out of 60) of cell type *aBO* where we have a Greek noun ending in *a*, functioning as an object and preceded by a Bulgarian lexical item.

The syntactic factor group *Subject-Object* apparently does not determine the speaker's choice. Rather it seems that Greek nouns with Bulgarian articles occur far more frequently (31 tokens) as objects in the clause, than with Greek articles (13 tokens). In the position of a subject, preferences are for using Greek definite article more than Bulgarian one.

After having made these observations on the structural aspect of code-switching and the definite article I now briefly look at the semantic aspect of definiteness. This topic is broadly discussed in the field of Balkan linguistics. Attempts to define definiteness gravitate toward one of two main camps – that of identifiability (Lyons 1999) and that of inclusiveness (Hawkins 1978). Additional definitions have been proposed also, such as familiarity (Christopherson 1939) and uniqueness (Russel 1905).

The identifiability-based theories argue that when a speaker uses a nominal plus a definite article he assumes that the hearer is able to identify the referent. That can be done in different ways: the referent was previously mentioned in the conversation (through direct or indirect anaphora); the referent is physically accessible in the situation or it can be deduced based on a shared knowledge of the world (Mladenova 2007).

The inclusiveness-based theories focus on the type of noun and the particular context. They view referents differently according to the type of noun. For example the referent is unique for personal names, while for mass and count nouns the referent can be uniquely relevant in the given situation or to a kind in general (Mladenova 1997). These two descriptions roughly correspond to Löbner's (1995) distinction between semantic and pragmatic definiteness, which I will follow in my discussion. Semantic definiteness is said to be that one where the referent of the definite noun is found to exist independently of the immediate situation or context of utterance. Definiteness here depends entirely on the semantics of the nouns. The pragmatic definiteness on the other hand is exactly the opposite – where the referent is directly linked and determined by the context and concrete situation. That can be demonstrated through deictic, anaphoric or endophoric uses of definite nouns. I do not discuss pragmatic definiteness but will just mention that is entirely demonstrated by Greek nouns plus Bulgarian definite articles.

Löbner's distinction corresponds to a greater extent to Napoli's (2009) Map of Definiteness where she prefers the term logical definiteness instead of semantic. Napoli states that languages across the world differ very little in terms of their pragmatic definiteness but greatly in their logical definiteness.

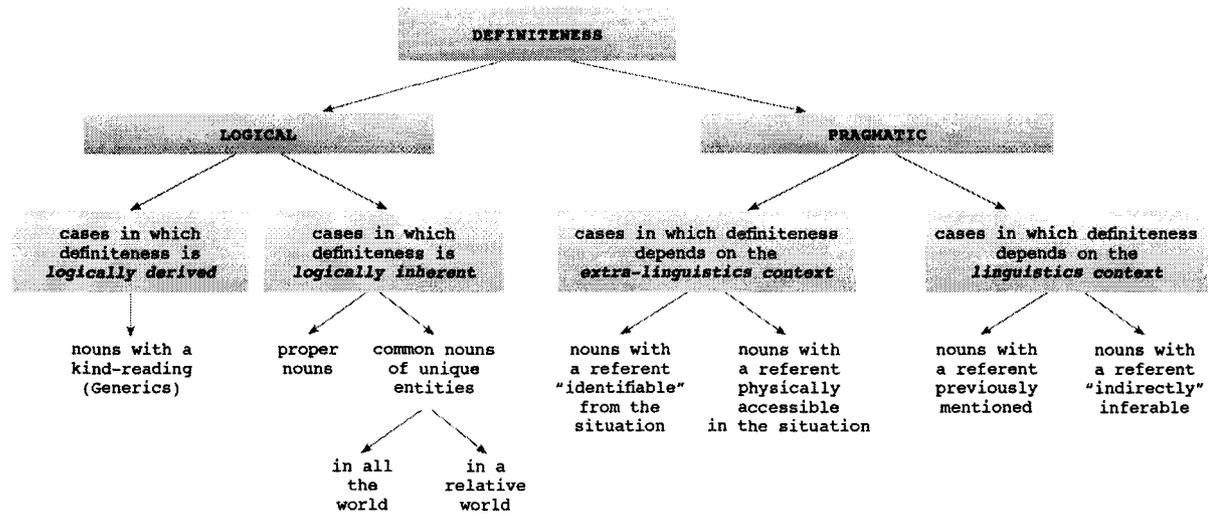


Figure 5. Map of definiteness (Napoli 2009)

