AZERBAIJANI-RUSSIAN CODE-SWITCHING AND CODE-MIXING: FORM, FUNCTION, AND IDENTITY

by

KENNETH ZUERCHER

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2009

ACKNOWLEDGEMENTS

It would be impossible for me to list everyone who has participated in and contributed to this dissertation. To start off, I would like to thank my wife Lis and our children Reiner, Laurence, and Claire. I appreciate your putting up with me while bringing this project to completion. I would also like to thank my committee: Laurel Stvan, David Silva, Jala Garibova, and Pete Smith. Your assistance and advice has been and will be invaluable. Thank you for your contributions to my present and future success. Next I would like to thank the Azerbaijani University of Languages in Baku for lending me your vice rector's time (Jala Garibova) and sponsoring my visas in 2006 and 2007. Thanks to Aynura Garabayli for your help in checking transcriptions and consulting on numerous points of interpretation. Without you this project would have been impossible. Thanks as well to all those who participated in the study. Your names have been changed in the text that follows, but I know who you are and I appreciate the time you spent recording data and discussing it with me. Lastly, I would like to thank the God of the universe who created such an amazing diversity of human beings and this 'curse' called language through which we live our social lives. Blessed be His Name.

September 24, 2009

ABSTRACT

AZERBAIJANI-RUSSIAN CODE-SWITCHING AND CODE-MIXING:

FORM, FUNCTION, AND IDENTITY

Kenneth Zuercher, PhD.

The University of Texas at Arlington, 2009

Supervising Professor: Laurel S. Stvan

From incorporation into the Russian Empire in 1828, through the collapse of the U.S.S.R. in 1991 governmental language policies and other socio/political forces influenced the Turkic population of the Republic of Azerbaijan to speak Russian. Even with changes since independence Russian use – including various kinds of code-switching and code-mixing – continues. This dissertation studies the language situation in Azerbaijan through a detailed analysis of naturally occurring conversational data. Approaches include corpus analysis of the transcribed data to show relative amounts of Azerbaijani and Russian, linguistic description of the types of code-switching and code-mixing, quantitative analysis of variation between subjects, and sequential analysis of a

few subjects to demonstrate ways in which code-switching/mixing can be used to construct social identities in contemporary Azerbaijan.

Subjects' use of Russian content words varied from 11.2% to 97.2%. While some conversational turns contained only Russian, code-switching/mixing within turns and clauses was common, with nominal insertion and peripheral alternation of adverbial elements occurring most frequently. Congruent lexicalization (Muysken 2000) also occurs in stative clauses with the data showing evidence for a zero copula in Azerbaijani as well as Russian. Russian and code-switching/mixing can be used to construct a range of social identities. The case studies in this dissertation show subjects avoiding Russian use to conform to social norms in some family domains and professional contexts, using substantial Russian and Russian code-mixing in private domains when appropriate for the situation and interlocutor, as well as using Russian to contest traditional gender roles and portray themselves as 'modern' and free of stereotypes.

The results of this analysis do not contradict recent theoretical and descriptive work on code-switching/mixing (Muysken 2000, Myers-Scotton 2002) but confirm their propositions with a new language pair. They also open the door to further research into language behavior in the former Soviet space by providing a data oriented description of language behavior and linguistic identity construction in Azerbaijan. While governmental language policy and planning firmly support the development and use of Azerbaijani, Russian use persists in some sectors of society.

The research for this dissertation was conducted under UT Arlington IRB protocol #06.95s following a pilot project conducted under protocol #05.276.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure Page

LIST OF TABLES

CHAPTER 1

AZERBAIJAN: SOCIO-HISTORICAL AND LINGUISTIC BACKGROUND[1]

The country that is now the Republic of Azerbaijan has been subject to a number of linguistic and political influences throughout history: Turkic, Persian, Arabic, Russian, and now (since the breakup of the U.S.S.R. in 1991) English. In order to give a fuller picture of the language situation in Azerbaijan, this dissertation analyzes the ways in which Azerbaijani and Russian interact within conversation. This chapter provides an overview of the historical processes which have influenced language in Azerbaijan. Chapter 2 summarizes the relevant literature for the present study, while Chapter 3 provides descriptions of both the Azerbaijani and Russian linguistic systems. The remainder of the study focuses on a detailed analysis of recorded conversations. This includes a description of the code-switching and code-mixing in Azerbaijan according to Muysken's typology (2000), a quantitative comparison of the language used by the recorded subjects, and qualitative case studies focusing on the ways that a few of the subjects use language to construct identity.

1.1 Historical context

Azerbaijan is a nation about the size of the U.S. state of Maine, located on the western shore of the Caspian Sea north of Iran and south of Russia. It borders Georgia and Armenia, as well as Turkey via the separate region of Nakhchivan. According to

[1] Much of the information in this chapter was adapted from Zuercher 2004: 1-14 and was also utilized for Zuercher 2009.

2001 governmental statistics, the total population of the country is slightly more than 8.1 million (Veliyev 2002: 13).

Figure 1-1 The Republic of Azerbaijan in its immediate geographic context

Although somewhat small in size and population, the history of Azerbaijan has produced a language situation that is worthy of study. With an ancient literary tradition of its own, Azerbaijan has been strongly influenced by Russian since its conquest by the Russian Empire in the early 19th century. It has been impacted by numerous empires and states, ancient and modern, and by the migration of various ethnic groups, particularly Turkic groups who provided the foundation for the modern Azerbaijani language. Azerbaijan was invaded by Cyrus the Great in the 6th century BC, and later by Alexander

the Great in the 4th century BC. The Romans appeared in the area in the 1st century BC and left what may be the easternmost Roman inscription near the capital city, Baku.

By the 4th century BC two states had emerged in the area: Caucasian Albania in the territory now occupied by the Republic of Azerbaijan, and Aturpatkan in what is now the state of Azerbaijan in northwestern Iran. Albania became Christian in the 4th century AD, but learned Islam from Arabic invaders in the 7th century. This state survived into the 9th century as a vassal state of the Persian Empire, and was succeeded by the state of Shirvan. Four dynasties of Shirvan Shahs ruled from the 9th century to around 1530. The Shirvan Palace in Baku is considered a national landmark. Shi'a Islam was adopted by the rulers of Shirvan in the 16th century and is the most common religion in Azerbaijan today.

The influx of Turks into this area occurred in various waves between the 4th and 11th centuries, but there is debate as to the date of their first arrival, and thus the ethnicity of the Caucasian Albanians or the inhabitants Shirvan.

> Ashurbeyli stated that "from antiquity" the Shirvan region had been a place where Caucasian-, Iranian-, and Turkish-speaking tribes mingled and argued that "in the 6th century intensive migrations of Turks in Aran, Shirvan, and Mughan occurred" (Altstadt 1992:5).

An alternative view is that there was a direct ethnic link between Albania and Turks. Whichever interpretation is taken, it is indisputable that complete Turkization had taken place by modern times.

By the end of the 17th century, the area was fully under Persian influence and part of the Persian Empire until the early 1800s, when it came under the Russian rule as a result of the Russo-Iranian wars (1804-1813 and 1826-1828). In 1828 the treaty of Turkmanchai set the border between the two empires as the Araxes River, leaving northern Azerbaijan under Russian control and southern Azerbaijan a part of the Persian Empire.

Since independence from the U.S.S.R. in 1991, modern nation building efforts have caused Azerbaijanis to look to the ancient indigenous states as ancestors of the current Republic. In particular, references to the Caucasian Albanians as distinct from Armenians are commonly used to refute claims by neighboring Armenia that the Nagorno Karabakh area was historically Armenian.

1.2 Immigration

With Russian political domination in the 19th century came large numbers of Russian and other non-Muslim immigrants. While some Russian peasants immigrated to the countryside as agricultural workers, the vast majority came as administrative workers and industrialists to the larger cities, especially Baku.

For Azerbaijan, the late 19th century was characterized by industrialization with Baku becoming the center of oil production for the Russian Empire. This brought many Russians and Armenians (also Russian speakers) to Baku as workers and administrators. Thus, Baku developed a culture quite apart from the rest of Azerbaijan. Given this combination of development and immigration, marriage between ethnic Azerbaijanis and other nationalities was fairly common, especially under the U.S.S.R.

Baku has continued to be more oriented toward the Russian language than the rest of the country. According to de Waal, probably based on observations from the late 1990s, "The city's lingua franca is still just as much Russian as Azeri, spoken with a gentle southern lilt that rises in intonation at the end of a sentence" (2003: 98). However, Russian is rarely heard in public any longer, though a high percentage of the population know it and may speak it in certain private contexts.

1.3 Language policy, planning, and development in Azerbaijan

According to Haugen, language planning is "the activity of preparing a normative orthography, grammar, and dictionary for the guidance of writers and speakers in a non-homogeneous speech community… Planning implies an attempt to guide the development of a language in the direction desired by the planners" (1959: 8). Based on a range of research, Hornberger identifies three types of language planning:

> Status planning – "those efforts directed toward the allocation of functions of languages/literacies in a given speech community"
>
> Acquisition planning – "efforts to introduce the allocation of users or the distribution of languages/literacies, by means of creating or improving opportunity or incentive to learn them, or both."
>
> Corpus planning – "those efforts related to the adequacy of the form or structure of languages/literacies" (2006: 28)

Moreover, each of these three types of planning can be approached from one of two perspectives (Hornberger 2006: 29): a Policy Planning perspective (focus on form –

Haugen 1983: 275), or from a Cultivation Planning perspective (focus on function – Haugen 1983: 275).

Examples of all three types of language planning abound through Azerbaijani history. Status planning (overt and covert) through both the Russian Empire and the Soviet era pushed Russian into 'high' domains while often attempting to keep Azerbaijani confined to 'low' domains such as home and bazaar. Acquisition planning most often took the form of schools where either Azerbaijani or Russian was the language of instruction. Though Russian was never the object of corpus planning in Azerbaijan, Azerbaijani has undergone constant development both through private and governmental efforts: the production of dictionaries, conscious development of vocabulary, publication of literature, etc.

1.3.1 The Russian Empire: 1828-1918

Under the Russian Empire, from the Treaty of Turkmanchai in 1828 to a short lived independence in 1918, Azerbaijan functioned as a colonial territory (the colonial model of integration - Laitin 1998: 66 ff). As the language of government and of much commerce, Russian experienced an increase in status. Though there were a few Russian language schools, they were not accessible to the vast majority of the Azerbaijani speaking population. In this period, Azerbaijani was not the object of either governmental status or acquisition planning. However, through the efforts of a cadre of educated intelligentsia with the backing of a number of Azerbaijani industrial philanthropists, the language did undergo significant development and regularization – corpus planning. "The period was characterized by a rediscovery of history, literature,

and philosophy and by a debate about politics, social change, religion and morality, and historical identity" (Altstadt 1992: 50-51). During this period a number of classic plays were written in Azerbaijani, and newspapers flourished. A few Azerbaijani schools were also established.

Table 1-1 Language planning summary – Russian Empire

	Russian	Azerbaijani
Status Planning	+	-
Acquisition Planning	limited	limited
Corpus Planning	n/a	+

Given the prevalence of Russian in administration and Azerbaijani in most 'low' domains, this period was, in all likelihood, characterized by diglossia without bilingualism (Fishman 1972). Russian speakers had little motivation to learn Azerbaijani, while the vast majority of the population did not have the means to acquire Russian. Thus the two populations remained separate for the most part with the exception of Azerbaijani political elites who acted as go-betweens for the two groups.

1.3.2 The U.S.S.R.: 1920 through 1991

The Soviet era was characterized by huge shifts in language policy and planning from encouraging the development of local languages to crushing centrism:

> the establishment of Soviet language politics in the 1920s meant a disruption of earlier discrimination against non-Russian languages during czarist… and 'bourgeois' Russian rule… It also implied the facilitation of the social functioning of all languages in the newly-founded Soviet Union (Haarmann 1995: 7).

It was not until the Stalin era that Russian use was emphasized and proponents of linguistic nationalism were liquidated.

Table 1-2 Language planning summary – U.S.S.R.

	Russian	Azerbaijani
Status Planning	+	gradual increase
Acquisition Planning	+	gradual increase
Corpus Planning	n/a	+

Much of the language politics of the U.S.S.R. was, however, covert in nature. According to Garibova:

> Although claiming to be the most democratic "voluntary" union of various nations and ethnicities, the Soviet government had, in its language policy, a well hidden agenda of promoting (often implicitly) the Russian culture and language through the idea of the "common Soviet culture" often metonymizing the concept of the "common Soviet nationality." (2009: 11)

According to Haarmann this "covert language-spread policy" (1992) took many forms throughout the Soviet era. These included three groups of factors:

Socio-demographic factors (1992: 116-117)

- Significant Russian populations in other Republics, particularly urban centers where they were usually skilled workers and specialists
- Non-Russian migration to Russian urban centers as part of the Soviet workforce
- Non-Russians educated at Russian universities, or at universities outside their home territories where Russian was the language of instruction

- Intermarriage of Russians and non-Russians both in non-Russian republics and urban centers in Russia

Status criteria (1992: 118-121)

- Dominance of Russian in Soviet public life
- Shift from Latin script to Cyrillic
- Russian only language with official status in all Soviet territories
- Russian dominance in publishing
- Russian dominance in translations of literature

Functional criteria (1992: 122)

- "voluntary" preference for Russian as a vehicle of intercommunication
- Russian as the all-Union language
- Russian associated with 'internationalization,' including lexical modernization of non-Russian languages by using Russian loan words

As in other Union republics, all of these factors affected the language situation in Azerbaijan.

As a result of both the officially and unofficially sanctioned development of Azerbaijani and these Russian language influences, the Soviet era probably produced a situation of diglossia with bilingualism (Fishman 1972) for many citizens of Azerbaijan. Russian was the expected language in most official/political domains, virtually all technical education, and other 'high' domains. Azerbaijani, on the other hand, was the home language for most Azerbaijanis and predominated in 'low' domains such as the bazaar.

Unlike some titular languages (languages for whom the Soviet Republics were titled), Azerbaijani did, however, experience a gradual increase in status and acquisition planning throughout the Soviet era. Though official rhetoric at the founding of the U.S.S.R. was favorable toward minority languages, Russian was the official language of the Azerbaijani S.S.R. in the 1936 constitution. By the efforts of Mirza Ibrahimov, this was amended in 1956 to make Azerbaijani the official language, though "de-facto implementation was not easy at that time, and Ibrahimov was soon removed from his position and labeled a 'nationalist'" (Garibova 2009: 14). Further success was achieved with the constitution of 1978 when the government was also charged with ensuring the implementation of Azerbaijani in state and public organizations as well as education. It was also charged with supporting its development. Throughout the Soviet era the availability of Azerbaijani language education mirrored this increase in official status.

Finally, the Soviet era ended with what Laitin (1998) describes as the "double cataclysm" for the Russian speaking populations of the non-Russian republics of the U.S.S.R. In 1989 language laws were passed in many of the republics which "threatened their security in speaking only the Soviet language of 'internationality communication'" (Laitin 1998: 85). Then in 1991, the Union itself dissolved leaving many Russian speakers culturally and linguistically stranded in non-Russian speaking independent countries.

1.3.3 The Republic of Azerbaijan: 1991 – present

Since independence from the U.S.S.R. in 1991, language policy and language practice in Azerbaijan have varied significantly. The first president of independent

Azerbaijan, Ayaz Mutalibov, did little to upset the linguistic status quo, allowing individuals and organizations to use whichever language they wished. He believed that "government policies ought to encourage everybody to master Azeri, [but] no steps should be taken to discriminate against those who do not know this language" (Landau & Kellner-Henkele 2001: 78). Thus, Russian was allowed to maintain its de facto status as the 'high' language. When Elbufez Elchibey was elected president in 1992 he made sweeping changes in government language policy. He insisted that Azerbaijani should be "employed in all spheres of politics, economics, society, scientific and cultural life, and should be taught to all other nationalities in Azerbayjan" (Landau & Kellner-Henkele 2001: 115). Thus, he sought to demote Russian to the same status as other "foreign" languages such as English, French, or German. When Heydar Aliyev came to power in 1993, he focused on national security and stability in response to the occupation of Azerbaijani territory by Armenian and Karabakh Armenian forces. Language policy issues were put on hold. Toward the end of his long presidency, however, his policies shifted toward Azerbaijani linguistic nationalism. Though a Latin version of the Azerbaijani alphabet was created shortly after independence in 1991 its use was not mandated until President Aliyev's 2001 declaration on language use (Dövlət Dilinin Təkmilləşdirilməsi 2001). This same declaration called for the reduction of foreign languages in advertising and increased use of Azerbaijani in broadcast entertainment. The language law that resulted from this declaration banned foreign languages from governmental proceedings and required that all radio and television programs in

Azerbaijan be broadcast in Azerbaijani. This last provision was later removed, however, and Russian language broadcasting continues to this day.

Table 1-3 Language planning summary – Republic of Azerbaijan

	Russian	Azerbaijani
Status Planning	-	+
Acquisition Planning	decreasing	+
Corpus Planning	n/a	+

Though change has been slower than some would have liked, Azerbaijani is clearly the dominant language in Azerbaijan today. While Russian is probably the most common 'foreign' language, it is facing significant competition from English. Today, Russian language education is available, but more and more parents are choosing to put their school children in Azerbaijani language sectors as these are becoming better equipped and Russian is no longer seen as a necessary prerequisite for future success.

1.4 Changes in script

One of the outcomes of the changes in language policy/planning priorities through these eras of Azerbaijani history and the dramatic shifts in national identity has been numerous changes in alphabet for the Azerbaijani language (see Hatcher 2009). Until the early 20th century, Azerbaijani was written with Arabic characters, like Persian. Although a Latin alphabet was proposed by some intellectuals in the 19th century, and was anticipated during the short period of independence from 1918-1920, it was not adopted until 1924. While this change was seen as a step toward modernization, the Bolshevik government also saw alphabet change as a means of separating northern from southern Azerbaijan (part of Persia), and cutting off the Turkic peoples of the U.S.S.R.,

including Azerbaijanis, from the now anti-Bolshevik Turkish Republic. It also prevented the spread of pre-revolutionary religious and "bourgeois" sentiments.

However, when the Turkish Republic and the Iranian Tajiks also adopted Latin alphabets, the Soviet government saw this as a source for anti-Bolshevik tracts. Accordingly, shortly before 1940 Azerbaijan and the other Turkic republics of the U.S.S.R. were encouraged to adopt versions of the Cyrillic alphabet (Altstadt 1992: 124). "During a relatively short time, different versions of the Cyrillic alphabet were introduced for writing the various languages, not only to fit the peculiarities of each but also probably to isolate related linguistic groups from one another" (Shahrani quoted in Landau & Kellner-Henkele 2001: 54). It was believed that the new alphabets would facilitate the learning of Russian. Since the old Azerbaijani intelligentsia had been destroyed by Stalinist purges, and previously published literature had been mostly destroyed or banned, when Azerbaijan became independent in 1991 there was virtually no one in the country who knew how to read Azerbaijani in a Latin alphabet. The collective memory of Azerbaijan had been captured in Cyrillic characters.

Shortly after independence from the U.S.S.R. in 1991, the parliament of Azerbaijan adopted a somewhat simplified version of the earlier Latin alphabet, presumably because of its association with the earlier independent Azerbaijani republic and in order to distance Azerbaijani from Russian and Soviet influence. This law also specified how the new alphabet's use should be phased in over the next few years. However, with the exception of elementary school books (Landau & Kellner-Henkele 2001: 133), little progress was made. Books, newspapers, and magazines continued to be

published in the Cyrillic alphabet until after the change of the millennium. This included high profile publications such as the *Azerbaijani-English Dictionary* put out by the Azerbaijan State Institute of Languages in 1996 (Musayev 1996) , and the *Explanatory Dictionary of the Azerbaijani Language* from the Academy of Sciences in 1997 (Akhundov 1997). It was not until President Aliyev's 2001 declaration that use of the Latin alphabet was made mandatory. The Azerbaijani Cyrillic script is now absent from public view in Azerbaijan, though it lives on in books published in the 1990s and before. All current books and newspapers are published in the Azerbaijani Latin script.

1.5 Ethnic conflict

While not central to this study, it would be misleading to discuss the history of Azerbaijan without touching upon the subject of ethnic conflict. The area which comprises present day Azerbaijan has been inhabited by various ethnic groups as far back as history is recorded, including Turkic peoples since shortly after the time of Christ (Altstadt 1992). Until the 20th century, however, conflicts in the area rarely took place along ethnic lines. Rather, conflicts were manifestations of empires or smaller states vying for control (see Swietochowski 1995). The beginning of the 20th century saw the growth of the concept of the nation state in the Caucus with both Armenian and Azerbaijani groups pushing for rationalized states of their own. In 1905 the Tartar-Armenian war claimed hundreds of lives. March 1918 saw the massacre of thousands of Azerbaijanis in Baku by predominantly Armenian forces, and in September of the same year Azerbaijanis, aided by Turkish troops, retaliated by killing Armenians. With the largely Armenian enclave of Nagorno Karabakh part of Azerbaijan, the foundation for

further conflict at the end of the century was laid (see de Waal 2003). With rising nationalism in the 1980s, thousands of Azerbaijanis were forced to leave Armenia, as were Armenians Azerbaijan. Open war became a reality with the disintegration the U.S.S.R. When a cease fire was declared in 1993, up to a million Azerbaijanis were either refugees from Armenia or Internally Displaced People from the 20% of Azerbaijani territory (including Nagorno Karabakh) still occupied by Armenia. Over 15 years later, many of these people have not integrated into other areas of Azerbaijani life but live in refugee camps, converted schools, and other forms of government housing, hoping to one day move back to their homes. The occupation of Azerbaijani territory by ethnic Armenian troops (whether under the auspices of the self-proclaimed Nagorno Karabakh Republic or Armenia proper) is still a very common topic in public and political discourse. Though there are occasional high level government meetings, neither side seems willing to compromise on any major issues, and a resolution to this 'frozen conflict' has yet to be brokered. Amazingly, no subjects in this study pointed to the war with Armenia or to Russian sponsorship of Armenia as a reason to avoid speaking or learning Russian.

1.6 Overview of the following chapters

Given the complex and changing language situation which has resulted from these socio/historical influences, this dissertation project uses a detailed analysis of spoken conversational data to provide a description of how Azerbaijani and Russian interact in conversation – code-switching and code-mixing. Chapter 2 contains an overview of the relevant literature. Chapter 3 gives descriptions of both the Azerbaijani

and Russian linguistic systems. Chapters 4, 5, 6, and 7 analyze the transcribed data from various perspectives, and Chapter 8 gives both a summary of the most important findings and provides directions for future research.

CHAPTER 2

REVIEW OF LITERATURE

Having attracted attention since the first half of the 20th century (see Haugen 1950), the literature relevant to the use of two languages within one conversation is vast. The aim of this chapter is to provide a brief overview of a few studies and perspectives that inform this research project, while providing sufficient background information to contextualize it.

2.1 Terminology and definitions: Code-switching versus code-mixing

In the linguistic literature a wide variety of language contact phenomena fall under the name 'code-switching.' These include language alternation at the societal level (diglossia - Ferguson 1959, Fishman 1972), the insertion of individual words or morphemes from one language into a clause or phrase from another language, and even the influence of one grammatical system upon clauses from another language.

Einar Haugen gave four arguments against using the term 'mixing' when discussing these phenomena (1950: 210-211):

1. It implies the creation of a new "concoction."
2. It implies the disappearance of the two previous entities.
3. It implies a haphazard combination.
4. It introduces inaccurate distinctions between 'mixed' and 'pure' languages.

Rather, Haugen claims that "in technical discussion it is more usefully replaced by the term 'borrowing'" (1950: 211) though he also uses the terms 'switch' and 'alternation' in describing bilingual language behavior.

Siding with Haugen in arguing against 'mixing' as an appropriate term, Myers-Scotton chooses the term 'codeswitching'. "[W]hen referring to the use of two languages in the same clause, I use the term 'codeswitching', not 'mixing'. Like Haugen (1950: 210) I think labeling such a phenomenon as 'mixing' has distinct disadvantages" (2002: 3).

Maschler (1998) and Serra (1998), however, argue for the term 'mixed code' in some cases. Citing a number of studies involving switching at the discourse marker level, Maschler argues that if these markers take on new meaning or expanded functions in a bilingual context, the term codeswitching is no longer appropriate. Rather, the speakers are now using a 'mixed code' apart from either of the two donor languages.

With the wide range of phenomena that result from bilingual/multilingual conversation, Muysken uses two separate terms: code-switching and code-mixing.

> I am using the term code-mixing to refer to all cases where lexical items and grammatical features from two languages appear in one sentence. The more commonly used term code-switching will be reserved for the rapid succession of several languages in a single speech event… (Muysken 2000: 1)

While acknowledging the concerns expressed by Haugen and Myers-Scotton, this dissertation will utilize Muysken's terminology. It will, however, use the clause (CP) as

the dividing line between the two phenomena since a 'sentence' can contain a number of clauses:

Code-switching - Change of language or alternation between languages either between conversational turns, or between clauses (CPs) within a single conversational turn.

Code-mixing - The presence of elements from more than one language within a clause (CP).

While it is possible that these mixed clauses represent a new "concoction," such as Maschler's mixed-code, the bilinguals who produce the utterances rarely confuse the two languages. Neither does the switching or mixing appear to be random or haphazard. Instead code-switching and code-mixing represent a range of linguistic resources available to bilingual speakers that are unavailable to monolinguals, and that can be used toward a wide range of sociolinguistic ends.

2.2 Language and identity

From its inception in the 1960s and 1970s the sub-discipline of sociolinguistics has been concerned with the ways that individuals use language to construct/project social identities and how society uses language behavior to identify individuals as belonging to a variety of social categories. In one of the earliest studies Labov (reprinted in 1972b: 43-69) found that lower class subjects in New York City omitted /r/ in a variety of phonological contexts more often than middle or upper class subjects. He also found that r-deletion was not categorical but that subjects had the ability to use /r/ in the same phonological contexts in formal speech settings or in emphasized speech. In his study of

language in Norwich, England, Trudgill (1974) found that lower class male subjects used high proportions of socially disapproved linguistic variants even though they claimed that they would rather avoid these forms. Trudgill proposed that these subjects accrued 'covert prestige' by using these forms to build identities of lower class masculinity. While all sociolinguistic research projects are concerned to some extent with identity construction, they cover a vast range of topics and use a large number of research methodologies.

McKay and Hornberger provide a useful way to conceptualize these various aspects of sociolinguistics. They arrange all the different methodologies/foci along two continua: level of linguistic analysis (micro to macro) and levels of social analysis (also micro to macro). See Table 2-1.

Table 2-1 Sociolinguistic methodologies according to macro and micro levels of linguistic and social analysis (based on McKay & Hornberger 1996: x)

Levels of Linguistic Analysis	Levels of Social Analysis: Macro	Levels of Social Analysis: Micro
Macro	*Language and Society* Language and attitudes, motivation, and standard Societal multilingualism World Englishes Language planning and policy	*Language and culture* Ethnography of communication Speech acts Literacy and literacies
Micro	*Language and variation* Regional and social variation Pidgins and creoles Language and gender	*Language and interaction* Ethnography microanalysis Interactional sociolinguistics Intercultural communication

Since they focus on micro level linguistic phenomena (the presence or absence of /r/, or the substitution of a glottal stop for /t/) and seek to correlate these with major demographic categories such as lower and middle class, the studies from Labov and Trudgill mentioned above would fall in the Micro Linguistic/Macro Social quadrant. Similarly, many studies of language and gender would fall in this same quadrant since they often look at issues such as word choice or phonological variation to construct identities recognized by the society at large. Studies of diglossia (Ferguson 1959, Fishman 1972) and language choice would fall into the Macro/Macro quadrant. With their focus on a small number of subjects but aiming to generalize principles to a general class of speech activity, most studies of politeness and speech acts fall into the Macro Linguistic/Micro Social quadrant. Finally with a focus on micro language choices made by a small group of subjects, most studies oriented toward Conversation Analysis (Sacks et al. 1974) would fall into the last quadrant. We could imagine studies involving bilingualism, language contact, and therefore code-switching/mixing which could fall into all four quadrants (see section 2.6.4).

Especially in regard to identity construction involving multiple languages, we must remember that a command of the various linguistic repertoires is a pre-requisite. According to According to LePage:

> The individual creates his systems of verbal behavior so as to resemble those common to the group or groups with which he wishes from time to time to be identified, to the extent that
>
> a) he is able to identify these groups

b) his motives are sufficiently clear-cut and powerful

c) his opportunities for learning are adequate

d) his ability to learn - that is, to change his habits where necessary - is unimpaired (1968: 192).

For individuals to use code-switching/mixing in identity construction they must not only have (at least a minimal) command of the languages involved, but also understand the norms for code-switching/mixing within the particular speech community of which they wish to appear members.

To understand code-switching/mixing as a social phenomenon, it is crucial to determine the reasons participants use it. Researchers in the tradition of social constructivism analyze how individuals access socially constructed Discourses (Gee 1999) in innovative or conventional ways through their spoken texts (discourses with a small 'd') to construct social identities of their own on an ongoing, dynamic basis. According to Talbot, Discourses are "structures of possibility and constraint" (1998: 151). Ochs claims that, to explain the relationship between language use and identity, we must understand "(1) how particular linguistic forms can be used to perform particular pragmatic work… and (2) norms, preferences, and expectations regarding the distribution of this work vis-à-vis social identities of speakers, referents, and addressees" (1992: 342). Thus, to understand the role code-switching plays in identity construction, we must determine its linguistic forms and the pragmatic functions that these forms serve in a given language situation, then understand how the society interprets them – the socially constructed Discourses that they invoke.

There are numerous examples of code-switching and code-mixing being used for identity construction. Early studies demonstrated how cs/cm was used to identify with a particular social group. Blom and Gumperz (1972) analyzed the use of two dialects of Norwegian in the town of Hemnesberget, Norway: Ranamål, and Bokmål. They found that speakers used Ranamål to show solidarity with the "local team" (1972: 419), while Bokmål was used as a prestige dialect associated with national issues, governmental office, and higher education. Finding similar motivations, Poplack (1980) observed her Puerto Rican subjects living in New York to use Spanish-English code-switching/mixing "in a way which minimized the salience of the switch points, and where the switches formed part of an overall discourse strategy to use both languages, rather than to achieve any specific local discursive effects…" (Poplack 1988: 230). Instead of using Spanish/English code-switching to communicate meaning in individual contexts, these speakers used an overall pattern of code-switching and mixing to construct identities as New York Puerto Ricans in opposition to Puerto Ricans in Puerto Rico, or native English speaking New Yorkers. Along the same lines, Nivens found that his subjects sometimes peppered their West Tarangan discourse with Maylay words which seemed to "represent an attempt by a speaker to index a Malay social identity or affect a Malayish style" (2002: 107). Other studies have shown that code-switching/mixing can be used in gender identity construction. Kulick, in his study of a small village in the Sepik region of Papua New Guinea, found that the local vernacular was associated with the cultural concept of *hed* - individualism, femininity, childhood, and badness, while Tok Pisin was associated with *save* - collectivism, masculinity, adulthood, and goodness (1992: 20). By mixing

and switching between these two languages, speakers could index a variety of identities (Stroud 1998). Similarly, Zuercher (2004, 2009) found that some social groups in Azerbaijan associate Russian use (most likely in Azerbaijani/Russian code-switching) with femininity and pure Azerbaijani with masculinity.

2.3 Theoretical approaches to code-switching and code-mixing

2.3.1 Constraints

One early theoretical approach to code-mixing involved the identification of syntactic constraints on the blending of two (or more) grammars. At the root of these constraints was the observation that code-mixing is not a random combining of languages, but rather that a small number of consistent patterns recur across numerous language contact situations. Therefore, code-mixing must be constrained or controlled in some way by Universal Grammar. While various constraints were proposed, the two constraint model proposed by Poplack (1980) and Sankoff & Poplack (1981) was quite influential.

According to this model, two constraints control the surface representation of code-switched utterances. These are:

a) The free morpheme constraint: a switch may not occur between a bound morpheme and a lexical form unless the latter has been phonologically integrated into the language of the bound morpheme.

b) The equivalence constraint: the order of sentence constituents immediately adjacent to and on both sides of the switch point must be grammatical with respect to both languages involved simultaneously. This requires some

> specification: the local co-grammaticality or equivalence of the two languages in the vicinity of the switch holds as long as the order of any two sentence elements, one before and one after the switch point, is not excluded in either language (Sankoff & Poplack 1981: 5-6).

While the Free Morpheme constraint holds across numerous language pairs, there are counter-examples (Bokamba 1989, Myers-Scotton 1993). The Equivalence constraint has also received challenges (Nartey 1982, Myers-Scotton 1993), and been defended, in part, by claiming that the counter-examples are not true examples of code-mixing (Poplack 1988, Poplack & Meechan 1995).

The data for this dissertation also have counter examples to both constraints:

Sevil #431	Республикансk	-ий	-də
	Republican	-m.s	-loc
	[rɛspublikansk	iy]	
	'At the Republic (hospital)'		

In this phrase the Azerbaijani locative post-position (a clitic) is affixed to the Russian adjective *Республиканский* [rɛspublikanskiy] 'Republican'. This is a violation of the equivalence constraint since Russian utilizes prepositions instead of the Azerbaijani post-positions. It is also a violation of the free morpheme constraint since the Russian adjective is not phonologically incorporated into the langauge of the bound locative morpheme: no native Azerbaijani words contain /nsk/, /ns/ or /sk/ as consonant clusters[2].

At issue with the constraints, and their purported counter-examples is the definition of code-switching/mixing. Poplack claims that code-switching is often

[2] Some surnames do contain these clusters, but they represent Russian/Russified naming conventions similar to the -ev/-eva/-ov/-ova name suffixes still common in Azerbaijan.

confused with other phenomena such as word borrowing – community wide, or idiosyncratic (1988: 216) – while Myers-Scotton (1993, 1995, 2002) believes that lexical borrowing and code-switching are very closely related phenomena. In this context, Nivens (2002) argues that a study of code-switching must include a lexical analysis of the data to determine individual and community wide patterns of borrowing, lexical choice, and co-locational influences upon word choice – a "lexical network analysis" of the data (2002: 215) – before counter examples to constraints can be reasonably considered.

Though the proposed constraints on code-mixing work in specific language contexts, interest in this way of explaining code-mixing has declined with time because the proposed constraints tended to be ad hoc. They were proposed to deal with specific code-mixing issues in a set of data, rather than growing from a larger, more inclusive theoretical framework. Therefore, they may or may not apply to data from a different set of languages. This dissertation will not make reference to a constraints model of code-mixing to analyze and explain its data.

2.3.2 The Matrix Language Frame model

One of the most influential theoretical frameworks for studying code-switching and code-mixing is Carol Myers-Scotton's Matrix Language Frame model (MLF). According to Myers-Scotton "…in 1988, I discovered that others were studying the grammar of codeswitching, but I found their constraints did not fit my data" (2002: xiii). She went on to develop her model in *Dueling Languages* (1993) and then further refined it in *Contact Linguistics* (2002). Her aim was to develop a model of code-switching

based on language production research (Levelt 1989) and Generative Syntax (Chomsky 1965, etc.).

At the heart of the MLF model are two basic oppositions: the Matrix Language (ML)/Embedded Language (EL) opposition, and the Content Morpheme/System Morpheme opposition. According to these:

1. The participating languages in code-switching do not contribute equally. Only one (the Matrix Language) contributes the overall structure of the clause.
2. Content morphemes are the main elements conveying semantic and pragmatic aspects of messages and system morphemes largely indicate relations between the content morphemes. Any language can contribute content morphemes, but only the Matrix Language can contribute system morphemes (Myers-Scotton 2002: 15).

This results in 3 different kinds of data:

1. ML+EL constituents – "a singly occurring EL lexeme in a frame of any number of ML morphemes" (Myers-Scotton 1993: 77).
2. ML Islands – "constituents consisting entirely of ML morphemes" (Myers-Scotton 1993: 78).
3. EL Islands – "They must also be well-formed constituents, but according to the EL grammar; they also must show internal structural dependency relations" (Myers-Scotton 1993: 78)

Thus the example mentioned in the previous section would be an ML+EL constituent:

Sevil #431	Республиканск	-ий	-də
	Republican	-m.s	-loc
	[rɛspublikansk	iy]	
	'At the Republic (hospital)'		

The Russian adjective is an EL content morpheme placed in an Azerbaijani ML frame shown by the system morpheme '–də'. In the following example, the English NP 'constructive criticism' is an EL Island in a CP with Azerbaijani as the ML:

Aygun #15	sən	'constructive criticism'	et-mə-din
	2.sg		do neg-past-2.sg.inf
	Literal –	'You did not do constructive criticism.'	
	Idiomatic –	'That criticism was not constructive!'	

While the difference between content and system morphemes seems intuitive and is easy to determine in the vast majority of cases, as originally conceived it lacked theoretical rigor. Myers-Scotton further defines this opposition with the 4-M Model which divides all morphemes into four categories based on three binary features (2002: 73):

[+/- conceptually activated]

[+/- thematic role receiver/assigner]

[+/- looks outside its immediate maximal projection for information about its form]

Figure 2-1 shows how these features produce the four classes of morphemes and Table 2-2 lists examples of each.

According to Myers-Scotton:

> [W]hen an element is [+conceptually activated] it is salient as soon as a speaker's intentions are encoded as language, at the lemma level in the mental lexicon.

Conceptual activation also means that such elements have semantic content – what speakers hone in on to convey their intentions (2002: 76)

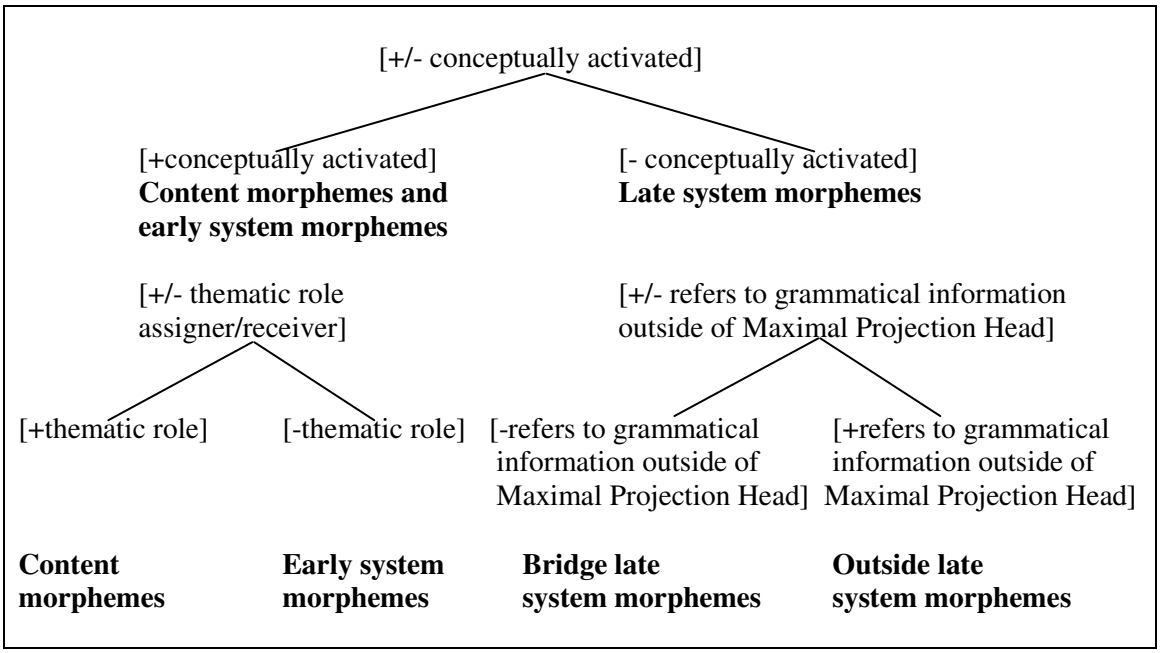


Figure 2-1 Feature-based classification of morphemes in the 4-M model (Myers-Scotton 2002: 73)

According to recent versions of Generative syntax, all nouns must receive a theta role before they can appear overtly in a clause. Most other morphemes do not receive theta roles. See section 3.1 for further explanation of theta role assignment. Bridge System Morphemes "integrate morphemes into larger constituents" (Myers-Scotton 2002: 78). They must be activated by their immediate maximal projection, and show hierarchical relationships. Along with the 4-M Model, Myers-Scotton further clarifies the Content/System morpheme opposition: "The principle does not state that *all* system morphemes must come from only one participating language (the Matrix Language); it states that only those that 'have grammatical relations external to their head constituent' are the system morphemes that must come from the Matrix Language" (2002: 87). Thus,

we would expect Content Morphemes to be the most commonly borrowed/switched items, while at the other end of the continuum Outside Late System morphemes would never be switched.

Table 2-2 Examples of morphemes predicted by the 4M model

Content morphemes	Nouns	Must have thematic roles to overtly appear in a CP. In the literature, nouns are the most commonly switched/borrowed.
Early system morphemes	Plural morphology	No thematic role. Activated at the conceptual level since it is part of the speaker's intention. Rarely switched.
	Adverbs	No thematic role. Normally activated at the conceptual level. Commonly switched but not to the extent that nouns are.
Bridge late system morphemes	Possessive 'of' and 's in English	Integrates the object N into a larger constituent. Indicates hierarchical relationship Very rarely switched.
Outside late system morphemes	Verbal agreement	Must look outside VP to subject for its morphological form. Should never be switched.

The 2002 version of the Matrix Language Frame model contains one final refinement – the Abstract Level Model. According to the Abstract Level Model, each lexical item has three levels of abstraction:

1. Lexical-conceptual structure (semantic/pragmatic features) – these are matched at the conceptual level with the intention the speaker intends to convey.

2. Predicate-argument structure – relations between theta role assigners and the arguments they map onto phrase structure units.
3. Morphological realization patterns – constituent orders required by well-formedness constraints (Myers-Scotton 2002: 96).

For a singly occurring Embedded Language morpheme to appear, it must have "passed checking for congruence at all three levels of grammatical abstraction" (Myers-Scotton 2002: 97). In discussing levels of abstraction, we must keep in mind that the "[m]otivation for code-switching is always at the level of lexical-conceptual structure" (2002: 97). An EL morpheme appears in the surface string because it conveys the intention that the speaker wants. To demonstrate the three levels of abstraction and the congruence required for an EL element to appear, let us return to a previous example:

Actual utterance:

Sevil #431	Республикансk	-ий	-də
	Republican	-m.s	-loc
	Radj		Apost-p
	[rɛspublikansk	iy]	

'Pure' Azerbaijani:

Respublika-nın		-da
Republic	-3.sg.poss	-loc
Anoun	-Aposs	-Apost-p

Here both the Russian adjective and the Azerbaijani possessive noun express the speaker's intention: 'The Republic's'. While the locative post-position *-da/-də* normally takes a noun as its argument, Azerbaijani commonly allows substantive adjectives to fill this slot, so there is also congruence at the level of predicate-argument structure. Lastly, since the locative suffix looks to the next previous vowel to determine its phonological form due to Azerbaijani vowel harmony, and Azerbaijani has an equivalent to the

Russian <и> [i], there is also congruence at the level of morphological realization patterns.