She includes three items in the semantic branch – instances in which the definiteness is logically derived, represented by nouns with generic reading, and instances in which definiteness is logically inherent. The last one is subdivided into proper nouns and common nouns of unique entities, which in turn are those in all world and those in relative world. Löbner discusses all these items in terms of Functional concepts. Every noun phrase is a “semantic definite if it represents a functional concept”. Functions link objects unambiguously to other objects. When the nouns execute these functions they become inherently unambiguous and as a result the referent is always identified unambiguously. An example for a functional noun, Löbner gives the word *wife* and compares it to *woman*. “It is the relationship of being married to a man, which determines the referent of ‘wife’ (Löbner1995:292). In this sense *wife* exemplifies a functional concept and *woman*, a sortal concept (it just has classifying function, exhibiting some property). All of the functional nouns have their argument(s). Löbner considers the situation as one of their arguments in all cases. On these grounds, he defines two types of functional concepts:

FC1 (one-place functional concepts) – a procedure assigning objects to a given situation;

FC2 (two-place functional concepts) – a procedure assigning objects to objects in a given situation.

In other words, candidates for semantic definiteness will be those nouns for which the referents can be identified. That, however, is possible due to semantic and other types of constraints. For example, typical representatives of FC1 nouns are those showing spatial location but having specific function for a particular district, which make them unique and easily identifiable to the hearer. Such locations in our data are *leshito* ('canteen'), *paraliata* ('beach'), and *limanito* ('port') among others.

43) Ama do **leshito** ne stigma li?
 But to canteen+DEF.ART.N.SG NEG reached.2SG INT PRT
 "What, didn't you make it to the canteen?"

44) Maikoo, minalata godina na **paralia-ta**
 mother SG.F.VOC last year on beach+DEF.ART. F.SG.
 "Oh my god, last year on the beach"

45) Vchera v **limani-to** biaha naglaseni
 yesterday in port+DEF.ART.N.SG were.3PL adorned
 "Yesterday in the port they were all dressed to kill"

In the examples above, all code-switched nouns can be identified immediately by the hearers due to the uniqueness of the referents. There is only one canteen in the university, only one port in the city where you can go for night life, and only one beach where book fairs are held.

Not all nouns having functional characteristics automatically become FC1 nouns. There are examples where nouns, usually exemplifying functional concepts, are used in a

more generic sense. To illustrate this point, here is an example with the common abstract noun *aravonas* ('engagement') and *gamos* ('marriage'). Even though these two words are functional nouns – requiring two arguments and referring to unique entity – their reading in the following mixed constituent is, in a sense, generic. The statement as a whole does not refer to a concrete, specific situation, but to an abstract one:

46) Grytskoto momiche v eparhia-ta za kakvo zhivee?
 Greek girl in eparchy+DEF.ART.F.SG for what lives
 "What does a Greek girl live for in the eparchy?"

47) **Aravon-yt,** **γamos-yt,** posle da rodi, da
 engagement+DEF. marriage+DEF.ART after SUB PRT give SUB PRT
 ART.M.SG .M.SG birth.3SG.PRES
 PERFVE
 stane **spitonikokira,** i da stane **koumpara**
 become.3SG.PRES landlady F.SG.NOM and SUB PRT become.3SG.PRES bridesmaid
 PERFVE PERFVE F.SG.NOM
 "(For) an engagement, a marriage, to give birth, to become a landlady, and to become a bridesmaid"

The most easily recognizable FC1 members are the proper names, for the simple reason that they denote specific objects, thus embodying functional concepts. The proper names in Greek (as in Albanian and Turkish in some grammatical cases) must necessarily be accompanied by a definite article. This is also a requirement for the other two items under Napoli's Logical Definiteness: generics and common nouns of unique entities. In Bulgarian (as well as Romanian), however, proper nouns do not take a definite marker. In isolated cases we find definite articles used with diminutives: *Sashe-to* (from *Sasbo*, a male name) or *Emmy-to* (from *Emmy*, female name). In our data we find all the proper Greek names to be used without either Greek or Bulgarian definite articles, but three instances of using a Bulgarian definite article with Greek proper names:

- 48) Nali kazvali, che **Mesoyio-to** i
 IINT PRT say.3PL.PPAA that Mediterranean+DEF.ART.N.SG and
Eyco-to bili naj-mrysnite moreta
 Aegean+DEF.ART.N.SG be.3PL.PPAA SPRLdirty seas
 “They said that the Mediterranean and Aegean seas were ones of most polluted seas”

- 49) Na **Maraki-to** sestra i det
 on Maraki+DEF.ART.N.SG sister 3SG.F.DAT.CLI that
 goli si trygnali ottuka.
 naked REFL PR leave.3PLPPAA from here
 “Maraki’s sister who left naked from here”

If a definite article serves as an indicator of non-ambiguity of reference (Christophersen 1939) why do the speakers use it with objects definitely unique in the world? There is only one Mediterranean sea in our physical world and the Aegean sea experiences the same semantic characteristic of uniqueness.

In the case of personal names, the use of the definite article cannot always serve as giving an identification status marker. There can be a few friends of the speakers and hearers that know girls with the name *Maraki* and the use of a definite marker does not really identify the referent unambiguously as to which *Maraki* the topic is about. In this sense, definiteness is not best conveyed by the use of the definite article.

The FC2, on the other hand, are from subdivided to FC2 with explicit arguments and those with implicit arguments. Examples of FC2 with explicit argument are the possessive constructions like in 50):

- 50) I toi kazva, tova e Boriana,
 and he says, this is
i kopela mou, moito momiche.
 GR DEF.ART.F.SG.NOM girl PR.SG.GEN my girl
 “And so he says, this is my girl, Boriana”

Example of FC2 with implicit argument is 51:

51)	Imalo	edno	vreme,	s'ena	ksehasmeno	ðasos,	
	have.3SG.N.	one	time	in + one	forgotten	wood	
	PPAA						
	kvo	beshe	tam	za	smyrfchetata,	ta	strofakia.
	what	was.2SG	there	about	Smurf	GR	dwarfs
						DEF.ART.PL.NOM	

“Once upon a time, in the woods faraway – how was it about the Smurfs?”

51) is an example of Hawkins’ larger situation use based on general knowledge. The location of the utterance belongs to a territory sharing common folk knowledge where the FC2 is assigned. In this sense the argument is implicit.

One general observation is that speakers tend to use Greek nouns plus Greek articles when it comes to FC2 with implicit arguments, but Greek nouns with Bulgarian articles are preferred when it comes to FC1. FC1 are the basic representatives of semantic definiteness, which is perhaps why their definite marking is drawn from the speakers’ native language, Bulgarian.

When we look at the type of nouns in terms of their abstract or concrete level, more prevalent are the concrete countable nouns with added Bulgarian definite marker, at 72%. The remaining 18% belong to the abstract countable nouns. The group of preferences of Greek nouns plus Greek definite marker are equally divided between abstract and concrete nouns, at 47% versus 53%.

CHAPTER 5: CONCLUSIONS

This study presented an examination of the code-switching patterns of Greek-Bulgarian bilinguals residing in the city of Thessaloniki, Greece. From a total of 1888 bilingual and monolingual CPs, 234 with switched elements—172 Greek lexical insertions and 62 phrasal Greek islands—were extracted and analyzed in terms of the contribution of content and morphological elements by the two languages.

The first major finding concerns the use of the definite article in nominal insertions. As discussed in Chapter 4, in 73% of the cases speakers use a Bulgarian definite article to accompany a Greek noun. The most significant factor affecting the choice is the preceding lexical item; in more than half of these cases the preceding lexical item is a Bulgarian preposition. It was also found that phonological factors can lead to gender or case mismatches between the two grammars; for example, a Greek neuter noun can be paired with a feminine Bulgarian definite article when the Greek neuter ending has the same form as Bulgarian feminine endings.

The second major finding is related to the semantic use of the definite article. Within the Balkan Sprachbund context, this grammatical feature shows different applications in the two languages with respect to the use of proper nouns. Proper nouns in Greek, as in Albanian and in some cases Turkish (Napoli 2009:588) are obligatorily accompanied by a definite marker. Proper nouns in Bulgarian (as in Romanian), however, do not take a definite marker. Subjects in this study used the Bulgarian definite article with Greek proper nouns. Such use is not allowed in standard Bulgarian—excluding isolated cases where definite articles are added to diminutives. Moreover, the use of a definite article was observed with

toponyms, which by definition are unique. It makes no sense to add a definite article in these cases, at least not in Bulgarian. Example 48 is repeated here for illustration:

48)	Nali IINT PRT	kazvali, say.3PL.PPAA	che that	Mesoyio-to Mediterranean+DEF.ART.N.SG	i and
	Ey eo-to Aegean+DEF.ART.N.SG	bili be.3PL.PPAA	naj-mrysnite SPRL.dirty	moreta seas	

“They said that the Mediterranean and Aegean seas were ones of the most polluted seas”

The third finding comes from the analysis of the verbs. No light constructions were found, in contrast to other code-switching studies. For example, Rezaeian’s (2009:59) Persian/English data for example provides a large number of those – 17% of a total 1088 embedded English elements. Table 8 in Chapter 4 shows evidential marking on Greek stems, a feature that is not part of Greek verbal morphology. Fully inflected verbal forms are assembled by using the aorist Greek verb stem in aorist, adding the suffix *-a* (which is a convergence between 1st Sg aorist in Greek and 3rd Sg aorist in Bulgarian) and the Bulgarian aspectual morpheme *-l/-li*.