Finally, we must examine one potential weakness of the Matrix Language Frame model. The MLF is only intended to explain code-mixing (code-switching in Myers-Scotton's terms) within the clause. With this model, Myers-Scotton does not intend to deal with language phenomena above the clause level. While she examines some of the same phenomena as researchers who take a pragmatic approach, they would include alternation in language between conversational turns, within conversational turns, or between CPs in the same 'sentence' in their research on code-switching, Myers-Scotton would not. Like most research in Generative syntax, the MLF model is only interested in clause level issues.

2.3.3 Muysken's typology of code-mixing

The third theoretical perspective on code-switching/code-mixing to be examined here is that of Pieter Muysken. In *Bilingual Speech* (2000), Muysken draws on a wide range of code-switching/mixing research. Rather than focusing only on one level of language, his typology encompasses both clause level and discourse level phenomena. Rather than trying to fit all the data into one theoretical rubric, he identifies three different types:

1. Insertion of material (lexical items or entire constituents) from one language into a structure from the other language.
2. Alternation between structures from languages.

3. Congruent lexicalization of material from different lexical inventories into a shared grammatical structure (2000: 3).

As mentioned earlier, he uses the term 'code-switching' only for alternation and 'code-mixing' for the other two. His argument for this is that "[t]he term code-switching is less neutral in two ways: as a term it already suggests something like alternation (as opposed to insertion), and it separates code-mixing too strongly from phenomena of borrowing and interference" (2000: 4). To this basic typology, Muysken adds an Adequacy Principle: "If in a code-mixed sentence two adjacent elements are drawn from the same language, an analysis is preferred in which at some level of representation (syntax, processing) these elements also form a unit" (2000: 61).

In discussing insertion, Muysken's terminology and analysis closely parallel that of Myers-Scotton. In insertion, a single lexical item or an entire constituent from one language is embedded into a structure from another language – EL versus ML. According to Muysken "[t]here is considerable variation in what is or can be inserted: in some languages this consists mostly of adverbial phrases, in others mostly single nouns, and in yet others again determiner + noun combinations" (2000: 5). Though there are exceptions, inserted elements are often morphologically integrated into the matrix language. This would be particularly expected in the case of Azerbaijani, an agglutinative language, and Russian with its rich nominal morphology:

> These examples, where a case affix from one language is attached to the nominal constituent from another one, are typical of code-mixing involving agglutinative languages such as Basque, Quechua, Finnish, Tamil, Maori, and Turkish. In these

> languages morphological elements such as case markers typically have phrasal scope, an invariant form (barring late phonological rules), and no stratal sensitivity (2000: 76).

Again, in agreement with Myers-Scotton, Muysken sees content words as the most commonly inserted lexical items.

According to Muysken alternation is the only item in his typology which can properly be called code-switching. "Alternation is [a] very common strategy of mixing, in which the two languages present in the clause remain relatively separate" (Muysken 2000: 96). It can take place at the discourse level: between conversational turns, between sentences, or between clauses within one utterance. It can also take place within a CP in peripheral elements. CP level adverbial modification, coordination, or left dislocation may all be cases of alternation (Muysken 2000: 100-101). Thus:

> Content words such as nouns and adjectives are likely to be insertions, while discourse particles and adverbs may be alternations. Sentence Grammar and Discourse Grammar may be relatively autonomous with respect to each other; there is very frequent language choice disparity between these systems (Muysken 2000: 97).

Alternation should be considered when there are several EL constituents in a row, non-nested A B A sequences, greater length and/or complexity of EL sequences, and alternation occurring at a major clause boundary.

The last type of code-mixing in Muysken's typology is congruent lexicalization. In this situation, it is impossible to determine which language is providing the structure

for the CP. Rather both languages can be seen as providing the CP structure together, with both providing the lexical items to fill out the surface string. Muysken proposes that congruent lexicalization may occur because:

a) There is an overabundance of homophonous words, dimorphs, that serve as bridges or triggers for the code-mix;

b) There is a general structural equivalence, both categorical and linear, making code-mixing possible, without there necessarily being any lexical correspondence (2000: 123).

Since many of the instances of code-mixing that Muysken classifies as congruent lexicalization involve either dialects of the same language or closely related language varieties, congruent lexicalization is much more closely related to style shifting or change in language register than insertional or alternational code-mixing/switching.

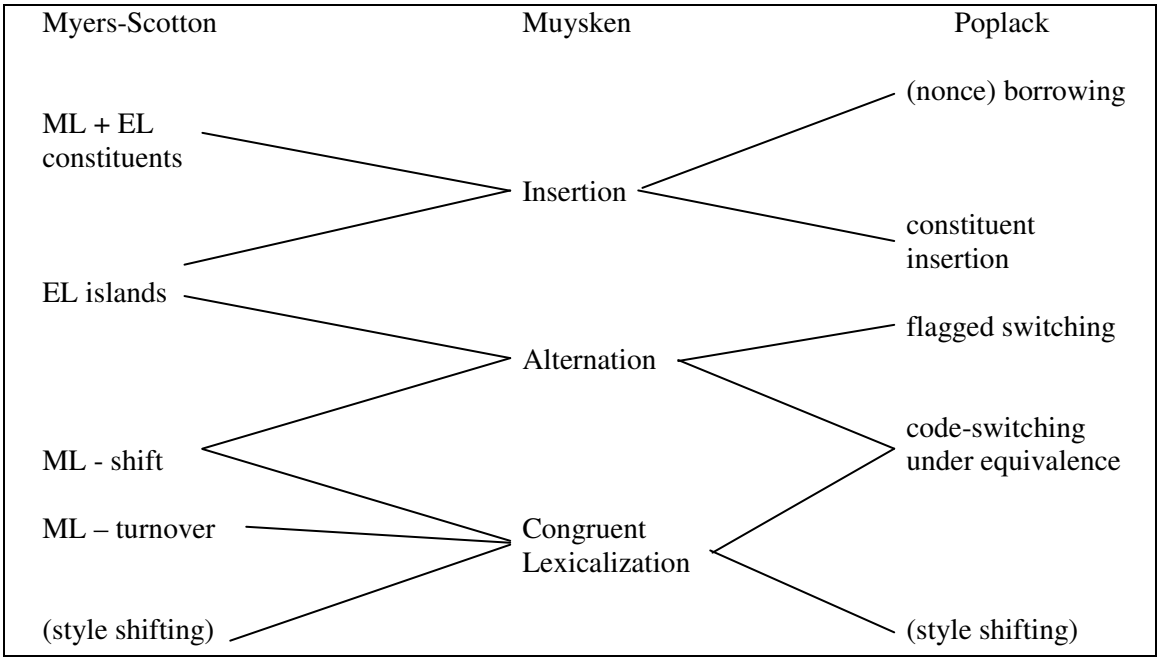


Figure 2-2 Muysken's model of code-switching/code-mixing correlated with Myers-Scotton and Poplack (based on Muysken 2000: 32)

Table 2-3 Diagnostic features of the three patterns of code-mixing (Muysken 2000: 230)

	Insertion	Alternation	Congruent Lexicalization
constituency	+	0	0
single constituent	-	+	0
several constituents	-	-	+
non-constituent	+	-	0
nested a b a	+	-	0
non-nested a b a	-	+	+
element switched			
diverse switches	-	0	+
long constituent	-	+	-
complex constituent	-	+	-
content word	+	-	-
function word	-	-	+
adverb, conjunction	-	+	-
selected element	+	-	+
emblematic or tag	-	+	0
switch site			
major clause boundary	0	+	0
peripheral	0	+	0
embedding in discourse	0	+	0
flagging	-	+	-
dummy word insertion	+	0	-
bidirectional switching	-	+	+
properties			
linear equivalence	0	+	+
telegraphic mixing	+	-	-
morphological integration	+	-	+
doubling	-	+	-
homophonous dimorphs	0	-	+
triggering	0	0	+
mixed collocations	0	-	+
self-correction	-	+	-

Figure 2-2 compares Muysken's typology of code-switching/mixing with Myers-Scotton's MLF model and the Constraint's model from Poplack, while Table 2-3 lists

diagnostic features which can be used to determine which kind of code-switching/mixing is present in a particular stretch of linguistic data.

In addition to his typology of code-switching/mixing, Muysken discusses sociolinguistic situations in which each of the three types can be expected:

Alternation	Stable bilingual communities. Tradition of language separation (diglossia).
Insertion	Colonial settings, recent migrant communities. Considerable asymmetry in speakers proficiency in the two languages. Language dominance shift = shift in the directionality of insertion.
Congruent Lexicalization	Second generation migrant groups. Dialect/standard and post-creole continua Bilingual speakers of closely related languages with roughly equal prestige and no tradition of language separation (based on Muysken 2000: 9).

2.4 Pragmatic approaches to code-switching and code-mixing

2.4.1 Contextualization and textualization

Moving away from syntactic/theoretical models, some of the earliest aspects of code-switching/code-mixing to arouse interest were the pragmatic significance of these phenomena and the functions they serve in conversation. Included in the list of functions that code-switching/code-mixing can serve are *contextualization* and *textualization* cues. In the contextualization cue, the switch in language provides conversational information in addition to the referential meaning of the words. The textualization cue, however, serves the more abstract function of emphasizing the structure of the spoken text.

The contextualization cue (Gumperz 1982) figures very prominently in literature on pragmatic/discourse level functions of code-switching. According to Auer,

contextualization "comprises all activities by participants which make relevant, maintain, revise, cancel… any aspect of context which, in turn, is responsible for the interpretation of an utterance in its particular locus of occurrence" (1992: 4). When code-switching functions as a contextualization cue, the languages in which the words are spoken provide inferences apart from the referential meaning of the words. According to Auer these may be interpreted either by "contrast or by inherent meaning potential" (1995: 124). When speakers intend utterances to be interpreted by contrast, they emphasizes the referential meaning by speaking them in another language. The meaning of the switch in language as a contextualization cue is to "indicate otherness… [and] the direction of the change is irrelevant" (Auer 1995: 124). When interpreted by inherent meaning, the switch itself carries meaning. Code-switching examples would include switching languages according to the situation or to serve metaphorical purposes such as indicating a change of speaker roles (see Blom & Gumperz 1972).

Brian Chan takes this concept one step further in positing the 'textualization cue' (Chan, B. 2004). He suggests that a common form of code-switching is to use EL discourse markers to structure the spoken text. He (2004) describes this phenomenon using numerous examples from the literature, such as code-switched connectives. Rather than invoking conversational implicatures (Grice 1975) or inferences (Gumperz 1982), these EL elements organize the spoken text by emphasizing the fact that a discourse marker has been spoken.

2.4.2 Conversation Analysis

Within the broad category of contextualization, Conversation Analysis looks at the details of talk-in-interaction to identify how participants do the 'work' of talk within the immediate speaker/hearer context. CA grew out of ethnomethodology in the 1960s and early 1970s through the collaboration of Harvey Sacks, Emanuel Schegloff, and Gail Jefferson. In their seminal article (Sacks et al. 1974) they analyzed detailed transcripts of 'normal' conversation to determine what principles interlocutors used to organize their talk. They identified 14 rules to account for turn taking in their conversational data (Sacks et al. 1974: 700-701).

While their analysis of turn taking is profound, what made a lasting impact on social science research at large was the methodology and assumptions they used both to arrive at and answer their research question. This is characterized by, among other things (see Goodwin & Heritage 1990): an interest in 'mundane' linguistic data - *parole* rather than *langue* (Saussure 1959); detailed transcripts showing speaker intonation, gaps in talk, overlap, variation in pronunciation, etc.; emphasis on local context (what these linguistic actions mean to these interlocutors in this situation); and, attention to sequential organization – adjacency pairs, first pair parts, preferred second pair parts, etc.

According to Goodwin and Heritage,

> this analytical approach - in which each conversational action is treated as both displaying an understanding of prior and projecting subsequent conversational actions - has enabled simultaneous analysis (a) of the organization of action and (b) of understanding in interaction (1990: 288).

Since Sacks, Schegloff, and Jefferson's early work, CA has been applied to talk-in-interaction in a wide range of contexts. Many researchers have applied it to corpora involving code-switching and code-mixing.

In 1984 Auer proposed a framework for using CA to understand the pragmatic/discourse functions of code-switching/mixing. Underlying this framework is an assumption that speakers have a preference for same language talk. Thus, changes in language or the insertion of elements from one language into constituents from another should bear meaning. However, Auer points out that what a linguist may see as two separate languages/codes may or may not represent different languages/codes to the bilingual speaker or bilingual participants in a given conversation: "The (bilingual) speaker may not make a distinction between two independent and strictly separated systems. Often, the varieties in the repertoires of bilingual speech communities show independent developments setting them off against the monolingual norms…" (1984: 26).

In this early version of Auer's framework, he sees two different types of language alternation: code-switching and transfer. Code-switching involves longer sections of language and would correspond to alternation between clauses or conversational turns in Muysken's typology. Transfer involves smaller units or individual words and would correlate with Muysken's insertion. However, in keeping with a participant orientation, Auer contrasts 'transfer' as a linguist would see it and 'transfer' as a participant would perceive it. The fact that a particular lexical item appears in the dictionary of one language and not in the dictionary of another may be enough for a linguist to classify it as

belonging to only one language. However, for the insertion/transfer of that item to be meaningful (the change of language itself to have pragmatic/conversational significance) the speaker/hearer must perceive it to be a change in language.

> The 'proof procedures' for transfer$_P$ [as perceived by the participant] and transfer$_L$ [as interpreted by a linguist] are therefore quite different. Usually, transfer$_L$ is the weaker alternative with which we have to content ourselves if we cannot demonstrate that the production of an other-language item has a function… It requires demonstrating how the participant *displays* a 'reason' for language alternation, in the way this alternation is produced, in visible-inspectable ways for his or her coparticipant(s); in short it requires a detailed sequential analysis (Auer 1984: 27).

To complete his framework Auer adds another category pair to his two types of code-switching/mixing: participant related language alternation, versus discourse related language alternation. Participant related language alternation would have to do with the speaker and his/her language preferences, while discourse related alternation would "'cue' the unfolding interaction" (Bani-Shoraka 2005: 58).

> For instance, participant related alternation may serve to either display speakers' own language dispositions, or to accommodate to those of the recipient(s)… Discourse-related alternation, on the other hand, may take part in the organization of the verbal episode by signaling a projected change of recipients of talk, or by bracketing certain sequences from the currently entertained activity, to name a few functions (Bani-Shoraka 2005: 58-59).

Since 1984, researchers have applied Conversation Analysis and CA inspired methodologies to a wide range of code-switching situations. Of particular interest to this study is Helena Bani-Shoraka's (2005, 2008, 2009) work with the Persian/Azerbaijani bilingual community in Tehran, Iran.

2.4.3 Perspectives on discourse

Studies of contextualization/textualization and the methodologies proposed by Conversation Analysis fit within the broader field of Discourse Analysis. From this field of linguistic and social science inquiry, Dell Hymes' 'etic grid' and Deborah Schiffrin's theoretical model are of particular use for this dissertation project.

In laying the groundwork for his work in the ethnography of communication, Hymes (1972) designed an etic grid to help researchers identify potentially salient aspects of any communicative event based on the acronym S.P.E.A.K.I.N.G. By describing a speech situation according to these eight categories a researcher can identify the important elements for analysis.

Table 2-4 Hymes' etic grid (Schiffrin 1994: 142)

S	setting/scene	physical circumstances
P	participants	speaker/sender/addressor
		hearer/receiver/audience/addressee
E	ends	purposes and goals
		outcomes
A	act sequence	message form and content
K	key	tone, manner
I	instrumentalities	channel (verbal, nonverbal, physical)
		forms of speech drawn from community repertoire
N	norms of interaction	specific properties attached to speaking
	and interpretation	interpretation of norms within cultural belief system
G	genre	textual categories

Rather than simply describing the speech situation, Schiffrin's theoretical model (1987) posits five levels or components of discourse:

- Participation Framework – relationships of participants to each other and the utterances they produce.
- Ideational Structure – at its simplest 'what the participants are talking about.'
- Action Structure – the actions participants accomplish or attempt through talk.
- Exchange Structure – the linear relationship of utterances to each other.
- Information State – the organization and management of knowledge and meta-knowledge – relates to speakers cognitive capacities (21-29).

The relationship between these is shown graphically in Figure 2-3.

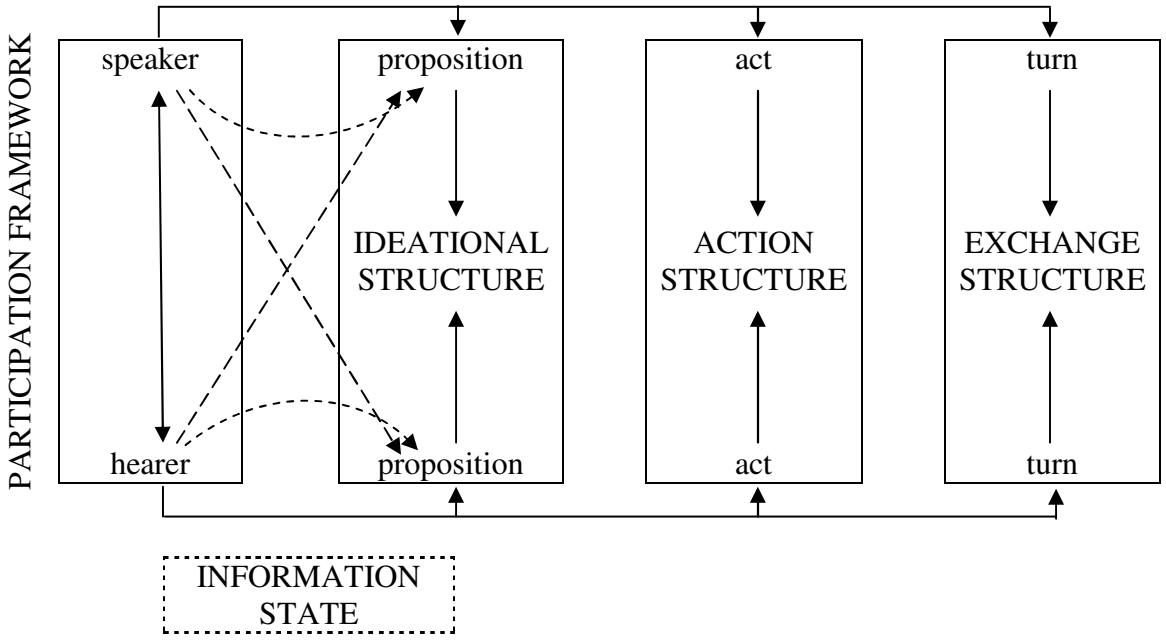


Figure 2-3 Schiffrin's discourse model (Schiffrin 1987: 25)

While most of these levels are easily understood, the information state requires a bit of clarification. As participants interact linguistically and socially, their actions are

constrained and facilitated by their knowledge, meta-knowledge (what they know and/or believe about their interlocutors' knowledge), and cognitive capacities. In the context of this dissertation, that involves each subject's knowledge of and ability to use each of the languages involved in the conversation, subjects' beliefs about the linguistic competence of their interlocutor/s, as well as their understanding of the subject at hand. While relating to the individual conversational participant, the information state is also co-constructed and constantly evolving because participants are constantly evaluating and revising their beliefs about their interlocutors' knowledge and abilities.

To summarize Schiffrin's model, the participants in the conversation (speaker and hearer in alternation) each provide propositions (ideational structure) which carry out the action of the conversation (see also Searle 1969 and Fasold 1990: 147-159) contained in the conversational turns (exchange structure). All of these levels are constrained and facilitated by the participants' linguistic knowledge, beliefs about their interlocutor's knowledge, and cognitive abilities as well as their knowledge of the subject under discussion (information state). When combined with Hymes' etic grid, Schiffrin's model facilitates an emic analysis of spoken discourse.

As Schiffrin correlates various discourse markers in English with discourse functions on these different levels, studies of code-switching/mixing could well be interpreted in terms of these underlying levels of discourse. By showing solidarity code-switching/mixing could relate to the Participation Framework. If words or phrases are inserted because they relate to a particular topic, they would relate to the Ideational Structure. If cs/cm functions as a contextualization cue it facilitates the Action Structure

of the conversation, or by organizing the stream of spoken discourse a textualization cue would relate to the exchange structure. Lastly, code-switching/mixing would be manifestations of the individual and co-constructed Information State – linguistic knowledge, co-knowledge, and cognitive abilities.

2.5 Methodological issues

2.5.1 Data collection

One of the biggest obstacles that a researcher working with 'live' language data must overcome is the Observer's Paradox. In coining this term, Labov says that "the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain these data by systematic observation" (1972b: 209). This is a particular problem when the researcher is collecting linguistic data from a speech community of which she or he is not a part. One obvious way to overcome this problem is to analyze conversational data collected while the 'foreign' researcher was not present either by having a research subject collect the data, or by using an insider in the speech community to participate as a partner in the research (see Labov 1972a). Another way would be for the researcher to spend enough time in the community to be considered a member before collecting the data – participant observation.

Thus, data for research into code-switching and code-mixing has been collected in a number of ways. Poplack (1980, 1988) collected tape recordings of large samples of naturally occurring conversation. Similarly, Piller (2002) collected recorded conversations by giving her bilingual subjects tape-recorders and a list of self-interview

questions. Blom and Gumperz (1972) observed public language behavior in Hemnesberget, Norway, but also placed subjects in structured interactional situations to observe how topic, setting, and interlocutor affected language use. Nivens (2002) recorded both naturally occurring and somewhat staged conversational data in West Tarangan, but also used interviews and surveys to determine which words his subjects considered to be borrowed into West Tarangan, and which they considered to be pure Malay. Similarly, Bani-Shoraka (2005, 2008, 2009) placed tape recorders in her subjects' homes and recorded the naturally occurring conversation around the house. But this was only possible "[a]fter the quantitative study based on questionnaires, during which [she] established many valuable contacts and had the opportunity of conducting extensive observations as well as different types of informal interviews…" (2005: 117). Given the wide range of phenomena that can be called code-switching or mixing, and the variety of ways that these can be used, a variety of data collection methods and analytic processes should be used to determine the role/s that these play within a given speech community.

2.5.2 Transcription theory

While the data of studies into code-switching/code-mixing are typically spontaneous spoken utterances, the analysis of them is almost always done via transcription. Before the linguistic data can be analyzed they must be transformed into written text. According to Ochs, a researcher should give serious thought to transcriptions:

a) because for nearly all studies based on performance, the transcriptions are the researcher's data;
b) because transcription is a selective process reflecting theoretical goals and definitions; and
c) because with the exception of conversation analysis (Sacks et al. 1974), the process of transcription has not been foregrounded in empirical studies of language behavior (Ochs 1979: 44).

Ochs goes on to discuss two bias issues in transcription: left/right bias – as Westerners we tend to see things to the left as first and therefore more important than things to the right, top/bottom bias – in the same way items at the top of a page appear more prominent than those at the bottom.

Essentially, a researcher must have the aim of the research project in mind before transcribing the data or re-transcribing if the direction of the project changes. A transcript must be designed that will include and highlight the salient features while not omitting other features that may prove important. While a 'classic' CA style transcript may seem superior because it preserves more details of the conversation, if issues such as intonation, lengthened vowels, overlap, etc. are not salient to the discussion, then documenting them in the transcript may well make it difficult to detect and analyze other features.

2.5.3 Regulation of research involving human subjects

One final issue affecting the methodology of this dissertation research project is U.S. Federal regulations covering research involving human subjects. As an institution

that receives Federal funding, The University of Texas at Arlington is subject to the regulations of the U.S. Department of Health and Human Services found at Title 45 Code of Federal Regulations, Part 46 (Code of Federal Regulations 2005) and follows the principles described in The Belmont Report (Department of Health, Education, and Welfare 1979). All research projects involving human subjects must be approved by the University's Institutional Review Board. One requirement of these regulations is that all research subjects who contribute any type of identifiable data (including voice recordings) must give informed consent for their participation in the research project. This is usually in the form of a signed consent document that explains the project and any potential negative consequences. More rigorous regulations are involved for vulnerable populations such as children or prisoners. Though these regulations are quite reasonable, they do put limitations on the scope of data collection. A method such as that used by Bani-Shoraka where a recorder was simply placed in a home and all conversation was recorded would be very difficult to utilize. The researcher would have to contact every person who came through the house and have them sign an informed consent form, and if any children were involved, both of their parents would have to give consent for the child's participation.

2.6 Application to this dissertation project

The aim of this dissertation project, then, is to document and describe the forms of code-switching and code-mixing between Azerbaijani and Russian in contemporary Azerbaijan (Chapter 4), show the ways that code-switching/mixing patterns vary among speakers (Chapters 5 and 6), analyze the functions that these can and do serve within

conversation (Chapter 7), and demonstrate ways in which code-switching and language mixing are used to construct social identities in that socio-cultural milieu (also Chapter 7).

2.6.1 Data types

To these ends, four types of data were collected: public observations of language behavior, informal interviews with groups and individuals, recorded conversational data, and surveys from each participant in the recorded conversations. Two different types of conversational data were recorded: Staged conversations, and Home recordings. In the Staged conversations subjects were paired by the researcher according to various demographic factors and given a list of topics to discus. In the Home recordings, subjects were given recorders and lavaliere microphones to take home so that they could record an extended stretch of their interaction in an informal setting. In the staged conversation each interlocutor was recorded, but in the Home recordings, only the subject with the recorder and microphone was recorded and filled out the surveys in Appendix A. The bulk of the analysis that follows is of the Home recorded data.

In neither type of recording was there any attempt to randomly select subjects, or select a subject population that replicates the demographics of Azerbaijan. In the Home recordings, subjects were selected based on demographic factors favoring code-switching/mixing, while subjects for the Staged recordings were selected from a variety of backgrounds to see how they would interact in conversation. As well, all subjects come from an urban cultural milieu – that of Baku which has long been oriented more toward Russian than the rest of Azerbaijan.

2.6.2 Transcription

Since the initial goal of this project is a linguistic description of code-switching/mixing, the transcriptions of the home recordings utilized 'normal' orthographic conventions for both Azerbaijani and Russian elements. While this type of transcript is not useful for detailed Conversation Analysis, it allows easy comparison between speakers and facilitates computer assisted corpus analysis. Where colloquial words were used and they clearly represented different lexical items from their formal counterparts, they were preserved. Colloquial pronunciation of 'normal' Azerbaijani words was lost since these were transcribed using regular spelling conventions. Thus, the full range of phonological and morpho-phonemic variation was not reflected in the transcriptions. No grammatical 'errors' were corrected, and the presence or absence of any morphology was preserved. After the initial transcription, each text was checked in detail by a native Azerbaijani speaker with near native fluency in Russian – also a participant in the Home recordings. The classification of elements as Azerbaijani (transcribed in Azerbaijani Latin script) or Russian (transcribed in Cyrillic script) was according to their presence in dictionaries of the two languages and the intuition of the research assistant who checked the transcriptions – $\text{transfer}_{\text{L}}$ in Auer's terminology (1984).

2.6.3 Theoretical orientation

Since this project aims to analyze and describe code-switching/mixing both from syntactic and pragmatic perspectives, it draws on a number of theoretical perspectives. Muysken's typology (2002) will provide the general framework to classify

Azerbaijani/Russian code-switching/mixing, though at the clause level Myers-Scotton's terminology and model (2002) will also be applied. The formal framework for all linguistic analysis will be the Principles and Parameters version of Generative syntax as explicated in Haegeman 1994. In determining the conversational functions of code-switching/mixing in Azerbaijan, reference will be made to broad theories of pragmatics including contextualization/textualization (Gumperz 1982, Chan, B. 2004, etc.). Though the Home recordings do not lend themselves to 'classic' Conversation Analysis (Sacks et al. 1974, Auer 1984, etc.) sequential analysis will be applied to some sections. Both Hymes' etic grid and Schiffrin's theoretical model will be used when analyzing the speech of individual participants in Chapter 7.

2.6.4 Micro/macro orientation

As section 2.2 of this chapter states, studies of code-switching and code-mixing can fall anywhere along the micro/macro continua of linguistic and social analysis (McKay & Hornberger 1996). In studying the language behavior of individuals as well as individual language behaviors, this study will fall into multiple quadrants of Table 2-1. By drawing on data from a few subjects but aiming to describe the forms and functions of Azerbaijani/Russian code-switching/mixing in general, this study would fall into the macro linguistic/micro social quadrant of Table 2-1 like studies into the ethnography of communication or speech acts. However, in looking at the pragmatics and meanings demonstrated by individuals in smaller conversational segments, part of the analysis will also fall into the micro linguistic/micro social quadrant. Finally by combining the analysis of the recorded conversational data with observations of public language

behavior in the context of other research, some general ways in which code-switching/mixing are used to construct identity in Azerbaijani culture at large will be posited – the micro linguistic/macro social quadrant.

CHAPTER 3

AZERBAIJANI AND RUSSIAN LINGUISTIC STRUCTURES

3.1 Introduction: Principles and Parameters grammar

In order to propose potential sites for and types of code-switching and code-mixing, this chapter provides an overview of the Azerbaijani and Russian linguistic systems. For brevity sake and to maintain the focus of the dissertation project, discussion of phonology is kept to a minimum. Moreover, Azerbaijani receives a more thorough treatment since Western linguistic analyses of this language are virtually non-existent, though one familiar with Turkic languages in general will recognize the similarities. The goal of the sections on the Russian Linguistic System is not to provide a comprehensive description of the language, but rather to provide an overview that accounts for the data to be analyzed in later chapters. Since Russian shows a large amount of morphological variation, only one or two paradigms are listed for each type of suffix. For a more thorough listing of forms, see Leed 1986. In the sections that follow, word level phenomena are first discussed then sentence level grammatical structures are covered.

The syntax sections are based on the Principles and Parameters (also known as Government and Binding) version of Generative Transformational Grammar (Haegeman 1994). P&P sees the grammar of a language as composed of 5 components. Theta Theory is concerned with the lexicon of a language, and the argument structures required by its nouns and verbs. Each NP (argument) in a sentence must be assigned a Theta

Role. In addition, each argument is assigned one and only one theta role, and each theta role is assigned to one and only one argument (the Theta criterion – Haegeman 1994: 54). X Bar theory (X') contains the phrase structure rules of the language. All phrase structure rules can be summarized by the following generalizations:

XP → Spec; X'
X' → X'; YP
X' → X; YP

X and Y represent the phrasal heads. Each language specifies the order of the constituents through parameters indicating whether the language is Spec (Specifier) first or last, and Head first or last. The Case component does not necessarily relate to morphological case, but rather concerns which NPs are licensed to appear overtly in a sentence. "Every overt NP must be assigned abstract case" (the Case filter – Haegeman 1994: 167). In most languages the case assigners are Infl (tense and inflection – the head of the sentence or IP), verbs, and pre- or postpositions. According to Principles and Parameters grammar, there are three kinds of NPs in the world's languages: R-expressions, pronouns, and anaphors (reflexive and reciprocal pronouns). R-expressions (Referential expressions) have independent reference. They are 'normal' nouns like 'table,' 'car,' or 'Fred.' Pronouns refer to something within the current discourse, but outside the pronoun's governing category (usually the clause). An anaphor, however, is bound to a subject within its governing category. The Binding component of P&P grammar deals with co-referencing anaphors with their antecedents. Lastly, Movement deals with the re-ordering of elements within the sentence.

Examples in this dissertation are written in the alphabet appropriate to the language of the utterance. Azerbaijani examples will appear in the current version of the Azerbaijani Latin Script. Since most of the letters correspond to English, phonetic transcriptions are not provided. The following characters are, however, either not in the English alphabet or pronounced differently than in English:

<ş>	[ʃ]	<ü>	[y]
<ç>	[tʃ]	<ö>	[ø]
<j>	[ʒ]	<ı>	[ɯ]
<c>	[dʒ]	<ə>	[æ]
<q>	[q or ɢ]		
<ğ>	[ʁ]		

Russian examples will be written in the standard Cyrillic alphabet with phonetic transcriptions in brackets. Please see Appendix B for a full listing of both alphabets. Brackets will be used as follows: pointed brackets (< and >) indicate regular orthography for the language, slashes (/ and /) indicate phonological transcription, and square brackets ([and]) hold IPA phonetic transcriptions except in sections dealing with syntax where they may be used to define phrase boundaries.

3.2 The Azerbaijani linguistic system[3]

3.2.1 Azerbaijani vowel harmony

Though not the focus of this dissertation, some mention of the Azerbaijani vowel harmony system is needed since it affects the phonological forms of all inflectional morphology in the language. Azerbaijani has two types of vowel harmony. The first involves the low vowels /æ/ and /ɑ/, and the second the high vowels /i/, /y/, /ɯ/ and /u/:

[3] The information in this section was originally part of a paper entitled "Azerbaijani Structures: Phonological, Syntactic, Discourse" submitted in fulfillment of the Non-Western Linguistic Structures requirement for the Ph.D. in Linguistics at UT Arlington.

Low Vowel Harmony

[gæl.mæk]	<gəl-mək>	'to come'
[ɢɑl.mɑq]	<qal-maq>	'to stay'

High Vowel Harmony

[ev.im]	<ev-im>	'my house'
[gyl.ym]	<gül-üm>	'my flower'
[bal.ɯm]	<bal-ım>	'my honey'
[ɢol.um]	<qol-um>	'my arm'

The vowels in all suffixes are underspecified for [back] and [round] with two possibilities:

I	A
[+high]	[+low]

One rule is necessary to assign the feature [-round] to the low vowels – a redundancy rule since Azerbaijani has no low, round vowels:

Low Vowel Feature Assignment: (insures that all [+low] vowels are [-round])

V

[+low] → [-round]

In vowel harmony, the underspecified features are copied from the next preceding vowel:

Front/Back Harmony: (attaches the [back] feature from the next preceding vowel to the vowel in a suffix)

V # V

|_____|

[±back]

Rounding Harmony: (attaches the [round] feature from the next preceding vowel to the vowel in a suffix that lacks a [round] feature assignment)

V # V

|_____|

[±round]

In reality, neither of these last two rules is required. They are both applications of the universal Right Spread Rule (Goldsmith 1990).

3.2.2 Word level phenomena: Azerbaijani morphology

Like all Turkic languages, Azerbaijani is agglutinative with a rich system of both nominal and verbal morphology. All of these suffixes are subject to vowel harmony and other phonological rules. In addition, several other morpho-phonemic rules are necessary to account for the deletion of consonants and vowels at the beginning of many suffixes. Table 3-1 lists the Azerbaijani personal pronouns. The verbal agreement and nominal possessive suffixes are very closely related to them.

Table 3-1 Azerbaijani pronouns

	Singular	Plural
1st	/mæn/	/biz/
2nd	/sæn/	/siz/
3rd	/o/	/onlar/

Table 3-2 Possible noun morphology

root	(plural)	(possessed)	(possessive)	(case/post-position)
<ev> 'house'	/-lAr/ <-lər> pl	see Table 3-3 <-im> 1.sg.possd	see Table 3-3 <-in> 3.sg.poss	see Table 3-4 <-ə> dat
<evlərimin ə> "…to my houses'…"				

As Table 3-2 shows, four sets of suffixes can be added to the noun root. In Azerbaijani, singular is unmarked while plural is marked by the suffix <-lar> or <-lər>. Both of these derive from the underlying phonological form /-lAr/ with the vowels /ɑ/ and /æ/ resulting from the Vowel Harmony rules detailed above. In Azerbaijani, possessed and possessive suffixes must be added to the noun root before other case or post-positional suffixes (Table 3-3). All of these contain the underlying vowel /I/ which becomes /i/, /u/, /ɯ/, or /y/ after Vowel Harmony. Table 3-4 shows the case/post-

positional endings which fill the last slot in the noun morphology. Like many languages, nominative is unmarked. All of these suffixes contain either /I/ or /A/ as the underlying vowel.

Table 3-3 Possessor and possessed suffixes

Possessor			Possessed		
	Singular	Plural		Singular	Plural
1st	/-Im/	/-Im/	1st	/-Im/	/-ImIz/
2nd	/-In/	/-In/	2nd	/-In/	/-InIz/
3rd	/-nIn/	/-nIn/	3rd	/-sI/	/-lArI/

Table 3-4 Case/post-positional suffixes

nominative	n/a
dative	/-jA/
accusative	/-nI/
locative	/-dA/
ablative	/-dAn/

Table 3-5 shows the five categories of morphology that can be added to a verb root.

Table 3-5 Possible verbal morphology

root	(passive, causative, reciprocal, or reflexive)	(negation)	(aspect)	(tense)	(agreement)
<yan> 'burn'	see Table 3-6 <-dır> cause	/-mA/ <-ma> neg	see Table 3-7 <-mış> perf	see Table 3-7 <-dı> past	see Table 3-8 n/a 3.sg
<yandırmamışdı> "he did not cause (it) to burn"					

Closest to the root is the slot for passive, causative, reciprocal, or reflexive. Each of these has the underspecified phoneme /I/ like the possessive/possessed suffixes above. Please

see section 3.2.3.1 below on Theta Theory below for a discussion of how these change the meaning of the verb.

Table 3-6 Passive, causative, reciprocal, and reflexive suffixes

passive	/-Il/
causative	/-dIr/
reciprocal	/-Iʃ/
reflexive	/-jIn/

In Azerbaijani, negation on the verb has only two surface forms, <-ma> and <-mə>. These are derived from the underlying form /-mA/ via Vowel Harmony.

Table 3-7 Aspect & tense forms

Aspect		Tense	
imperfective	/-jIr/	present	/-jIr/
perfective	/-mIʃ/ or /-jIb/	past	/-dI/
conditional	/-sA/	definite future	/-jAʤAQ/
		indefinite future	/-jAr/

As Table 3-7 shows, tense and aspect have significant overlap in Azerbaijani. The suffix /-yIr/ is normally understood as present tense with the implication of ongoing action, though like English it can also mean habitual action or immediate future. When combined with past tense, where it fills the Aspect slot, it implies ongoing action at the time of some other event. Similarly, the perfective suffix /-mIʃ/ is normally used as a past tense much like the 'conversational past' in German, while the past tense morpheme /-dI/ functions as a narrative past. When the two are combined they have a past-perfective meaning. While both the /-mIʃ/ and /-jIb/ forms are possible, /-jIb/ is almost invariably used for 3rd person subjects except in very formal written contexts. /-mIʃ/ is used for 1st and 2nd person subjects, with the exception of informal, spoken contexts when

/-jIb/ may be used for a 2nd person subject. The Definite Future occurs more frequently than the Indefinite Future.

Table 3-8 shows the verbal agreement suffixes. The verb must agree with the subject in person and number. 3rd person singular is unmarked, and there is no overlap in these suffixes. Vowel Harmony derives the surface vowel forms from the underspecified /I/ and /A/ phonemes, and the rule of '/k/ and /q/ alternation' fills in the features for the underspecified /Q/ in the 1st person plural suffix.

Table 3-8 Agreement forms

	Singular	Plural
1st	/-jIm/	/-jAQ/
2nd	/-sAn/	/-sInIz/
3rd	n/a	/-lAr/

Imperative mood is shown in Azerbaijani by the absence of Tense and Aspect. Passive, causative, and reflexive are possible in imperatives, though they are rare. Unlike English, agreement is also possible. If no agreement is present, the subject is 3rd person singular. 1st person plural agreement on an imperative verb has a jussive meaning, while 1st person singular indicates 'I must.' The 2nd person plural agreement morpheme is used for polite imperatives.

The infinitive suffix /-mAQ/ can follow passive, causative, or reflexive, but does not co-occur with tense, aspect, or agreement. This same form also functions as a noun forming suffix, similar to gerunds in English. Nominalized verbs can fill the root slot and be followed by all of the noun morphology described above.

In addition to the phonological rules described in the previous section, three rules are required to account for a number of changes in the nominal and verbal suffixes.

When a suffix begins with a <y> (phonologically /j/) or /n/, it is always deleted when the suffix follows a consonant, either in the noun/verb root or in a preceding suffix. If a suffix begins with a vowel, it is deleted when the suffix follows a vowel, either in the noun/verb root or in a preceding suffix. Lastly, the /s/ at the beginning of the 3rd person singular possessed suffix is deleted when the suffix follows a consonant. This rule does not apply to any other suffixes which begins with an /s/ such as the indefinite future tense /-sA/, the second person singular agreement /-sAn/, or the second person plural agreement /-sInIz/. I propose the following rules to deal with this variation:

V deletion: (deletes a vowel at the beginning of a suffix when it follows another vowel)

$$V \rightarrow \varnothing \; / \; V \; \# \; \_\_\_$$

/j/ and /n/ deletion: (deletes /j or /n/ at the beginning of a suffix when it follows a consonant)

$$\begin{bmatrix} C \\ +\text{coronal} \\ +\text{sonorant} \\ -\text{lateral} \end{bmatrix} \rightarrow \varnothing \; / \; C \; \# \; \_\_\_$$

3rd person singular possessed /s/ deletion: (deletes the /s/ at the beginning of the 3rd person singular possessed morpheme when it follows a consonant)

$$/s/ \rightarrow \varnothing \; / \; C \; \# \; \_\_\_$$
$$[3^{rd}.\text{sing.possd}]$$

These rules would cluster with a number of other phonological rules, but they would only apply one time. Thus, the /j/ could be deleted from the beginning of the definite future tense suffix, but not the /A/ that follows it.

3.2.3 Azerbaijani syntax

Moving beyond word level phenomena, this section will describe the Azerbaijani language according to the Principles and Parameters version of Generative

Transformational Grammar. Azerbaijani is head last (other than CP), but spec first. In most cases, heads occur at the ends of their phrases, but movement is toward the front of the sentence.

3.2.3.1 Theta theory

According to the Theta Criterion, each argument in a sentence must be assigned one and only one theta role, and each theta role must be assigned one and only one argument (Haegeman 1994: 73). In Azerbaijani, verbs, nouns, and post positions have the ability to assign theta roles. Here, the suffix /-dIr/ is shown as a copular verb that has become cliticized.