These findings can provide insight for several issues in the literature of code-switching. As already demonstrated in Chapter 4, Greek lexical insertions show morphological integration into the Bulgarian framework, but not a phonological one. Greek words preserve their original pronunciation, which means keeping sounds like [ɣ], [θ], [ð] – sounds that Bulgarian lacks. This fact and the absence of semantic integration with proper nouns mentioned above argue for looking at the lexical insertions in the present data as part of a continuum and not belonging to a binominal distinction – borrowing versus code-switching.

Second, my data provides support for Myers-Scotton’s MLF (Myers-Scotton 2002) in terms of the distinction between system and content morphemes. Looking at the sequence of cases with Bulgarian preposition—Greek noun—Bulgarian definite article the schematic representation will appear as follows:

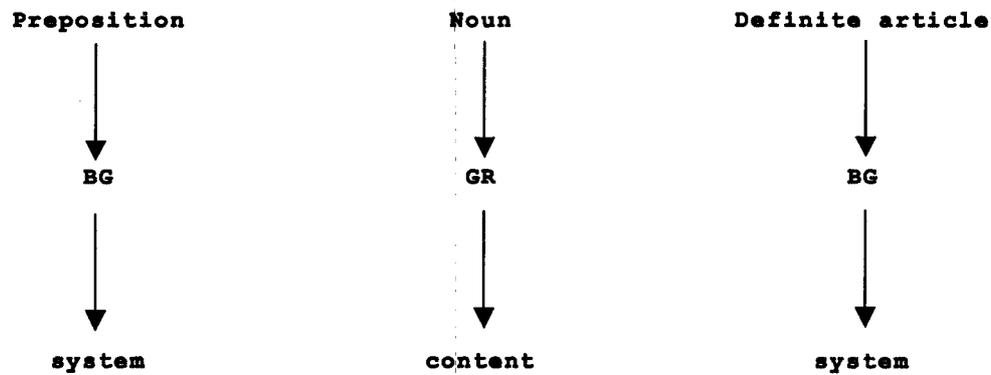


Figure 6. Sequence of system and content morphemes.

This supports her theory at least with respect to classifying definite articles and prepositions (albeit not all) as *system* morphemes, i.e. they neither assign nor receive thematic roles, and also supports her prediction that they should come from the matrix language (ML). The underlying idea of the MLF is that one language in a code-switching situation is dominant—defined as a matrix language—while the other, less dominant participating language is defined as an embedded language. Myers-Scotton (2002) proposes a 4-M model, where different types of morphemes are described and classified according to the abstract level of linguistic competence and production. In this model, system morphemes form a ‘nest’ for the insertion of content morphemes, and my data support that prediction. As to the type of prepositions, most frequently occurring are those with most abstract and general relationships in Bulgarian (Asenova 2002): *v* (‘in’), *na* (‘of’, ‘on’) and *za* (‘about’, ‘for’).

The findings of this study also revealed that morphological and syntactic integration of the embedded elements was in agreement with the matrix language and as a result no ill-formed utterance was found: nouns and adjectives appear in the same syntactic slots where the Bulgarian counterparts would be, and are morphologically adapted to the host language. The same phenomenon is observed with verbs, with the only difference of some being “modified” by adding Bulgarian aspectual morphemes.

At the same time the data propose challenges to MLF, providing examples with no clear-cut conclusions about code-switching patterns. In Chapter 4, I discussed an example (*ia paratisal*, ‘and he left her’) in which it is impossible to determine which is the matrix language, Greek or Bulgarian. In the example below, no explanation can be offered why the complementizer *otan* (‘when’) in the beginning of the sentence switched. According to the MLF, this type of morpheme must come from the matrix language, in this case from Bulgarian.

52)	Otan	tuka	doidoh	si	uchih
	when	here	came.1SG	REFL PR	studied.1SG
	samo	kvoto	triabvvashe	da	ucha
	only	what	must.3SG IMPRF	SUB PRT	study.1SG
	“When I came here, I studied only what I was supposed to study”				

Another challenge is the use of a definite article with proper names in Bulgarian. If Bulgarian is a matrix language and “dictates” the rules of setting the morphosyntactic frame, then no definite marker would be needed. One can ask whether this is a nascent structural and semantic transfer due to language contact represented by code-switching utterances. The small dataset notwithstanding, the observed tokens serve to motivate a future investigation of the definite article’s behavior in terms of its semantic use.

And lastly, the switches between a lexical stem and a bound morpheme as in the case of a Greek noun and a Bulgarian article are additional counterexamples against Poplack's (1980) notion of the Free morpheme constraint. It is one of the early attempts to determine allowed and prohibited mechanisms of code-switching, which states that only *free*, and not *bound* morphemes, can be switched in bilingual utterances. An exception is the cases when the stem has been 'phonologically integrated into the matrix language of the morpheme' (Winford 2003:128). The group of switched verbs with evidential marking also provides counterexamples to the Free morpheme constraint. Further support for this violation is provided by Eliasson (1995) on Maori-English, Zabrodskaia (2008) on Estonian-Russian, and Myers-Scotton (1993) on Swahili-English, among others.

Due to the small number of participants and tokens, these conclusions can only be suggestive. Additional cross-linguistic quantitative studies of code-switching situations (including other Balkan languages pairs) are required to verify these suggestions. A sufficient number of tokens will allow for more group factors to be included, revealing interesting correlations possibly related to the Balkan Sprachbund features. Such studies might be able to show if non-occurrence of code-switched prepositions is accidental or is a subject of code-switching restrictions, such as the class of words that can be switched, and in what position.

From a semantic aspect, it may be worth looking into the interaction between the semantics of the preposition and the semantics of the definite article. It was shown in Chapter 4 that these form a "frame" for the content morphemes to be inserted; it should therefore be interesting too look for mechanisms through which one determines the other.

APPENDIX

Interview Questionnaire for application No 38892
Researcher: Emanuela Yancheva Mileva
MA student
Department of Linguistics

A. PROFILE QUESTIONS

1. What is your name?
2. How old are you?
3. Where were you born?
4. What do you do for a living?
5. How many years of schooling have you had?
6. (If a university student) Where did you attend university?
7. Where do you live now?
8. How many years have you lived here?
9. Where did you live before coming here?
10. Do you like living here or would you like to move to a different area?
11. Do you live alone?

B. ELICITATION QUESTIONS

12. What is your fondest childhood memory?
13. Can you tell us a fairy-tale that you know?
14. How did you meet your spouse/boyfriend/girlfriend?
15. What is the profession you like the most?

REFERENCES

- ASENOVA, PETJA. 2002. *Balkansko ezikožnanije*. Veliko Tarnovo: Faber.
- AUER, PETER. 1995. The pragmatics of code-switching: A sequential approach. *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, ed. by Lesley Milroy and Pieter Muysken, 115–135. Cambridge: Cambridge University Press.
- AVGUSTINOVA, TANIA. 1994. On Bulgarian verbal clitics. *Journal of Slavic Linguistics* 2.29–47.
- BACKUS, AD. 2003. Units in code switching: Evidence for multimorphemic elements in the lexicon. *Linguistics* 41.83–132.
- BALDWIN-EDWARDS, MARTIN. 2004. *Statistical data on immigrants in Greece: An analytic study of available data and recommendations for conformity with European Union standards*. Athens: Mediterranean Migration Observatory.
- BAN (BALGARSKA AKADEMIJA NA NAUKITE). 1989. *Gramatika na savremen balgarski knižboven ezik*. Sofia: Izdatelstvo na balgarskata akademija na naukite.
- BERK-SELIGSON, SUSAN. 1986. Linguistics constraints on intrasentential codeswitching: A study of Spanish-Hebrew bilingualism. *Language in Society* 15.313–348.
- BLOM, JAN-PETTER, and JOHN GUMPERZ. 1972. Social Meaning in linguistic structures: Code-switching in Northern Norway. *Directions in sociolinguistics*, ed. by John Gumperz and D. Hymes, 407 – 434. New York Holt, Rinehart and Winston.
- BOJADJIEV, TODOR; IVAN KUCAROV; and JORDAN PENEV. 1999. *Savremen balgarski ezik*. Sofia: Petar Beron.
- CAMPBELL, LYLE. 1998. *Historical Linguistics*. Edinburgh: Edinburgh University Press.
- CAVOUNIDIS, JENNIFER. 2004. Migration to Greece from the Balkans. *South-Eastern Europe Journal of Economics* 2.35–59.
- CIVJAN, TATJANA. 1965. *Imja sušestvitelnoje v balkanskih jazykab*. Moscow: Nauka.
- CIVJAN, TATJANA. 1979. *Sintaktičeskaja struktura balkanskogo jazykovogo*. Moscow: Nauka.
- CHOLEZAS, IOANNIS, and PANOS TSAKLOGLOU. 2008. The economic impact of immigration in Greece: Taking stock of the existing evidence. Discussion paper no. 3754. Germany: Discussion Papers Series, Institute for the Study of Labor. Online: <http://ftp.iza.org/dp3754.pdf>
- CRAWHALL, NIGEL. 1990. Shona/English data. Unpublished manuscript.