Azerbaijani has verbs with at least 10 different types of theta grids:

1. /-dIr/ 'be': V

NP	NP/PP/AP
Topic	Comment

2. <durmaq> 'to stand': V

NP
Agent

3. <görmək> 'to see': V

NP	NP
Agent	Patient

4. <bilmək> 'to know': V

NP	NP
Agent	Phenomenon

5. <getmək> 'to go': V

NP	NP
Agent	Goal

6. <qalmaq> 'to stay': V

NP	NP
Agent	Location

7. <eşitmək> 'to hear': V

NP	NP
Experiencer	Communiqué

8. <gətirmək> 'to bring': V

NP	NP	NP
Agent	Patient	Goal

9. <demək> ‘to say’: V

NP	NP	NP
Speaker	Theme	Addressee

10. <vermək> ‘to give’: V

NP	NP	NP
Agent	Recipient	Patient

The theta grids for nouns are much less complicated. No nouns occur with complement arguments, but presumably every non-proper noun may occur with a possessor. Thus most Azerbaijani nouns would have theta grids like this:

<ev> ‘house’: N

(NP)
Possessor

In addition, many Azerbaijani nouns have mandatory classifiers. Unlike some languages in Africa or Asia, there are only a few classifiers in Azerbaijani:

<dənə> countable thing (very common)

<baş> ‘head’ as of livestock (limited distribution)

<nəfər> ‘person’ (parallel to <dənə> but rarely used)

Thus, Azerbaijani has at least four classes of nouns, described by the following features:

[± count] identifies whether the noun does/does not take the plural marker /-lAr/
[± class] identifies whether the noun takes a classifier such as <dənə>
[± arg.] identifies whether the noun occurs with an external argument (usually Possessor)

Features:			Examples:
[-count]	[-class]	[-arg.]	<Könül> a female name
[+count]	[-class]	[+arg.]	<yol> ‘way or road’
[+count]	[+class]	[+arg.]	<alma> ‘apple’
[-count]	[-class]	[+arg.]	<süd> ‘milk’

<Könül> a name +N +[_]	<yol> ‘way or road’ +N +[(det)(numb)(adj)_]	<alma> ‘apple’ +N +[(det)(numb)(class)(adj)_]	<süd> ‘milk’ +N +[(det)(adj)_]

Theta grids:

<Könül> a name: N

<yol> 'way or road': N

(NP)
Possessor

<alma> 'apple'

(NP)
Possessor

<süd> 'milk'

(NP)
Possessor

Azerbaijani has a complex system of possessor/possessed suffixes. The possessive noun is followed by the possessive suffix, which agrees with it in person and number, while the possessed noun is followed by the possessed suffix which agrees in person and number with the possessive noun. The following rule accounts for these suffixes:

SD:	[NP	X	$N]_{np}$	$\rightarrow$	[NP	-poss	X	N	$-possd]_{np}$
SI:	1	2	3		1		2	3	
SC :	α person					α person			α person
	β number					β number			β number

When an NP appears in [Spec, NP] it receives the possessive suffix which agrees with it in person and number, while the head N receives the possessed suffix which agrees with the NP in [Spec, NP].

Post-positions also assign theta roles in Azerbaijani, but unlike other languages, such as Russian, their arguments do not receive surface case endings. Some post-positions appear as separate words, while others are clitics. Here are a few of the most common post-positions in Azerbaijani:

1. <kimi> 'like/as': P

NP
Measure

2. <üçün> 'for': P

NP
Beneficiary

3. <görə> ‘for/because of’: P

NP
Agent (?)

4. <tərəf> ‘toward’: P

NP
Goal

5. /-dAn/ ‘from’: P

NP
Origin

6. /-dA/ ‘in’: P

NP
Location

Azerbaijani manifests passivization in which the Patient argument occurs as the S-Structure subject, and the Agent occurs (optionally) with the /-dAn/ post-position. The verb is marked morphologically with the suffix /-Il/. The following lexical rule is proposed to capture the relationship between the active and passive verbs:

V: | NP | NP | + Pass → V_{pass}: | NP | (NP + <-dan>) |
|---|---|---|---|---|
| X | Patient | | Patient | X |

Azerbaijani also manifests causativization in which a one argument verb is changed to a two argument verb. The original argument becomes the S-Structure object (with accusative case), and an external Agent argument is added. The verb is marked morphologically with the suffix /-dIr/. The following lexical rule is proposed to capture the relationship between the active and causative verbs:

V: | NP | + Caus → V_{caus}: | NP | NP |
|---|---|---|---|
| X | | Agent | X |

Other constituents in the Azerbaijani lexicon would include:

determiners (D)	numbers (Numb)	plural marker
<bu> ‘this’	<bir> ‘one’	<-lar>
<o> ‘that’	<iki> ‘two’	
<bir az> a little’	<üç> ‘three’	
<az> ‘little’	etc.	

3.2.3.2 X bar theory

Phrasal heads in Azerbaijani include:

P	head of PP (Post-Position Phrase)
A	head of AP (Adjective Phrase)
Adv	head of AdvP (Adverb Phrase)
D	head of DP (Determiner Phrase)
Numb	head of NumbP (Number Phrase)
N	head of NP (Noun Phrase)
V	head of VP (Verb Phrase)
I	head of IP (Infl Phrase)
C	head of CP (Complementizer Phrase)

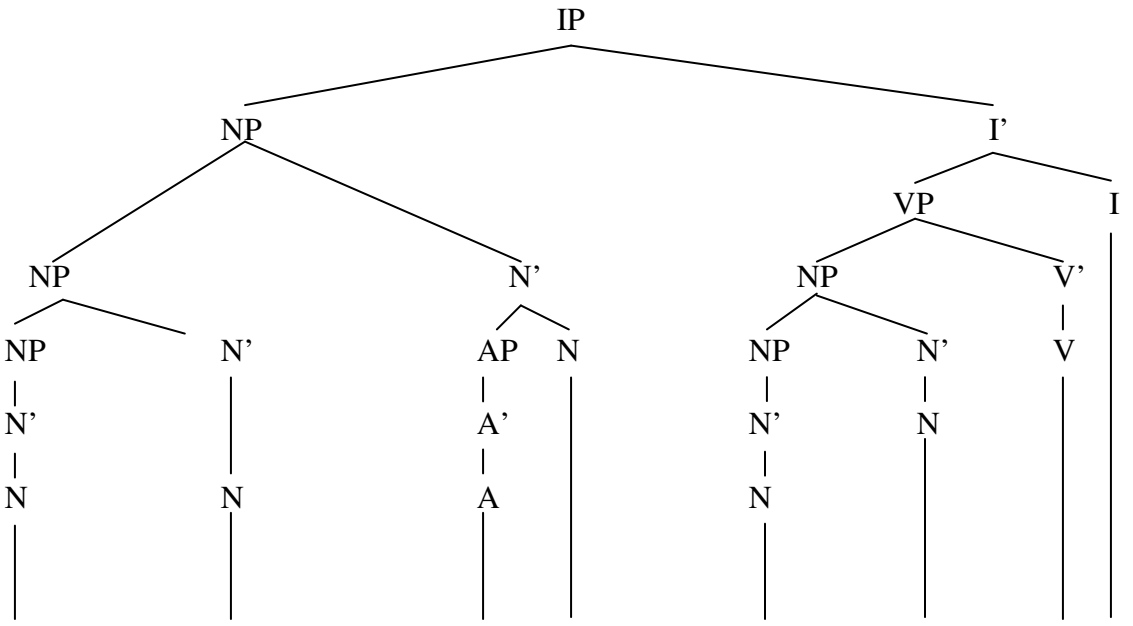
The following phrase structure rules seem to account for the vast majority of Azerbaijani sentences:

PP:	PP	→	NP P	ev-də house-in	‘in (the) house’	
AP:	AP	→	(AdvP) A’	çox yaxşı	‘very good’	
	A’	→	(A’) A	much good		
AdvP:	AdvP	→	(Intens) Adv’	lap çox	‘very much’	
	Adv’	→	(Adv’) Adv	most much		
DP[4]:	DP	→	D (NumP)	bu bir that one	‘that one’	
NumbP:	NumbP	→	Numb (Clas)	bir dənə one class	‘one countable thing’	
NP:	NP	→	{ (DP) / (NP) } N’	bu iki dənə alma that two class apple	‘those two apples’	
	N’	→	(AP) N’			
	N’	→	(IP) N	[park-da ol-an]$_{IP}$ maşın park-in be-part car	‘the car in the park’	

[4] Though the rules for DP, NumbP, and NP presented here are sufficient to account for the data in this dissertation and facilitate comparison of the Azerbaijani and Russian linguistic systems, an analysis that posits NP as a daughter of DP could provide additional insight into issues such as agreement between possessed and possessor nouns and classifiers in Azerbaijani.

VP:	VP	→	(Spec)	V’	(Spec position posited for movement)
	V’	→	(NP)	V’	
	V’	→	{ (AdvP) (PP) (AP) (NP) }	V’	
	V’	→	V	(CP)	
IP:	IP	→	NP	I’	
	I’	→	VP	I	
CP:	CP	→	(Spec)	C’	(Spec position posited for movement)
	C’	→	IP	C’	
	C’		C	IP	

The following tree diagram demonstrates a number of these rules:



‘The Republic of Azerbaijan’s state language is Azerbaijani’ (Dövlət Dilinin Təkmilləşdirilməsi 2001)

Azerbaijani uses relativization less often than other languages. Rather, it prefers to use IP in NP. For example, Azerbaijanis prefer the second construction below to the first, though they will admit that the first is possible:

'That is the dog that bit me.'

1.	[[Bu	it	-dir]$_{IP}$	[ki [*e*	mən-i	tut-ub	]$_{ip}$]$_{c'}$]$_{cp}$
	that	dog	-be	C	1.sg.acc	bite-perf	

2.	[[Bu	[PRO$_i$	mən-i	tut-an]$_{IP}$	it$_i$]$_{NP}$	-dir]$_{IP}$
	that		1.sg.acc	bite.part	dog	be

In Azerbaijani, the verb in a finite clause must agree with the subject of the clause. This can be accounted for by the following rule:

SD:	NP	[X	V	]$_{I'}$
	1	2	3	
	[α person]			
	[β number]			
SC:	1	2	3	
			[α person]	
			[β number]	

Like all languages, Azerbaijani has a number of verb classes. These are examples of the most common:

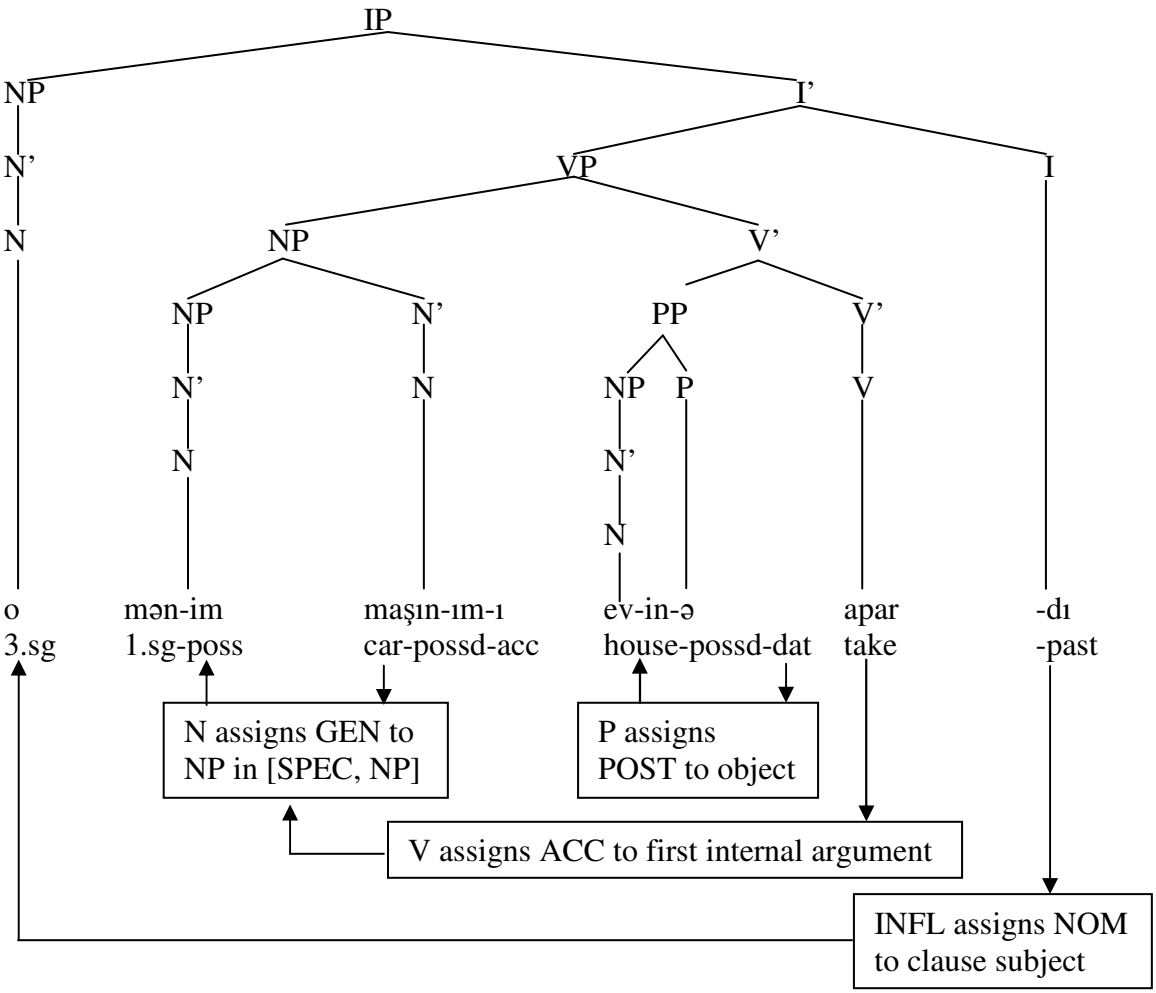
Intransitive	Transitive	Di-transitive
<durmaq> 'to stand' +V [__]	<bilmək> 'to know' +V +[NP/IP__]	<gətirmək> 'to bring' +V +[NP NP__]
<getmək> 'to go' +V [(NP)__]	<-dir> 'be' +V +[NP/PP/AP__]	<demək> 'to say' +V +[NP NP__]

The internal argument of <getmək> 'to go' will be assigned morphological dative case due to its theta role as [Goal]. The same is true of the second internal argument of <gətirmək> and <demək>.

Negation in Azerbaijani takes two forms. Non stative verbs are negated by the morpheme /-mA/ which is appended to the verb stem before any other suffixes. Stative clauses with the copula <-dir> are negated with the word <deyil> 'not' which fills the AP slot immediately preceding the cliticised verb.

3.2.3.3 Case

In Azerbaijani, abstract case is assigned by INFL (finite), V, P, and N. Finite INFL ([+Tense], [+Aspect]) assigns NOM case to clause subject NPs when it M-commands them, and there are no intervening barriers. Adjacency is not required.



'He took my car to his house.'

V assigns ACC to the NP in [NP, VP] position. Again, it must M-command the NP and there must be no intervening barriers. Adjacency is not required. P assigns POST (post-positional case) to its complement NP. M-command and adjacency are required. There must be no barriers. N assigns GEN to NPs in [Spec, NP] either by a peripheral rule, or by inherent case assignment. Azerbaijani does not seem to have any ECM verbs. As is common in languages with a rich system of morphological agreement, Azerbaijani is a *pro*-drop language. Thus, the subject NP that receives NOM case is often non-overt *pro*. See the example on the previous page.

3.2.3.4 Binding

As in all languages, in Azerbaijani an R-expression must be free everywhere, a pronoun must be bound in its governing category, and an anaphor must be bound in its governing category. The pronouns listed in Table 3-1 take case suffixes like nouns. The only anaphor in Azerbaijani is the reflexive pronoun <öz>+agr. It must agree with its antecedent in person and number utilizing the same agreement suffixes as Azerbaijani nouns (see Table 3-3 and Table 3-4). The agreement morphology may be followed by case or post-positional suffixes as required by the meaning of the sentence. In all cases, the governing category is the entire clause, containing the anaphor itself, the verb which governs it, and the subject of the verb to which it is bound. For example:

1. Elçin$_i$ kitab-ı öz$_i$-ü üçün[5] al-dı
   name book-acc self-(3.sg)-possd for buy-past-(3.sg)
   'Elchin bought the book for himself.'

[5] Though the suffix on the antecedent for *üçün* has been labeled as "possessed", there is some irregularity and variation especially in the 3rd person indicating a need for further research.

2. Elçin$_i$ kitab-ı o_j-nun üçün al-dı
   name book-acc 3.sg-possd for buy-past-(3.sg)
   'Elchin bought the book for him/her/it.'

In 1 the anaphor <öz> must be co-indexed with the clause subject, Elchin, while in 2, the pronoun <o> must not be co-indexed with the clause subject.

Like the English 'I myself', the reflexive pronoun in Azerbaijani also plays an emphatic role. This can be accounted for by a rule such as:

$$\begin{bmatrix} \text{N} \\ \alpha\ \text{pers} \\ \beta\ \text{numb} \end{bmatrix} + \text{emphatic} \rightarrow \text{N} + \text{öz} \begin{bmatrix} \alpha\ \text{pers} \\ \beta\ \text{numb} \end{bmatrix}$$

For example:

Mən öz-üm kitab-ı al-mış-am
1.sg self-1.sg book-acc buy-perf-1.sg
'I myself bought the book'

Azerbaijani also shows evidence of PRO – the non-overt subject of non-finite clauses. PRO is proposed as the subject of a clause where there is no finite INFL to assign NOM, *pro* is proposed where there is finite INFL, but an overt subject is lacking. PRO_{arb} may be posited where a generic expression is intended, and no co-indexation is possible with an overt NP in the entire sentence. It appears that all verbs in Azerbaijani are subject control so PRO must always be bound by the subject of the matrix clause. For example:

1. *pro* ged-ir-əm.
   go-pres-1.sg
   'I am going.'

2. [PRO$_i$ mən-i tut-an]$_{IP}$ it$_i$ yol-dan ged-ir.
   1.sg-acc bite-part dog road-abl go-pres-(3.sg)
   'The dog that bit me is going down the road.'

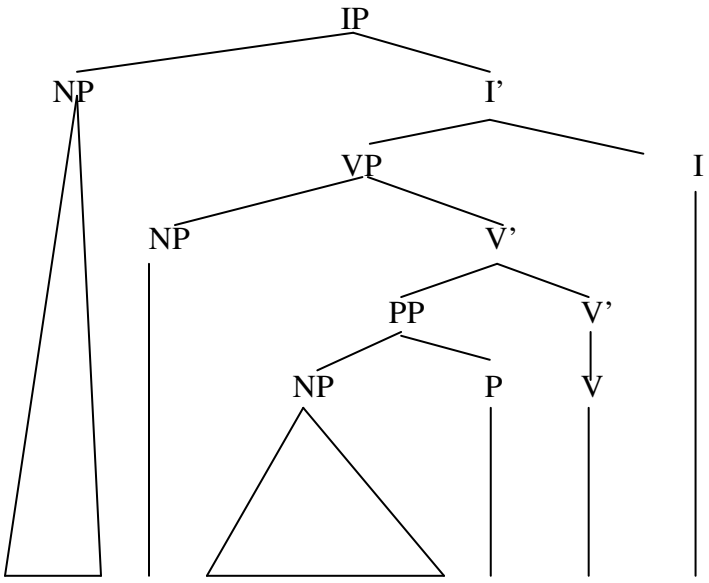
3. [PRO_{arb} yat-maq]$_{IP}$ yaxşı-dır
sleep-infin good-be
'Sleeping is good.'

However, example 3 is somewhat problematic. It is unclear if the suffix /-mAQ/ indicates a non-finite verb since it also functions as a noun forming suffix. While Example 3 may be an example of PRO_{arb}, the non-finite form of the verb also appears in contexts where it functions like a noun:

[Mən-im yat-mağ-ım]$_{NP}$ yaxşı-dır
1.sg-poss sleep-infin-possd good-be
'My sleep is good.'

It is quite likely that there are two /-mAQ/ suffixes in the Azerbaijani lexicon. One of them marks non-finite verbs, while the other is a noun-forming derivational suffix.

3.2.3.5 Movement



[Bu kitab]$_i$ t_i mən-im tərəf-im-dən oxu-nul -ur
det book 1.sg-poss side-possd-abl read-pass -pres-(3.sg)
'That book is being read by me.'

Azerbaijani exhibits both NP movement and WH movement. If we assume that like English, Azerbaijani passive verbs lose the ability to assign abstract ACC case to their internal arguments, then NP movement is required to move an NP from a position inside the VP headed by a passive verbs (where they cannot receive abstract case) to an empty [Spec, IP] position where finite INFL assigns abstract NOM. In the example on the previous page the surface subject <Bu kitab> has been moved from [Spec, VP] to [Spec, IP] where it can be assigned case.

However, as a result of other movement rules, word order is quite free in Azerbaijani. While the main verb is normally sentence final, other constituents can be moved to [Spec, CP] or [Spec, VP]. Thus, all of the following are possible:

1. Elçin adam-lar-ı Gəncə-də gör-dü
   name person-pl-acc Ganja-in see-past-(3.sg)
   'Elcin saw people in Ganja.' (basic word order)

2. Gəncə-də Elçin adam-lar-ı gör-dü.
   Ganja-in name person-pl-acc see-past-(3.sg)
   (adjunct PP moved to [Spec, CP])

3. Adam-lar-ı Elçin Gəncə-də gör-dü
   person-pl-acc name Ganja-in see-past-(3.sg)
   (object NP moved to [Spec, CP])

4. Elçin Gəncə-də adam-lar-ı gör-dü.
   name Ganja-in person-pl.-acc see-past-(3.sg)
   (adjucnt PP moved to [Spec, VP])

Presumably, constituents are moved for emphasis but native speaker intuitions as to the English translations of these sentences vary since much of the interpretation of Azerbaijani sentences is based on intonation. While there is no obligatory movement of question words in Azerbaijani, they are subject to the same free word order. To account

for these word-order variations, (Spec) positions were posited in the rules presented in section 3.2.3.2 . Thus, in 2 above the adjunct PP is moved to [Spec, CP]. A partial D-structure would be:

2. [[Gəncə-də]$_{PP\text{-}i}$[Elçin$_{NP}$ [*e* [[adam-lar-ı]$_{NP}$ [t_i gör]$_{V'}$] $_{V'}$]$_{VP}$ -dü]$_{C'}$]$_{CP}$
   Ganja-in name person-pl.-acc see -past-(3.sg)

In 3, the internal argument NP is moved out of the VP to [Spec, CP]:

3. [[adam-lar-ı]$_{NP\text{-}i}$ [Elçin$_{NP}$ [*e* [t_i [[Gəncə-də]$_{PP}$ gör]$_{V'}$]$_{V'}$]$_{VP}$-dü]$_{C'}$]$_{CP}$
   person-pl.-acc name Ganja-in see -past-(3.sg)

In 4, the adjunct PP is moved to [Spec, VP]:

4. [*e* [Elçin$_{NP}$ [[Gəncə-də] $_{PP\text{-}i}$ [[adam-lar-ı] $_{NP}$ [t_i gör]$_{V'}$]$_{V'}$]$_{VP}$-dü]$_{C'}$]$_{CP}$
   name Ganja-in person-pl.-acc see- past-(3.sg)

3.3 The Russian linguistic system

3.3.1 Morphological variation in Russian

Though Russian does not have a vowel harmony system like Azerbaijani, it does exhibit a tremendous amount of phonological variation in its inflectional morphology. For example, the Russian infinitival suffix exhibits 4 forms in the data collected for this project:

-ить	[i^{j}t]	пить	[pi^{j}t]	‘to drink’
-еть	[ɛ^{j}t]	смотреть	[smotrɛ^{j}t]	‘to look’
-ть	[a^{j}t]	играть	[igra^{j}t]	‘to play’
-ти	[ti]	идти	[idti]	‘to go one direction’

All inflectional morphology in Russian exhibits similar variation. The correct form usually depends on the phonology and class of the preceding stem, but can also be idiosyncratic. For the sake of brevity, only one of the possible forms will be listed in the inflectional paradigms in the following sections.

3.3.2 Word level phenomena: Russian morphology

3.3.2.1 Derivational morphology

While Azerbaijani exhibits some derivational morphology with suffixes to form nouns, adverbs, and adjectives, Russian makes more extensive use of derivation. According to Patrick, in Russian "prefixes alter the meanings of words, while suffixes alter their functions" (1989: 2). Most, if not all, prefixes in Russian have the same phonological form as prefixes in the language. Many times, these prefixes serve to narrow the definition of the word stem:

играть	[igrat]	'to play'
проиграть	[proigrat]	'to lose'
думать	[dumat]	'to think'
выдумать	[vɯdumat]	'to invent'

While Azerbaijani uses suffixes to show tense and aspect, Russian uses both prefixes and vowel changes to indicate perfective aspect. These changes are not predictable, and many times use the same prefixes are identical to the derivational prefixes mentioned above:

делать	[dʲɛlat]	'to do'
сделать	[sdʲɛlat]	'to do - perfective'
понимать	[ponimat]	'to understand'
понять	[ponyat]	'to understand - perfective'

In his section entitled "List of the Most Important Suffixes" Patrick details over 70 non-inflectional suffixes which can be added to a word root to change its function (1989: 3-8). For example:

Noun Forming			
	ездить	[jɛdit]	‘to ride’
	ездок	[jɛdok]	‘rider’
	петь	[pɛt]	‘to sing’
	певец	[pɛvets]	‘singer (masc.)
Diminutive			
	мама	[mama]	‘mama’
	мамаенька	[mamajɛnka]	‘dear mama’
	нож	[noʃ]	‘knife’
	ножик	[noʃik]	‘little knife’
Adjective Forming			
	берег	[bɛrɛg]	‘coast (noun)’
	береговой	[bɛrɛgovoj]	‘coastal (adj.)’
	болеть	[bolɛt]	‘to be sick (verb)’
	больной	[bolnoj]	‘sick (adj.)

3.3.2.2 Inflectional morphology

In Russian, verbs, nouns, and adjectives all take a variety of inflectional suffixes:

- Verbs – tense, agreement (person and number), and gender (past tense only)
- Nouns – number, and case
- Adjectives – gender, number, and case (to agree with the noun head of the NP)

Verb conjugation

The set of past tense suffixes is much simpler than other inflectional paradigms for Russian. However, unlike the present tense conjugation, the past tense suffixes agree with the clause subject in gender.

Table 3-9 Russian past tense suffixes

			Example with быть [bɯʸt] ‘to be’		
Masc. Sing	-л	[l]	был	[bɯl]	‘he was’
Fem. Sing.	-ла	[la]	была	[bɯla]	‘she was’
Neut. Sing.	-ло	[lo]	было	[bɯlo]	‘it was’
Plural	-ли	[li]	были	[bɯli]	‘they were’

The past tense suffixes appear most often with perfective verbs, though it is possible for them to occur with imperfective verbs to show ongoing action in the past.

The present tense verb conjugation paradigm closely resembles other Indo-European languages, encoding both person and number. Like other tables in this section, Table 3-10 gives an abbreviated list of possible present tense verb suffixes. In practice, the suffix will depend both on the phonology and verb class of the preceding stem.

Table 3-10 Russian present tense verb conjugation

1st Sing.	-ю/-у	[ju] or [u]	1st Plural	-ём	[jom]
2nd Sing.	-ёшь	[joʃ]	2nd Plural/Formal	-ёте	[jotɛ]
3rd Sing.	-ёт -	[jot]	3rd Plural	-ют/-ат	[jut] or [at]

Example with пить [pit] 'to drink'			
	пью	[pju]	'I drink'
	пьёш	[pjoʃ]	'you (sing.) drink'
	пьёт	[pjot]	'he/she/it drinks'
	пьём	[pjom]	'we drink'
	пьёте	[pjotɛ]	'you (pl./form.) drink'
	пьют	[pjut]	'they drink'

Adjective declension

Russian has two sets of adjective declensions: short form and long form. The short form declension appears when the adjective is in attributive position: AP is complement of VP. Short form adjectives agree with the subject of the IP in gender and number, but they do not show case.

Table 3-11 Russian short form adjectival suffixes

			Example with белый [belɯy] 'white'		
Masc. Sing	-	n/a	бел	[bjɛl]	'he (is) white'
Fem. Sing.	-а	[a]	бела	[bjɛla]	'she (is) white'
Neut. Sing.	-о	[o]	бело	[bjɛlo]	'it (is) white'
Plural	-ы	[ɯ]	белы	[bjɛlɯ]	'they (are) white'

Since Russian nouns can carry masculine, neuter, or feminine gender, and long form adjectives agree with the noun they modify in gender, number, and case, the paradigms for long form adjective declension are quite complex. This combined with the Russian tendency for irregularity based on noun class, there is a tremendous amount of morphological variation in adjectival suffixes – Leed lists 8 paradigms for adjectives and several others for 'special' adjectives such as possessives and demonstratives.

Table 3-12 Russian long form adjectival suffixes (unstressed basic and palatalized paradigms based on Leed 1986)

	Basic unstressed paradigm based on белый [belɯy] 'white'							
	Masc. Sing.		Neut. Sing.		Fem. Sing.		Plural	
Nominative	белый	[bɛlɯj]	белое	[bɛlojɛ]	белая	[belaja]	белые	[belɯjɛ]
Accusative	*		=masc.		белую	[bɛluju]	*	
Genitive	белого	[bɛlovo]	=masc.		белой	[bɛloj]	белых	[bɛlɯx]
Prepositional	белом	[bɛlom]	=masc.		"		"	
Dative	белому	[bɛlomu]	=masc.		"		белым	[bɛlɯm]
Instrumental	белым	[bɛlɯm]	=masc.		"		белыми	[bɛlɯmi]
	Palatalized unstressed paradigm based on синий [sinij] 'blue'							
	Masc. Sing.		Neut. Sing.		Fem. Sing.		Plural	
Nominative	синий	[sinij]	синее	[sinjɛjɛ]	синяя	[sinjaja]	сине	[sinjɛ]
Accusative	*		=masc.		синюю	[sinjuju]	*	
Genitive	синего	[sinjɛvo]	=masc.		синей	[sinɛj]	синих	[sinix]
Prepositional	синем	[sinjɛm]	=masc.		"		"	
Dative	синему	[sinjɛmu]	=masc.		"		синим	[sinim]
Instrumental	синим	[sinim]	=masc.		"		синими	[sinimi]
* If the noun is inanimate the accusative form is the same as the Nominative. If it is animate, it is the same as the Genitive.								

3.3.3 Russian syntax

Russian is head first and spec first (other than VP). Heads occur at the beginning of their phrases, but movement can be both toward the front and back of the sentence.

3.3.3.1 Theta theory

In Russian, verbs, nouns, and prepositions have the ability to assign theta roles. However in many stative sentences, there is no overt verb. Rather than viewing other elements as being predicates, the present analysis posits a non-overt copula. Here is a partial list of verb theta grids for Russian:

1. Ø 'be': V (non-overt copula)

NP	NP/PP/AP
Topic	Comment

2. быть [bɯ^{y}t] 'to be': V

NP	NP/PP/AP
Topic	Comment

3. плыть [plɯ^{y}t] 'to swim': V

NP
Agent

4. покупать [pokupa^{y}t] ': V

NP	NP
Agent	Patient

5. давать [dava^{y}t] 'to give': V

NP	NP	NP
Agent	Patient	Goal

In the data to be analyzed in the following chapters, no Russian nouns seem to require complement arguments, but non-proper nouns may occur with a possessor. Thus, most Russian nouns would have theta grids like this:

дом [dom] 'house':N

(NP)
Possessor

Since Russian has both count and non-count nouns, at least 3 noun classes are possible based on the features [count] and [arg.]:

[± count] identifies whether the noun does/does not have a plural form
[± arg.] identifies whether the noun occurs with an external argument (usually Possessor)

Features:		Examples:
[-count]	[-arg.]	Саша [saʃa] a name
[+count]	[+arg.]	яблоко [jabloko] ‘apple’
[-count]	[+arg.]	молоко [moloko] ‘milk’

Саша [saʃa] a name +N +[_]	яблоко [jabloko] ‘apple’ +N +[(det)(numb)(adj)_]	молоко [moloko] ‘milk’ +N +[(det)(adj)_]

Theta grids:

Саша [saʃa] a name: N

молоко [moloko] ‘milk’: N

(NP)
Possessor

яблоко [jabloko] ‘apple’: N

(NP)
Possessor

Lastly, prepositions also assign theta roles to their complement NPs. Here are theta grids for four of the most common prepositions:

1. в [v] ‘in’: P

NP
Location

2. у [u] ‘at’: P

NP
Location

3. на [na] ‘on’: P

NP
Location

4. с [s] ‘with’: P

NP
Instrument

Unlike English, German, and Azerbaijani, Russian does not have special morphology or syntax to encode passives. Rather, passive meaning is conveyed by either 3rd person constructions and/or reflexives (Borras and Christian 1969: 168-169).

3.3.3.2 X bar theory

Phrasal heads in Russian include:

P	head of PP (Prepositional Phrase)
A	head of AP (Adjective Phrase)
Adv	head of AdvP (Adverb Phrase)
D	Head of DP (Determiner Phrase)
N	head of NP (Noun Phrase)
V	head of VP (Verb Phrase)
I	head of IP
C	head of CP

The following phrase structure rules seem to account for the vast Russian sentences to be analyzed:

PP: PP → P NP — в Америке [vamerikyɛ] ‘in America’

AP: AP → (AdvP) A’ — очень хороший [otʃɛn xoroʃij] ‘very good’
A’ → A (A’)

AdvP: AdvP → (Intens) Adv’
Adv’ → Adv (Adv’)

DP: DP → D’ (Num)
D’ → D

NP: NP → (DP) N’
N’ → (AP) N’
N’ → N { NP, PP, IP }

эта	очень	красивая	девушка
[ɛta	otʃɛn	krasivaya	dʲɛvoʃka]
that	very	pretty	girl

VP: VP → V’ (Spec) — (Spec position posited for movement)
V’ (AdvP) V’
V’ → V’ { (PP), (NP), (IP), (CP) }
V’ → V { (NP), (AP) }

IP:	IP	→	(Spec)	I’	(Spec position posited for movement)
	I’	→	NP	I’	
	I’	→	I	VP	
CP:	CP	→	(Spec)	C’	(Spec position posited for movement)
	C’	→	C’	(IP)	
	C’		C	IP	

Since these rules place I to the left of the verb, a rule is needed to copy the [tns] feature from I and put it on the verb. Moreover, finite verbs in Russian must agree with the subject in number, person, and for past tense gender. Two rules are required:

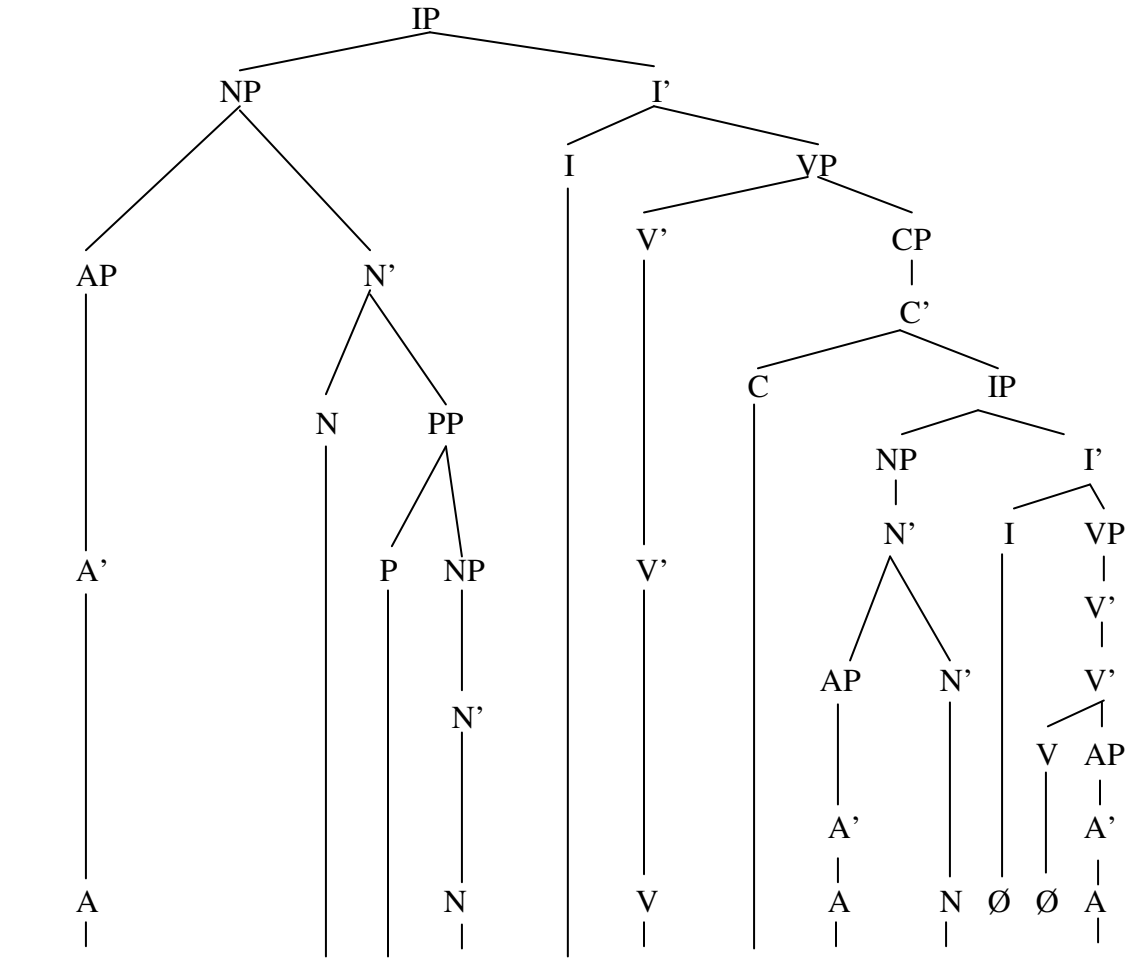
<u>Tense copying</u>

SD:	I	[	X	V	]$_{VP}$
	1		2	3	
	[α tns]				
SC:	1		2	3	
				[α tns]	

<u>Verb agreement</u>

SD:	NP	[	X	V	]$_{I'}$
	1		2	3	
	[α person]				
	[β number]				
	[γ gender]				
SC:	1		2	3	
				[α person]	
				[β number]	
				[γ gender]	

This example demonstrates a number of these rules:



As shown by the theta grids above, Russian has a number of verb classes. The following syntactic frames are examples of a few of the possible classes:

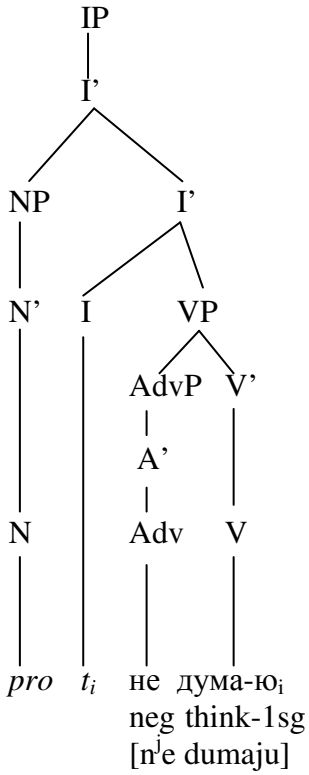
Intransitive	Transitive	Di-transitive
плыть [plɯt] ‘to swim’ +V [__]	знать [znat] ‘to know’ +V +[__NP/IP]	давать [davat] ‘to give’ +V +[__ NP NP]
смотреть [smotrjet] ‘to look’ +V +[__ (PP)]	Ø ‘be’ +V +[__NP/PP/AP]	сказать [skazat] ‘to say’ +V +[__NP PP]

The second internal argument of давать [davat] would receive inherent dative case due to its Goal theta role.

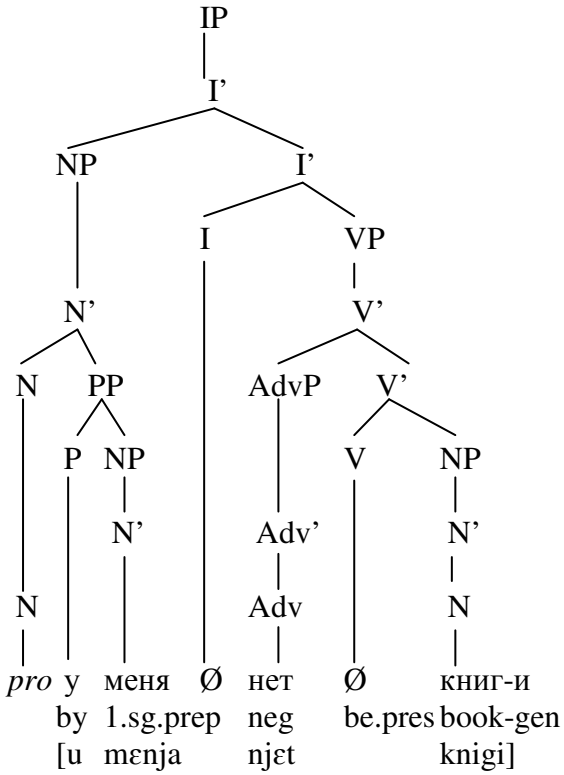
Russian has two forms of negation. The particles не [nʲe] and нет [nʲet] function as Adverbs but do not co-occur with any intensifiers:

не [nʲe] 'not' +Adv [__]	нет [nʲet] 'not' +Adv [__]

These two, however do not occur in the same contexts. Не [nʲe] occurs in clauses with full verbs, whereas нет [nʲet] occurs either alone or in clauses with a non-overt copula. For example:



'I do not think so.'



'I do not have any books.'