- CHRISTOFERSON, PAUL. 1939. The articles. A study of their theory and use in English. Oxford: Munksgaard, Copenhagen and Miford.
- DISCIULLO, ANNA-MARIA, MUYSKEN, PETER and R. SINGH. 1986. Government and code-mixing. *Journal of Linguistics* 22.1-24.
- ELIASSON, STIG. 1989. English-Maori language contact: Code-switching and the free-morpheme constraint. *Reports from Upsala University Department of Linguistics* 18.1–28.
- ELIASSON, STIG. 1995. Grammatical and lexical switching in Maori-English “grasshopper speech”. *Summer School on Code-switching and Language Contact, Ljouwert/Leeuwarden, The Netherlands, September, 14-17, 1994*. 45–57. Ljouwert/Leeuwarden: Fryske Academy.
- ERASMOS PROGRAMME. Online: <http://www.interel.uoa.gr/socrates-erasmus/index.html>
- FRIEDMAN, VICTOR. 2006. The Balkans as linguistic area. *Elsevier Encyclopedia of Language and Linguistics*. Vol. ed. by Keith Brown. 657–672. Oxford: Elsevier.
- FRIEDMAN, VICTOR. 2004. The typology of Balkan evidentiality and areal linguistics. *Balkan Syntax and Semantics*, ed. by Olga Tomić and Aida Martinovic-Zic, 101–134. Amsterdam: John Benjamins Publishing.
- GALL, SUSAN. 1979. *Language shift: Social determinants of Linguistic change in bilingual Austria*. New York: Academic Press.
- GARDNER-CHLOROS, PENELOPE. 2004. How code-switching mediates politeness: Gender-related speech among London Greek-Cypriots (with Katerina Finnis). *Estudios de Sociolingüística*, vol. 4 (2) ‘Language and gender: An interdisciplinary perspective.’ University of Vigo, Spain. V. Acuña-Ferreira & S. Álvarez-López, (Eds.), University of Vigo, Spain, pp. 505-533.
- GARDNER-CHLOROS, PENELOPE. 1995. Code-switching in community, regional and national repertoires: the myth of the discreteness of linguistic systems. *One Speaker, Two Languages*, ed. By Lesley Milroy and Pieter Muysken, 68 – 90. Cambridge: University Press.
- GOLĄB, ZBIGNIEW. 1964. *Conditionalis typu bakaskiego w językach południowosłowiańskich ze szczególnym uwzględnieniem macedońskiego*. Kraków: Polska Akademia Nauk.
- GUMPERZ, JOHN. 1982. *Discourse strategies*. Cambridge University Press.
- HAUGEN, EINAR. 1950. The analysis of linguistic borrowing. *Language* 26.210–231.
- HAWKINS, JOHN. 1978. *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. London: Croom Helm.
- JACOBSON, ROMAN. 1985. *Verbal art, verbal sign, verbal time*.

- JOSEPH, BRIAN. 1983a. *The synchrony and diachrony of the Balkan infinitive. A study in areal, general, and historical linguistics. (Cambridge studies in linguistics, supplementary series no. 1).* Cambridge: Cambridge University Press.
- JOSEPH, BRIAN. 1983b. Language use in the Balkans: The contributions of historical linguistics. *Anthropological Linguistics* 25.275–287.
- JOSEPH, BRIAN. 1987. A fresh look at the Balkan *Sprachbund*: Some observations on H.W. Shaller's *Die Balkansprachen*. *Mediterranean Language* 3.105–114.
- JOSEPH, BRIAN, and PANAYOTIS PAPPAS. 2002. On some recent views concerning the development of the Greek future system. *Byzantine and Modern Greek Studies* 26.247–273.
- JOSEPH, BRIAN. 2003. The Balkan languages. *International encyclopedia of linguistics, 2nd edition*, ed. by William J. Frawley, 153–155. Oxford: Oxford University Press.
- JOSHI, ARAVIND. 1985. Processing sentences with intra-sentential code switching. *Natural language Parsing*, ed. by D. R. Dowty, L. Karttunen, and A. Zwicky 190–205. New York: Academic Press.
- KAMWANGAMALU, NKONKO. 1989. Some morphosyntactic aspects of French/English-Bantu code-mixing. *Parasession on language contact: Regional meeting of the Chicago Linguistic Society*, ed. by Bradley Music et al, 157–70. Chicago: CLS.
- KLAVANS, J. E. 1983. The syntax of code-switching: Spanish and English. *Selected papers from the 13th Linguistics Symposium on Romance Languages*, ed. by L.D. King and C.A. Matey, 213–231. Chapel Hill, N.C.
- KOPITAR, JERNEJ (BARTOL). 1829. *Albanische, walachische u. bulgarische Sprache*. In: *Jahrbücher der Literatur*. Band 46. Wien, 1829. S. 59–106. Nachdruck in: *Jerneja Kopitarja Spisov*. II. Del, 2. Knjiga. Ljubljana, 1945. S. 227–272
- KÖPPE, REGINA, and JÜRGEN MEISEL. 1995. Code-switching in bilingual first language acquisition. *One speaker, two languages*, ed. by Lesley Milroy and Pieter Muysken, 276–301. Cambridge: Cambridge University Press.
- LABOV, WILLIAM. 1984. Field methods of the project on linguistic change and variation. *Language in use: Readings in sociolinguistics*, ed. by John Baugh and Joel Sherzer, 28–54. Englewood Cliffs, NJ: Prentice-Hall.
- LANCE, DONALD M. 1975. Spanish-English code-switching. *El lenguaje de los chicanos*, ed. by Eduardo Hernandez-Chavez, Andrew D. Cohen, and Anthony F. Beltramo, 138–153. Arlington, VA: Center for Applied Linguistics.
- LINDSTEDT, JOUKO. 2000. Linguistic balkanization: Contact-induced change by mutual reinforcement. *Languages in Contact (Studies in Slavic and General Linguistics, 28)*, ed. by Dicky Gilbers, John Nerbonne, and Jos Schaecken, 231–246. Atlanta, GA: Rodopi.

- LYONS, CHRISTOFER. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- MACSWAN, JEFF. 1999. *A minimalist approach to intrasentential code switching*. New York: Garland Publishing.
- MARKOVA, EUGENIA. 2003. Bulgarian economic migration in Greece. *Bulgarski Glasove* 1.
- MLADENOVA, OLGA. 2007. *Definiteness in Bulgarian. Modeling the process of Language change*. New York: Mouton de Gruyter.
- MUYSKEN, PETER. 2000. *Bilingual speech: A typology of code-mixing*. Cambridge: University Press.
- MUYSKEN, PETER. 1997. Alternation, insertion, congruent lexicalization. *Language Choices: Conditions, Constraints and Consequence*, ed. by Martin Putz, 361-80. Amsterdam Benjamins.
- MYERS-SCOTTON, CAROL. 2002. *Bilingual encounters and grammatical outcomes*. Oxford: Oxford University.
- MYERS-SCOTTON, CAROL, and JANICE JAKE. 2001. Explaining aspects of codeswitching and their implications. *One mind, two languages: Bilingual language processing*, ed. by Janet Nicol, 84–116. Oxford: Blackwell.
- MYERS-SCOTTON, CAROL. 1999. Putting it all together: The matrix language and more. *Language encounters across time and space*, ed. by B. Brenemoen, E. Lanza and E. Ryen. 13–28. Oslo: Novus Press.
- MYERS-SCOTTON, CAROL. 1997. “Matrix language choice” and “Morpheme sorting” as possible structural strategies in pidgin/creole formation. *Pidgins and creoles: Structures and status*, ed. By Spears, Arthur and Winford, 151–174. Amsterdam: Benjamins.
- MYERS-SCOTTON, CAROL. 1993. *Dueling languages: Grammatical structure in code-switching*. Oxford: Clarendon Press.
- MIKLOSICH, FRANZ. 1861. Die slavischen Elemente im Rumunischen. *Denkschriften der Kaiserlichen Akademie der Wissenschaften, Philosophisch-Historisch Klasse, Vol.12*. Wien: K.K. Hof- u. Staatsdruckerei.
- MILROY, LESLEY, and PIETER MUYSKEN. 1995. *One speaker, two languages. Cross-disciplinary perspectives on code-switching*. Cambridge: Cambridge University Press.
- MIRCHEV, KIRIL. 1978. *Istoricheska gramatika na balgarskija ezik*. Sofia: Nauka i izkustvo.
- NASEH-LOTFABBADI, LEYLA. 2002. *Disagreement in agreement: A study of grammatical aspects of codeswitching in Swedish/Persian bilingual speech*. Stockholm: Stockholm University dissertation.