Question: чай хоч-ешь? "Do you want (some) tea?"
[tʃaj xotʃ-ɛʃ]
tea want-2.sg.inf

Reply 1:	не	хоч-у	“I do not want (any).”
	neg	want.1sg	
	[nʲɛ xotʃu]		
Reply 2:	нет		“No”
	neg		
	[nʲɛt]		

3.3.3.3 Case

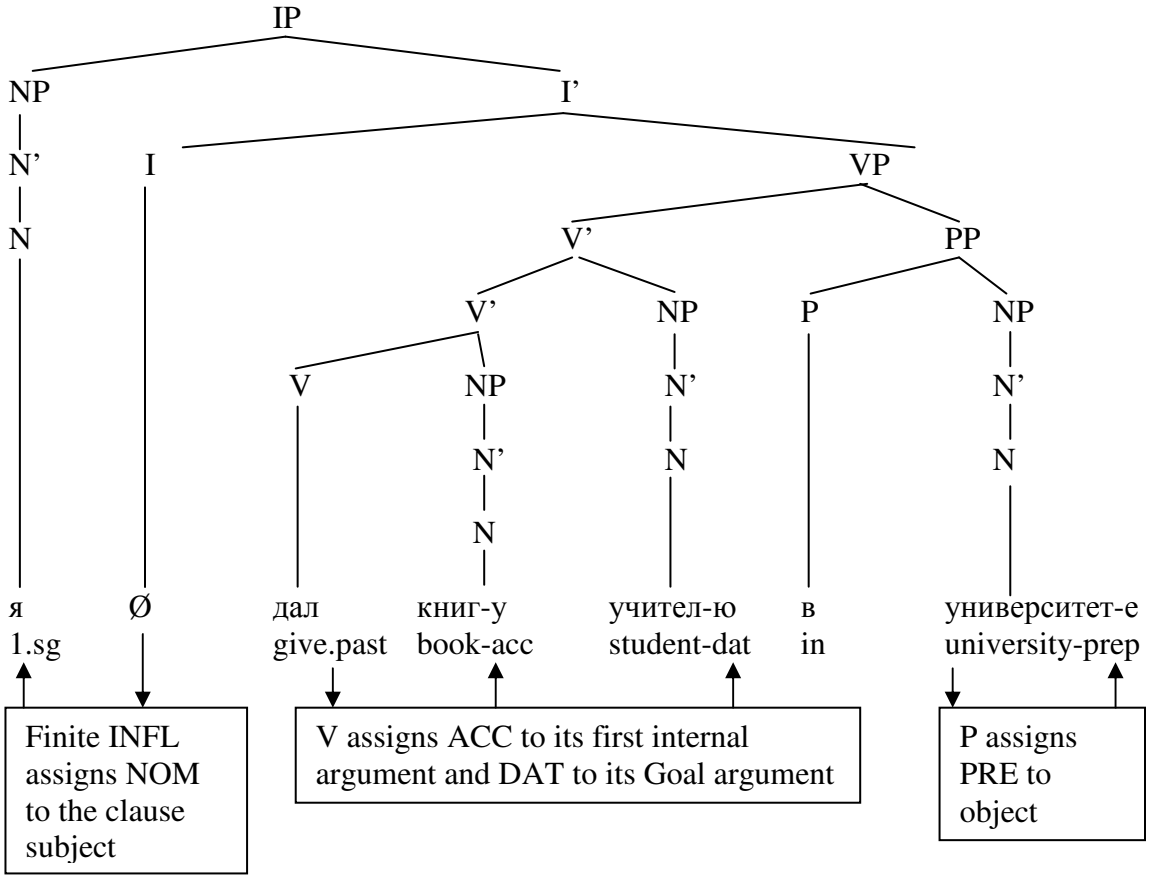
In Russian INFL (finite), V, P, and N assign abstract case. Finite INFL ([+Tense]) assigns NOM case to clause subjects when it M-commands them, and there are no intervening barriers. Adjacency is required. V assigns ACC to its internal argument NP. It must M-command the NP and there must be no intervening barriers. Adjacency is required. P assigns abstract PRE (prepositional case) to its NP object. M-command and adjacency are required. N assigns GEN to its internal NP Possessor argument either by a peripheral rule, or by inherent case assignment. Three-argument verbs such as давать [davat] ‘to give’ mentioned above also exhibit inherent case, assigning DAT (dative case) to the second internal Goal argument. Unlike Azerbaijani, NP complements of prepositions in Russian show morphological case. Which morphological case a particular preposition assigns would be part of the individual lexical entry. The data to be analyzed for this dissertation do not show evidence of ECM verbs in Russian. As is common in languages with strong systems of morphological agreement, Russian is a *pro*-drop language. Thus, the subject NP that receives NOM case is often non-overt *pro*.

Determiners and adjectives must agree with the head N of NP in number, gender, and case. The following rules account for this:

Determiner Agreement						Adjective Agreement						
SD: [	D	X	N	]$_{NP}$		SD: [	X	A	X	N	X	]$_{NP}$
	1	2	3				1	2	3	4	5	
			[α number]							[α number]		
			[β gender]							[β gender]		
			[γ case]							[γ case]		
SC:	1	2	3			SC:	1	2	3	4	5	
	[α number]							[α number]				
	[β gender]							[β gender]				
	[γ case]							[γ case]				

This last rule applies equally to all adjectives within the NP.

Example:



[ja dal knigu utʃitɛlju v univɛritʲɛtʲɛ]
'I gave the book to a student at the university.'

3.3.3.4 Binding

As in all languages, in Russian an R-expression is free everywhere, a pronoun must be bound in its governing category, and an anaphor must be bound in its governing category. Russian appears to have two anaphors:

-ся/-сь [cja/^{j}c] a cliticized anaphor which is attached to the verb following all tense/person morphology. It is not declined like other NPs, has both reflexive and reciprocal implications, and is always co-referenced with the clause subject.

себя [sebya] is declined like noun for case and number (see Leed 1986: 211). It only appears as an internal argument in VP, and is always co-referenced with the clause subject. Boris and Christian refer to this as 'the reflexive pronoun' (1984: 286-287).

Though often translated as himself/herself/itself, the 'special adjective' сам [sam] does not appear to be an anaphor. It is declined like an adjective (Leed 1986: 208, 315). Since it always occurs inside NP and agrees with the head N in number, gender, and case, it should be classified as an adjective, probably with a literal meaning of 'the same':

сам	король	мне	сказал	'The king himself told me.'
same	king	1sg.dat	say.past	(Boris & Christian 1984: 317)

The following examples demonstrate the co-reference requirements for -ся/-сь [cja/^{j}c] and себя [sebja]:

1. они$_i$ целова-л-и-сь$_i$
   3pl.nom kiss-past-pl-ref
   [oni tselovalis]
   'They kissed each other.'

2. он$_i$ целова-л его$_k$
   3sg.m.nom kiss-past 3sg.m.acc
   [on tseloval ^{j}evo]
   'He kissed him.'

3. учител-я$_i$ да-л-и книг-и себе$_i$
   student-pl.nom give-past-pl book-acc.pl self.dat
   [utʃitelja dali knigi s^{j}eb^{j}e]
   'The students gave books to themselves.'

4. учител-я$_i$ да-л-и книг-и ним$_k$
   student-pl.nom give-past-pl book-acc.pl 3.pl.dat
   [utʃitelja dali knigi nim]
   'The students gave books to them'

In 1 -ся/-сь [cja/c] must be co-indexed with the clause subject они [oni] 'they' whereas in 2 the direct object and subject must not be co-indexed. Similarly, in 3 себя [s^{j}ebja] must be co-indexed with the clause subject учителя [utʃitelja] 'students' whereas in 4 the direct object pronoun ним [nim] 'to them' must not be co-indexed with an NP within the sentence.

Like Azerbaijani, Russian appears to have PRO which functions as the non-overt subject of non-finite clauses. This analysis proposes PRO as the subject of any clause where there is no finite INFL to assign NOM; *pro* is proposed where there is finite INFL but no overt NP subject is present. PRO$_{arb}$ is posited where generic expression is intended and no co-indexation with an overt NP in the sentence is possible. Only subject control verbs are present in the data for this project, so this analysis assumes that PRO is always bound by the subject of the matrix clause. For example:

1. *pro* ид-у дом-ой
   go-1sg house-dat(irreg)
   [idu domoj]
   'I am going home.'

2. я$_i$ хоч-у [PRO$_i$ спа-ть]$_{IP}$
   1.sg want-1.sg sleep-infin
   [ja xotʃu spat]
   'I want to sleep.'

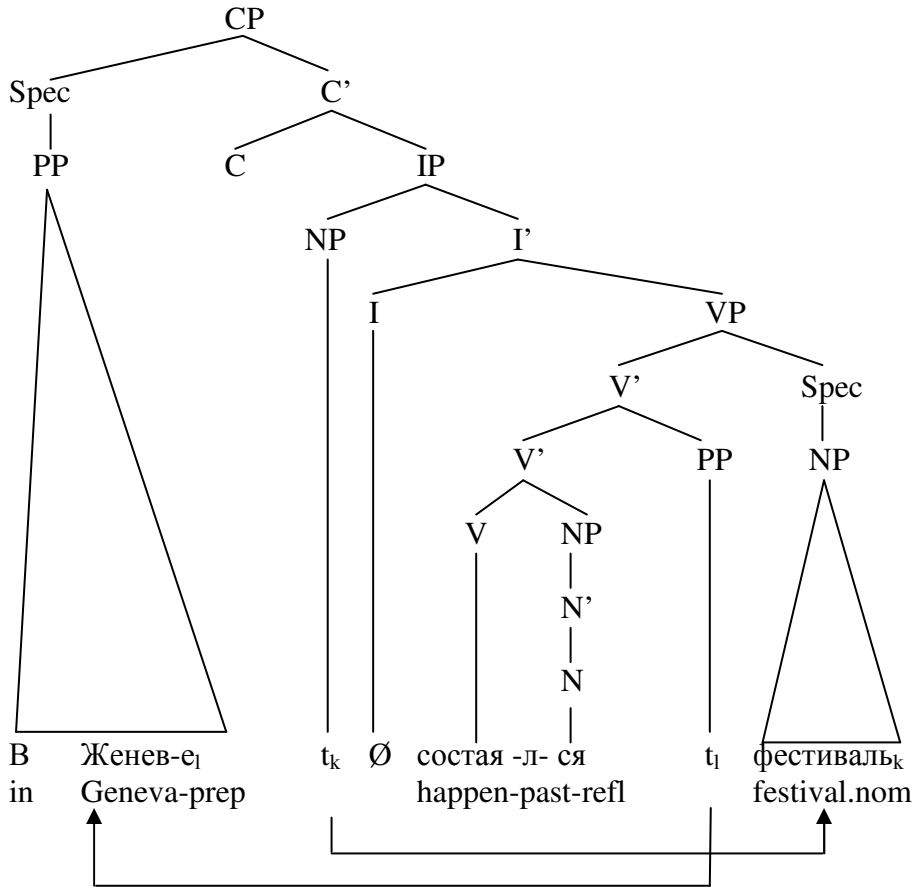
3. [PRO$_{arb}$ спа-ть]$_{IP}$ Ø хорошо
   sleep-infin be good
   [spat xoroʃo]
   'It is good to sleep'

3.3.3.5 Movement

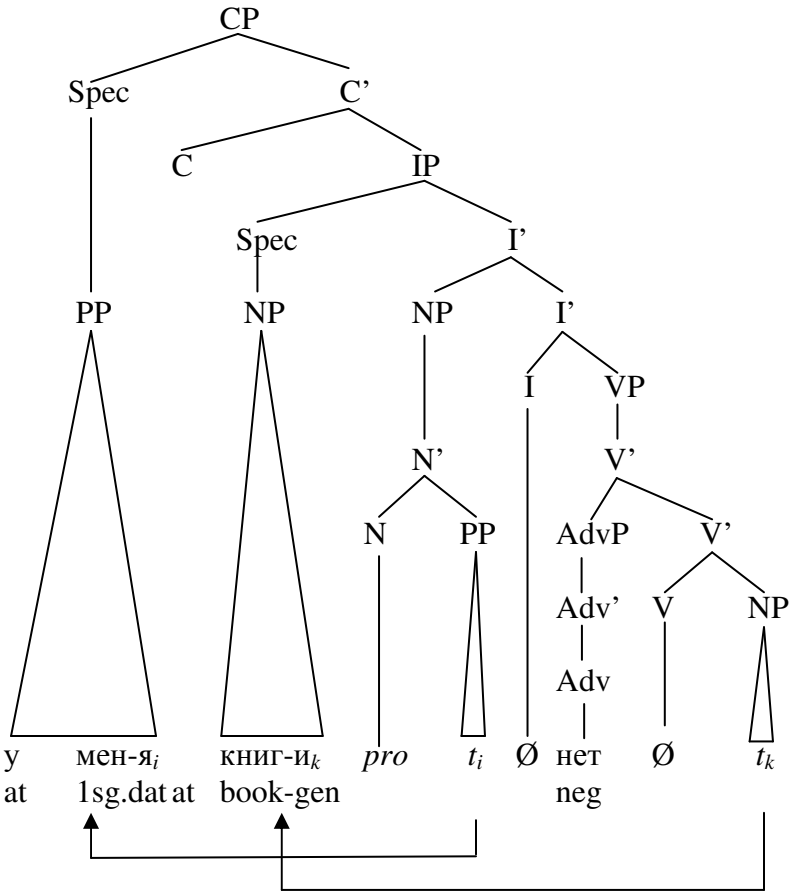
With extensive inflectional morphology, Russian word order is quite free. Thus, Movement is common in Russian. Since it does not appear to have any verbs without the ability to assign abstract ACC case to internal arguments (typical of passive verbs in English and Azerbaijani), all of this must be explained as Wh movement. As shown earlier, the X Bar rules for VP, IP, and CP contain Spec positions as landing spots for moved constituents:

VP:	VP	→	V'	(Spec)	(Spec position posited for movement)
IP:	IP	→	(Spec)	I'	(Spec position posited for movement)
CP:	CP	→	(Spec)	C'	(Spec position posited for movement)

While these positions will allow virtually any sentence word order, according to Wade SVO is the most common arrangement (2000: 525). However, unlike Azerbaijani, information that is to be brought into focus (information new to the discourse) is moved to the end of the sentence. "The order of 'given' information + 'new' information (with less essential preceding essential new items) is standard in a Russian sentence" (Wade 2000: 522). Wade lists the following example (here diagrammed according to the phrase structure rules discussed above):



[v ʒjenevje sostajalsja festival]
‘A festival took place in Geneva.’



[u menja knigi nʲet]
'I do not have any books.'

Thus in Russian, any NP, PP, can be moved into any of the three positions: [Spec, CP], [Spec, IP], or [VP, Spec] depending on whether the discourse deems the information in the phrase to be 'new' or 'given'.

3.4 Congruence, contrast, and potential code-switching and code-mixing

The purpose of the preceding sections outlining the Azerbaijani and Russian linguistic systems was to provide the necessary foundation to predict ways in which these two languages might interact within a conversation, within conversational turns and within clauses (CPs). A preliminary list of the ways in which this may happen follows

organized according to Muysken's typology of code-switching/mixing: alternation, insertion, and congruent lexicalization.

3.4.1 Alternation between Azerbaijani and Russian

In a nation like Azerbaijan with a long history of language contact, language alternation within a conversation should be expected. There is nothing about the two linguistic systems that would inhibit a switch in linguistic code between conversational turns. Rather than a reflection of linguistic issues, the presence or absence of switches between Azerbaijani and Russian corresponding to turn changes would be either a result of political realities, and/or demonstrations of local identity construction in progress. Moreover, since many in Azerbaijan are fluent in both languages, alternation between CPs within a conversational turn could also be common.

More closely related to the two linguistic systems would be alternation within CPs. Muysken lists a number of features of language alternation (2000: 231). Those that pertain to alternation within the CP cluster into four groups:

1. Peripheral elements: emblematic switching or switching of tags, switching adverbs or conjunctions, and flagging.
2. Large sections: several constituents in a row, and switching long or complex constituents
3. Revision: switching as self-correction, and doubling
4. Linear equivalence: when both languages call for the same ordering of clause and/or phrase constituents, alternation could occur without being able to determine which language provides the basic framework.

The switching of peripheral elements in CPs is quite possible for Azerbaijani and Russian. Elements such as adverbs (ну [nu] 'well', vay 'oh no', etc.), tags (значет [znatʃɛt] 'that means', bilersən 'you know', etc.), and conjunctions are often analyzed as filling Spec positions. Both Azerbaijani and Russian are Spec first languages and have initial [Spec, CP] and [Spec, IP] positions. In addition, Russian has a CP final [Spec; VP] position. The congruence between the two linguistic systems would favor this type of alternation, though the position of the peripheral elements would be of interest since Azerbaijani would only allow such switches to occur before the sentence, while Russian would allow them sentence finally as well (the [Spec; VP] position).

The switching of complex constituents or several constituents in a row is less likely for Azerbaijani and Russian than for alternation in peripheral elements. While both languages allow IP in NP and VP, Azerbaijani's avoidance of relative clauses would limit the frequency of this type of alternation. We would expect switched non-finite clauses, and complex NPs to be most frequent while relative clauses would occur less frequently. In most cases the non-finite clauses and NPs would be best analyzed as insertion rather than alternation.

Alternation as revision both for self-correction and doubling would be possible in Azerbaijan. If a Russian speaker 'accidentally' produced a Russian element in conversation with a non-Russian speaker, they might well re-cast it in Azerbaijani to facilitate communication. In a situation where both speakers are sufficiently proficient in Russian, the same behavior could serve as a contextualization cue to emphasize the content of the doubled element. This would not, however, fit with the phrase structure

rules of either language. Presumably, the doubled element would be inserted as a right sister of the element it mimics.

Though both Azerbaijani and Russian allow extensive movement which could result in surface linear equivalence, the inflectional morphology of the two languages should make it possible to determine which language is providing the overall structure of the sentence (Matrix Language). Moreover, since Azerbaijani is a head last language and Russian is head first, no equivalence could ever occur within pre/post-positional phrases, or within complex NPs or VPs.

3.4.2 Insertional mixing of Azerbaijani and Russian

Of the three types of code-switching/mixing listed by Muysken, insertion appears most likely to occur within Azerbaijani/Russian CPs. According to Muysken, features to indentify insertion are (2000: 231): switching of a single constituent or content word, dummy word insertion, and switching of a selected element (larger than a constituent). Nested A B A sequences and morphological integration are also indicators of insertional mixing. Muysken also classifies 'telegraphic mixing' as insertional.

The insertion of a content word or whole constituent from one language into a CP from the other language is very likely. Since the phonological forms of the Azerbaijani inflectional morphemes depend only on syntactic function and the phonology of the preceding word stem, non-Azerbaijani words should easily integrate into Azerbaijani CPs. However, the more complex and irregular Russian morphological system would make it less likely for Azerbaijani elements to be incorporated into Russian CPs. While the form of most Russian suffixes is dependent on the phonological form of the preceding

stem, most suffixes have multiple variants which are also influenced by the word class that precedes them. To integrate an Azerbaijani element, a speaker would have to search the mental lexicon for an analogous Russian word in order to determine the correct forms for inflectional morphology. This added processing expense should make Azerbaijani insertions less likely, but not impossible. Phonological integration into Russian would solve this problem, but would also be a signal of more permanent borrowing either at the societal or individual level.

Other forms of insertion should occur equally in both directions. Dummy word insertion ('do X') would not be dependent on morphology, nor would insertion of larger elements or nested A B A sequences. Telegraphic mixing, where words or constituents from both languages occur adjacent to each other without a clearly defined clause structure, is quite likely in conversation. However, since these mixed pieces of speech are not clauses, they would not qualify as code-switching to researchers such as Myers-Scotton (2002) who see the CP as the appropriate domain of analysis.

3.4.3 Congruent lexicalization

In congruent lexicalization, it is impossible to determine which language is providing the structure for the CP. Rather both languages can be seen as providing the CP structure together while both provide the lexical items to fill out the surface string. Muysken (2000: 123) proposes that congruent lexicalization may occur because:

a) There is an overabundance of homophonous words, dimorphs, that serve as bridges or triggers for the code-mix;

b) There is a general structural equivalence, both categorical and linear, making code-mixing possible, without there necessarily being any lexical correspondence.

While Azerbaijani has borrowed vocabulary from Russian resulting in numerous homophones, these constitute a fairly small percentage of the words in Azerbaijani. As mentioned earlier, movement rules make a certain amount of linear equivalence possible between Azerbaijani and Russian, but the divergent morphological systems make it unlikely that an analyst will have difficulty determining which language is providing the grammatical matrix for a CP. Thus, it seems congruent lexicalization is unlikely between Azerbaijani and Russian.

Thus, in the analysis of recorded conversation that follows, a range of code-switching and code-mixing behavior would be expected. This may include alternation between conversational turns, between CPs within the same turn, and in peripheral elements of CPs. Insertion should also be common, and will probably involve a range of word classes. Lastly, given the verbal morphology of both languages, congruent lexicalization should not occur.

CHAPTER 4

CODE-SWITCHING AND CODE-MIXING IN AZERBAIJAN

4.1 Data collection and methodology

4.1.1 Data types

Several types of data were collected in preparation for this dissertation project: recorded conversations, group and individual interviews, as well as observations of public language behavior. However in order to narrow the scope, this analysis will focus predominantly on the recorded conversational data. This includes two types: recordings of staged conversations and of freely occurring language mostly in a home setting.

4.1.1.1 Staged conversations

Staged conversational data were collected on May 30, 2007. Students from the Azerbaijan University of Languages were asked to meet in the American Studies Center at the University to participate in this research project. Some of the students knew each other, while others did not. The students filled out demographic surveys, and were given a list of suggested conversation topics. Since previous research (Zuercher 2004 and 2009) showed that language variation in Azerbaijan was based on gender and language of education, the aim of this session was to record conversations between subjects from the different groups defined by those factors: male/Azerbaijani educated, male/Russian educated, female/Azerbaijani educated, and female/Russian educated. The students were assigned to conversational pairs; each pair was given a digital recorder, and asked to talk

for 15 to 20 minutes based on the list of conversational topics. While these were not fully transcribed, they provide data to analyze code-switching between conversational turns.

4.1.1.2 'Home' recordings

In order to collect naturally occurring language data, a number of subjects were asked to take a digital recorder and microphone home with them and record themselves interacting with their family and friends.

Table 4-1 Summary of subjects and data collected in the Home recordings

Name	**Sex**	**Age**	**Context**	**Total usable recorded time**
Aner	male	47	at home with wife and family	28 min (14 recordings, 4 usable)
Aygun	female	28	making/eating dinner with roommate and sister	45 min (3 recordings)
Agil	male	32	various contexts at work	1 hour 2 min (7 recordings)
Lale	female	33	at home with her daughter (9 years old) and mother	43 min (1 recording)
Mahir	male	32	at home with wife and family	26 min (8 recordings, 7 usable)
Shahin	male	20	playing video games with a friend at an Internet club	20 min (1 recording)
Sevil	female	44	at home with her young daughter (5 years old)	1 hour 53 min (1 recording)
Tarana	female	32	making/eating dinner with her sister (Aygun) and roommate	1 hour 15 min (1 recording)

This data collection method has a number of advantages: it provided a large corpus for analysis, numerous conversational topics and roles are evident in the recordings, and since most of the recordings are quite long they can be assumed to represent 'natural' language as the subjects seem to have forgotten that the recorder was running. It does, however, also have some disadvantages. Only one side of the conversations was recorded though in many cases the other speaker can be generally identified by voice

quality (male, female, adult, child, etc.). One subject's family (Shahin) forbade him from bringing the recorder home so he recorded himself playing video games with a friend in an internet club. All of the subjects involved in this phase of data collection were educated in Russian, though their language backgrounds vary in other ways. Table 4-1 summarizes the subjects and data collectedHere is a summary the subjects and data collected:

4.1.2 Analytical methodology

4.1.2.1 Staged conversations

Since the Staged conversations were not fully transcribed (other than one segment discussed below) the analysis procedure for them was quite simple. A file was created for each conversation noting the speaker and language for each conversational turn. If any type of code-switching or language mixing was detected in a turn, a note was made. Code-switching or language mixing within conversational turns was, however, quite rare in these recordings. The information from the demographic survey was coded and input in Microsoft Excel.

4.1.2.2 Home recordings

Because the bulk of the analysis in this dissertation is based on the Home recordings, they were analyzed in much greater detail. The initial step was to fully transcribe each recording. As described in Chapter 2, this was done using standard orthography and spelling for each language. Thus colloquial pronunciation of Azerbaijani words or phonological variation was lost, but the presence or absence of all

morphology was preserved. After the initial transcription, each text was checked in detail by a native Azerbaijani speaker with near native fluency in Russian.

The next step in analysis was to convert the checked transcripts into tagged text. Each text was parsed using SIL's Toolbox program. This allowed all morphemes to be regularized (each variant converted into a base form) and labeled according to language and type (noun, verb, agreement morpheme, etc.). By exporting to plain text and using Excel, the Morpheme Break (\mb) and Part of Speech (\ps) lines of the Toolbox output were merged. The following excerpt from Aner's conversation shows the stages in this process:

Original Text: крах olub, elə deyil?

Toolbox Output:

\ref	Aner.149					
\tx	крах	olub	,	elə	deyil	?
\mb	крах	ol -miş	,	elə	deyil	?
\ge	crash	be -perf	,	thus	not	?
\ps	Rnm	Av -Aasp	,	Aadv	Aneg	?

Tagged Text: крах_Rnm ol_Av -miş_-Aasp ,_, elə_Aadv deyil_Aneg ?_?

Thus, <крах> [krax] is tagged as 'Rnm' to show that it is a Russian language item, a noun, and carries masculine gender. The Azerbaijani verb stem <ol> meaning 'be' is tagged as 'Av': Azerbaijani language, and a verb. The variant form of the Azerbaijani perfective morpheme <-ub> is converted to the base form <-miş> and tagged as 'Aasp': Azerbaijani aspect. <Elə> the Azerbaijani adverb meaning 'thus' is tagged as 'Aadv' and the negative <deyil> is tagged as 'Aneg.' While such a tagged text may not be easily used for conversation analysis or examining the use of morphological variants (like <-ub> and <-miş>), it is very easy to use for corpus analysis.

As an intial analysis, Wordsmith was used to determine the relative frequency of Azerbaijani and Russian content words (noun, verb, adjective, and adverb) in each recording. The results appear in Table 4-2.

Table 4-2 Initial analysis: Azerbaijani and Russian content words

Name	**Azerbaijani content words**	**Russian content words**
Aner	54.7% n=268	44.7% n=219
Aygun	75.4% n=1278	21.0% n=356
Agil	52.2% n=609	46.2% n=539
Lale	46.2% n=493	52.9% n=564
Mahir	45.8% n=261	53.2% n=303
Shahin	1.5% n=5	97.2% n=318
Sevil	88.8% n=1412	11.2% n=178
Tarana	53.7% n=447	52.4% n=537

The next stage of analysis was to take the tagged text for each recording andbreak conversational turns into clauses (CPs) and analyze these for code-switching. For this and the subsequent stages of analysis, Excel was used so that different types of CPs could be labeled and sorted according to various factors. The tagged text was pasted into Microsoft Excel. Each CP within a conversational turn was then broken out into a separate row. These CPs were labeled Y or N based on the presence or absence of code-switching/code-mixing (cs/cm). The CPs with code-switching/mixing were then examined individually and labeled according to the type/s of switching or mixing present.

4.1.2.3 Ten minute samples

After initial analysis, it became clear that a more balanced and thoroughly tagged sample from each participant was needed to allow for analysis of certain clause types, movement, and comparison between subjects. Thus, a ten minute sample from each participant in the Home recordings was selected. Optimally a segment toward the end of

a recording was chosen in order to get the most natural data, but since some subjects made several shorter recordings instead of one long one, the length of time from the beginning of the recording to the beginning of the sample varied greatly. Table 4-3 lists the subjects, recordings, and section selected.

Table 4-3 List of seletions for the ten minute samples corpus

Name	**Recording number**	**Minutes selected**
Aner	Recording 6 of 14	Minutes 10 to 20
Aygun	Recording 1 of 3	Minutes 10 to 20
Agil	Recording 6 of 7	Minutes 5 to 15
Lale	Recording 1 of 1	Minutes 25 to 35
Mahir	Recording 6 of 8	Minutes 0 to 10
Shahin	Recording 1 of 1	Minutes 10 to 20
Sevil	Recording 1 of 1	Minutes 1:30 to 1:40
Tarana	Recording 1 of 1	Minutes 50 to 1:00

To these samples from the Home recordings, a full transcription of one participant from the Staged conversations was added. This was necessary since all participants in the Home recordings were Russian educated and had other significant predictors of Russian language use and code-switching/mixing. If the samples from only these participants were analyzed it would be impossible to say if the phenomena in question were due to Russian language influence, or not. The subject chosen was Tamam: an Azerbaijani educated female with no Russian relatives who claimed to know only a few Russian words and phrases.

4.2 Code-switching: Language alternation

The first type of code-switching to be examined is alternation. According to Muysken (2000) this can occur between conversational turns, between CPs within a conversational turn, or with a CP itself.

4.2.1 Alternation between conversational turns

To examine alternation between conversational turns, the Staged conversational data will be used. The Home recordings would not be well suited to this since only one side of the conversation was recorded. Language alternation between conversational turns could be expected in two situations: full fluency in both languages by both speakers, or differential fluency with competition or negotiation for linguistic dominance. Given the vast differences in linguistic background for the subjects in these conversations, both of these situations could be expected.

The range of these backgrounds can be seen by examining the subjects' answers to the language background section of the demographic survey (see Appendix A). This consisted of five questions:

What language did you learn first as a child?
What language did you speak at home with your family as a child?
What sector did you attend in elementary school?
What sector did you attend in high school?
What sector did you attend in university?

Table 4-4 Staged conversation participants from most Azerbaijani (A) to Russian (R) influence

Name	**Sex**	**Age**	**First Lang**	**Home Lang**	**Elem Lang**	**HS Lang**	**Univ Lang**
Kamala	female	22	A	A	A	A	A
Tamam	female	21	A	A	A	A	A
Vugar	male	24	A	A	A	A	A
Adil	male	25	A	A	R	A	A
Vahid	male	23	A	A	R	R	R
Shahin	male	20	R	Both	R	R	R
Leyla	female	19	R	R	R	R	R
Zarifa	female	18	R	R	R	R	R

In Table 4-4 the eight participants in the Staged recordings are arranged from most Azerbaijani influence to most Russian influence.

Language alternation between conversational turns would be most expected when the participants are on opposite ends of this scale. However, alternation between turns is rare in these conversations. The majority of them begin with a brief discussion of which language the interlocutors should speak, and then they proceed in that language. Twelve conversations were recorded. Given the language influences of the speakers, these conversations could fall into four categories:

1. Those between speakers with high Azerbaijani influence,
2. Those between speakers with high Russian Influence,
3. Those between speakers at opposite ends of the scale, and
4. Conversations involving Vahid who claimed Azerbaijani as both first and home language, but attended all Russian schools.

Of the twelve, two conversations were between speakers with high Azerbaijani influence: Adil/Kamala who spoke Azerbaijani exclusively with one mixed turn by Adil, and Tamam/Vugar who spoke only Azerbaijani. Only one conversation was between speakers at the bottom of Table 4-4: Shahin and Zarifa spoke a few turns in English then switched to Russian for the remainder of the recording.

Five conversations were recorded between speakers at opposite ends of Table 4-4, as listed in Table 4-5. Rather than exhibiting competition for linguistic dominance, these conversations overwhelmingly show linguistic cooperation and accommodation. In four of these five conversations, the Russian speakers defer to the Azerbaijani speakers by

speaking Azerbaijani. This linguistic accommodation is difficult for Zarifa when talking with Tamam, but she makes it work. In the case of Vugar and Zarifa, the two speakers simply agreed to speak a third language: English.

Table 4-5 Description of conversations between subjects at opposite ends of Table 4-4

Speakers	**Description of Conversation**
Adil (male) Leyla (female)	Spoke Russian at the beginning, then switched to Azerbaijani. Both had a few mixed turns
Shahin (male) Tamam (female)	All turns in Azerbaijani. Tamam had two turns with Russian insertions – listing words and phrases she knew.
Shahin (male) Vugar (male)	All turns in Azerbaijani.
Kamala (female) Leyla (female)	All turns in Azerbaijani. Leyla had one turn with Russian insertions.
Tamam (female) Zarifa (female)	English at the beginning both with the researcher then and then with each other trying to negotiate the language. Finding that Tamam's English was weak, they switched to Azerbaijani for the remainder of the recording with Zarifa speaking haltingly and with a few English and Russian insertions.
Vugar (male) Zarifa (female)	All turns in English.

Three of the twelve conversations involved Vahid who showed mixed language influences.

Table 4-6 Description of conversations involving Vahid

Speakers	**Description of Conversation**
Adil (male) Vahid (male)	Very short conversation. All turns in Russian with some discussion of which language to speak at the beginning.
Kamala (female) Vahid (male)	All turns in Azerbaijani.
Leyla (female) Vahid (male)	A few turns in English and mixed Azerbaijani/Russian at the beginning. Switch to Russian, then back to Azerbaijani for the last third of the turns.

As expected from his answers to the language background survey questions, Vahid shows a great deal of variation in his language use. Apparently Adil is either comfortable in

Russian as well as Azerbaijani or simply wants to show off his language skills. With Vahid he spoke only Russian. With Kamala, a speaker with high Azerbaijani influences, Vahid speaks only Azerbaijani. Though Leyla is at the bottom of Table 4-4, she also has great variation in her language use. With Kamala she spoke Azerbaijani, with Adil Russian, and in this conversation with Vahid she switches from a mixture of Azerbaijani and English, to Russian, then back to Azerbaijani. It is interesting to note that Vahid was the first to speak Russian, initiating the first switch, but then also the first to speak Azerbaijani initiating the second.

In all twelve of the staged conversations, alternation of languages between conversational turns is much less frequent than might be expected. This may be due to several factors. All of the participants are students at the same university, though from different language sectors. Thus, they may feel a great degree of solidarity. Moreover, this was a fairly formal setting since many of the interlocutors had not met each other ahead of time, it was on the university campus, and an older foreign researcher was observing. Alternation and linguistic competition might be more common in less formal and monitored settings where each speaker feels free to try and force his or her language ideology on the other speaker. Lastly, in contemporary Azerbaijan public language policy greatly encourages Azerbaijani usage. Thus Russian speakers may defer to Azerbaijani speakers because that is what is generally expected of them in the society at large.

4.2.2 Alternation between CPs within conversational turns

In addition to alternation between speakers across conversational turns, alternation is likely within conversational turns. All eight subjects in the Home recorded data exhibited this type of code-switching, but to different extents.

Table 4-7 Number of turns with alternation between CPs

Name	**Total turns**	**Turns w/ CS/CM**	**Turns w/ alternation between CPs**	**Percentage of total turns with alternation between CPs**
Aner	204	20	3	1.5%
Aygun	366	119	38	10.4%
Agil	373	72	8	2.1%
Lale	404	134	23	5.7%
Mahir	219	73	9	4.1%
Shahin	182	16	1	0.5%
Sevil	652	86	10	1.5%
Tarana	336	108	41	12.2%

The speaker with the highest rate of alternation between CPs in a conversational turn was Tarana at 12.2% of her total turns. Of 108 turns with some kind of code-switching/mixing, 41 (38%) contained CPs from different languages. Shahin showed the lowest rate of alternation between CPs in conversational turns at 0.5% of his total turns. However, the percentage of turns with alternation between CPs is not a very meaningful measure. For this type of alternation to occur, at least two CPs are required within the turn, so it is more likely in long turns than in shorter ones. The length of a conversational turn depends on a number of factors other than subject demographics or identity construction: topic, age of interlocutor, type of discourse (narrative, interrogation, etc.), conversational role being played out, etc. Though we do not have enough information from the data presented above to draw conclusions about why alternation within a

conversational turn occurs, or what pragmatic roles it might play, we can see that it is common. Here are two examples:

Lale #223	CP1	И	успокои-л-а-сь,			
		and	calm.down.perf-past-f-refl			
		[i	uspokojlas]			
		'And calm down'				
	CP2	sonra	Азюка-нынbir dənə	игрушечн- ый	зайчик-i	var
		later	Name-possoneclass	toy- m.s.nom	hare-poss	be
			[azjuka]	[igruʃɛtʃnɯj	zajtʃik]	
		'Later, Aziyuka's (daughter), there will be a toy hare.'				
Agil #41	CP1	mən-də	şükür Allah-a	yaxşı-lıq-dır,		
		1.sg-loc	thanks God-dat	good-noun.form-be		
		'With me everythings good, thank God'				
	CP2	всё	нормальн-о			
		all.pl.nom	normal-adv			
		[vsyo	normalno]			
		'Everything's good'				

4.2.3 Alternation within CPs

Muysken (2000) suggests at least four ways that alternation can occur within clauses: peripheral elements in a different language than the main part of the clause, large sections of a clause in a different language, revision or self correction in a different language, and linear equivalence. All but the last occur in the current corpus.

4.2.3.1 Peripheral elements

Code-switching of peripheral elements takes place at the edges of the clause: the switched elements do not participate in the argument structure of the clause, but rather fill [Spec; X] slots where they modify the entire clause to some extent. Peripheral alternation could take three forms:

a) emblematic switching or switching of tags,

b) switching adverbs or conjunctions,

c) and flagging

The first and last of these are rare in the Home recordings corpora. Subjects almost never oriented to their code-switches by flagging them with hedges, pauses, or adverbial elements, but make smooth, fluid transitions between languages. While emblematic switching or switching of tags is present, it is also not common. Eight times давай [davaj] 'allow (it)' precedes an Azerbaijani clause, three times значит [znatʃit] 'it means', and once the stative clause ой это что-то [oj ɛto ʃto to] 'Oh! that's something':

Sevil #1057	дава-й	gətir
	allow-imper	take
	[davaj]	
	'OK, take (it)!'	

Aner #248	знач-ит,	manat-dır
	mean-3sg	manat-be
	[znatʃit]	
	'That means it is manat.'	

Aygun #21	ой	это	что-то,	on-a	görə	bela	bax-ır
	Oh!	that.sg.nom	something	3sg.dat	because	thus	look-pres
	[oj ɛto ʃtoto]						
	'Oh that's something, that's why he/she/it looks like that'						

In contrast, alternation in conjunctions and peripheral adverbs is quite common. In the Home recorded corpus as a whole, conjunctions appear in a contrasting language 36 times: 35 times a Russian conjunction precedes an Azerbaijani CP, and once an Azerbaijani conjunction precedes a Russian clause. Peripheral adverbs appear in a contrasting language 73 times: 29 times an Azerbaijani adverb is associated with a Russian clause, and 44 times a Russian adverb is appended to an Azerbaijani clause.

Unlike the preceding types of peripheral alternation, however, adverbs both precede and follow the clauses to which they are appended.

Table 4-8 Count of Azerbaijani and Russian adverbs with Azerbaijani and Russian CPs

	A adv with R CP	**R adv with A CP**
precede	25	34
follow	4	10
total	29	44

Examples:

Lale #305 или heç bax-ma-dı
or nothing see-neg-past
[ili]
'Or he/she/it did not see anything.'

Aygun #695 уже qurtar-dı
now finish-past
[uʒɛ]
'Now it is finished.'

Tarana #437 ye-din уже?
eat-2.inf.past now
[uʒɛ]
'Did you eat already?'

Mahir #113 hə я пи-ть хоч-у
yes 1sg pour-inf want-1sg
[ja pit xotʃu]
'Yes I want to pour (something out).'

The fact that Azerbaijani adverbs both precede and follow Russian clauses is no surprise since Russian has a [Spec, CP] slot at the beginning of the clause and a [Spec, VP] slot at the end to allow for the various forms of movement that the language exhibits. However, Russian adverbs appearing at the end of Azerbaijani clauses present a problem for the X' rules presented in Chapter 3.

Azerbaijani movement rules allow for the fronting of discourse salient information to a [Spec, CP] position – the position also occupied by code-switched peripheral elements such as adverbs. Therefore moved elements and peripheral adverbial elements should only appear at the front of the sentence. If adverbial elements appear after the verb, then it may be appropriate to posit a CP final Spec position. This raises several questions. Are other elements moved to the end of the sentence? Would the backing of discourse salient information in clauses with Azerbaijani as the Matrix Language is itself a form of language mixing? Would this represent interference between Russian and Azerbaijani or is it evidence for a blended Azerbaijani/Russian language in the mental grammars of bilingual Azerbaijani/Russian speakers? To answer these questions we must see whether backing occurs when no Russian lexical items are involved, and if a speaker with very low Russian influence (such as Tamam) presents the same type of movement.

Table 4-9 Azerbaijani CPs with movement past the verb

Name	**A CPs with no R lexical items**	**Verb not final**	**Percentage**
Aner	34	9	26.5%
Aygun	62	15	24.2%
Agil	n/a	n/a	n/a
Lale	36	8	22.2%
Mahir	30	10	33.3%
Shahin	n/a	n/a	n/a
Sevil	68	11	16.2%
Tarana	32	10	31.3%
Tamam	124	12	9.7%

To examine this phenomenon we must use the more thoroughly tagged 10 minute samples from each Home recording participant and from Tamam in the staged

recordings. Table 4-9 lists the number of CPs with Azerbaijani as the matrix language and with no Russian lexical items, and the number of these which have an element other than the verb in final positions. Excluded from this number are CPs with final relative clauses, tag questions, or vocatives. Two subjects (Agil and Shahin) had no CPs with Azerbaijani as the matrix language in the 10 minute sample, so they contribute nothing to this examination. All the other subjects, however, including Tamam showed movement of various CP components to final position. These included subject NPs, object NPs, adjuncts (NP and PP), and adverbs. Tamam shows the same types of movement with object NPs, adjuncts (NP and PP), and adverbs in CP final position. Thus, the Azerbaijani X' rules presented in Chapter 3 must be modified to accommodate movement both to the front and back of the CP, and therefore peripheral alternation after the verb. The best way to accomplish this would be to modify CP to allow both Spec first and Spec last positions:

CP:	CP	$\rightarrow$	(Spec)	C'	– Spec position proposed in Chapter 3
	CP	$\rightarrow$	C'	(Spec)	– Spec position required for rearward movement and peripheral alternation at the end of the clause

Though the Azerbaijani X' rules must be modified to allow movement of this type, we cannot rule out the possibility that, since the speakers with high Russian influence show this type of movement to a greater extent, it could also represents interference from Russian. It is entirely possible that since this type of movement is much more common in Russian discourse than in Azerbaijani discourse, it could be over-used in Azerbaijani by speakers accustomed to Russian language norms. This topic will be examined in Chapter 6 (Russian-Like Movement - 6.6).

Another phenomenon closely related to code-switched adverbs and conjunctions is the appearance of discourse markers in a contrasting language. In the Home recording corpus, the Azerbaijani discourse markers <ay> 'oh!' and <ey> 'look here' precede Russian clauses three times and the Russian discourse markers ну [nu] 'well' and ой [oj] 'oh!' precede Azerbaijani CPs three times as well.

4.2.3.2 Large sections

In the Home recording data there are a few examples of alternation between main clause and relative clause. This appears to be the only place that CP internal alternation of large sections occurs. To analyze this phenomenon we turn first to the corpus of 10 minute samples. In this data set, there are 47 relative clauses including 6 from Tamam. Of these, 4 show alternation of language between main clause and relative clause. In 3 cases a Russian relative clause follows an Azerbaijani matrix clause, and once an Azerbaijani relative clause follows a Russian matrix clause. Some relative clauses are preceded by an overt complementizer (n=36), but others are not (n=11). Table 4-10 summarizes these by direction of language alternation.

Table 4-10 Summary of relative clauses in the 10 minute samples

Matrix clause → Relative clause	**Number**	**Complementizer**
Azerbaijani → Azerbaijani	15	14 with <ki>, six of these from Tamam
Azerbaijani → Russian	3	no complementizers
Russian → Russian	28	21 with complementizers, 11 of these use что [ʃto]
Russian → Azerbaijani	1	что [ʃto]

In these 47 examples of relative clauses in the 10 minute samples, a number of verbs recur:

Verb	Gloss	Occurrences with Rel C in 10 min samples	
demək	'to say'	7	4 Azerbaijani, 3 Russian
говорить [govorit]	'to say (imperfective)'	2	both Russian
сказать [skazat]	'to say (perfective)'	4	1 Azerbaijani, 3 Russian
знать [znat]	'to know (imperfective)'	7	all Russian

However, this is only an analysis of a small data sample. If we fill out the list of word types likely to occur with relative clauses to include semantic and syntactic equivalents in both languages, a corpus analysis of all the Home recordings transcriptions is possible. Since Russian verbs occur in pairs (imperfective and perfective) we should add the perfective equivalent of *знать* [znat] 'to know' to the list: *узнать* [uznat] 'to know (perfective)'. We should also add the equivalent verb in Azerbaijani: *bilmək* 'to know'. By examining these six verbs and two complementzers (*ki* and *что* [tʃo]) we can see how common language alternation across a relative clause boundary is in the Home recordings corpora.

Table 4-11 Alternation across relative clause boundaries in the Home recordings

Word	Gloss	Type	N	N w/ alternation
ki		A comp	131	13
что [tʃo]		R comp	53	1
demək	'to say'	A verb	302	12
bilmək	'to know'	A verb	137	1
говорить [govorit]	'to say (imp)'	R verb	33	0
сказать [skazat]	'to say (perf)'	R verb	20	1
знать [znat]	'to know (imp)'	R verb	48	0
узнать [uznat]	'to know (perf)'	R verb	1	0

Thus we see that alternation between matrix and relative clause is possible, but not common. It is more common, however for a Russian relative clause to follow an

Azerbaijani matrix clause. Moreover, a Russian relative clause is much more likely to follow the Azerbaijani complementizer than the other way around.

4.2.3.3 Revision

While duplication of peripheral elements either by way of self correction, revision, or emphasis should be possible, it occurs very rarely in the Home recordings data. In the following example, it appears that Aner repeats phrases meaning 'why' in different languages to express incredulity:

Aner #478	Почему,	nə-yə	görə,	скаж-и	свою	мысль
	why	what-dat	according.to	say-2.sg.imp	own.f.s.acc	thought
	[potʃɛmu]			[skaʒi	svoju	mɯsl]

'What? What? Say what you think!'

4.2.3.4 Linear equivalence

Under linear equivalence both languages would call for the same ordering of clause and/or phrase constituents. In such a situation, it could be difficult or impossible to determine which language provides the overall matrix structure for a CP, and alternation could occur between the two languages in a somewhat random fashion. Given the morphological systems of Azerbaijani and Russian, such a situation would be unlikely when an overt verb is present. However, numerous stative clauses appear either with or without overt copular verbs and with lexical items from one, the other, or both languages. Though this may represent linear equivalence, this dissertation will treat it as congruent lexicalization. It is discussed later in this chapter.

4.3 Code-mixing: Insertion

In the Home recordings corpora, insertion is by far the most common form of code-mixing. This includes insertion of individual lexical items, verbs with 'dummy'

supporting verbs in the matrix language, and multiple word constituents. There are also a few cases of telegraphic mixing.

4.3.1 Insertion of single content words

While insertion occurs more frequently than other forms of code-switching/code-mixing, by far the most common form of insertion in this corpus is that of single content words. The Home recordings corpus includes both examples of Russian lexical items being inserted into Azerbaijani CPs and vice-versa. Table 4-12 shows the frequency of such insertions.

Table 4-12 Insertion of single content words

Type of Word	**Ru insertions into Az CPs**	**Az insertions into Ru CPs**
Common Noun	110	21
Name	7	15
Pronoun	14	2
Adjective	11	2
Number	3	n/a
Adverb (non-perif)	44	1

It is worthy of note that other than names, Russian insertions into Azerbaijani CPs far exceed Azerbaijani insertions into Russian CPs.

While both languages would call for insertions to be morphologically incorporated into the matrix clause, inserted items occur both with and without ML morphology. The following examples illustrate the various possibilities:

Russian noun inserted into an Azerbaijani CP with Azerbaijani case morphology

Lale #213	bun-un	da	зайчик-dən	ürəy-i	get-di
	this-3.sg.poss	therefore	hare-abl	heart-acc	go-3.sg.past
			[zaytʃik]		

'This one, then, the heart is going out of the hare (beating wildly).'

Russian noun inserted into an Azerbaijani CP without Azerbaijani case morphology

Agil #151	bir	<u>история</u>	danış-ım	
	one	story	say-1.sg.im	(Accusative morphology absent)
		[istorija]		
	'I am going to tell a story.'			

Azerbaijani noun inserted into a Russian CP with Russian nominal morphology

Aner #298	эти	представител-и	не	спрашива-ют
	these	representative-pl	neg	ask-3.pl
	[ɛti	prɛdstavetli	nɛ	spraʃivajut
	сколько	<u>müəllim</u>-ы	получа-ют?	
	how.many	teacher-pl	receive-3.pl	
	skolko	-ɯ	polutʃayut]	

'These representatives did not ask how many teachers they would get?'

One possibility missing from the data is an Azerbaijani noun inserted into a Russian CP without case morphology. While Azerbaijani nouns are inserted into Russian CPs 21 times in the Home recording corpora, when they appear without nominal morphology, they are in positions where no morphology would be expected: either as the subject, or inanimate nouns as objects (which do not receive any accusative morphology).

4.3.2 Dummy word constructions – insertion of non-finite verbs

Another form of insertional code-mixing common in the literature is dummy word constructions. In these constructions a non-finite verb from one language is inserted into a verb phrase from the other with a 'dummy' supporting verb. In the Home recordings corpus non-finite Russian verbs are inserted into Azerbaijani verb phrases 13 times. 9 times these occur with the Azerbaijani helping verb *etmək* 'to do', once with a colloquial variant *eləmək* 'to do', and three times with olunmaq 'to make.' There are, however, no

examples of Azerbaijani non-finite verbs being inserted into Russian verb phrases. For example:

Mahir #63	nə	ed-ir	bu	Samirə	bax,	<u>бомб-ить</u>	ed-ir?
	what	do-cont	this	name	look.imper	bomb-infin [bombit]	do-cont

'Look Samira, what is this one doing, is he bombing?'

Agil #711	просто [prosto]	bu this	mal goods	<u>занос-ить-ся</u> [zanositsja]	olun-ur make-cont	ora there

'These products are simply carried there.'

4.3.3 Insertion of multiple word constituents

Analogous to single content word insertion is the insertion of multiple word constituents. In a number of cases larger units (phrases) are inserted into CPs governed by the other language. Though less common than single word insertion, the pattern is quite similar. Russian phrases are inserted into Azerbaijani clauses more often than Azerbaijani phrases are inserted into Russian CPs.

Table 4-13 Insertion of multiple word constituents

Type of Phrase	**Ru insertions into Az CPs**	**Az insertions into Ru CPs**
Adjective Phrase	2	1
Noun Phrase	16	5
Prepositional Phrase	6	2

4.3.4 Telegraphic mixing

Muysken also lists telegraphic mixing as a case of insertional code-mixing. A number of conversational turns in the Home recordings corpus are not CPs since they do not have either overt or implied verbal heads. Rather they are a series of words and/or phrases salient to the conversation strung together. These often contain lexical items from both Azerbaijani and Russian. Here are two examples:

Sevil #114	sonra,	kaşa	sonra	günorta	bəs	обед	vaxt-ı
	after	porridge	after	mid.day	oh	dinner	time-3.sg.possd
						[obɛd]	

'later, porridge after mid-day, oh dinner time'

Lale #181	mən	уже
	1.sg	now
		[uʒɛ]

'now me'

4.4 Congruent lexicalization: Stative/copular clauses

Thus far the sections of this chapter have been in line with the predictions presented in chapter 3 based on Muysken's typology of code-switching/mixing. That chapter predicted that congruent lexicalization of Azerbaijani and Russian was unlikely. Nevertheless, congruent lexicalization does appear in the data to a limited extent.

The two situations in which Muysken predicted congruent lexicalization were where:

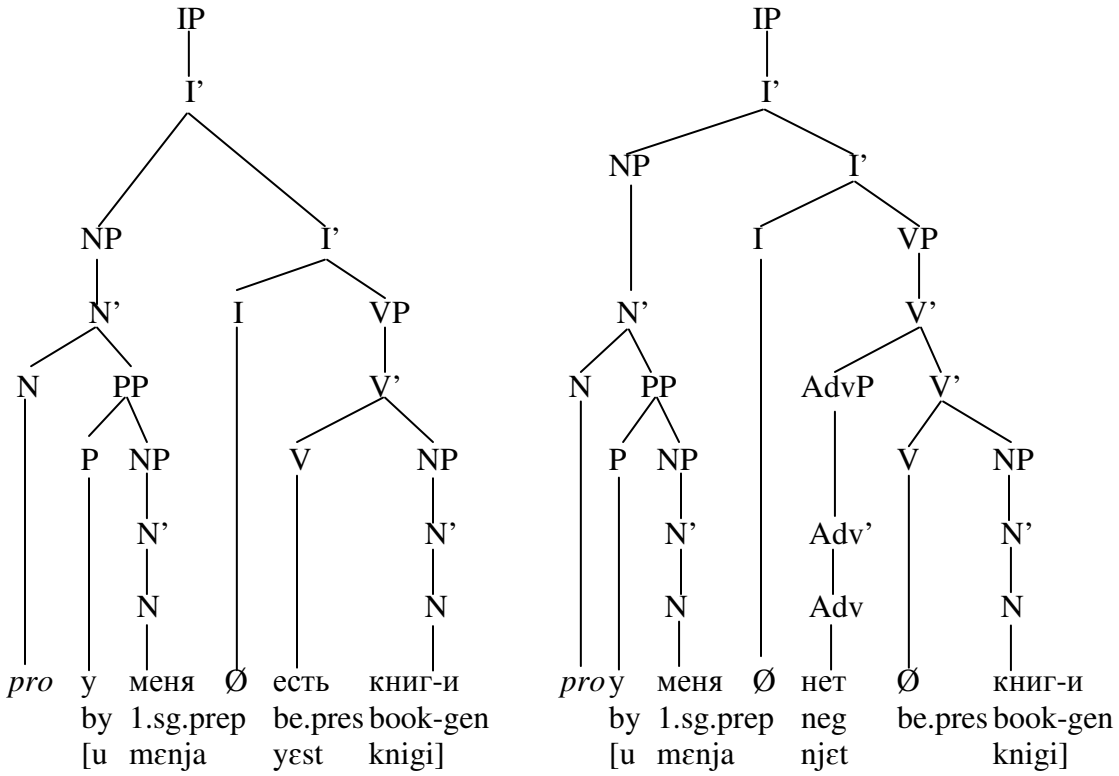
a) There is an overabundance of homophonous words, dimorphs, that serve as bridges or triggers for the code-mix;

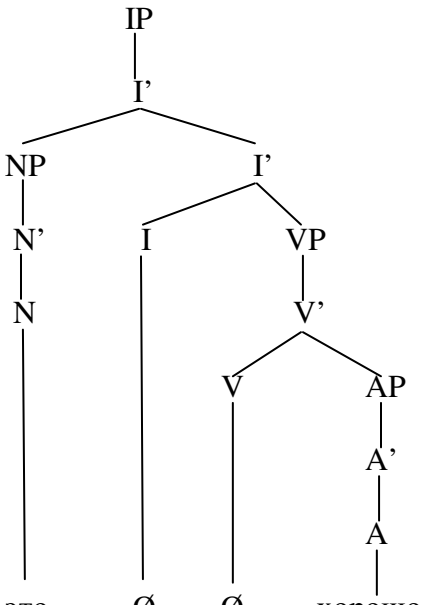
b) There is a general structural equivalence, both categorical and linear, making code-mixing possible, without there necessarily being any lexical correspondence (2000: 123).

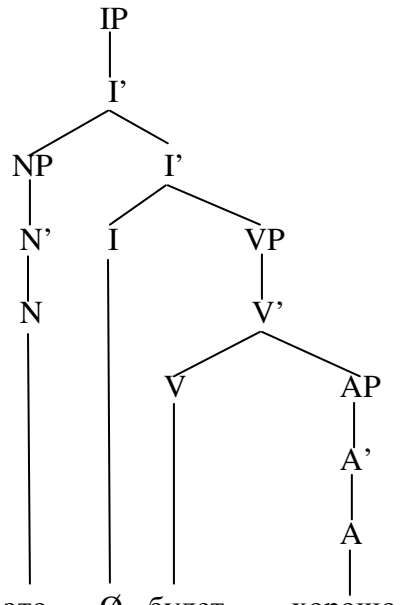
Though Azerbaijani has been in contact with Russian for several centuries, lexical items originating from Persian and Arabic are far more common than those from Russian. While movement rules can result in the same surface ordering of constituents, the morphological systems of the languages should make it clear which language is providing the matrix of the CP. However, a close examination of the Home recording data shows

some ambiguous cases. In particular, it is difficult to assign a matrix language to some stative or copular clauses

In the present tense, Russian often utilizes a non-overt copula. Though a present tense be verb, *есть* [jest], exists, its use is restricted to set phrases, or as a form of emphasis. In contrast, past tense and future forms of *есть* [jest] are obligatory. For example:







Descriptions of formal Azerbaijani, on the other hand, always call for an overt verb in each clause. Azerbaijani has three different stative verbs with slightly different theta grids and therefore different co-location restrictions:

/-dIr/ 'be': V

NP	NP/PP/AP
Topic	Comment

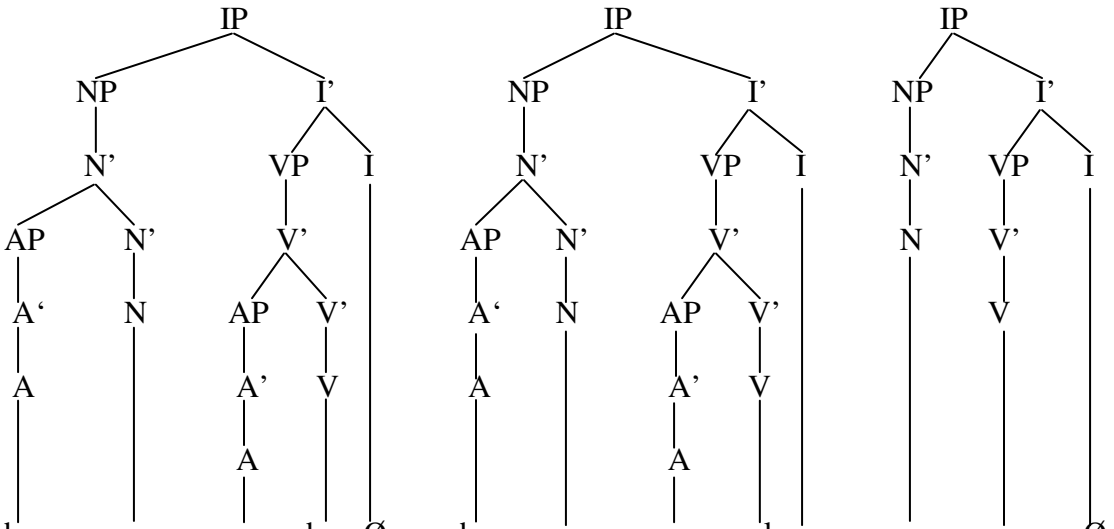
<ol-> 'be': V

NP	NP/PP/AP
Topic	Comment

<var> 'there is': V

NP
Topic

/-dIr/ only occurs in present tense while <ol-> has a wider distribution including some set phrases such as *sağ ol* 'thank you' (literally 'be healthy'), future tense, and past tense. For example:



Based on this analysis we would expect to find the following types of stative clauses

1. All Russian words with an overt Russian copula
2. All Russian words with a non-overt Russian copula
3. All Azerbaijani words with an overt copula
4. Mixed Azerbaijani and Russian with an overt copula from Russian or Azerbaijani which would reveal the matrix language
5. Mixed Azerbaijani and Russian without an overt copula. Since only Russian has a non-overt copula, it would be assumed to be the matrix language

If clauses occurred entirely in Azerbaijani without an overt copula, Russian could be assumed the matrix language with Azerbaijani lexical items filling all the terminal nodes – a kind of covert code-mixing. However, we would only expect subjects with access to the Russian lexicon to use the Russian non-overt copula, so we would only expect

Russian speakers to produce this type of CP. In looking at the more thoroughly tagged 10 minute samples corpus for the fluent Russian speakers, we do find all the types of clauses expected plus stative clauses with all Azerbaijani lexical items and no overt copula:

Table 4-14 Stative clauses for Home recordings subjects in 10 minute samples

	With an overt verb	**Without an overt verb**	**% without an overt verb**
All Az words	71	50	41.3%
All Ru words	37	141	79.2%
Both Az and Ru words	18 Az 4 Ru	26	54.2%

However, Tamam, a non-Russian speaker, also delivers stative clauses entirely in Azerbaijani without an overt copula, though to a lesser extent:

With an overt verb	48
Without an overt verb	7
% without an overt verb	12.7%

Thus, to the list of Azerbaijani stative/copular verbs we must add a non-overt copula, probably the equivalent of /-dIr/. Since it is absent from formal analyses of Azerbaijani, it is probably more likely to occur in informal conversation than in formal speech. Matrix languages must be assigned, then, to the stative clauses in the Home recording corpora as follows:

- Clause with Azerbaijani copula - Azerbaijani is the matrix language
- Clause with Russian copula - Russian is the matrix language
- Clause with all Azerbaijani words but no overt copula - Azerbaijani is probably the matrix language
- Clause with all Russian words but not no overt copula - Russian is probably the matrix language
- Mixed Azerbaijani and Russian with no overt copula - Matrix language cannot be determined (congruent lexicalization)

4.5 Conclusions

From this analysis we see that actual Azerbaijani/Russian code-switching/mixing behavior both agrees with and contradicts the predictions made in Chapter 3. While alternation occurs between conversational turns, in the data for this study it is rare. The subjects seem more concerned with cooperation and linguistic accommodation that with the competition for linguistic dominance that inter-turn alternation would imply. Alternation between CPs also occurs but for information about the pragmatic functions it might serve, we must look to more detailed analyses such as those presented in Chapter 7. We also find that alternation in peripheral elements is common, but its occurrence at the end of Azerbaijani clauses required modification of the Azerbaijani X' rules to allow a CP final Spec position. In accordance with Myers-Scotton's 4M model (2002), nominal insertion of common nouns is the most common form of code-mixing in this corpus.

The evidence for congruent lexicalization presented above directly contradicts the conclusions from Chapter 3. Although formal analyses of Azerbaijani call for an overt verb in every clause, we see evidence for a non-overt copula in the 10 minute samples corpus. This, in turn, opens the door to an analysis of mixed Azerbaijani/Russian CPs where it is impossible to determine the matrix language.

The value of this description of Azerbaijani/Russian code-switching and code-mixing is that it opens the door for inquiry into the social significance of these behaviors. Since it has identified the possible range of behavior, we can go on to examine variation

between subjects, the pragmatic functions of different types of code-switching/mixing, and the ways in which it can be used to construct social identities in Azerbaijan.

CHAPTER 5

ANALYSIS OF SURVEY DATA

As mentioned previously, four types of data were collected for this dissertation project: observations of public language behavior, informal interviews (group and individual), survey information, and recorded conversations (Staged and Home). Chapter 4 drew predominantly on the recorded conversational data to detail the kinds of code-switching/mixing present in contemporary Azerbaijan. This chapter will discuss each section of the survey instrument and analyze the data gathered from each. Though due to sampling and instrument design issues, these data do not show a meaningful statistical relationship with the data discussed in following chapters, they do point to ways in which future research can help us more thoroughly understand the language situation in Azerbaijan.

Three types of data were collected in the surveys: basic demographic information with a focus on influences toward Azerbaijani and Russian use, six questions regarding language philosophy, and seven questions to determine how much affinity the subject felt toward Russian culture (Appendix A).

5.1 Demographic information

Since previous research (Zuercher 2004, 2009) showed age, gender, and language of education to correlate with and/or influence Russian language choice, and since early

language influences can be assumed to continue throughout life, the first section of the survey collected the following information:

- Name (a firm indicator of sex/gender in Azerbaijan)
- Age
- Nationality (equivalent to ethnicity in a Western context)
- Father's nationality
- Mother's nationality
- First language learned in childhood
- Language/s spoken in the home as a child
- Elementary school language
- High school language of education
- University language of education

Table 5-1 summarizes the responses to this portion of the survey for participants in both the Home and Staged recordings by category. Other than name and age, the demographic data collected fall into three categories: nationality, childhood language, and language of education.

Chronbach's alpha was used to examine the internal consistency of the responses for each category, with 0.7 as the threshold to show a statistically significant level of internal reliability – a "modest" level of reliability (Nunnally & Bernstein 1994: 265). Moreover, surveys collected in the summer of 2006 were added to the 15 participants in the Home and Staged recordings to yield a total sample size of 56. The responses for the three nationality questions from this larger sample showed a high level of reliability.

Table 5-1 Demographic survey results

Name	Category	Sex	Age	Nationality	Father's Nationality	Mother's Nationality	First Language	Home Language	Elementary Language	H.S. Language	University Language
Tarana	Home - w/Adults	F	32	Az	Az	Az	Az	Az	Ru	Ru	Az
Lale		F	33	Az	Az	Az	Az	Az	Ru	Ru	Ru
Aygun		F	28	Az	Az	Az	Az	Az/Ru	Ru	Ru	Ru
Mahir		M	32	Az	Az	Az	Az	Ru/Geor	Ru	Ru	Ru
Aner		M	47	Az	Az	Germ	Ru	Az/Ru	Ru	Ru	Ru
Sevil	Home - Other	F	44	Az	Az	Az	Az	Az	Az	Az	Ru
Agil		M	32	Az	Az	Az	Az	Az	Ru	Ru	Az/Ru
Shahin*		M	20	Az	Az	Az	Ru	Az/Ru	Ru	Ru	Ru
Kamala	Staged	F	22	Az	Az	Az	Az	Az	Az	Az	Az
Tamam		F	21	Az	Az	Az	Az	Az	Az	Az	Az
Vugar		M	24	Az	Az	Az	Az	Az	Az	Az	Az
Adil		M	25	Az	Az	Az	Az	Az	Ru	Az	Az
Vahid		M	23	Az	Az	Az	Az	Az	Ru	Ru	Ru
Leyla		F	19	Az	Az	Az	Ru	Ru	Ru	Ru	Ru
Zarifa		F	18	Az	Az	Az	Ru	Ru	Ru	Ru	Ru

* Shahin participated in both the Home and Staged recordings

The alpha for all three questions was 0.8912, and if Mother's Nationality was omitted, the alpha was 0.9843. This indicates that all three questions are indeed examining the same factor. All 15 subjects in the Home and Staged recordings claimed Azerbaijani nationality, and all had Azerbaijani fathers. Only Aner had a non-Azerbaijani mother – she was German. The answers to the childhood language questions also show a statistically significant level of consistency. If the answers are coded into two groups (group 1 = only Azerbaijani; group 2 = includes anything other than Azerbaijani) there is a high level of reliability between the responses (α=0.8843). Thus 'first language' and 'home language' likely represent one factor. All 15 subjects in the Home and Staged

recordings claimed either Azerbaijani or Russian as their first language. Aner, Aygun, and Shahin claimed that both Azerbaijani and Russian were spoken in their homes, while Mahir lists Russian and Georgian as his home languages. For the larger 56 member sample, the three education language questions also show a statistically significant level of agreement ($\alpha = 0.9214$; without Elementary Language, increases to 0.9324). Thus, Chronbach's alpha shows that the three categories of questions are most likely measuring one factor each: nationality, first/home language, and education language. A logical next question would be whether any two of these categories are measuring the same factor. Checking for reliability of questions from different groups yields no statistically significant results. Home Language (as operationalized above) and University Language come closest with an alpha of 0.6492 - below the 0.7 confidence level.

While the internal reliability of the three sets of questions speaks for the validity of the questions, the results are in no way surprising. Parents' nationality is most likely a causative factor in the subject's nationality, and home language is probably the biggest determiner of which language an individual will learn first. Lastly, if students attend school in a particular language, they are most likely to continue with education in the same language. They are better equipped to do so than if they were to switch language sections.

Correlating this survey data with language behaviors exhibited in the Home recordings is more difficult. The first issue to be addressed is that of comparability. Though the eight longer, non-directed recordings are named 'Home' they did not all take place in equivalent contexts.

Table 5-2 Summary of subjects and data collected in the Home recordings (repeated from Chapter 4)

Name	**Sex**	**Age**	**Context**	**Total usable recorded time**
Aner	male	47	at home with wife and family	28 min (14 recordings, 4 usable)
Aygun	female	28	making/eating dinner with roommate and sister	45 min (3 recordings)
Agil	male	32	various contexts at work	1 hour 2 min (7 recordings)
Lale	female	33	at home with her daughter (9 years old) and mother	43 min (1 recording)
Mahir	male	32	at home with wife and family	26 min (8 recordings, 7 usable)
Shahin	male	20	playing video games with a friend at an Internet club	20 min (1 recording)
Sevil	female	44	at home with her young daughter (5 years old)	1 hour 53 min (1 recording)
Tarana	female	32	making/eating dinner with her sister (Aygun) and roommate	1 hour 15 min (1 recording)

Agil's and Shahin's recordings clearly represent different speech situations from the others since they took place outside the home environment – Agil at work and Shahin in an Internet club. The other six recordings took place at home, however Sevil's recording is not comparable to the other five. While it was recorded in a home context, the only interlocutor present for the vast majority of the time was her 5 year old daughter. Thus if we are to compare the frequency of the various linguistic behaviors discussed in the previous chapter, we must narrow the sample to 5 subjects: Aner, Aygun, Lale, Mahir, and Tarana. All of these participants recorded their conversational contributions in a home context with other adults present. Unfortunately, as Table 5-1 shows, there is not enough variation in the demographic backgrounds of these subjects to look for correlations with language behavior: all claim Azerbaijani nationality, all but Aner claim

Azerbaijani as their first language, and all but Tarana attended Russian language university.

5.2 Language philosophy questions

The next section of the survey consisted of six questions concerned with attitudes toward Azerbaijani and Russian:

1. Everyone in Azerbaijan should speak Azerbaijani.
2. Everyone in Azerbaijan should speak Russian.
3. Azerbaijani is good for every topic.
4. Russian is better for some topics.
5. I like to hear Azerbaijanis speak Russian.
6. I do not like to hear Azerbaijanis speak Russian.

Each question elicited a response on a five point Likert scale:

Totally disagree	coded 1
Somewhat disagree	coded 2
Unsure	coded 3
Somewhat agree	coded 4
Totally agree	coded 5

Table 5-3 shows participant responses from the Home and Staged recordings to these questions.

The items in this section of the survey were intended to function as three sets of paired questions with each pair measuring a different language attitude:

Questions 1 and 2	support for or opposition to the current orientation toward Azerbaijani linguistic nationalism
Questions 2 and 3	attitude toward the Azerbaijani focused language planning that has taken place since the breakup of the U.S.S.R.
Questions 4 and 5	affective response to Azerbaijanis speaking Russian

Table 5-3 Language philosophy survey results

Name	**Category**	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**
Aner	Home - w/Adults	5	3	4	5	5	1
Aygun		5	3	5	4	4	5
Lale		5	3	4	5	4	1
Mahir		5	1	5	5	3	5
Tarana		2	4	3	5	5	1
Agil	Home - other	1	1	4	2	3	3
Shahin*		2	1	3	5	3	3
Sevil		3	3	1	1	1	1
Adil	Staged	1	1	4	4	1	2
Kamala		4	1	5	4	4	2
Leyla		2	2	5	1	5	1
Tamam		2	2	n/a	3	1	3
Vahid		1	1	1	4	4	4
Vugar		5	1	4	2	1	5
Zarifa		5	4	2	4	5	1

If these questions were functioning as intended, then the responses to question 1 should show a high reliability score with reverse coded answers from question 2, etc. However, using the 56 member sample discussed in the last section, none of the alpha scores for the question pairs is over the 0.7 threshold. The closest is the relationship between question 5 and the reverse coded question 6 (α=0.6718). This indicates that each question is measuring a different factor, rather than pairs of questions measuring the same factor.

Thus, there is no statistical support for combining the responses to any of the six questions together.

Like the demographic information discussed above, the attempt to correlate the six Language Philosophy questions with language behavior in the Home recordings corpus is not productive. If we take only the 5 subjects who recorded conversation at home with adults, and divide the responses into High (4 and 5) and Low (1 and 2), the only question that splits the 5 subjects into groups is question 6. The other five questions would result in one of the 'groups' having only one member. Graphs of these two groups (High and Low based on question 6) and measures of overall language use show promising results. Three measures were examined: ratio of Russian to Azerbaijani content words, percent of Azerbaijani CPs with Russian insertion, and percent of Azerbaijani CPs with Russian peripheral alternation. Though the trends in all three cases are in the anticipated direction, none of the differences are statistically significant. In all three cases, an independent samples t test comparing the means of the High and Low groups results in a probability well above the .05 confidence level used in this study.

5.3 Russian cultural affinity questions

The last section of the survey was a series of seven questions designed to determine the subjects' affinity to Russian culture. These questions were adapted from Crandall's Social Distance Questionnaire (Crandall 1991 as summarized in Robinson, Shaver, and Wrightsman 1999: 341-343):

1. I generally like Russians.
2. I like Russians for close friends.

3. I like Russians as neighbors.
4. I like to work with Russians.
5. I resemble Russians.
6. I would like Russians as relatives.
7. I avoid contact with Russians (designed to be reverse coded).

Since this section was not included in the survey administered in the summer of 2006, only the 15 subjects in the Home and Staged conversations recorded in the summer of 2007 can be included in this analysis. Coded according to the same Likert scale as the Language Philosophy questions, the participant responses to the Russian Cultural Affinity questions are shown in Table 5-4.

Table 5-4 Russian cultural affinity survey results

Name	**Category**	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**
Aner	Home - w/Adults	5	4	4	4	1	1	1
Aygun		5	5	5	5	4	3	4
Lale		1	3	4	3	1	1	1
Mahir		1	1	3	3	1	1	1
Tarana		2	4	4	4	4	4	2
Agil	Home - other	4	2	2	2	2	2	3
Shahin*		3	4	2	3	1	3	1
Sevil		2	5	5	4	4	1	1
Adil	Staged	4	4	4	4	1	2	1
Kamala		3	4	2	1	4	1	3
Leyla		2	4	1	1	1	1	3
Tamam		4	n/a	3	3	n/a	1	3
Vahid		5	5	5	5	4	2	1
Vugar		5	5	5	5	1	3	1
Zarifa		2	4	4	n/a	n/a	3	1

With question 7 reverse coded, Chronbach's alpha was used to check for internal reliability between the responses to the questions to determine if any were measuring the

same factor. The only two that showed a statistically significant result were questions 3 and 4 (α=0.9587). Thus, in Azerbaijan the neighborhood and work domains are very closely related.

These 7 questions can be arranged to form an implicational scale. If question 7 is reverse coded, and the responses for all questions converted into H and L (4 and 5 = H, 1 and 2 = L, and 3=·), then the columns and rows arranged so as to minimize errors, the implicational scale shown in Table 5-5 emerges. As expected from their level of reliability, questions 3 and 4 occur next to each other while the reverse coded question 7 lands at the left of the chart. This is not surprising since the reverse coding of "I avoid contact with Russians" would show a general attitude of non-avoidance.

Table 5-5 Russian cultural affinity survey results as an implicational scale

Name	**7 rev**	**2**	**3**	**4**	**1**	**5**	**6**
Agil	·	L	L	L	H	L	L
Leyla	·	H	L	L	L	L	L
Mahir	H	L	·	·	L	L	L
Tamam	·	n/a	·	·	H	n/a	L
Lale	H	·	H	·	L	L	L
Kamala	·	H	L	L	·	H	L
Shahin*	H	H	L	·	·	L	·
Zarifa	H	H	H	n/a	L	n/a	·
Aner	H	H	H	H	H	L	L
Sevil	H	H	H	H	L	H	L
Adil	H	H	H	H	H	L	L
Aygun	L	H	H	H	H	H	·
Vugar	H	H	H	H	H	L	·
Tarana	H	H	H	H	L	H	H
Vahid	H	H	H	H	H	H	L

*Implication line drawn according to the number of H responses in each row.

At first glance, it may seem odd that question 1 lands between questions 4 and 5 in Table 5-5. This can, however, be explained in terms of general versus specific associations:

2.	I like Russians for close friends.	specific	'My Russian friends'
3.	I like Russians as neighbors.	specific	'My Russian neighbors'
4.	I like to work with Russians.	specific	'My Russian co-workers'
1.	I generally like Russians.	non-specific	Russians as a whole
5.	I resemble Russians.	non-specific	Russians as a whole
6.	I would like Russians as relatives.	non-specific	Hypothetical Russians relatives

Given the history and demographics of Azerbaijan it is quite possible that most of the participants in this study either have or have had Russian friends, neighbors, and co-workers. With these positive associations High answers to questions 2, 3, and 4 are more likely. High answers to the more hypothetical/general questions are less likely.

While Table 5-5 shows that the Russian Cultural Affinity results represent an implicational trend, the Coefficient of Reproducibility (C_{rep}) is just under the usual 0.90 confidence level at 0.895 (see Hatch and Lazaraton 1991: 210). This result does not count blanks and 'unsure' answers as errors. If these are counted as errors the C_{rep} drops to 0.80.

The results to these questions allow us to divide the five subjects who recorded conversation at home with children into two groups: High Russian cultural affinity, and Low Russian cultural affinity:

High		Low	
Aner	5 H answers	Lale	2 H answers
Aygun	5 H answers	Mahir	1 H answer
Tarana	6 H answers		

As with the Language Philosophy questions, three measures were examined to see if there is a meaningful relationship between them and Russian cultural affinity: ratio of Russian to Azerbaijani content words, percent of Azerbaijani CPs with Russian insertion, and percent of Azerbaijani CPs with Russian peripheral alternation. In all three cases the results are counter intuitive. Lale and Mahir, who claimed a low level of Russian cultural affinity, used a higher proportion of Russian content words than did Aner, Aygun, and Tarana, and inserted Russian lexical items into Azerbaijani CPs more often. While Tarana showed the highest proportion of Azerbaijani CPs with Russian peripheral alternation, when viewed as High and Low groups, the Low group still showed a higher overall percentage of Azerbaijani CPs with Russian peripheral elements.

5.4 Relationship with conversational data

The difficulties in using the survey data to analyze the conversational data in this study stem from both instrument design and sample size and problems. While the demographic information shows high internal reliability for each section, the five member sample of subjects who recorded data at home with adults present is too small and does not show enough variation in demographic background to allow meaningful correlations between linguistic variation and demographic factors. The same problems plague the Language Philosophy section of the questionnaire: the sample size of

comparable speech situations is too small, and there is not enough variation in the subjects' responses. Additionally, this section of the survey suffers from design issues. Before linguistic variation can be correlated with language attitudes in Azerbaijan, a survey instrument must be designed that accurately measures the appropriate attitudes and demonstrates an acceptable amount of internal reliability. Since the Language Philosophy questions do not show sufficient internal reliability, it is difficult to identify the attitudes that they are measuring. As Table 5-5 shows, the Russian Cultural Affinity questions form a fairly dependable implicational scale, with the sub-.9 C_{rep} being due, in part, to the small sample size of 15. However when the attempt is made to correlate subjects' location on this scale with actual language use, the results are counter-intuitive. This could indicate one of three things: the survey instrument is not measuring the right attitudes; the intuitions mentioned are wrong and that subjects who feel an attraction to Russian culture actually do speak less Russian and vice-versa; or, that the results from this small sample are being skewed by individual idiosyncrasies and a larger sample size would yield more accurate results. The third option seems most likely.

In addition to problems with instrument design and sample size, this study is impacted by cultural issues related to doing survey based research in Azerbaijan. Unlike the West, Azerbaijan and probably the entire former Soviet space does not have a cultural affinity for this type of research. Given a long background of political oppression and social control, subjects are hesitant to respond to surveys. Subjects who do choose to participate are not accustomed to answering survey questions; they seem to exhibit a greater amount of idiosyncratic variation (resulting in statistical error) than do subjects in

the West (see also Zuercher 2004: 51). A larger sample size would allow these idiosyncratic variations to balance each other out, thereby allowing the survey instrument to measure the intended underlying social constructs. Given the very small sample size for the subjects who recorded conversational data at home with adults, it is not surprising, that the results are either counter-intuitive, not statistically significant, or both. While a particular subject's answers to the survey may provide useful information in interpreting that particular subject's language behavior, the survey data is not useful in analyzing overall language use patterns.

CHAPTER 6

QUANTITATIVE ANALYSIS OF VARIATION

Moving away from the survey data, this chapter will provide a quantitative analysis of code-switching/mixing behavior. While the surveys provide us with little useful information to interpret the variation in language use, the subjects in the Home recordings corpora (plus Tamam in some cases) do use language differently. This chapter will examine eight measures of overall language use:

- frequencies of Russian and non-Russian content words,
- frequencies of Russian and non-Russian clauses (CPs),
- frequency of Russian CPs with and without Azerbaijani insertion
- frequency of Azerbaijani CPs with and without Russian insertion
- frequency of Russian CPs with and without Azerbaijani peripheral alternation
- frequency of Azerbaijani CPs with and without Russian peripheral alternation
- frequency of Azerbaijani stative clauses with and without overt verbs, and
- frequency of Azerbaijani CPs exhibiting and not exhibiting Russian-like movement.

While most of the charts and graphs show all subjects in the Home recording corpora, with the addition of Tamam in the last two, all of the statistical analyses look at variation only between the five subjects who recorded conversation at home with adults present:

Aner, Aygun, Lale, Mahir, and Tarana. Though the statistical tests look at relative frequencies of the language use items, the figures presented show percentages to facilitate visual comparison.

Throughout this chapter and the next, the chi-square test is used for statistical analysis. Like all statistical tests chi-square has a number of underlying assumptions: (Hatch & Lazaraton 1991: 406-410):

1. The data must consist of frequencies
2. The categories must form a logical classification
3. Whenever the frequency of an event is counted, the frequency of nonoccurrence must also be counted
4. The data must be independent (no data are included in more than one cell)
5. The sample size must be large enough to obtain an expected cell frequency of five
6. When the number of degrees of freedom equals 1 (as is true for a 2x2 table), apply the Yates' correction factor which reduces the magnitude of the expected frequencies (see also Woods et al. 1986: 146)

The chi-square tests the independence of the cells in a contingency table by comparing the actual frequencies with their expected frequencies. This provides a way of determining "whether the differences observed could be due simply to sampling variation, that is, we have two samples drawn from the same population; or whether they indicate a real difference, that is, the two samples are actually from different populations" (Woods et al. 1986: 140). In all cases the null hypothesis will be that the samples are not independent. For example, the null hypothesis for the data in Table 6-1 would be:

H_0: The observed frequencies of Russian content words and non-Russian content words represent the linguistic output of a single homogeneous population of speakers.

If the chi-square is large enough to be statistically significant at the confidence level set by the study (here, alpha = .95), it shows that there is a statistically significant probability that the samples are *not* drawn from the same population, but that they represent different populations.

Table 6-1 shows how the chi-square is calculated. The expected frequency for each cell is calculated by multiplying the row total by the column total then dividing by the grand total. The number of deviances for each cell is calculated by subtracting the expected frequency from the observed frequency, multiplying the result by itself, then dividing by the expected frequency. The deviances for each cell are then added together, and the result is the chi-square statistic. In Table 6-1 the chi-square equals 435.713.

Before we can determine if this chi-square shows the samples to be from independent populations, we must determine the degrees of freedom (df). According to Woods et al. "the degrees of freedom can be considered in a sense as the number of independent pieces of information we have on which to base the test of a hypothesis" (1986: 138). The value for df is calculated by subtracting one from the number of columns then multiplying by the number of rows minus 1: (columns – 1)*(rows – 1). Thus, the degrees of freedom for the data in Table 6-1 is 4.

Throughout this study the chi-square test is used to determine whether the samples are likely to have been drawn from the same population, assuming a 95%

confidence level. Therefore when the probability (p) is less than 0.05, the chi-square for the data is considered statistically significant; as such we reject the null hypothesis (that the samples are from one population) and embrace the alternative: that each sample represents a distinct population.

Table 6-1 Chi-square contingency table: Russian content words from five 'at home with adults' subjects

a) Observed frequencies

Name	**Russian Content**	**Non-Russian Content**	**Row Totals**
Aner	219	271	490
Aygun	356	1339	1695
Lale	564	503	1067
Mahir	303	267	570
Tarana	537	487	1024
Column Totals	1979	2867	*Grand Total = 4846*

b) Expected frequencies
(Column Total*Row Total)/Grand Total

Name	**Russian Content**	**Non-Russian Content**
Aner	200.105	289.895
Aygun	692.201	1002.799
Lale	435.739	631.261
Mahir	232.775	337.225
Tarana	418.179	605.821

c) Deviances
(Observed – Expected)^2/Expected

Name	**Russian Content**	**Non-Russian Content**
Aner	1.784	1.232
Aygun	163.292	112.715
Lale	37.754	26.060
Mahir	21.186	14.624
Tarana	33.762	23.305

Total Deviances (chi-square) = 435.713

6.1 Russian content words

As mentioned in chapter 4, one of the first phases of analysis, after transcription, was to compare the total number of content words in the languages used by each subject in the Home recordings corpora. There are two major reasons for this: content words were tagged for language in previous stages of analysis so Wordsmith could quickly provide tallies; and, a count of content words is less subject to influence by the morphological systems of the languages than either an overall word count or an overall morpheme count. These would be less informative because of the morpho-syntactic differences between Azerbaijani and Russian. For example, Russian uses prepositions while Azerbaijani mostly uses postpositions. While both languages have overt pronouns Azerbaijani rarely uses them in favor of nominal agreement morphology. Both Azerbaijani and Russian add inflectional morphology to nouns and verbs but the number and nature of morphemes added to each stem vary between the two languages. For example, Azerbaijani adds morphology for aspect, while Russian has different verb stems to indicate perfective and imperfective action. Lastly, function words (determiners, prepositions/postpositions, conjunctions, etc.) normally appear prominently in lists of most frequent words because they provide the linguistic structure of the discourse. The meaning is carried by content words. For this analysis, only the following word classes were considered content words: nouns, verbs, adjectives, and adverbs. Therefore, comparing the percentage of Russian content words gives us a good idea of overall language usage, though it does not give us specific information as to whether these were used in code-switching/mixing, or into which of Muysken's three types of code-

switching/mixing they would fall. Table 6-2 and Figure 6-1 show the total number and percentage of Russian content words used by each subject in the Home recordings.