- NISHIMURA, MIWA. 1986. Intrasentential code switching: The case of language assignment. *Bilinguals: Psycholinguistic and Neuropsychological Perspectives*, ed. by J. Vaid, 123–43. Hillsdale, NJ: Erlbaum.
- NORTIER, JACOMINE M. 1990. *Dutch-Moroccan Arabic code-switching among young Moroccans in the Netherlands*. Dordrecht: Foris.
- PASHOV, PETER. 1990. *Blgarska gramatika*. Sofia: Nauka i izkustvo.
- PAFF, CAROL. 1979. Constraints on language mixing: Intra-sentential code-switching and borrowing in Spanish/English. *Language* 55.291–318.
- POPLACK, SHANA. 1978. Dialect acquisition among Puerto Rican bilinguals. *Language in Society* 7.89–103.
- POPLACK, SHANA. 1980. „Sometimes I'll start a sentence in English Y TERMINO EN ESPANOL”: Towards a typology of code-switching. *Linguistics* 18.581–618.
- POPLACK, SHANA; DAVID SANKOFF; and CHRISTOPHER MILLER. 1988a. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26.47–104.
- POPLACK, SHANA. 1988b. Language status and language accommodation along a linguistic border. *Language spread and language policy*, ed. by Peter H. Lowenberg, 90–118. Washington, DC: Georgetown University Press.
- REZAEIAN, FRERESHTEH. 2009. Structural and social aspects of codeswitching among Iranian/Canadian bilinguals. Unpublished Master thesis.
- RUSSEL, BERTRAND. 1905. On Denoting. *Mind* 14.479-493.
- SANDELD, KRISTIAN. 1930. *Linguistique balkanique. Problèmes et résultats*. (Collection linguistique publiée de la Société de Linguistique de Paris 31.) Paris: Librairie ancienne Honoré Champion.
- SANKOFF, DAVID; SHANA POPLACK; and SWATHI VANNIARAJAN. 1986. The case of nonce loan in Tamil. *Technical report 1348*. Montreal, QC: Centre de recherches mathématiques, University of Montreal.
- SANKOFF, DAVID; SHANA POPLACK; and SWATHI VANNIARAJAN. 1990. Dramatically contrasting language mixture strategies in two communities of fluent Arabic-French bilinguals. Paper presented at NWAVE 19, October 1990, University of Pennsylvania.
- SANKOFF, DAVID; SHANA POPLACK; and SWATHI VANNIARAJAN. 1990. The case of the nonce loan in Tamil. *Language Variation and Change* 2.71–101.

- SAVILLE-TROIKE, MURIEL. 1997. The ethnographic analysis of communitive events. *Sociolinguistics: A Reader and coursebook*. ed. Nikolas Coupland and Adam Jaworski. 126–145. Palgrave: New York.
- SCHLEICHER, AUGUST. 1852. *Les langues de l'Europe moderne. Traduit de l'allemand par Hermann Ewerbeck*. Paris: Ladrangé & Garnier frères.
- SHAFFER, DOUGLAS. 1978. The place of code-switching in linguistic contacts. *Aspects of Bilingualism*, ed. by Michel Paradis, 265–274. Columbia, SC: Hornbeam Press, Inc.
- SOLTA, GEORG RENATUS. 1980. *Einführung in die Balkanlinguistik mit besonderer Berücksichtigung des Substrats und des Balkanlateinischen*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- SPYROPOULOS, VASSILIOS, and MARIA-ANNA TILIOPOULOU. 2004. Greek and Turkish in contact in Asia Minor: Definiteness and case in Cappadocian Greek. Paper presented at GLOW'04, Thessaloniki. Online:
<http://web.auth.gr/del/GLOW2004/abstracts/Spyropoulos-Tiliopoulou.pdf>
- TAGLIAMONTE, SALI A. 2006. *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.
- THOMASON, SARAH. 2001. *Language Contact*. Edinburgh: Edinburgh University Press.
- TOMIĆ, OLGA MIŠESKA. 2006. *Balkan sprachbund morpho-syntactic features*. Dordrecht: Springer-Verlag.
- TORIBIO, ALMEIDA. 2004. Convergence as an emergent property in bilingual speech. *Bilingualism: Language and cognition* 7.165-173.
- TREFFERS-DALLER, JEANINE. 1991. *French-Dutch language mixture in Brussels*. Unpublished PhD thesis, University of Amsterdam.
- TROUBETZKOY, NIKOLAI. 1928. Studies in general linguistics and language structure. First International Congress of Linguistics, Actes, 17f. Leiden. Proposal 16.
- VOGT, HANS. 1954. Language contacts. *Word* 10.365–374.
- WEIGAND, GUSTAV. 1928. Texte zur vergleichenden Syntax der Balkan Sprachen. *Balkan-Archiv* 4.53–70.
- WINFORD, DONALD. 2003. *An introduction to contact linguistics*. Oxford: Blackwell Publishing.
- ZABRODSKAJA, ANASTASSIA. 2008. Code-switching as a mechanism of contact-induced language change in Estonia's Russian. *Conference on Cognitive and Functional Perspectives on Dynamic Tendencies in Languages, University of Tartu, Estonia, May 29 to June 1, 2008*.

ZENTELLA, ANNA. 1982. Code-switching and interaction among Puerto Rican children.
Spanish in the United States: Sociolinguistic aspects, ed. By J. Amastae and L. Ellas-Olivares,
354-385. Cambridge: Cambridge University Press.