Table 6-2 Percentage Russian content words

Name	**Total Content Words**	**Russian Content Words**	**Percentage Russian Content**
Sevil	1590	178	11.2%
Aygun	1695	356	21.0%
Aner	490	219	44.7%
Agil	1167	539	46.2%
Tarana	1024	537	52.4%
Lale	1067	564	52.9%
Mahir	570	303	53.2%
Shahin	327	318	97.2%

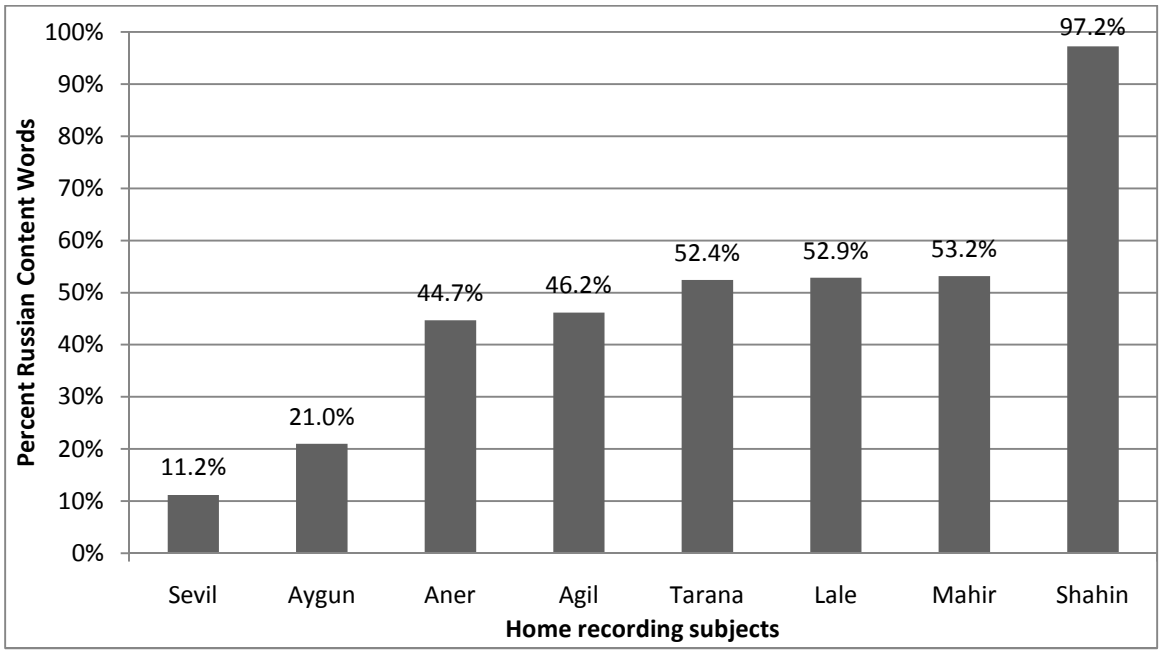


Figure 6-1 Percentage Russian content words

A chi-square test of the frequencies of Russian and non-Russian content words for the five subjects who recorded at home with adults present (underlined in Table 6-2) shows

that they represent more than one independent group – they are not a homogenous population (chi-square=435.713, df=4, p=0.000).

6.2 Percent Russian CPs

Though still not a measure of code-switching/mixing, looking at the number of clauses (CPs) in each language provides both a measure of overall language use and an indication of the levels of code-switching between turns and between CPs within turns. Table 6-3 and Figure 6-2 show the number of CPs contributed by each participant, the number of CPs with Russian as the Matrix Language, and the percentage of Russian language CPs. Like the Russian content words, a chi-square test of the frequencies for Russian and non-Russian headed CPs shows that the five 'at home with adults' subjects (underlined in Table 6-3) represent multiple populations (chi-square=102.829, df=4, p=0.000).

Table 6-3 Percentage Russian CPs

Name	**Total CPs**	**Russian CPs**	**Percentage Russian CPs**
Sevil	878	88	10.0%
Aygun	734	150	20.4%
Aner	250	87	34.8%
Tarana	493	196	39.8%
Agil	564	242	42.9%
Lale	542	233	43.0%
Mahir	308	139	45.1%
Shahin	188	186	98.9%

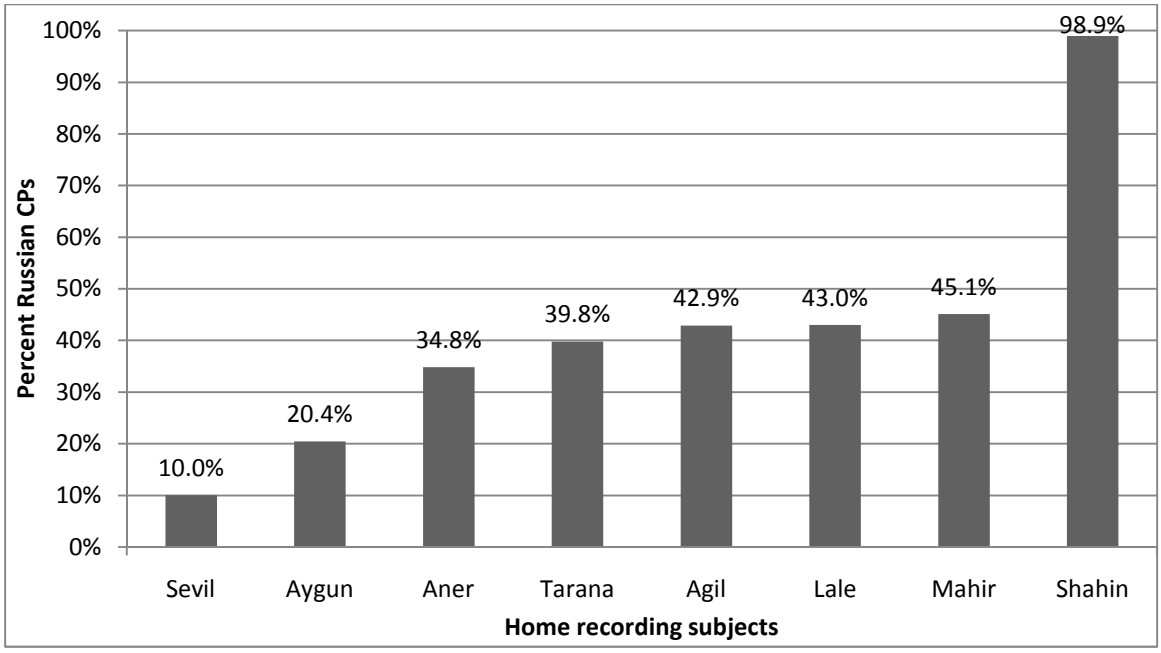


Figure 6-2 Percentage Russian CPs

6.3 Insertional code-mixing

In the Home recordings corpora, insertion is the most common form of code-mixing. This can, however take many different forms. Table 6-4 and Figure 6-3 show the types and relative frequencies of insertions.

Table 6-4 Relative amounts of insertional code-mixing

Inserted Element	**Ru in Az CPs**	**Az in Ru CPs**
Nominal Elements*	184	18
Nominal Modifiers	26	6
Adverbs	48	2
P Phrase	8	2
Non-finite Verb	17	1

* Excluding names, *mama*, and *papa*

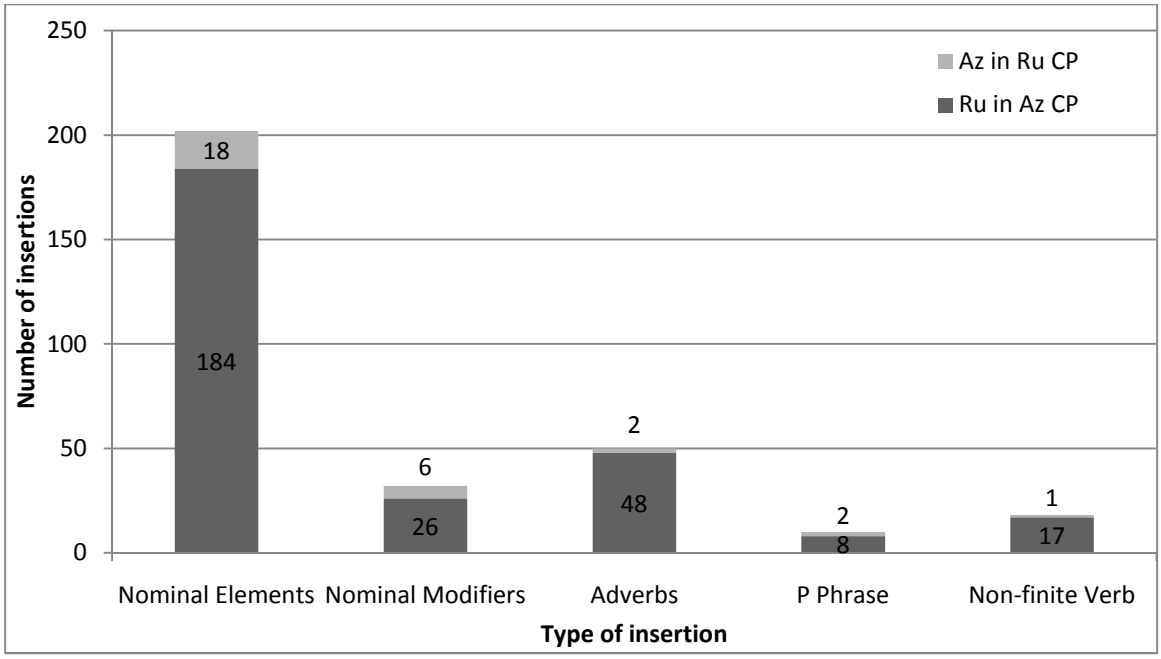


Figure 6-3 Relative amounts of insertional code-mixing

Personal names and the nouns *mama/мама* and *papa/папа* were excluded from this analysis. Their wide distribution makes it doubtful that they represent code- mixing in the minds of the speakers (transfer$_P$ – Auer 1984). As in many other studies of code-switching/mixing the insertion of nominal elements (nouns, noun phrases, pronouns) is the most common form of code-mixing. The relative distribution of insertion within Azerbaijani and Russian CPs reflects overall language dominance in the recordings. Azerbaijani is the most common language, thus the insertion of Russian elements into Azerbaijani CPs is much more common than the reverse.

6.3.1 Insertion of Azerbaijani elements into Russian clauses

Table 6-5 and Figure 6-4 show that the subjects in the Home recordings corpora do not insert Azerbaijani elements into their Russian CPs at the same rate. A chi-square test of the frequency of Russian CPs with and without Azerbaijani insertion for the five

subjects who recorded conversations at home (underlined in Table 6-5), however, does not prove that they represent different populations (chi-square=3.177, df=4, p=0.529).

Table 6-5 Percentage Russian CPs with Azerbaijani insertion

Name	**Total Russian CPs**	**Russian CPs with Azerbaijani insertion**	**Percentage Russian CPs with Azerbaijani insertion**
Sevil	88	0	0.0%
Aner	87	2	2.3%
Agil	242	9	3.7%
Shahin	186	8	4.3%
Mahir	139	7	5.0%
Lale	233	12	5.2%
Aygun	150	10	6.7%
Tarana	196	14	7.1%

This could either be because the distribution is due to random chance, or the low frequency of such insertion could simply not provide a large enough sample for accurate statistical analysis.

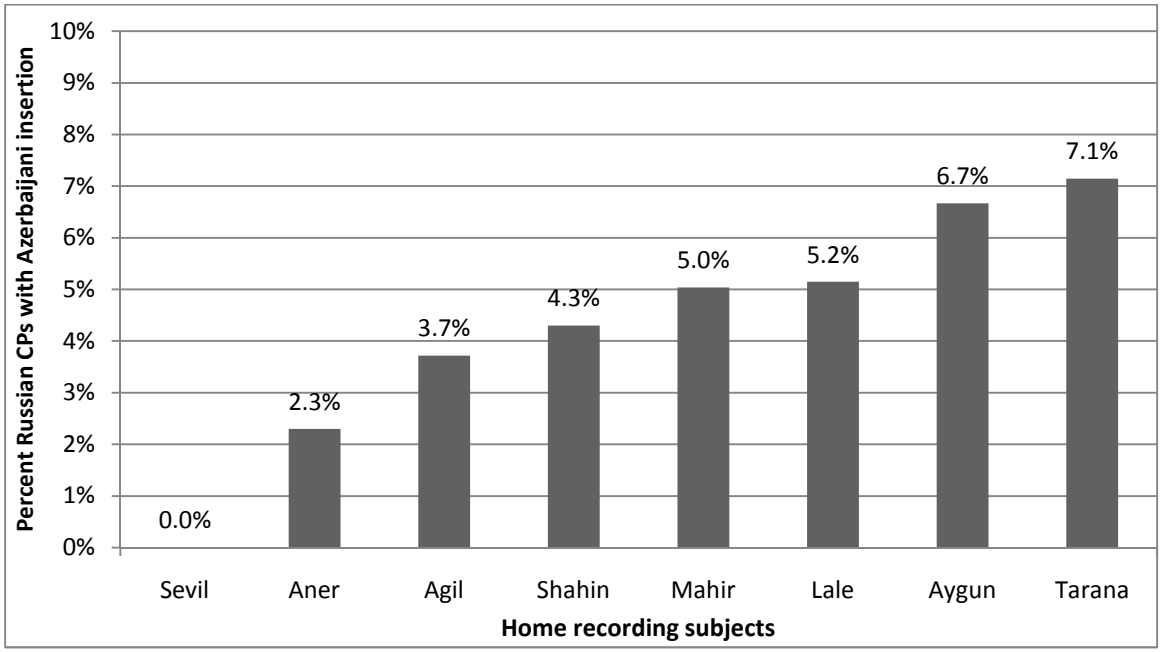


Figure 6-4 Percentage of Russian CPs with Azerbaijani insertion

Most of the other measures of overall language usage have some subjects with more than 10% of the phenomenon under consideration, while the highest percentage of Russian CPs with Azerbaijani insertion is 7.1% (Tarana).

6.3.2 Insertion of Russian elements into Azerbaijani clauses

The rates of Russian insertions into Azerbaijani CPs do, however, show statistically significant variation. Table 6-6 and Figure 6-5 show the rates of Russian insertion for all subjects in the Home recordings corpora. A chi-square test of Azerbaijani CPs with and without Russian insertion shows that the five 'at home with adults' subjects (underlined in Table 6-6) do indeed represent independent populations (chi-square=58.075, df=4, p=0.000).

Table 6-6 Percentage of Azerbaijani CPs with Russian insertion

Name	**Total Azerbaijani CPs**	**Azerbaijani CPs with Russian insertion**	**Percentage Azerbaijani CPs with Russian insertion**
Shahin	2	0	0.0%
Aner	163	6	3.7%
Aygun	545	25	4.6%
Sevil	780	36	4.6%
Agil	297	39	13.1%
Tarana	267	37	13.9%
Lale	295	52	17.6%
Mahir	152	28	18.4%

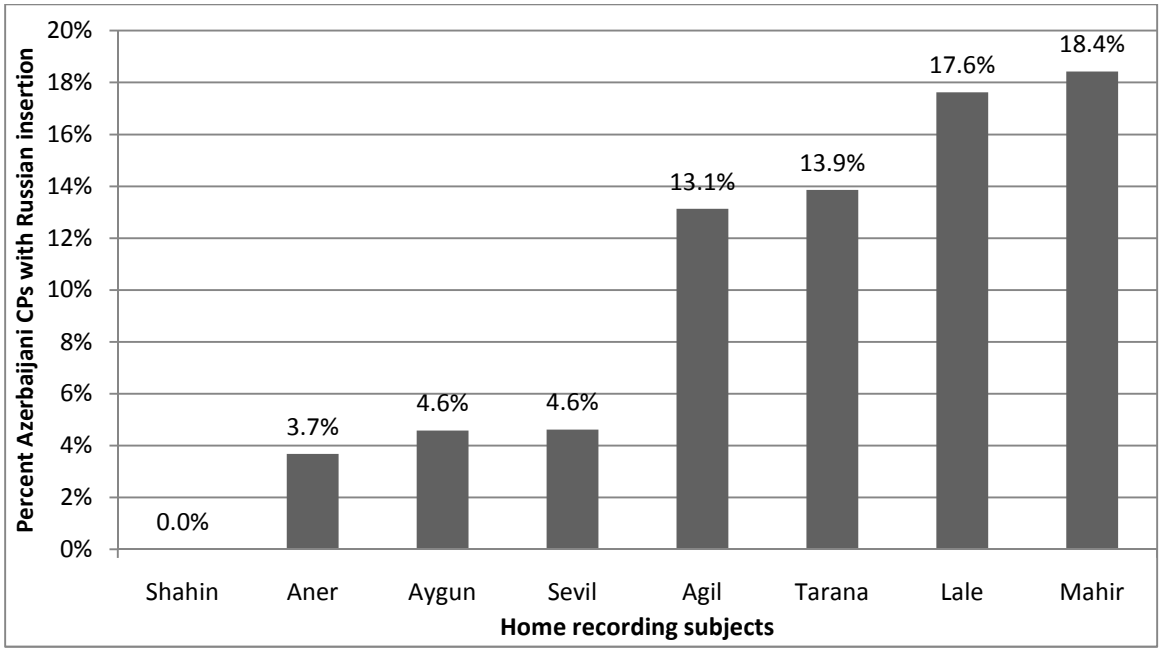


Figure 6-5 Percentage Azerbaijani CPs with Russian insertion

6.4 Peripheral alternation

Though less common than insertion, peripheral alternation is still frequent in the Home recordings corpora. Table 6-7 and Figure 6-6 show the types and relative amounts of peripheral alternation present. Like the analyses of nominal insertion ignored personal names, *mama/мама* and *papa/папа*, vocatives are not be included in the analysis that follows, though their frequency is shown Table 6-7 and Figure 6-6. It would be very difficult to determine which mental language lexicon contains any particular personal name, and while *mama/мама* and *papa/папа* appear in Russian but not Azerbaijani dictionaries, their wide distribution indicates that they are probably included in both mental lexicons for the recorded subjects. While vocatives that are clearly language specific and interjections such as *ay qız* ('hey girl' in Azerbaijani) may correlate with

subject specific personal factors, they are proably best studied in their conversational context, rather than by using the corpus approach of this chapter.

Table 6-7 Relative amounts of peripheral alternation

Element	Az on Ru CPs	Ru on Az CPs
Perif Adv	31	57
Vocative	15	47
Conjunction	2	35
Relative Clause	4	24
Discourse Marker	2	3
Perif PP	0	5
Perif AdvP	0	4
Perif N/NP	2	3
Perif CP	0	3
Tag	1	2
Dup Adv	1	0
Perif Pro	0	1

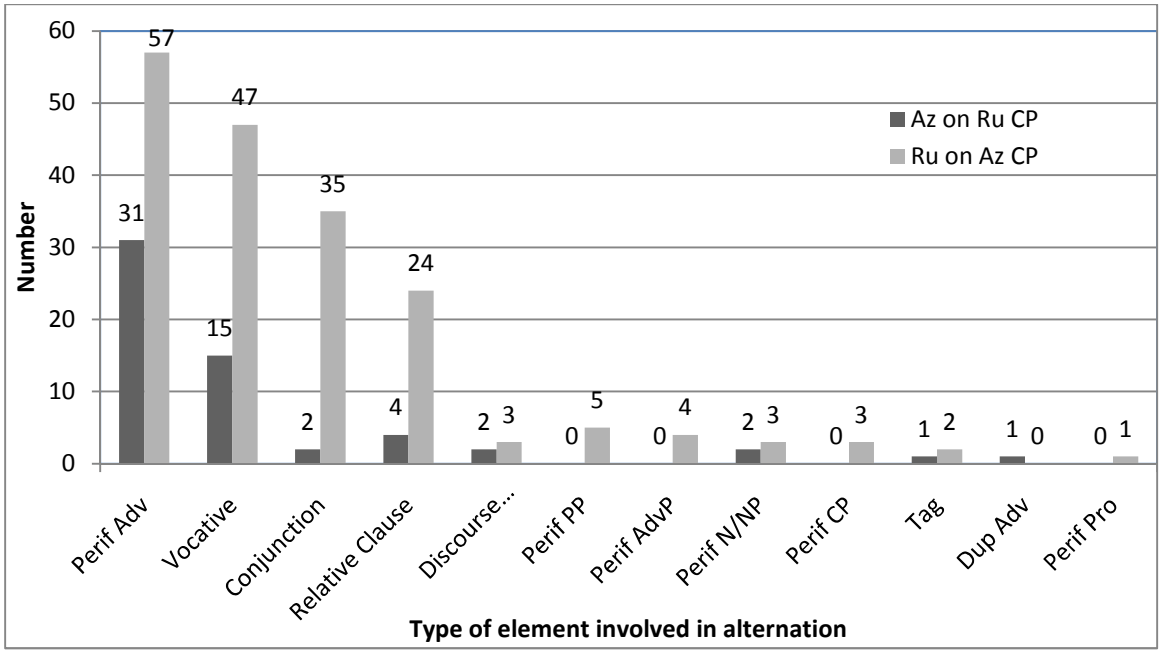


Figure 6-6 Relative amounts of peripheral alternation

6.4.1 Russian clauses with Azerbaijani peripheral alternation

Table 6-8 and Figure 6-7 show the levels at which the subjects in the Home recordings corpora used Azerbaijani peripheral alternation (other than vocatives) on Russian CPs.

Table 6-8 Percentage of Russian CPs with Azerbaijani peripheral alternation

Name	**Total Russian CPs**	***Russian CPs with Azerbaijani alternation**	**Percentage of Russian CPs with Azerbaijani alternation**
Tarana	196	3	1.5%
Shahin	186	3	1.6%
Agil	242	5	2.1%
Sevil	88	3	3.4%
Aner	87	3	3.4%
Mahir	139	5	3.6%
Lale	233	9	3.9%
Aygun	150	12	8.0%

*Vocatives not included

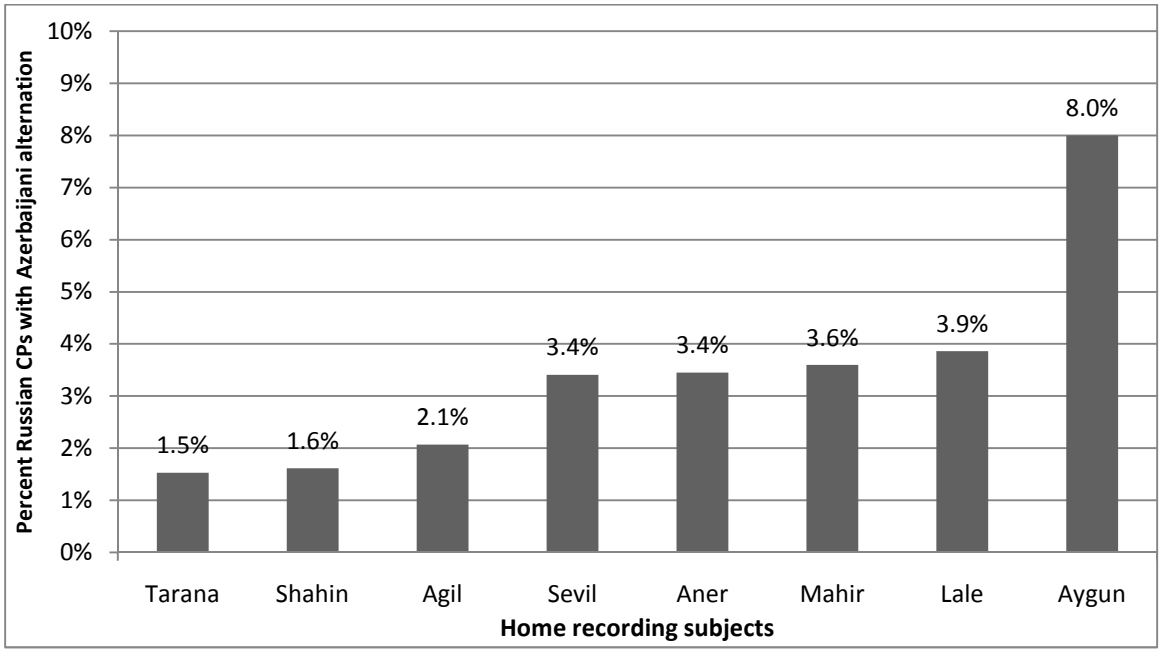


Figure 6-7 Percentage of Russian CPs with Azerbaijani peripheral alternation

Like a number of the measures above, a chi-square test for the number of Russian CPs with and without Azerbaijani alternation shows that there is a greater than 95% chance that the five 'at home with adults' subjects (underlined in Table 6-8) represent independent populations. This is true even though the chi-square is quite a bit smaller than elsewhere (chi-square=9.557, df=4, p=0.049; one cell has an expected count of less than 5).

6.4.2 Azerbaijani clauses with Russian peripheral alternation

Since there are many more Azerbaijani than Russian clauses in the Home recordings corpora, it is not surprising that there are many more Azerbaijani CPs with Russian peripheral alternation than the reverse. Table 6-9 and Figure 6-8 show the relative amounts for all 8 participants in the Home recordings, though Figure 6-8 omits Shahin. He only recorded two CPs with Azerbaijani as the Matrix Language and one of these contained an Azerbaijani peripheral element. A chi-square test shows that the five 'at home with adults' subjects (underlined in Table 6-9) represent independent populations (chi-square=32.535, df=4, p=0.000).

Table 6-9 Percentage of Azerbaijani CPs with Russian peripheral alternation

Name	**Total Azerbaijani CPs**	***Azerbaijani CPs with Russian alternation**	**Percentage of Azerbaijani CPs with Russian alternation**
Aner	163	1	0.6%
Mahir	152	3	2.0%
Sevil	780	24	3.1%
Aygun	545	24	4.4%
Agil	297	27	9.1%
Lale	295	29	9.8%
Tarana	267	28	10.5%
Shahin	2	1	50.0%

*Vocatives not included

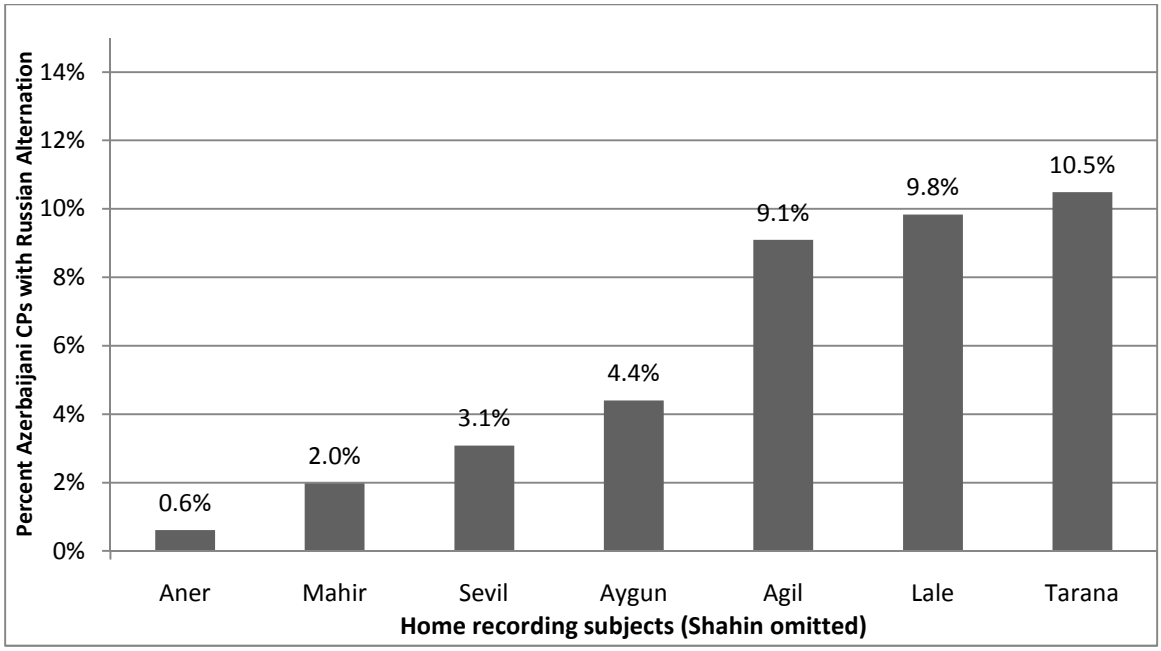


Figure 6-8 Percentage of Azerbaijani CPs with Russian peripheral alternation

6.5 Stative clauses (10 minute samples)

The next two sections discuss phenomena which may well be code-mixing, but are quite different from the other measures of overall language use discussed in this chapter since they do not involve the mingling of lexical information from the two languages. Thus an analysis of Azerbaijani stative clauses without overt verbs and Russian-Like Movement more closely resemble studies of L1 transfer or interlanguage than other studies of code-switching/mixing. In studying the English production of native Chinese students in Hong Kong, Alice Chan found that "the typical interlanguage output of the Chinese ESL learners in Hong Kong was strongly reminiscent of the usual or normative sentence structure in their L1, Chinese" (A. Chan 2004: 65-66). Chan's ESL learners produced ungrammatical structures in English because of syntactic transfer from Chinese. Here some subjects in the Home recordings corpora appear to

overproduce structures which may be considered grammatical in Azerbaijani, but are much more common in Russian. These results are similar to Schachter and Rutherford's (1979) finding that Japanese learners of English overused extraposition in English. Rather than producing ungrammatical sentences, they found that Japanese speakers produced sentences with extraposed components at a much higher rate than native English speakers or speakers of other languages - "Japanese learners are overproducing these constructions" (Schachter & Rutherford 1979: 3). While the results of the current study are similar in some ways, the languages are reversed. In examining Azerbaijani stative clauses without verbs and Russian-Like movement in Azerbaijani CPs, we see Russian (L2?) influence on Azerbaijani (L1?) production.

As detailed in the last section of Chapter 4, we must add a non-overt copula to the list of Azerbaijani stative verbs since all the speakers in the 10 minute samples, including Tamam who knows virtually no Russian, produced clauses with all Azerbaijani lexical items and no overt verbs. They did so, however, to varying extents.

Table 6-10 Percentage of Azerbaijani CPs without verbs (10 minute samples)

Name	**Azerbaijani stative CPs with verbs**	**CPs with all Azerbaijani morphemes and no verb**	**Percentage of Azerbaijani stative CPs without verbs**
Tamam	49	8	14.0%
Aygun	27	6	18.2%
Tarana	11	4	26.7%
Mahir	9	4	30.8%
Sevil	11	7	38.9%
Lale	11	9	45.0%
Aner	18	17	48.6%
Agil	0	1	100.0%
Shahin	0	2	100.0%

Table 6-10 and Figure 6-9 show the rates at which participants produced Azerbaijani stative clauses with and without overt copulas in the 10 minute samples. A chi-square test of the frequencies of such clauses for the five 'at home with adults' (underlined in Table 6-10) shows that they represent more than one population (chi-square=27.806, df=4, p=0.000, two cells have an expected count of less than 5).

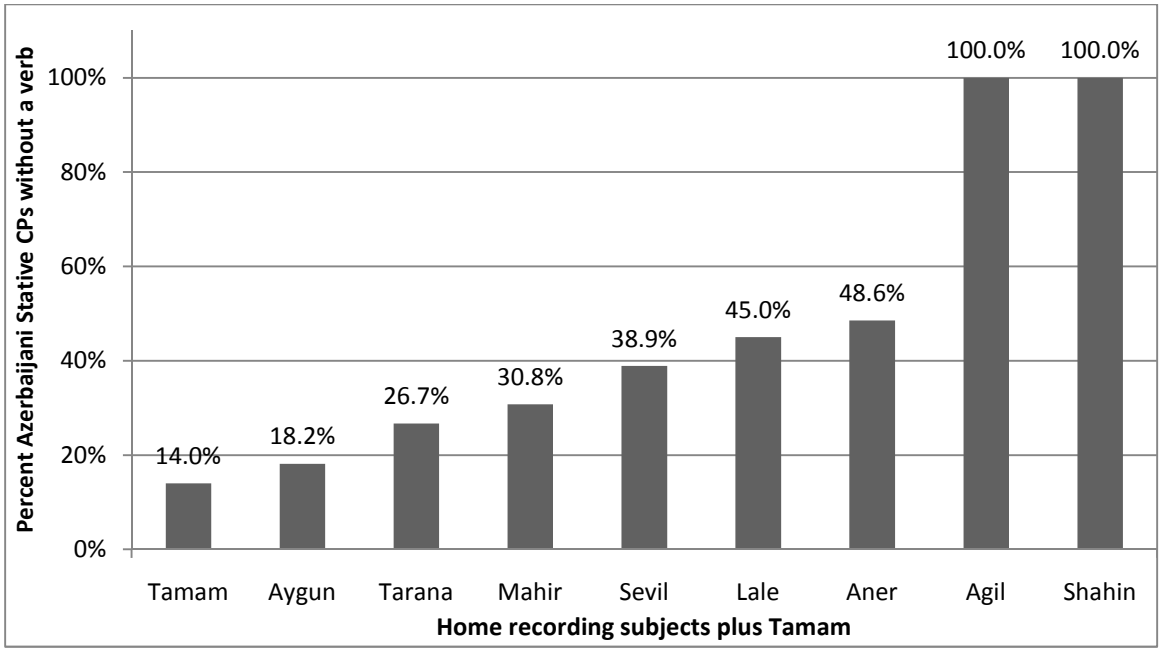


Figure 6-9 Percentage of Azerbaijani stative CPs without verbs (10 minute samples)

6.6 Russian-like movement

As Chapter 3 shows, both Russian and Azerbaijani have fairly free sentence word order. Both languages have CP-initial and -final landing spots for moved components (see section 4.2.3.1 for the modification to the Azerbaijani X' rules to create this position), but movement to the end of the clause is fairly rare in Azerbaijani. In Russian on the other hand, discourse salient information is commonly moved to the clause final slot. All of the subjects with clearly Azerbaijani headed clauses in the 10 minute

samples, however, show some evidence of Russian-Like Movement. For this analysis, Russian-Like Movement is operationalized as the movement of subject NPs, object NPs, and/or adjunct NPs or PPs past the verb in Azerbaijani CPs. Excluded from this analysis are Azerbaijani stative clauses without overt verbs. Relative clauses, adverbial particles, and tags were not counted as Russian-Like Movement. Relative clauses normally appear following the matrix clause, and adverbial particles and tags are assumed to originate outside I' and would therefore be less constrained by clause level movement restrictions than arguments or adjuncts. Table 6-11 and Figure 6-10 show the relative percentages of Azerbaijani CPs with Russian-Like movement in the 10 minute samples – Agil and Shahin did not produce any Azerbaijani ML clauses with overt verbs in their conversational excerpts.

Table 6-11 Percentage of Azerbaijani CPs with Russian like movement (10 minute samples)

Name	**Total Azerbaijani CPs***	**Azerbaijani CPs with Russian Like Movement**	**Percentage of Azerbaijani CPs with Russian Like Movement**
Tamam	124	7	5.6%
Sevil	72	5	6.9%
Tarana	38	3	7.9%
Lale	58	8	13.8%
Mahir	42	6	14.3%
Aygun	87	13	14.9%
Aner	36	6	16.7%

*Excluding Azerbaijani stative CPs without verbs.

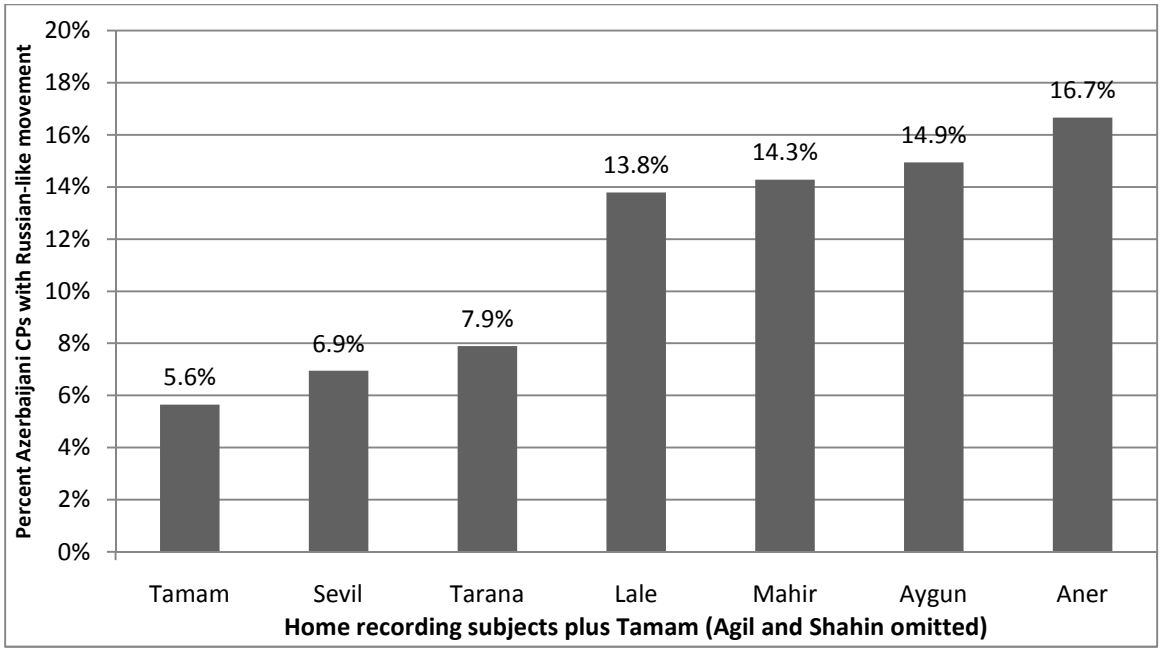


Figure 6-10 Percentage of Azerbaijani CPs with Russian like movement (10 minute samples)

Though these subjects apply Russian-Like movement at different rates, a chi-square test of the frequencies for Azerbaijani CPs with and without it for the five 'at home with adults' subjects (underlined in Table 6-11) does not show them to represent independent populations (chi-square=1.467, df=4, p=0.832, one cell has an expected count of less than 5). If, however the test is done only for Tamam and Aner (the two ends of the continuum) it shows that their speech is not homogenous but does represent two different populations (chi-square=4.540, df=1, p=0.033, 1 cell has an expected count less than 5). This is clearly a topic worthy of further investigation with a larger and more diverse population.

6.7 Comparison of five 'at home with adults' subjects

As the previous sections demonstrate, six measures of overall language use show statistically significant variation among the five subjects who recorded conversation at

home with adults present. Table 6-12 shows the rank order (least to greatest) of these five subjects for each measure that showed the subjects to represent multiple populations.

Table 6-12 Rank order of 'at home with adults' subjects for each statistically significant measure of overall language use

Percent Russian content words	**Percent Russian CPs**	**Russian insertion into Azerbaijani CPs**	**Russian CPs with Azerbaijani peripheral alternation**	**Azerbaijani CPs with Russian peripheral alternation**	**Azerbaijani stative clauses without verbs**
Aygun	Aygun	Aner	Tarana	Aner	Aygun
Aner	Aner	Aygun	Aner	Mahir	Tarana
Tarana	Tarana	Tarana	Mahir	Aygun	Mahir
Lale	Lale	Lale	Lale	Lale	Lale
Mahir	Mahir	Mahir	Aygun	Tarana	Aner

Percent Russian Content Words and Percent Russian CPs show exactly the same ordering, while Russian insertion into Azerbaijani CPs has Aygun and Aner reversed. The other three measures show re-ordering to a greater extent. Spearman's rho (non-parametric rank order correlation) confirms that there is a statistically significant relationship between the first three measures (95% confidence level).

Table 6-13 Rank order correlations for three measures of overall language use

	Correlation Coefficient	**Significance (p=)**
% Russian Content and % Russian CPs	1.000	0.000
% Russian Content and % Azerbaijani CPs with Russian insertion	0.900	0.037
% Russian CPs and % Azerbaijani CPs with Russian insertion	0.900	0.037

There are however no significant rank-order correlations among any of the other measures of overall language use. It is not surprising that subjects who produce more Russian headed clauses also produce more Russian content words overall. Neither is it

remarkable that these same subjects would produce the most common forms of code-mixing at a higher rate than those who use Russian less overall. Since the other measures do not correlate significantly with high levels of overall Russian use they must either have to do with idiosyncratic language use habits or with forms of social identity construction apart from those associated simply with Russian language use.

6.8 Discussion

Although the variation in answers to the survey questions do not correlate in significant ways with the variation in language use patterns (Chapter 5), these patterns by themselves do show significant variation. In fact, the variation in some measures of overall language use is quite large: from 11.2% Russian Content words for Sevil to 97.2% for Shahin, and from 10% Russian CPs for Sevil to 98.9% for Shahin. Contrary to what some interview subjects claimed, the data in this study show us that Russian use is quite common for some Azerbaijani subjects, as is code-switching and code-mixing. What they do not tell us, though, is whether this variation correlates with other macro level sociolinguistic factors such as social class (Labov 1972b, Trudgill 1974) or gender (Gal 1979, McDonald 1989). We can see, however, that the speech situations in this study varied dramatically: from Sevil at home speaking primarily with her 5 year old daughter, to Shahin playing computer games in an Internet Club with a friend. Clearly some of the variation in language use could be due to the different speech situations thus all the statistical analysis in this chapter included only subjects in comparable situations – those who recorded conversation at home with adults present.

Another potential source for the variation in language use patterns is individual identity construction. In terms of language of education and nationality, the subjects in the Home recordings form a very homogeneous group, yet even when controlling for speech situation by analyzing only those who recorded conversation at home with adults present, they show statistically significant variation for a number of the measures of overall language use. For these particular subjects, the variation cannot be explained based on gender or age alone. It is very likely, then, that the variation in language use patterns is due to individual identity construction efforts apart from the broad categories addressed in the survey data.

CHAPTER 7

CASE STUDIES OF LANGUAGE USE

7.1 Introduction

As the previous chapters show, subjects in the Home recordings corpora used Russian, Azerbaijani, and various types of code-switching/mixing to different extents. These data, however, do not tell us what functions these variations in language use serve. Therefore we must move away from an aggregated corpus approach toward one that is more sensitive to the specific discourses in which the language occurs. To that end, this chapter contains three case studies:

- Sevil – Change in conversational role
- Aygun and Tarana – A comparison of two sisters
- Aner – Soviet language influence

While paying close attention to sequential dependencies (as in Conversation Analysis – Sacks et al. 1974, etc.) these case studies will utilize the two broader perspectives of discourse described in Chapter 2: Hymes etic grid (Hymes 1972, Schiffrin 1994) to describe the overall speech situations and identify potential areas of interest, and Schiffrin's theoretical model (1987) to show ways in which code-switching/mixing correlates with underlying discourse issues.

7.2 Shift in conversational role: Sevil

Sevil's Home recording differs from many of the other recordings in that it is almost two hours of continuous verbal output. She does not turn off the recorder, delete sections, or remove the microphone from beginning to end. Over this time period, Sevil plays out a number of conversational roles. These cluster into two groups according to interlocutor: roles played out while talking with her daughter, and roles played out in the six phone conversations that take place during the recording. At first glance these sections of the recording seem to vary on two dimensions: amount of Azerbaijani and Russian, and amount of conversational support offered by Sevil. Table 7-1 shows the relative amounts of Azerbaijani and Russian content words in these two major sub-groups of conversation within Sevil's recording. Sevil did not use any other language content words.

Table 7-1 Distribution of Azerbaijani and Russian content words in talk between Sevil and daughter and in phone conversations

	Az content	Ru content	% Ru content
Talk with daughter	676	148	18.0%
Phone conversations	738	46	5.9%

A chi-square test of these frequencies shows that Sevil's talk with her daughter and in phone conversations do indeed represent independent 'populations' of language use data, and therefore may represent different genres of conversation (chi-square=55.387, df=1, p=0.000). As Table 7-1 shows Sevil used more Russian in speaking with her daughter than with adults in the phone conversations.

In addition to differing in terms of language, these differ in terms of conversational support. In all sub-sections of the recording Sevil contributes both turns

containing language content (Non-Backchannel Turns) and turns which only contain supporting morphemes such as *hm*, *uhuh*, *hə* 'yes' in Azerbaijani, etc. (Backchannel Turns). Rather than contributing content to the conversation, the later show that Sevil is playing a supportive role at the time – a listener role rather than a contributor. However, we must keep in mind that conversational support can serve opposite ends. It can show either engagement with the topic, or it can show boredom and disengagement since it provides only minimal conversational involvement. Table 7-2 shows the distribution of Sevil's conversational support across the two different types of conversation.

Table 7-2 Distribution of non-backchannel and backchannel turns in talk between Sevil and daughter and in phone conversations

	Non-backchannel turns	**Backchannel turns**	**% Backchannel turns**
Talk with daughter	291	40	12.1%
Phone conversations	184	139	43.0%

Like the frequencies of Azerbaijani and Russian content words, a chi-square of backchannel and non backchannel turns shows that Sevil's talk with her daughter and talk over the phone are independent from each other (chi-square=78.771, df=1, p=0.000).

7.2.1 Phone conversations

Table 7-3 and Table 7-4 summarize the six phone conversations in Sevil's recording according to both Hymes' etic grid and Schiffrin's theoretical model of discourse. In all cases, the interlocutors call Sevil's land line and she picks up. In the hour and 53 minutes of the recording, she never makes outgoing calls or talks on a cell-phone. In the first brief conversation, Sevil talks with a male caller (she asks about his wife toward the end). He is calling to arrange to meet the next day. She gives him

driving directions and arranges a location to leave the key since he will arrive before she gets home from work. Thus, it is probably a male relative coming for a visit, though it is possible that it is someone coming to conduct some business. If the later is true, they seem to be close acquaintances since she asks how his wife and family are.

Table 7-3 Summary of Sevil's phone conversations according to Hymes' etic grid

Setting	Sevil – at home Others – unknown
Participants	Sevil – 44 year old, female, university professor Conv 1 – male (she asks about his wife) Conv 2 – female, teacher, similar age? Conv 3 – male, her husband Conv 4 – older male relative (uncle) Conv 5 – co-worker at the university Conv 6 – her husband
Ends	Conv 1 – making plans for a visit the next day Conv 2 – Sevil trying to find translators for a government office Conv 3 – chat about the day and discussing family needs (groceries) Conv 4 – making plans to host foreign group, catching up on family news Conv 5 – complaining about work problems Conv 6 – review of grocery needs
Act sequence	conversational turn taking; can't identify overlap or interruption since only Sevil's side of the conversations were recorded
Key	Conv 1 – calm, informative (giving directions) Conv 2 – cheerful, business-like Conv 3 – calm, informative (telling story, etc.) Conv 4 – cheerful, supportive Conv 5 – animated and/or upset Conv 6 – clear voice, but very brief conversation – no full CPs
Instrumentalities	exclusively verbal since they are phone conversations
Norms	to be determined
Genre	Conv 1 – probably family talk Conv 2 – work talk, sales pitch Conv 3 – family talk Conv 4 – family talk Conv 5 – work talk Conv 6 – family talk

Table 7-4 Summary of Sevil's phone conversations according to three levels of discourse from Schiffrin's theoretical model

Information State	all participants are adults and appear to have equivalent cognitive abilities and linguistic skills
Participation Framework	Conv 1 – female relative to male relative? Conv 2 – female teacher to female teacher Conv 3 – wife to husband Conv 4 – niece to uncle Conv 5 – co-worker to co-worker Conv 6 – wife to husband
Ideational Structure	Conv 1 – "Come over tomorrow; I will leave the key" Conv 2 – "I've got a job for one of your students" Conv 3 – "Story from the day" and "Grocery needs" Conv 4 – Family talk Conv 5 – Work talk Conv 6 – Groceries – a reprise

In the second conversation a female teacher calls, apparently returning an earlier call from Sevil. Though the key is cheerful, it is predominantly a business call where Sevil tries to find a student to work as a translator for a governmental office. The third call seems to be from Sevil's husband. She names him as *Azər* on the phone, while in conversation with her daughter she repeatedly refers to her as *Azərin* 'Azer's (daughter)'. Sevil talks about what she did earlier in the day and they talk about what groceries she would like him to buy on the way home from work. In the fourth phone conversation an older male relative calls. She refers to him as *əmi* 'father's brother'. In the next few days Sevil's university department will be hosting a group of foreign visitors and she apparently called earlier to ask for his help. She briefly outlines the plans and how he can help. She then asks for news about a number of relatives. In the fifth conversation a co-worker from the university calls and they discuss some problems in the office. Though the call is fairly short, Sevil contributes some lengthy turns and seems quite animated and

or upset. The sixth phone conversation is very brief and contains no full CPs. It appears that Sevil's husband has called to review the grocery items needed. She simply agrees with the items as he lists them. This conversation does not contain enough turns or content words for the chi-square test and it includes no Russian content words, so it is not included in most of the discussion that follows.

Table 7-5 Distribution of backchannel turns in phone conversations

Conversation	**Non-backchannel turns**	**Backchannel turns**	**% backchannel turns**
1	19	1	5.0%
2	21	10	32.3%
3	27	5	15.6%
4	90	123	57.7%
5	27	0	0.0%

Table 7-6 Chi-square tests for non-backchannel versus backchannel turns in five phone conversations. Here (and beyond), statistical significance ($p \leq 0.05$) is indicated by a *.

Conversations	**Chi-square**	**df**	**Significance**
All Five	62.280	4	.000*
1/2	5.339+	1	.021*
1/3	1.361#	1	.243
1/4	20.432	1	.000*
1/5	1.379#	1	.240
2/3	2.401	1	.121
2/4	7.090	1	.008*
2/5	10.524+	1	.001*
3/4	19.784	1	.000*
3/5	4.609#	1	.032*
4/5	31.983	1	.000*

+ 1 cell has an expected value of less than 5
# 2 cells have an expected value of less than 5

Table 7-5 shows the relative distribution of non-backchannel and backchannel turns among these conversations while Table 7-6 contains the results of chi-square tests to show whether or not the differences in frequency are enough to indicate that they are

independent from one another. In terms of Sevil's conversational support, the five conversations fall into three groups: Sevil very supportive (number 4), balance between support and linguistic contribution (numbers 2 and 3), and conversations where Sevil maintains a focus on communicating information (numbers 1 and 5). In conversation number four Sevil takes on a predominantly supportive role. After she discusses plans for her foreign visitors, she listens supportively while her uncle talks about family affairs. It appears that she is engaged with the topics because she often laughs as well as providing backchannel cues. When speaking with a fellow teacher in conversation 2 Sevil listens supportively while her interlocutor speaks, but also keeps the focus on why she called the lady earlier. Similar but to a lesser extent, Sevil discusses her day and family needs with her husband in conversation 3. The conversations where Sevil contributes the highest proportion of information giving turns are 1 and 5. In 1 she is giving directions to her house and making arrangements for the following day, while in 5 she is talking with a co-worker. It appears that both she and her co-worker are speaking quickly and if we could hear the other side of the conversation we would probably hear significant overlap judging from the number of times that Sevil stops in the middle of words and clauses.

Moving away from conversational support to code-switching/mixing, Table 7-7 shows the relative amounts of overall language measures in each of the phone conversations. However, since Sevil uses low levels of Russian, most of these do not lend themselves to statistical analysis. Accordingly, Table 7-8 shows the results for chi-square tests only for Azerbaijani and Russian content words in the five conversations.

Table 7-7 Distribution of overall language use measures among phone conversations

	Content words			Russian headed CPs		Insertion		Alternation	
Conv	**Az**	**Ru**	**% Ru**	**CPs**	**Ru CPs**	**Ru in Az CP**	**Az in Ru CP**	**Ru on Az CP**	**Az on Ru CP**
1	66	0	0.0%	33	n/a	n/a	n/a	n/a	n/a
2	69	1	1.4%	38	n/a	1	n/a	1	n/a
3	155	15	8.8%	77	4	8	n/a	3	1
4	254	18	6.6%	139	8	2	n/a	5	2
5	190	12	5.9%	98	1	6	n/a	5	n/a

Table 7-8 Chi-square tests for Azerbaijani versus Russian content words in five phone conversations

Conversations	**Chi-square**	**df**	**Significance**
All Five	9.533#	4	0.049*
1/2	0.950#	1	0.330
1/3	6.219+	1	0.013*
1/4	4.613+	1	0.032*
1/5	4.105+	1	0.043*
2/3	4.358+	1	0.037*
2/4	2.857+	1	0.091– trend
2/5	2.326+	1	0.127
3/4	0.737	1	0.391
3/5	1.140	1	0.286
4/5	0.090	1	0.765

+ 1 cell has an expected value of less than 5
# 2 cells have an expected value of less than 5

While the chi-square test for all five conversations shows that they represent more than one 'population', chi-square tests for individual pairs of conversations are not significant in many cases. They fall into two groups: low Russian usage (conversations 1 and 2), and moderate Russian usage (numbers 3, 4, and 5). However, this categorization is not perfect since the chi-square tests between conversations 2 and 4 as well as 2 and 5 are not significant at the 95% confidence level.

Through the differences in Russian use, we can see the interplay of interlocutor, topic, and identity being played out. Previous studies have shown a relationship between

language choice in Azerbaijan and gender construction, with Russian associated with some femininities and Azerbaijani with masculinity (Zuercher 2004, 2009). Even in Soviet times, with strong influences toward Russification, Azerbaijani was associated with the home domain. Lastly, in contemporary Azerbaijan, Azerbaijani is the preferred language in virtually all public domains, and is required in all governmental proceedings.

In the first and second conversations we see several of these factors talked out. In the first, Sevil is presumably speaking with a male relative about a family visit. The interlocutor and topic do not apparently call for any Russian linguistic elements. In the second conversation, Sevil is speaking with a fellow teacher because she is trying to find someone to do some translation work for a governmental agency. Though Russian is appropriate for some femininities, it is not appropriate when speaking about government work. In the brief conversation, Sevil only uses one Russian lexical item, an adverb inserted into an Azerbaijani CP:

Sevil #330	çox	maraq-lı-dır.	həmişə	hamı	iş	iş	axtar-ır
	much	interesting-be	always	everyone	work	work	search.for-3.sg.cont
	mən	indi	iş	ver-mək	istə-yir-əm.		
	1.sg	now	work	give-infin	want-cont-1.sg		
	наоборот	deyirlər.	yox?				
	backwards [naoborot]	say-cont-3.pl	no				

'It is very interesting. Everyone is always looking for work, but I am trying to give work away. Don't they say that's backwards?'

But what function does saying this Russian adverb serve? It is clearly not related to the ideational structure of the conversation. The topic does not call for any other Russian lexical items. Since it is the lone Russian item in the conversation it is also unlikely that

it is realted to the participation framework – this pairing of interlocutors does not evoke Russian langauge anywhere else. It is possible that the langauge of the adverb serves as a contextualization cue, relating to the action structure of the utterance. By speaking it in Russian instead of Azerbaijani (the conversation's default language), Sevil could be emphasizing the unusualness of the situation. Lastly, Sevil does not orient to the 'foreign' lexical item in any way: no pauses, self correction, raised voice, etc. Thus, it is possible that it is simply a case of $transfer_L$ rather than $transfer_P$ (Auer 1984). Sevil may not have realized that she spoke the word in Russian, making it simply a manifestation of her own linguistic and cognitive abilities (information state in Schiffrin's model).

In the third phone conversation Sevil talks with her husband, Azər. At 8.8% this conversation had the highest ratio of Russian content words – 15 of 170 words. Nine of these Russian content words occur in a pair of turns where Sevil tells a story about seeing a family get the news that their child needed an operation. These two turns are repeated in full on the following two pages.

Sevil #431

(1.) hə de-d-im hara-da , Республикансk-ий-də bir iki dəfə zəng elə-d-im ev-ə,
yes say-past-1.sg where-loc republican-m.s.n/a-loc one two occasion ring do-past-1.sg house-dat
[rɛspublikansk-iy]
'Yes, I told them there at the Republic (hospital), once or twice I called to the house'

(2.) ded-i yox-dur hələ bir şey yox-dur hələ bir şey yox-dur sən iş-də ol-an-da
say-past no-be present one thing no-be present one thing no-be 2.sg work-loc be-ing-loc
'He said there is nothing now, there is nothing now (for) you, from being at work…'

(3.) прямо gəl-d-im həyət-ə gör-d-üm hamısı ora-da-dır-lar, Республикансk-in-in
straight come-past-1.sg garden-dat see-past-1.sg everyone there-loc-be-3.pl republican-possd- poss
[prjamo] [rɛspublikansk]
'I came directly to the garden. Everyone was there (at) the Republic's (hospital)…'

(4.) əmi-si, əmi-si-nin yoldaş-ı nə həyət-in-də
uncle-3.sg.possd uncle-3.sg.possd-3.sg.poss comarade-3.sg.possd what garden-3.sg.possd-loc
'his uncle, his uncle's friend, and others in the garden'

(5.) bil-im yoldaş-ın-ın baldız-lar-ı nə-yi-sə, hə,
know-1.sg comarade-3.sg.possd-3.sg.poss sister.in.law-3.pl.possd what-acc-poss.fut yes
'I know, the friends sisters-in-law, whatever, yes'

(6.) elə mən ora-da ikən, umm, içəri-dən xəbər ver-di-lər ki ol-du,
thus 1.sg there-loc gathering bk-chnl inside-abl news give-past-3.pl rel be-past
'So, me, there (was this) gathering, um, so from inside they gave the news, what happened.'

(7.) sonra ora-da o həkim var onların tanış-ı , telefon-a çək-miş-di
after there-loc 3.sg doctor be 3.pl.poss acquaintance-3.sg.possd, telephone-to pull-imperf-past
'After that there was a doctor, their acquaintance, who made a phone call'

(8.) çıx-dı çöl-ə hə elə Azər-in-i necə çək-ib-lər-sə операционн-ı-da,
go.out-past field-to yes thus Azer-3.sg.poss-acc how pull-imperf-3.pl-poss.fut operating-3.sg.possd-loc
[opɛatsion]
'Then went out to the open lot, you know; like when they took out Azerin's (something) in the operation.'

(9.) bun-u da eləcə bütün klip kimi çəkib bütün
this-acc therefore simply complete short.video like pull-imperf complete
'So this was totally like it was just shot for a video.'

(10). hamsı-nı по эпизод-ам, как реж-ут живот как вытаскива-ют,
all-acc by episode-m.p.dat how cut-3.pl stomach how pull.out-3.pl
[po ɛpizodam kak rɛʒut ʒivot kak vɯtaskevajut]
'Everything (was like) in a (soap opera) episode, 'How they cut the stomach and take (it) out [appendectomy].'

(11.) uşağ-ı da hamsı-nı, uh, hə, göstər-dı hamsı-nı bizə, mm,
child-acc therefore all-acc bk-chnl yes show-past all-acc 1.pl-to bk-chnl
'So the child everything, uh, yeah, (he) showed us everything

(12.) arvad da ora-dakı həkim var onlar-ın tanış-ı-dır
wife therefore there-in.which doctor there.is 3.pl.poss acquaintance-3.sg.possd-be
'So the wife there, there was a doctor who they knew…'

Sevil #433

(13.) mən de-di-m ürəy-imiz xarab ol-acaq uşağ-ın səs-in-i eşid-ən-də
1.sg say-past-1.sg heart-1.pl.possd rotten be-fut child-3.sg.poss voice-3.sg.possd-acc hear-ing-loc
'I said our hearts will break for hearing the child's voice.'

(14.) adam üçün xoş-dur da, hə, да уже она key kimi idi
person because nice-be therefore yes yes now 3.s.f.n numb like be. past
[da uʒɛ ona]
'Well because of that person being nice, yeah, yes now she was like numb.'

(15.) yat-mış-dır öz-ün-ü gör-mə-di-m çıx-dı-m gəl-di-m,
sleep-imperf-be self-3.sg.possd-acc see-neg-past-1.sg go.out-past-1.sg come-past-1.sg
'She has gone to sleep herself, I did not see (it), I left and then came (home).'

(16.) hə, qız-ın şahmat-ı var-dı gəl-di-m get-di-m
yes girl-3.sg.poss chess-acc there.is-past come-past-1.sg go -past-1.sg
'Yeah, there was the girl's chess game, I came and left.'

The instances of code-switching/mixing within this section of the conversation cover the full range of Muysken's typology:

Insertion

Russian nouns in Azerbaijani CPs with Azerbaijani morphology (lines 1, 3, and 8)

Республиканск-ий-də	Республиканск- ий -in	операционн-ı-da
republican-m.s.n/a-loc	republican-m.s.n/a-3.sg.poss	operating.room-3.sg.possd-loc
[rɛspublikansk-iy]	[rɛspublikansk-iy]	[opɛatsion]

Russian adverb inserted into an Azerbaijani CP without any morphological incorporation (line 3)

sən	iş-də	ol-an-da	прямо	gəl-di-m	həyət-ə
2.sg	work-loc	be-part-loc	straight	come-past-1.sg	garden-dat
			[prjamo]		

'You were at work (so) I went to the court yard right away.'

Alternation

Russian CPs followed in the turn by Azerbaijani CPs (line 10 followed by 11)

как	реж-ут	живот	как	вытаскива-ют
how	cut-3.pl	stomach	how	pull.out-3.pl
[kak	rɛʒut	ʒivot	kak	vɯtaskevajut]

'How they cut the stomach and take (it) out.' (probably referring to appendectomy)

Russian peripheral alternation (twice) and a Russian pronoun as the subject (line 14):

hə,	да	уже	она	key	kimi	idi
yes	yes	now	3.s.f.n	numb	like	be.past
	[da	uʒɛ	ona]			

'Yeah, yes, now she was like numb.'

Congruent lexicalization

Russian and Azerbaijani elements in a stative clause without an overt verb (line 10).

hamısı-nı	по	эпизод-ам
everyone--acc	by	episode-m.p.dat
	[po	ɛpizodam]

'Everyone (was like) in a (soap opera) episode.'

The four instances of insertion seem to relate to at least two different levels of discourse in Schiffrin's model. The two insertions of *Республиканск-ий* [rɛspublikansk-iy] would require an information state where the Russian insertions are appropriate to Sevil's linguistic knowledge as well as fitting her understanding of the listener's knowledge. These refer to the Republic Hospital in Baku, an institution dating from Soviet times:

Russian	Azerbaijani
Республиканская Болница [rɛspublikanskaja bolnitsa]	Respublika Xəstəxanası

Though both names would have been used in official documents in Soviet times, the Russian version would have been most commonly used. Since the fall of the U.S.S.R., and especially since the President's Declaration on Language Use, the Azerbaijani name would be the only one used in official documents. At 44 years of age in 2007, when these data were collected, Sevil would have completed all of her education under the U.S.S.R., including Russian language university. Though she spoke Azerbaijani at home, it is most likely that she learned the name of this hospital in Russian. While she would have the ability to use the Azerbaijani name when necessary, she would only do so to achieve some particular conversational end. These two instances of Russian insertion almost certainly represent transfer$_L$ rather than 'true' code-mixing, and thus they relate to the conversation's Information State. Moreover, their use relates to the interlocutor. Traditionally, husbands are a few years older than wives in Azerbaijan, so we can expect Azər to be in his mid-forties. Thus, he would have experienced very similar cultural linguistic influences. There was no need, in this conversation then, to switch from the

Russian to the Azerbaijani name for the hospital. So, as well as relating to Sevil's unique Information State, the use of these Russian lexical items also relates to his cognitive and linguistic abilities, and to the hearer/speaker Participation Framework. Sevil's use of the Russian word for 'operating room' would be somewhat different. As a technical term, she probably learned this term in adolescence or even at university. Under the U.S.S.R., medical science was a Russian oriented domain, thus she would have learned the word in Russian. Though she might know the Azerbaijani word, she almost certainly uses it very rarely since she does not work in a medical field. She probably chose the word because it best represented her intentions (Myers-Scotton 2002), thus it is a reflection of her own linguistic knowledge and abilities (information state). Lastly, the insertion of the Russian adverb *прямо* [prjamo] 'straight' is quite different. While it is made possible by the same combination of information states and participation framework that brought about the nominal insertions, it must certainly be acting as a contextualization cue in this utterance since the Azerbaijan equivalent, *duz*, is a very common word, but she does not choose it. Thus, by saying the adverb in Russian, Sevil is emphasizing the directness with which she went to the court yard of the hospital to see the family.

This section of phone conversation number three contains two different types of alternation: alternation between CPs within a conversational turn, and peripheral alternation within a CP. As shown in lines 10 and 11 above, Sevil utters two full CPs in Russian which are followed by Azerbaijani CPs. These Russian CPs do not contain any Azerbaijani lexical items, reflecting the overall avoidance in all the Home Recording data of code-mixing when Russian is the matrix language. These would only make sense in

the conversation if both speaker and hearer were fairly proficient in Russian, reflecting the same combination of information state and participation framework discussed above. They would also relate to the ideational structure of the conversation, since medicine was and to some extent still is a Russian dominated field. In describing the procedure to be performed, Sevil switches to Russian, though she does not use any technical terminology. Moreover, these two Russian CPs follow the one instance of congruent lexicalization (beginning of line 10) discussed below. That clause ends with Russian lexical items, so these Russian CPs may have also been triggered by the immediately preceding context. After these two Russian CPs however, she switches back to clauses in Azerbaijani with no code-switching/mixing.

Though syntactically similar, the Russian peripheral alternation in line 14 seems to have a very different relationship to the layers of discourse in Schifrin's model than the alternation between CPs. This sentence is the only one in which Sevil overtly orients to the code-switched/mixed elements. She starts the CP with an Azerbaijani peripheral adverb, *hə* 'yes', then follows it with the semantically equivalent item in Russian, да [da] 'yes'. This, in turn, is followed by the Russian peripheral adverb уже [uʒɛ] 'now'. The matrix language of the CP is established by the overt Azerbaijani stative verb *idi* 'was (3rd singular, past)'. Thus this Azerbaijani CP has three [C, CP] slots filled by both Azerbaijani and Russian lexical items. In contradiction to Muysken's Adequacy Principle, these three peripheral adverbs are followed by a Russian pronoun filling the subject position of the clause – there is no way to syntactically interpret the three Russian items as part of a single larger constituent. As with the other instances of code-

switching/mixing in this conversation, these Russian alternations and insertion are conditioned by the combination of information state and participation framework. Their unusual nature, however, indicates that they relate to the action structure of the conversation. They serve as contextualization cues to highlight this particular CP.

Lines 13 and 14 appear to correspond to the narrative peak of this brief story. This is marked firstly by a change in perspective: "Our hearts will break from hearing the child's voice." Moreover, the very unusual stacking of adverbial elements in multiple [C, CP] slots serves as further rhetorical underlining (Longacre 1996) as does the Russian subject pronoun *она* [ona] 'she'. This sentence occurs toward the end of this conversational section, and it contains the last occurrences of code-switching/mixing for this story. Thus, these adverbs and pronoun serve to emphasize the peak of the story by emphasizing the affect that this bad news has had on the mother. Since the Russian pronoun is marked for gender, it also provides a more efficient way of specifying the subject than the unmarked Azerbaijani third person pronoun.

The single occurrence of congruent lexicalization in this narrative, while conditioned by the information state and participant framework of the conversation, is somewhat ambiguous. The stative clause contains both an Azerbaijani NP and a full Russian PP. Since Sevil does not orient to the change in languages by pauses, repetition, etc. we must assume that this is a case of transfer$_{L.}$ Though both halves of the clause contain morphology appropriate to the two languages, we are forced to interpret this simply as a manifestation of the mixed-code that Sevil sees to be appropriate for the speech situation.

In the fourth conversation Sevil gets a call from an older relative. Table 7-7 shows that the code-switching/mixing differs from the other conversations in that there is a higher number of Russian CPs. The majority of the Russian lexical items counted in that table occur in these 8 CPs:

Sevil #762	hə,	дай	Бог,	дай	Бог	
	yes,	give.2sg.inf.imper	God,	give.2sg.inf.imper	God	
		[daj	bog	daj	bog]	
	‘Yes, may God give (it), may God give (it).’					
Sevil #854/860	молодец					
	well.done					
	[molodɛts]					
Sevil #894	неужели	belə	это	реально		
	really!	thus	this-n.n/a	real		
	[nɛuʒɛli		ɛto	realno]		
	‘Really! It is really like this.’					
Sevil #906	əşi	yaxşı	ey,	поздно	буд-ет ,	поздно буд-ет
	now.then	good	look.here,	late	be-3.sg,	late-s.n be-3.sg
				[pozdno	budɛt	pozdno budɛt]
	‘So good then, it will be late, it will be late.’					
Sevil #912	но но,	то	же	сам-ое,	да	
	but but,	that.n.n/a	emph.	same-n.s.n/a,	yes	
	[no no	to	ʒɛ	camoʲɛ	da]	
	‘But, but, that is really the same, yes.’					

This analysis does not count *давай* [davaj] (literally) ‘give (it)’ as a separate CP, since in practice it is used as an adverb parallel to the English ‘OK’. All eight of these Russian CPs occur in the second half of this conversation where Sevil is playing a supportive conversational role while her uncle talks about family business. Thus, this portion of the conversation is characterized by alternation rather than the insertion common in the rest of the Home recordings. Since Sevil was in her mid-forties when this recording was

made, we can assume that her uncle was likely in his late fifties or early sixties. Sevil's responses to his family news indicate that he must have a high level of Russian proficiency, and that he may have been speaking in Russian to her. Thus the code-switching in this conversation relates to both the conversation's information state (Sevil's linguistic abilities as well as her understanding of her uncles linguistic abilities) and the participation framework – by accommodating his linguistic choices she chooses to build solidarity with him.

Though conversation 5 is fairly short (27 turns by Sevil) its emotional Key is more intense than that of the other conversations. In this phone conversation Sevil discusses some work issues with a female co-worker. Given the moderate number of Russian content words, we should assume that this interlocutor also knows Russian. While there is one Russian CP, it is more of a 'micro-clause':

Sevil #939	ничего
	nothing
	[nitʃɛvo]
	'(It is) nothing.'

In contrast to conversation 3 that uses both Russian CPs and Russian insertion into Azerbaijani CPs, or conversation 4 where Sevil utters a number of Russian CPs in support of her uncle's talk about family business, this conversation uses Russian insertion and peripheral alternation to emphasize the topics under discussion and heighten the emotional tension. Twice Russian nouns are inserted into Azerbaijani CPs, once with morphological integration and once without, and once a Russian NP – *как раз* [kak raz] 'that time' – fills the subject slot in an Azerbaijani CP. The Russian demonstrative *это* [ɛto] 'this' occurs in an Azerbaijani CP:

Sevil #949	bu	gəl-ir	это
	this	come-cont	this-n.n/a
			[ɛto]
	'It is coming.'		

Here Sevil repeats the subject in both languages, emphasizing its immediacy. In this conversation, Sevil has two CPs with Russian adverbial elements inside Azerbaijani CPs:

Sevil #949	ay	qız-ım	indi	дипломатично	да,	de-di-m
	hey	girl-1.sg.possd	now	diplomatically	yes,	say-past-1.sg
				[diplomatitʃno	da]	
	'Hey my girl, now yes, I said (it) diplomatically.'					

Sevil #949	institut-dan	уже	mən	çıx-a	bil-m-ir-əm
	institute-abl	now	1.sg	go.out-part	know-neg-cont-1.sg
	'Now I cannot leave the Institute.'				

In both cases the Russian elements appear in the middle of the Azerbaijani CPs, but they may well represent cases of Russian peripheral alternation with extensive movement. In the second example, the Russian adverb *уже* [uʒɛ] 'now' could have occupied a [C, CP] position, but the Azerbaijani object PP, *institutdan*, was moved to the left of it for emphasis, filling an additional [C, CP] slot. In the first the Azerbaijani CP *dedim* 'I said (it)' could simply be preceded by four [C, CP] slots filled by an Azerbaijani vocative *ay qızım* 'Hey my girl', an Azerbaijani adverb *indi* 'now', a Russian adverb *дипломатично* [diplomatitʃno] 'diplomatically', and then by the Russian adverb *да* [da] 'yes'. Clearly in both cases the movement and code-switching/mixing serve as contextualization cues, heightening the emotional tension of the conversational turn.

Like the insertion described above, the instances of peripheral alternation in this conversation seem to serve as contextualization cues adding emphasis to the clauses they

modify. In only one case is the peripheral adverb a single word. In all other cases there are either multiple clause level modifiers, or they are multi-word units:

Sevil #949	уже	в	общ-е	pis	şey	alın-ır
	now	in	general-m.s.prep	bad	thing	be.brought-cont
	[uʒɛ	v	obʃɛ]			
	'Now in general bad things are happening.'					

Sevil #959	абсолютно	абсолютно	bu	Məcid-in	dekanlığ-ı-dır-sa
	absolutely	absolutely	this	Majid-3.poss	dean's.job-3.possd-be-cond
	[absoljutno	absolyutno]			
	'Absolutely, absolutely, if this is Majid's job as dean…'				

Sevil #973	во перв-ых,	qoy	get-sin	konkurs-dan	keç-sin
	in first-p.g/a/pput	go-2.inf.imper		competition-abl	pass-2.inf.imper
	[vo pɛrvɯx]				
	'First off, go put yourself in the running.'				

Sevil #983	давай, давай	давай	yaxşı	sağ	ol
	let's, let's	let's	good	health	be
	[davaj davaj	davaj]			
	'OK, OK OK, good, thanks.'				

Like the previous conversations, these instances of code-switching/mixing are conditioned by both an information state and participation framework allowing for Russian language use. Unlike those conversations, however, these instances of Russian insertion and peripheral alternation have more to do with the action structure of the conversation than with anything else. Rather than simply adding vocal emphasis, or using Azerbaijani syntactic structures to strengthen her statements, Sevil as a bilingual Azerbaijani/Russian speaker uses code-switching/mixing to construct and emphasize the heightened emotional key of this conversation.

In these phone conversations we see Sevil playing a number of different conversational roles, each of which seems to call for different norms of language use. In

the first, she speaks with a male relative about a family visit. Her role as 'hostess' the facts that the interlocutor is a male, and the conversation related to the home domain, call exclusively for Azerbaijani language items. While the fourth conversation also relates mostly to the family domain, something about that interlocutor called for increased Russian language use. It is quite likely that her uncle was speaking Russian in those segments, and she was accommodating his language choices. That would be consistent with her role as supportive younger relative in that half of the conversation. In the second conversation, Sevil plays the role of government contractor by trying to secure employees for translation work. Though she uses one Russian lexical item, this role clearly calls for proficiency in Azerbaijani. Her conversational role as wife in the third conversation as well as her reported role as supportive friend, both call for a moderate level of Russian use, both because of the medical topic under discussion for part of the conversation, and because of the co-constructed Information State: she has adequate Russian ability and she believes that her husband does as well. In the fifth conversation Sevil acts out a co-worker role. Since she works in a Russian oriented academic domain, Economics, her interlocutor should also have fairly high Russian proficiency. This role, the information state created by the conversational participants, and the emotional key of the conversation all work together to encourage Sevil to use Russian insertion and alternation as contextualization cues.

Through all of these phone conversations, we see Sevil displaying proficiency in shifting her language behavior according to the situation. She uses the linguistic

resources available to her to effectively negotiate a very complex system of changing social norms with great intersituational fluidity.

7.2.2 Talk with daughter

Sevil's home recording lasts just under two hours, and the bulk of this time was not spent on the phone. Rather, she talked with her five year old daughter about a number of things. This talk was accompanied by long stretches of silence where Sevil was doing household chores while her daughter watched TV and/or played. Unlike the phone conversations, however, some of the themes of Sevil's talk with her daughter were spread out and in some cases mixed together rather than having clear boundaries.

Table 7-9 Summary of Sevil's talk with her daughter according to Hymes' etic grid

Setting	at home in the afternoon
Participants	Sevil – mother, 44 years old, university professor Daughter – 5 years old
Ends	Chess theme – playing chess and teaching daughter rules of chess Food theme – discussing what the daughter ate at school, what to eat, and what they are eating Song theme – teaching daughter a Soviet children's song in Russian TV theme – discussing what to watch on TV and what they are watching
Act sequence	conversational turn taking, often with large gaps
Key	for each theme, various keys were present ranging from jovial mother/daughter chat to emphatic correction.
Instrumentalities	only verbal instrumentality recorded but since mother and daughter were both present, in many cases there must have been other forms of interaction as well
Norms	to be determined
Genre	Chess theme – teaching and game oriented chat Food theme – mostly informal chat Song theme – teaching TV theme – informal chat

Table 7-10 Summary of Sevil's conversations with her daughter according to three levels of discourse from Schiffrin's theoretical model

Information State	Sevil – 44 year old university professor who attended both Azerbaijani and Russian schools. Educated level of both Azerbaijani and Russian. Daughter – 5 years old, probably attending Russian school. Appropriate native speaker proficiency in Azerbaijani for a native speaking child, but with low Russian proficiency.
Participation Framework	mother to young daughter
Ideational Structure	Chess theme– playing chess and teaching daughter rules of chess Food theme – discussing what daughter ate at school, what to eat, and what they are eating Song theme – teaching daughter a Soviet children's song in Russian TV theme – discussing what to watch on TV and what they are watching

Four major themes occur in Sevil's conversation with her daughter: talking about TV programs, talking about Food, teaching a Song, and playing a game of Chess. These topics occupy 278 of the 330 conversational turns that Sevil contributes in talking with her daughter. Table 7-9 and Table 7-10 summarize Sevil's talk on these themes according to Hymes' etic grid and Schiffrin's theoretical model.

One major difference between the four themes themselves and between these and the phone conversations discussed in the previous section is boundaries. Phone conversations make convenient units of analysis because they have clear beginnings and endings. They begin when the phone rings and end when it is hung up. In contrast, the Food and TV themes occur throughout the recording, providing a conversational backdrop to the other topics of discussion. The Song and Chess themes are more clearly bounded. Chess only has two breaks: one for the very brief sixth phone conversation, and the other where Sevil returns to the Song theme for one turn: her daughter has

continued singing the song while they play chess and Sevil corrects one word. Accordingly, the one break in the Song theme is because Sevil briefly returns to this topic while playing chess.

Like the phone conversations, these conversational themes differ both in terms of language use and conversational support. Table 7-11 summarizes both the gaps in each conversational theme and summarizes the rate of conversational support that Sevil provides in each. Table 7-12 shows the results of chi-square tests both for all four themes under discussion as a group and between each pair.

Table 7-11 Distribution of backchannel turns in child conversational themes

Theme	**Breaks**	**Total turns**	**Backchannel turns**	**% BC turns**
Chess	2	136	16	11.8%
Food	6	54	0	0.0%
Song	1	36	5	13.9%
TV	9	52	7	13.5%

Table 7-12 Chi-square tests for non-backchannel versus backchannel turns in Sevil's conversation with her daughter

Theme	**Chi-square**	**df**	**Significance**
All Four	7.717+	3	0.052– trend
chess/food	6.937+	1	0.008*
chess/song	0.120+	1	0.729
chess/TV	0.101	1	0.751
food/song	7.941#	1	0.005*
food/TV	7.783#	1	0.005*
song/TV	0.003+	1	0.954

+ 1 cell has an expected value of less than 5
# 2 cells have an expected value of less than 5

In considering the differences in Sevil's level of conversational support for each conversational theme, we must examine how Backchannel turns are assigned to each theme. Since only one side of the conversation was recorded, we cannot be certain that

any particular Backchannel turn was supporting the statements made by the daughter pertaining to any specific conversational topic. We can only assign a Backchannel turn to a theme if it is bounded by information relating to that topic. For Food and TV this is problematic. These themes are spread over the entire recording and turns or sections of the recording were assigned to them based on the information in Sevil's talk. Thus, some levels of conversational support in Table 7-11 may be artifacts of the research methodology rather than reflections of reality. Since the boundaries for the Chess and Song themes are more clearly defined, the information on conversational support is also more reliable. Thus conclusions based on Sevil's number of Backchannel turns in speaking with her daughter should be viewed as tentative.

As Table 7-12 shows, the chi-square test for Backchannel and Non-Backchannel turns among all four conversational themes is not statistically significant at the 95% confidence level; whether this finding stems primarily from a dearth of relevant data is a matter to be verified in future research. Nevertheless, the range of p-values across the various pairs, along with the 0.052 p-value for the larger chi-square test (suggestive of a trend) indicates the need for some further analysis. One notes, for example, that when viewed in pairs, only the Food theme and the other three show themselves to be independent. Thus the themes fall into two groups. In the Food theme, Sevil offers all content-oriented turns, while for the other three themes she shows a consistently moderate level of conversational support.

As Table 7-1 shows, Sevil used about three times as many Russian content words with her daughter, than in the phone conversations. Table 7-13 shows the relative amounts of each measure of overall language use in the four themes under discussion.

Table 7-13 Distribution of overall language use measures among child conversational themes

	Content words			Russian CPs			Insertion			Alternation	
Theme	**Az**	**Ru**	**% Ru**	**CPs**	**Ru CPs**	**% Ru CPs**	**Ru in Az CP**	**% Ru ins.**	**Az in Ru CP**	**Ru on Az CP**	**Az on Ru CP**
Chess	317	32	9.2%	245	26	10.6%	6	2.7%	n/a	6	n/a
Food	110	46	29.5%	79	20	25.3%	7	11.9%	n/a	2	n/a
Song	33	53	61.6%	54	33	61.1%	2	9.5%	n/a	1	n/a
TV	112	9	7.4%	74	1	1.4%	6	8.2%	n/a	n/a	n/a

Of these, Russian content words, Russian CPs and the insertion of Russian elements into Azerbaijani CPs have sufficient tokens for chi-square tests as shown in the following three tables.

For all three of these measures, the chi-square shows that there is a statistically significant chance that the four themes represent independent linguistic populations (the first line in all three tables).

Table 7-14 Chi-square tests for Azerbaijani versus Russian content words in Sevil's conversation with her daughter

Theme	**Chi-square**	**df**	**Significance**
All Four	141.184	3	0.000*
chess/food	34.079	1	0.000*
chess/song	120.771	1	0.000*
chess/TV	0.338	1	0.561
food/song	23.691	1	0.000
food/TV	20.819	1	0.000*
song/TV	70.361	1	0.000*

For Russian content words all of the chi-square values between the pairs of themes are significant with the exception of Chess and TV. When looking at Russian versus

Azerbaijani CPs all of the chi-square tests are significant, even that between Chess and TV.

Table 7-15 Chi-Square tests for Non-Russian versus Russian CPs in Sevil's conversation with her daughter

Theme	**Chi-Square**	**df**	**Significance**
All Four	95.035	3	0.000*
chess/food	10.603	1	0.001*
chess/song	71.241	1	0.000*
chess/TV	6.291	1	0.012*
food/song	17.145	1	0.000*
food/TV	18.532	1	0.000*
song/TV	57.154	1	0.000*

While the frequency of Russian insertion in Azerbaijani CPs is fairly small for all four themes, when added to the Russian lexical items in the Russian CPs, the insertions in the TV theme are enough to make the levels of Russian content words for the Chess and TV themes almost equal.

Table 7-16 Chi-square tests for Azerbaijani CPs with and without Russian insertion in Sevil's conversation with her daughter

Theme	**Chi-square**	**df**	**Significance**
All Four	8.104#	3	0.044*
chess/food	7.537+	1	0.006*
chess/song	2.430+	1	0.119
chess/TV	3.745+	1	0.053– trend
food/song	0.068+	1	0.794
food/TV	0.399	1	0.527
song/TV	0.030+	1	0.863

+ 1 cell has an expected value of less than 5
# 3 cells have an expected value of less than 5

Though the overall effect is significant in Table 7-16, only the chi-square for Chess and Food is significant at the 95% confidence level. With a 94.7% chance that Chess and TV represent independent linguistic groups, the difference in the frequencies of Russian and

Azerbaijani content words probably represents a statistical trend – one that merits further investigation.

Though each conversational theme has a different level of Russian insertion, and there is variation in terms of peripheral alternation, most of the variation in code-switching/mixing among the four conversational themes is in the frequency of Russian language clauses. The Song theme has the highest proportion of Russian CPs followed by the Food theme. Chess and TV have much lower levels, though the difference is significant.

In the Song theme, Sevil is teaching her daughter a Russian language children's song dating from Soviet times: *Солнечный круг* [solnɛtʃnɯj krug] 'The Circle of the Sun' (see Appendix C for complete lyrics). The majority of the Russian CPs in this section are from the song that Sevil is teaching. She also encourages her daughter seven times with the Russian idiom *молодец* [molodʲɛts] 'well done!' Each of these was counted as a Russian CP. As the following example shows, many of the Russian CPs are housed in a two long turns where Sevil is singing the song for or with her daughter. In both cases the song is framed by Azerbaijani CPs:

Sevil #1033	səhv	oxu-yur-san,	səhv	oxu-yur-san,
	wrong	sing-cont-2.sg,	wrong	sing-cont-2.sg,
	'You're singing it wrong, you're singing it wrong'			

istə-yir-sən	düzəl-d-im?
want-cont-2.sg	correct-past-1.sg
'Do you want me to correct (it)?'	

пусть	всегда	буд-ет	солнце,
let	always	be-3.sg	sun
[pust	vsɛgda	budɛt	solntsɛ]
'Let there always be sun'			

пусть	всегда	буд-ет	небо,
let	always	be-3.sg	sky
[pust	vsɛgda	budɛt	nʲɛbo]

‘Let there always be sky’

пусть	всегда	буд-ет	мама,
let	always	be-3.sg	mom
[pust	vsɛgda	budɛt	mama]

‘Let there always be mama’

пусть	всегда	буд-у	я
let	always	be-1.sg	1.s.n
[pust	vsʲɛgda	budu	ja]

‘Let me always be’

A further example from this theme gives us some idea of Sevil’s daughter’s Russian proficiency:

Sevil #1025	Azər-in	ora-da	səhv	dey-ir-sən
	Azer-3.sg.possd	there-loc	wrong	say-cont-2.sg

‘Azer’s (daughter) you are saying (it) wrong there’

“нарисабат”	уох	нарисова-л	eşit	bir	də
???	no	draw.perf-past	hear	one	therefore
[narisabat]		[narisoval]			

‘It’s not narisobat (but) narisoval, so listen once’

de	нарисова-л	он	на	листке
say	draw.perf-past	3.s.m.n	on	paper-n.s.prep
[narisoval		on	na	listkɛ]

‘Say ‘it is drawn on paper’’

Here we get the impression that the daughter is not thinking about what the words mean, but is rather repeating the sounds – learning the song by rote rather than focusing on the content. Thus it seems that Sevil has two motivations in this section of the conversation: to teach her daughter this particular song, but also to teach her Russian in general.

In Table 7-13 the Food theme shows the next highest ratio of Russian to Azerbaijani clauses. Twenty of the seventy nine CPs in this section have Russian as the

matrix language. However, all but five of these occur in a very short section where Sevil switches entirely to Russian for eight turns:

Sevil #393
что?
what
[ʃto]
'What?'

Sevil #395
ты	перв-ая?	молодец,	я	буд-у	перв-ая
2.s.n	first-f.s.nom	well.done	1.s.n	be-1.sg	first-f.s.nom
[tɯ	pɛrvaja	molodɛts	ja	budu	pɛrvaja]

'You first? Very well. I will be first.'

Sevil #398
покуша-л-а?
eat.perf-past-s.f
[pokuʃala]
'You already ate?'

Sevil #400
молодец,	принцесс-ы	свою	тарелку	сюда
well.done	princess-f.s.g/a	own-f.s.g/a	plat-f.s.n/a	here
[molodɛts	printsessɯ	svoju	tarɛlku	sjuda]

'Well done. The princesses own plate is here.'

Sevil #402
спасибо
thanks
[spasibo]
'Thank (you).'

Sevil #404
молодец	а	помидорку	не	покуша-ешь?
well.done	and	little.tomato	not	eat-perf-2.sg
[molodɛts	a	pomidorku	nɛ	pokuʃaʲɛʃ]

'Well done, and don't you want a little tomato?'

Sevil #406
а	хорошо	я	покуша-ю	иди	рук-и	помой,
and	well	1.s.n	eat.perf-1.sg	go..imper	hand-f.p.n/a	wash.perf.imper
[a	xoroʃo	ja	pokuʃaju	idi	ruki	pomoj]

'Well good. I will eat. Go wash (your) hands.'

иди,	иди	быстро,,	нет,	не	там,	здесь
go.imper	go.imper	quickly	no	not	there	here
[idi	idi	bɯstro	nʲɛt	nʲɛ	tam	zdɛs]

'Go, go, quickly. No. Not there, here.'

Sevil #409	вымы-л-а	рук-и,	молодец,	чай	хоч-ешь,,	
	wash-past-s.f	hand-f.p.n/a	well.done	tea	want-2.sg	
	[vɯmɯla	ruki	molodɛts	tʃaj	xotʃɛʃ]	

‘Did you wash (your) hands? Well done. Do you want tea?’

Açka,,	чай	буд-ешь?
Achka	tea	be-2.sg?
	[tʃaj	budɛʃ]

‘Achka, will you have tea?’

This covers just over three minutes of the recording and contains two very large gaps: 47 seconds between #395 and #398, and one minute twenty three seconds between #406 and #409. Over this time, the background noise indicates that Sevil is either washing dishes or cooking. What motivation is there for Sevil to switch to Russian for this segment of the recording? From the Song theme, we see that her daughter probably has fairly low Russian proficiency, so it was not function of the conversation’s information state since both participants do not have equal Russian proficiency. Rather it seems to relate to the action structure of the discourse. Sevil wants her daughter to learn Russian, so she switches to Russian. The same is true of an earlier instance in the same theme where she uses a Russian proverb to encourage her daughter not to talk with her mouth full:

Sevil #290	Azər-in,	когда	ем	я	глух	и	нем,
	Azer-3.sg.possd	when	eat	1.s.n	deaf	and	dumb
		[kogda^{j}	ɛm	ja	glux	I	n^{j}ɛm]

‘Azer’s (daughter), when eating I am deaf and dumb.’

adam	yemək	ye-yən-də	söz	danış-maz	ki!
person	food	eat.part-loc	word	speak-must.not	rel!

‘Thus, a person when eating food should not say a word!’

Here, the teaching is more overt since Sevil recasts the Russian proverb in Azerbaijani to ensure comprehension.

In the Chess theme we see some tension and or shift in language use. This portion of the recording can be divided into two parts. In the first part (34 of 136 turns) Sevil is talking about the chess pieces and how to set them up. In the remainder, Sevil and her daughter actually play the game. This section starts with a Russian clause:

Sevil #1124	давай	ты	начина-й
	let's	2.s.n	begin-imper
	[davaj	tɯ	natʃinaj]
	'OK, you start.'		

15 of the Russian clauses in this theme occur in the first portion. Sevil consistently refers to the pieces by their Russian names, but only refers to the name of the game once in Azerbaijani at the beginning of the theme. For example:

Sevil #1099	эт-а	пешущ-ая,,	a	сколько	там	пешек,
	this-f.n	pawn--f.s.nom	and	how.many	there	pawn
	[ɛta	pɛʃuʃtʃaja	a	skolko	tam	pɛʃɛk]
	'This is a pawn, and how many pawns are there?'					

сколько	там	пешек,	нет,	сколько	пешек?
how.many	there	pawn	no	how.many	pawn
[skolko	tam	pɛʃɛk	nʲɛt	skolko	pɛʃɛk]
'How many pawns are there? No, how many pawns?'					

Thus, the Russian orientation of the 'teaching' role for Sevil carries over from teaching the Song (the immediately preceding topic) to teaching her daughter chess. However, once the teaching portion is over, Sevil switches back to Azerbaijani. Since the game of chess is strongly associated with Azerbaijani culture and history, this is not surprising. For the remainder of the Chess portion of the recording, Sevil speaks almost exclusively in Azerbaijani. While she sometimes switches to Russian when discussing the pieces, she also uses their Azerbaijani names. She uses the Russian idiom *давай* [davaj] 'OK'

(literally ‘give it’) seven times – sometimes as a peripheral adverb on Azerbaijani CPs, and sometimes alone.

Sevil	#1212	давай	давай	başqa	şey	oyna,	oyna,	düz	oyna
		OK	OK	different	thing	play	play	correct	play
		[davaj	davaj]						

‘OK, OK move something else, play, play correctly.’

In this portion of the recording, it seems that Sevil uses alternation for emphasis – as a contextualization cue relating to the action structure of the discourse. For example, in this turn, Sevil corrects her daughter’s incorrect move:

Sevil #1261	bil-m-ir-sən	bura	at	gəl-ir,	
	know-neg-cont-2.sg	here	horse	come-cont	

‘Don’t you know, the knight goes here.’

bir	iki	üç	dörd	всё,	say	düzgün,	
one	two	three	four	all.n.s.n/a	,	number	correctly
				[vs^{j}o]			

‘One, two, three, four. That’s all. Count correctly.’

Of the four conversational themes under discussion, TV has the least Russian usage. It has the lowest proportion of Russian content words, and only one Russian headed clause. It has no peripheral alternation, but does have 6 Azerbaijani CPs with Russian elements inserted. With a total of 52 turns and 9 breaks, it serves as a backdrop to the other conversational themes. The TV is constantly running while Sevil and her daughter are cooking, cleaning, eating, playing chess, etc. Thus, it probably contains the least marked language use for Sevil. Therefore it is interesting to note that at 7.4% it still has a higher proportion of Russian content words than all 6 phone conversations except number 3 (8.8% Russian content words). The only Russian CP in this theme is this Russian idiom:

по щуч-ьему велени-ю
by own-f.s.dat will-n.s.dat
[poʃtʃutʃʲɛmu vɛlɛniyu]
‘(It appeared) as if by magic.’

The remainder of the Russian lexical items are nouns inserted into Azerbaijani CPs. Thus this combination of speakers (co-constructed information state and participation framework) and topic (ideational structure) call for more Russian lexical items than the majority of Sevil’s talk with adults. She is a Russian speaker who wants her daughter to learn Russian, so there is no reason for her to avoid Russian lexical items when they come to mind. This may or may not be true when talking with adults. However, in this recording when Sevil is talking about TV programs with her daughter she does not shift into Russian teaching mode.

7.2.3 Discussion

In Sevil’s hour and 53 minutes of recording we see her playing a number of conversational roles. Each role calls for different norms of language use. See Table 7-17. More important to this study, Sevil utilizes a wide range of code-switching/mixing behavior throughout her recording. These relate to at least four of the levels of discourse in Schiffrin’s model – Table 7-18.

Table 7-17 Sevil's conversational roles and their language norms

Section	**Role**	**Language use norms exhibited**
Phone Conv 1	-Female relative/hostess	-All Azerbaijani
Phone Conv 2	-Government contractor	-Azerbaijani, though Russian cs/cm possible for emphasis
Phone Conv 3	-Wife	-Azerbaijani, Russian, and cs/cm as appropriate to the family's language skills
Phone Conv 4	-Supportive younger relative	-Switch to Russian to accommodate/support older relatives language use
Phone Conv 5	-Agitated co-worker	-Mostly Azerbaijani, but cs/cm appropriate to express strong emotion
With Daughter – Chess	-Teacher -Chess opponent	-Speak Russian with recasting in Azerbaijani possible - Speak Azerbaijani, but cs/cm possible for emphasis
With Daughter – Food	-Mom -Teacher	-Azerbaijani with some Russian insertion -Speak Russian with recasting in Azerbaijani possible
With Daughter – Song	-Teacher	-Speak Russian with recasting in Azerbaijani possible
With Daughter - TV	-Mom	-Azerbaijani with some Russian insertion

From Table 7-18 we see that there is not a one to one relationship between the different levels of discourse and the types of code-switching/mixing in Muysken's typology. Each type of switching can be related to different levels of discourse and each level of discourse can invoke different kinds of code-switching. The biggest exception to this appears to be peripheral alternation. While it would require an appropriate combination of information state and participation framework, its functions relate mainly to the action structure of the discourse serving as contextualization cues to heighten the emotional tone of the conversation, emphasizing the content of the CPs it precedes.

Table 7-18 Types of code-switching/mixing in relation to Schiffrin's model of discourse

Discourse Level	**Conditioning Factors**	**Types of cs/cm**	**Where prevalent**
Information State	Sevil's bilingual proficiency and that of her interlocutor	- Alternation of CPs - Insertion – N/NP	- Phone Conv 4 - Phone Convs 3/4/5, all Child themes
Participation Framework	Sevil feels that cs/cm is appropriate with this interlocutor	- Alternation of CPs - Insertion – N/NP	- Phone Convs 3/4, Chess/Song themes - Phone Convs 3/5, all Child themes
Ideational Structure	Topic associated with a Russian oriented domain	- Insertion – N/NP - Alternation of CPs	- Phone Conv 3 - Phone Conv 3
Action Structure	Need for contextualization combined with appropriate information state and participation framework	- Insertion – N/NP/Adv - Peripheral alternation - Alternation of CPs	- Phone Conv 2/3/4 - Phone Conv 3/5 - Song/Chess/Food themes (teaching)

While still relating to the action structure of the conversation, the alternation of CPs in the Song, Chess, and Food themes serves a very different function. Rather than emphasizing the content, Sevil's choice of Russian for these sections of the recording tie in to cultural norms above the level of this conversation (Discourse with a big 'D' – Gee 1999). Here Sevil takes on the role of Russian teacher for her five year old daughter, and positions the daughter as student. This would in turn, cue up a number of cultural assumptions: Sevil is a qualified Russian teacher in this context, the daughter is able to learn Russian, Russian is a valuable language to learn, etc.

Therefore, by making the language choices we see above, Sevil is constructing a number of identities for herself, and co-constructing identities with her various interlocutors. By speaking Russian with her daughter she is both presenting herself with the identity of 'Russian Speaker' (see Laitin 1998) and assisting her daughter in constructing a similar identity. As Zuercher 2004 and 2009 show, such identities are more associated with femininity in Azerbaijan than they are with masculinity. If Sevil's goal as a mother is to assist her child in gaining the linguistic repertoires necessary to build socially accepted gender identities, the conversation might have been quite different if Sevil were speaking with a five year old son instead of a daughter. She could well have focused on Azerbaijani language use rather than Russian. By accommodating her uncle's choices to speak Russian, she presents herself as a product of Soviet education and culture and allows him to do the same. However, in the second phone conversation, Sevil proves she is able to move in the current socio/political milieu where Azerbaijani is the appropriate language for all official government interaction – she is a modern Azerbaijani free from the linguistic effects of Soviet language oppression. Though many other examples are possible, we see that Sevil uses her language proficiencies – Azerbaijani, Russian, and various combinations of the two – to navigate a complex system of socially approved identities with a high level of intersituational fluidity.

7.3 Two sisters: Aygun and Tarana

Sevil's case study was presented first for several reasons. Her contribution was arguably the most 'natural' of the Home recordings corpora since it consisted of one recording almost two hours in length. The recordings from other subjects were much

shorter. Moreover, Sevil had the lowest overall Russian use in the Home recordings. Thus each occurrence of Russian was easier to analyze in its sequential context. Lastly, in Sevil's recording we find clear transitions from one cluster of conversational roles and identity factors to another given the numerous phone conversations and shifts in topic in talk with her daughter. As a result, the two following case studies are much more brief and general in nature.

In the next case study we will look at two sisters in the Home recordings: Aygun and Tarana. They are both from the second largest city in Azerbaijan, Ganja, where both attended Russian language schools. Since Aygun was 28 years old in 2007 when the recordings were made and Tarana was 32, neither finished secondary school under the U.S.S.R.. Tarana made one recording of one hour and fifteen minutes where she was talking with Aygun and possibly Aygun's roommate.

Table 7-19 Summary of Aygun and Tarana's talk according to Hymes' etic grid

Setting	All recordings took place in Aygun's apartment before, during, and after an evening meal.
Participants	Tarana's recording – Tarana, Aygun, and possibly Aygun's roommate Aygun's recordings – Aygun and her roommate
Ends	Various ends for each recording including discussion about food/dinner, telling stories from the day, and casual chat
Act sequence	Conversational turn taking, can't identify overlap or interruption since only one side of each conversation was recorded.
Key	Cheerful and friendly
Instrumentalities	Only verbal instrumentality recorded but since, other than brief phone conversations, all participants were present there must have been other forms of interaction as well
Norms	To be determined
Genre	Informal chat

Aygun made three recordings, all of which appear to be with her roommate though they may be on different evenings: #1 - 22 minutes, #2 - 8 minutes, and #3 - 15 minutes. Table 7-19 and Table 7-20 summarize these recordings according to Hymes' etic grid and Schiffrin's theoretical model of discourse.

Table 7-20 Summary of Aygun and Tarana's talk according to three levels of discourse from Schiffrin's theoretical model

Information State	All participants are young female adults and appear to have similar linguistic proficiencies including high levels of Azerbaijan and Russian. Aygun and Tarana also have very good command of English, while Aygun's roommate appears to be an English learner.
Participation Framework	Tarana's recording – sister in her 30s to her slightly younger sister and roommate (similar age). Aygun's recordings – female in her late 20s to a female peer
Ideational Structure	Casual talk in the evening while preparing, eating, and cleaning up after dinner.

These two sisters show a much higher Russian usage than Sevil. One reason for this would be that all interlocutors have both high Russian and Azerbaijani proficiency resulting in an overall information state that favors code-switching/mixing. Table 7-21 shows the distribution of Russian content words, Russian CPs, insertional code-mixing, and peripheral alternation in the recordings from Aygun and Tarana.

Though both women were raised in the same home and are of similar ages, Tarana shows significantly higher levels of Russian than Aygun for each measure. In all cases, a chi-square test shows that with respect to these measures of language use, Aygun and Tarana appear to be behaving independently.

Table 7-21 Distribution of overall language use measures in Aygun and Tarana's recordings

Aygun	**Content Words**		**CPs**	
	Az Content	1278	Az CPs	545
	Ru Content	356	Ru CPs	150
	Total Content	1695	Total CPs	734
	% Ru Content	*21.0%*	*% Ru CPs*	*20.4%*
	Insertion		**Peripheral Alternation**	
	Ru w/ Az Ins	10	Ru w/ Az Alt	3
	Az w/Ru Ins	25	Az w/ Ru Alt	28
	% Az CPs w/ Ru Ins	4.6%	% Az w/ Ru Alt	10.5%
Tarana	**Content Words**		**CPs**	
	Az Content	447	Az CPs	267
	Ru Content	537	Ru CPs	196
	Total Content	1024	Total CPs	493
	% Ru Content	*52.4%*	*% Ru CPs*	*39.8%*
	Insertion		**Peripheral Alternation**	
	Ru w/ Az Ins	14	Ru w/ Az Alt	3
	Az w/Ru Ins	37	Az w/ Ru Alt	28
	% Az CPs w/ Ru Ins	*13.9%*	*% Az w/ Ru Alt*	*10.5%*

Table 7-22 Chi-square tests for overall measures of Russian use in Aygun and Tarana's recordings

Measure	**Chi-square**	**df**	**Significance**
Russian Content Words	286.055	1	0.000*
Russian CPs	54.372	1	0.000*
Azerbaijani CPs with Russian Insertion	21.838	1	0.000*
Azerbaijani CPs with Russian peripheral alternation	11.064	1	0.001*

7.3.1 Tarana

Tarana's language use is distinguished both by a high percentage of Russian CPs and by an even higher percentage of Russian content words. In Sevil's recording we see that a change in conversational language (shown by CPs and turns in the same language) corresponded to a change in identity with Russian corresponding with a 'teacher'

identity. While Tarana uses a much higher percentage of Russian CPs than Sevil (39.8% as opposed to 10.0%), Tarana does not show categorical changes of language over sections of the conversation like Sevil. Also, Tarana never appears to instruct her sister or roommate. Rather Tarana rapidly switches languages within and between conversational turns as well using peripheral alternation and insertional code-mixing. Another feature of Tarana's recording is fairly short conversational turns. She has very few extended turns where we can examine her language use in its immediate context. One exception is the example on the following page where Tarana discusses some problems she has encountered while getting a house remodeled.

Tarana #367

(1.) а это-т мужик, котор-ый прише-л размер окон сним-ать, не окон ey, двер-ей,
and this-m.n/a peasant which-m.s.n/a arrive.to-past size window take-infin not window look.here door-f.p.g/a
[a ɛtot muʒnik kotorɯ j priʃjol] razmɛr okon snimat nɛ okon] [dvɛrɛj]
'Well, this worker that came to take the size of the windows, oh, not windows of the doors'

(2.) dey-ir yalan dey-ir-lər dey-ir birinci dəfə de-di-lər ki, filan et-di-lər
say-cont falsehood say-cont-3.pl say-cont first occasion say-past-3.pl rel such.and.such do-past-3.pl
'He says that they told me wrong, he says the first time they said that they would do such and such'

(3.) ki, прекрати-л-и, а сейчас, говор-ит, сам-и специально прос-ят,
rel, cease.perf-past -pl and now speak-3.sg same-m.p.n/a specially request -3.pl
[prɛkratili a sɛjtʃas govorit sami spɛtsialno prosjat]
'Well, they stopped and now they say that they are specially requesting the same thing.'

(4.) государственн-ые деньги люд-ям в карман ид-ет,
governmental-p.n/a money people-m.p.dat in pocket go-3.sg ,
[gosudarstvɛnɯjɛ djɛngi ljudjam v kraman idjot]
'Governmental fees go into people's pockets.'

(5.) dey-ir-əm, три-ста двадцать три manat-a, эт-о почти четыре-ста доллар-ов ey дела-ет,
say-cont-1.sg three-hundred twenty three manat-n.p.a this-n.n/a almost four-hundred dollar-m.p.g look.here do-3.sg
[trista dvadtsat tri] [a] [ɛto potʃti tʃɛtɯresta dolarov] [dɛljɛt]
'I am saying, three hundred twenty three manat! That makes almost four hundred dollars, see.'

(6.) dey-ir kim-in-sə borc-un-u sənin üst-ün-ə yaz-ıb-lar,
say -cont who-3.sg.poss-if debt-3.sg.possd-acc 2.sg.poss upper-3.sg.poss-to write-perf-3.pl
'It says, if someone has a debt they write it higher.'

(7.) dey-ir-əm yaz-sın-lar da, kim-dir onlar-a pul ver-ən?
say -cont-1.sg write-2.inf.imper-3.pl therefore who-be 3.pl-to money give-ing
'So I say, y'all write it. Who is it who is giving (that) money?'

(8.) Mehdi də dey-ir ki, a вс-е равн-о sən-dən al-acaq-lar dey-ir o pu- (incomprehensible)
Mehdi therefore say -cont rel and all-p.n/a equal-s.n 2.sg-abl take-fut-3.pl say -cont 3.sg …
[a vsjo ravno]
'So, Mehdi says that, well, everything is the same. They take it from you, he says that…'

One of the challenges in transcribing and analyzing Tarana's recording is her rapid change in topic. This can be seen in this excerpt. In 1 through 3 she is specifically talking about remodeling, but then throws in a sentence complaining about government corruption in 3 through 8. While the two are probably related (taxes and fees for government inspection of construction, etc.), Tarana gives no transition for the topic change. This seems to mirror the thought patterns and her general mode of conversation, at least with her sister Aygun.

In terms of content words the sample on page 205 reflects the rest of Tarana's recording almost perfectly, though the ratios of other measures of overall language use do not follow the trends in her recording as a whole.

Table 7-23 Distribution of overall language use measures in Tarana's speech sample

Content Words		**CPs**		**Insertion**		**Periph. alternation**	
Az	20	Az CPs	12				
Ru	22	Ru CPs	4	Ru w/ Az I	2	Ru w/ Az A	1
Ttl	42	Ttl CPs	16	Az w/ Ru I	0	Az w/ Ru A	1
% Ru	*52.4%*	*% Ru*	*25.0%*				

We also see the quick alternation of language characteristic of Tarana's speech. Russian CPs are spread over lines 1, 3, 4, 5, and 8 (as a relative clause), while Azerbaijani CPs are on lines 2, 6, 7, and 8. Since in this section Tarana is telling a story about dealing with workmen and by implication government officials, it is not surprising that Tarana uses words for 'say' numerous times. What may be surprising, however, is how frequently Azerbaijani and Russian are used to express the concept. The Azerbaijani word *dey-* occurs 9 times (lines 2, 5, 6, 7, and 8), while the Russian говор- [govor-] only occurs once on line 3. Thus Azerbaijani provides most of the narrative frame for the story, while

the switch to Russian serves as a contextualization cue to emphasize the content of the CP – relating to the action structure for this section of the conversation.

In this speech sample we also see a number of discourse markers both in Azerbaijani and Russian. The Azerbaijani marker *ey* 'look here' occurs in both lines 1 and 5. In line 1 it precedes a self correction, so it serves a repair function. In Line 5, however, it seems to be present simply for emphasis. In both cases, this Azerbaijani discourse marker occurs in Russian CPs immediately preceding the verb with the language of the marker further emphasizing it. All of these relate to the conversation's action structure but in different ways. In the first line we see the first occurrence of the Russian connective *а* [a] 'and' or 'while' marking the opening of the story. In line 3 we see this marker acting as a simple connective to join events on the story line: "They stopped and now they say…" In line 1 *а* [a] also serves to connect events but in a less direct way. In Tarana's previous two turns she was talking about problems with air conditioning. The connective at the beginning of 1 then ties the following turn together with the previous turn indicating that it is on the same topic – home remodeling. In both cases, *а* [a] occurs adjacent to Russian lexical items. As with the 'say' verbs both occurrences *а* [a] relate to the conversation's Action Structure. One of the most common discourse markers in Azerbaijani appears alternately as *da* and *də*. While there are 365 occurrences in the Home recordings corpora, it only appears twice in this selection (lines 7 and 8). In both cases it is backward looking, emphasizing the previous lexical item. Of the three discourse markers discussed here, *ey*, *а* [a], and *da/də*, only the Azerbaijani *ey*

is inserted into other language CPs making it the only one associated with code-switching/mixing.

In this selection we also see two separate uses of the Azerbaijani relativizer *ki*. In lines 2 and 8 we see *ki* introducing relative clauses – in Azerbaijani on line 2 and Russian on line 8. However, in 3 we see *ki* being used as a discourse marker. It would not make sense to assign *ki* as a relativizer following the verb on the previous line *et-* 'make/do' since that two argument verb already has an object preceding it – *filan* 'such and such'. Rather it functions as a discourse marker indicating that the following clause is a further development of the story with the alternation of language acting as a contextualization cue serving as emphasis.

Clearly, Tarana is using Azerbaijani and Russian to construct an identity or identities quite apart from those constructed by Sevil. In Sevil's recording we see Russian language use associated with a 'teaching' identity in conversation with her daughter, but also associated with an identity as a 'supportive younger relative' to accommodate the language choices of her uncle. Azerbaijani was associated with home and government domains to contribute to identities such as 'government contractor' and 'female relative'. Sevil uses code-switching/mixing to contribute to a heightened emotional key. Tarana's language choices and habits are quite different. While she uses code-switching/mixing for emphasis, she does so much more often than Sevil. Also, while Sevil shows categorical changes of conversational language, Tarana favors quick alternation between languages both between and within turns. As reflected in their language choices, these two women live in very different personal worlds. Sevil, though

a product of the U.S.S.R., pursues identities approved and encouraged by her culture: wife, mother, and university professor – a profession encouraged for women since Soviet times (see Heyat 2002). Tarana, on the other hand, has never been married (though in her 30s) dates foreign men, likes to sleep until late in the day, and depends on family and her foreign boyfriend to provide her finances rather than working in a culturally approved profession. In mixing Russian and Azerbaijani together, Tarana chooses to distance herself from traditionally approved stereotypes of femininity and build an identity as modern and free.

7.3.2 Aygun

As mentioned earlier Aygun made three recordings all of which appear to be with her roommate, though they may be on different occasions. Like Tarana's recording these all take place in the evening surrounding a meal. Thus a common topic is food, with casual chat about a number of other topics filling the gaps. Overall, Aygun uses quite a bit less Russian than Tarana, though still more than Sevil. However, the vast majority of Aygun's Russian content words occur within Russian CPs. She shows a fairly low rate of Russian insertion into Azerbaijani CPs as well as Russian peripheral alternation. The excerpt on the following page reflects these overall trends. In this section of the recording, Aygun is providing evidence for her observation that men like to eat meat more than women, by discussing her father.

Aygun #47

(1.) если честно, мы не люб-им мясо в семь-е, моя мама всегда готовила мясо,
if honestly 1.p.n not love-1.pl meat in family-n.s.prep 1.s-f.n mom always to.prepare-past-s.f meat
[ɛsli tʃɛstno mɯ nʲɛ ljubim myaso v sɛmʲɛ moja mama vsʲɛgda gotovila mjaso]
'To be honest, we do not like meat in my family. My mom always cooked meat…'

(2.) и после того, как мой папа , ооо, bu nə-dir, ət, hh
and after that-m/n.g/a that 1.s-m.n/a father , oh this what-be meat (breath)
[i poslɛ tovo kak moj papa]
'and after that, my dad 'Oh, what's this, meat!''

(3.) gör-dü ki ye-m-ir-di ye-məy-i heç kim, qal-ır-dı de-yə xarab ol-ur-du
see-past rel eat-neg-cont-past eat-infin-acc not.at.all who remain-cont-past say-to rotten be-cont-past
'He saw that (everyone) did not eat it. No one eating it. It stayed and, I am saying, it got rotten.'

(4.) bircə o ye-yir-di, bir nəfər-lə nə ol-ur hh папа-m da heç yad-ım-da deyil
only 3.sg eat-cont-past one person-with what be-cont (breath) father-1.sg.possd so not.at.all memory-1.sg.possd-loc not
[papa]
'He was the only one eating it, one person. That's what happened! Then dad did not remember anything at all.'

(5.) ət ye-yə икр-у он е-л всегда, , hə , bu da belə dey-ir
meat eat-to caviar-f.s.acc 3.s.m.n eat-past always yes , this therefore thus say-cont
[ikru on ʲɛl vsʲɛgda]
'Eating meat, he always ate caviar, yeah, so he talks like this.'

(6.) görə qal-mış-am belə on-a görə ət-dən heç nə bişir-m-ir-əm
according.to remain-imper-1.sg thus 3.sg-to according.to meat-abl not.at.all what cook-neg-cont-1.sg
'So, because of that I stay this way. I do not cook any meat at all.'

(7.) qal-acaq or-da xarab ol-acaq, amma qardaş-ım biz-ə gəl-ən-də ...
remain-fut 3.sg-loc rotten be-fut but brother-1.sg.possd 1.pl-to come-ing-loc
'It will stay there and get rotten, but when my brother comes over…'

While this example has a higher proportion of Russian content words, the percentage of Russian CPs is almost identical to the rest of Aygun's recordings.

Table 7-24 Distribution of overall language use measures in Aygun's speech sample

Content words		**CPs**		**Insertion**		**Periph. alternation**	
Az	28	Az	11				
Ru	14	Ru	3	Ru w/ Az I	0	Ru w/ Az A	1 (?)
Ttl	42	Ttl	15	Az w/ Ru I	0	Az w/ Ru A	0
% Ru	*33.3%*	*% Ru*	*20.0%*				

In this excerpt, we see the same easy flow from one language to the other that we see in Tarana's recording, but without other language insertion or discourse markers. The Russian CPs are spread over lines 1, 2, and 5. Aygun opens the turn in Russian, but switches to Azerbaijani in the middle of line 2. While there is no relativizer or 'say' verb, this is clearly reported speech. Aygun is mimicking her father's reaction when her mother cooked meat. The alternation from Russian to Azerbaijani could be for one of two reasons. Either her father would have actually said such words in Azerbaijani, so her use of Azerbaijani reflects actual historical language use, or the switch from Russian to Azerbaijani could serve as a contextualization cue to set the quotation apart from the frame in which it occurs. This relates then to the conversation's action structure though it must also have been conditioned by the overall information state favoring code-switching/mixing. Though the quote finishes on line 2, line 3 continues in Azerbaijani rather than utilizing the Russian frame that preceded it. This is not surprising since Azerbaijani is clearly the default language in Aygun's recordings. The last Russian CP of the excerpt occurs in the first half of line 5. Here we see the Azerbaijani participial phrase *ət yeyə* 'eating meat' preceding a Russian clause. While this could be analyzed as

a stative clause with a non-overt copula, it seems most likely that it is serving an adverbial function and fills a [C, CP] slot for the Russian CP. This would be in line with Russian discourse norms. As given information, it should occur at the beginning of the clause. Since Aygun does not orient to the switch to Russian with any repetition, pause, or self-correction, we must assume that the switch is a contextualization cue to emphasize the addition of caviar to the list of meat products that her father was in the habit of eating. The double commas after *всегда* [vsʲɛgda] 'always' in line 5 indicates a longer than normal pause preceding the switch back to Azerbaijani.

In this excerpt from Aygun's recordings we see only alternational code-switching. This is mostly between CPs within a conversational turn, but in line 5 we also see peripheral alternation. In at least two of Sevil's phone conversations we see insertion and peripheral alternation being used to heighten the emotional key of the conversation. This may also be true in contrasting the excerpts from Tarana and Aygun. In following the general theme of 'food' in Aygun's recordings, she is giving background about food in her family. This is very casual cheerful section of the conversation, so the emotional key is not particularly heightened. The increased use of insertion and peripheral alternation in Tarana's excerpt could be because she is upset about the problems she has been having in getting work done on her apartment, so she uses code-switching/mixing to heighten the key.

Unlike Tarana, Aygun does change conversational matrix language for one short section to take on a teacher identity. Here Aygun switches to English to correct a few English errors by her roommate. This is a very natural identity for Aygun since she was

educated as an English teacher and works part-time as an English tutor for university students. While for Sevil the teacher role calls for Russian, Aygun's teacher speaks English.

In Aygun's recording, as in Tarana's and Sevil's, we see identity being talked out. While Tarana uses quick alternation and seemingly random switches to construct an overall identity as modern and free, Aygun exhibits much more 'proper' language usage. Though she uses a fairly high percentage of Russian content words, these occur mostly in Russian clauses. While she is a modern Azerbaijani woman with multiple language proficiencies, her language use is regular and much more controlled than Tarana's as are her personal and professional lives. Outside the family domain it is likely that her language use would resemble Sevil's – choosing the appropriate language for a given situation to convey an identity as stable, professional, and educated with a great degree of intersituational fluidity.

7.3.3 Relative clauses and reported speech

One issue that arises in a minor way in the excerpts from both Tarana and Aygun is code-switching in reported speech and across relative clause boundaries (alternation of large sections - 4.2.3.2). In both languages relative clauses follow the matrix clause preceded by a relativizer, so there is structural congruence making this a likely spot for code-switching. Moreover, code-switching has been found in reported speech and quotation by other researchers (see McClure & McClure 1988, Poplack 1988, etc.). This could be for at least two reasons: the reported words were actually spoken in the code-switched language ("inherent meaning potential" - Auer 1995: 124), or the switch served

as a rhetorical device to contextualize or emphasize the reported utterance ("contrast" - Auer 1995: 124).

We see code-switching across relative clause boundaries in Aygun's example on line 2 and in Tarana's on line 8, both times corresponding to reported speech. In the Home recordings we see three choices of relativizer: Azerbaijani *ki*, Russian ЧТО [tʃo], and a non-overt Ø. All of these occur with both reported speech and relative clauses that occur in other contexts. Table 7-25 shows how these relativizers occur with matrix/relative clause languages for Aygun and Tarana for reported speech (RepSp) and other relative clauses (Other).

Table 7-25 Relativizers and matrix/relative clauses in Aygun and Tarana's recordings

	Relativizer	Az matrix with Az RelC (RepSp/Other)	Az matrix with Ru RelC (RepSp/Other)	Ru matrix with Ru RelC (RepSp/Other)	Ru matrix with Az RelC (RepSp/Other)
Tarana	ki	5/0	2/0	·	·
	что [tʃo]	·	·	1/3	·
	Ø	10/3	1/0	1/1	·
Aygun	ki	34/9	·	·	·
	что [tʃo]	·	·	2/0	1/1
	Ø	22/4	1/0	2/0	·

This does, however bring up a number of difficulties in classifying the CP matrix language. These are listed below with the solution:

Az matrix followed by stative clause with non-overt copula but all Az lexical items	Counted as Az matrix with Az RelC
Ru matrix followed by stative clause with non-overt copula but all Az lexical items	Counted as Ru matrix with Az RelC

Az matrix followed by English relative clause	Not included in Table 7-25 but counted as 'with CS' in Table 7-26.

While both use relative clauses most often with reported speech, they also are used in other contexts. 74% of Tarana's relative clauses represent reported speech (20 of 27), as do 82% of Aygun's (62 of 76).

As Table 7-25 shows, code-switching across relative clause boundaries is possible in Aygun's and Tarana's talk, but it is not terribly common. The Azerbaijani relativizer *ki* only occurs with Azerbaijani matrix clauses, while the Russian ЧТО [tʃo] occurs exclusively with Russian matrix clauses. The zero relativizer can occur with both languages, but it occurs most frequently following Azerbaijani matrix clauses. All three relativizers can occur at a code-switch boundary. Table 7-26 summarizes the occurrences of relative clauses with and without code-switching for these two speakers. While Aygun has 4 occurrences of code-switching to Tarana's 3, Aygun also has almost three times as many relative clauses total.

Table 7-26 Summary of relative clauses with and without code-switching for Aygun and Tarana

	Without CS (RepSp/Other)	With CS (RepSp/Other)
Tarana	17/7	3/0
Aygun	56/17	3/1

Even though Tarana has a higher proportion of relative clauses with code-switching, a chi-square test to compare their frequencies does not yield statistically significant results (chi-square=1.115, df=1, p=0.291)

Table 7-27 shows the frequency of code-switching across relative clause boundaries for all the other Home recordings participants.

Table 7-27 Code-switching across relative clause boundaries for all other Home recordings subjects

	Relativizer	Az matrix with Ru RelC (RepSp/Other)	Ru matrix with Az RelC (RepSp/Other)
Agil	ki	2/1	·
	что [tʃo]	1/0	·
	Ø	·	·
Lale	ki	2/0	·
	что [tʃo]	·	·
	Ø	6/0	·
Mahir	ki	·	·
	что [tʃo]	·	·
	Ø	·	·
Shahin	ki	·	·
	что [tʃo]	·	·
	Ø	·	·
Sevil	ki	·	·
	что [tʃo]	·	·
	Ø	1/0	·

Mahir and Shahin never switch languages across a relative clause boundary. Aygun is the only subject to switch from a Russian matrix clause to an Azerbaijani relative clause, but she only does this twice. Agil is the only subject who uses *что* [tʃo] following an Azerbaijani matrix clause to precede a Russian relative clause. In all other cases the overt relativizer agrees with the language of the matrix clause. Lastly, Lale switches from Azerbaijani to Russian across a relative clause boundary with a non-overt relativizer more often than any other subject, all six times in reported speech. Thus, while code-switching in relative clauses is used by these subjects it is neither consistent across the subjects, nor very common. In all cases, it probably functions as a contextualization cue to emphasize the content of the relative clause.

7.4 Soviet influence: Aner

The last case study in this chapter is Aner. As a male and as the oldest subject in the Home recordings corpora, his language use is worth separate consideration. Zuercher 2004 and 2009 show that, while women's Russian language use in Azerbaijani seems to decline with age, men's Russian language changes very little. Thus, in her 40s Sevil probably uses much less Russian than she did as a university student in her early 20s. Aygun and Tarana may well be at the point where there language use is changing to be more Azerbaijani dominant. By examining Aner's language use we are more likely to get a picture of language use in Azerbaijan during the Soviet period. At 47 years of age in 2007 when these data were collected, he would have finished his university studies well before the nationalistically oriented language laws of the late 1980s.

Of the participants in the Home recordings, Aner seems to have been the least comfortable using the digital recorder. He made a total of 14 recordings, but of these only 4 included any usable data with only one being of any significant length. The majority were very short with either no talk, or no talk by Aner (no male voice recorded). It seems that he just kept turning the recorder on and off, as if he was not sure whether or not it was working.

All of these recordings appear to have taken place in Aner's home in the evening. Table 7-29 and Table 7-30 summarize the four usable recordings according to Hymes' etic grid and Schiffrin's theoretical model, while Table 7-31 shows the relative amounts of each measure of overall language use in all four recordings together.

Table 7-28 Summary of Aner's recordings

Recording	**Length**		**Participants**
#1	4.5 seconds	not transcribed (no male voice)	
#2	36.5 seconds	not transcribed (no male voice)	
#3	21.3 seconds	not transcribed (no male voice)	
#4	14.5 seconds	not transcribed (no male voice)	
#5	2 minutes, 20.2 seconds	transcribed	Aner Adult female (probably his wife)
#6	25 minutes, 23.0 seconds	transcribed	Aner Adult female (probably his wife) Adult male Child (probably his grandchild)
#7	10.2 seconds	not transcribed (no talk)	
#8	22.9 seconds	transcribed	Aner Child (probably his grandchild)
#9	15.6 seconds	transcribed	Only Aner
#10	13.2 seconds	not transcribed (no talk)	
#11	.5 seconds	not transcribed (no talk)	
#12	1.9 seconds	not transcribed (no talk)	
#13	.6 seconds	not transcribed (no talk)	
#14	58.4 seconds	not transcribed (no male voice)	

What distinguishes Aner from other subjects in the Home recordings corpora is his fairly high percentage of Russian content words combined with his very low rates of insertion and peripheral alternation. Like Aygun, the vast majority of Aner's Russian lexical items occur within Russian CPs.

Table 7-29 Summary of Aner's talk according to Hymes' etic grid

Setting	At home in the evening
Participants	Aner, an adult female (presumably his wife), an adult male (presumable son or son-in-law), and a child (presumably his grandchild)
Ends	Various including discussing and filling out the surveys for this research project, deciding what TV program to watch, discussing the program, etc.
Act sequence	Conversational turn taking, can't identify overlap or interruption since only one side of each conversation was recorded.
Key	Cheerful and friendly
Instrumentalities	Only verbal instrumentality recorded but since, other than brief phone conversations, all participants were present there must have been other forms of interaction as well
Norms	To be determined
Genre	Informal chat

Table 7-30 Summary of Aner's talk according to three levels of discourse from Schiffrin's theoretical model

Information State	All participants appear to have similar linguistic proficiencies including Azerbaijani and Russian, including the child (Aner speaks Russian and Azerbaijani to him/her in conversational contexts)
Participation Framework	Man in his late 40s to his wife, a younger male relative, and grandchild (possibly early teens)
Ideational Structure	Casual talk in the evening surrounding this research project, television programs, etc.

Table 7-31 Distribution of overall language use measures in Aner's recordings

Content Words		**CPs**		**Insertion**		**Periph. alternation**	
Az	268	Az	163				
Ru	219	Ru	87	Ru w/ Az I	2	Ru w/ Az A	3
Ttl	490	Ttl	250	Az w/ Ru I	6	Az w/ Ru A	1
% Ru	*44.7%*	*% Ru*	*34.8%*	*% Az w/Ru I*	*3.7%*	*% Az w/ Ru A*	*0.6%*

Unlike Aygun, however, Aner changes language within conversational turns very rarely. Of Aner's 203 conversational turns, he only exhibits code-switching and/or mixing in 19. Additionally, his Russian language clauses are concentrated in two sections of the

recording where most of his turns are entirely in Russian. Thus, Aner avoids mixing languages together within turns, keeping the same language for fairly large chunks of conversation. The following example is one of the few turns where Aner changes languages:

Aner #2	Mövzu	götür-ər-ik	danış-ar-ıq	hə
	topic	take-indef.fut-1.pl	speak-indef.fut-1.pl	yes

'We pick a topic and we talk (about it), yes?'

как	раз	их	интересу-ет	их	итересу-ет	что?
that	time	3.pl.a/g	interest-3.sg	3.pl.a/g	interest-3.sg	what
[kak	raz	ix	intɛrɛsuʲɛt	ix	intɛrɛsuʲɛt	tʃo]

'It is just (that) they are interested, they are interested, right?'

Aner's tendency to keep languages separated would be expected in an environment such as the U.S.S.R. where bilingualism was common, but where there was a strong language hierarchy. Russian was the language of government, education, and science, while Azerbaijani was associated with 'low' domains such as home and bazaar, as well as with culture specific domains like literature and folk-lore. In 2003 several interview subjects commented that mixing Russian words into Azerbaijani sentences was characteristic of poorly educated people under the U.S.S.R. since they felt it made them sound like they knew some Russian. A more educated person, such as Aner, would keep his languages separated to prove that he was proficient in both. It appears, then, that code-switching/mixing such as that seen in other recordings is characteristic of the post-Soviet period when the language hierarchy is less clearly established. Though Azerbaijani is the language now receiving official support, there are speakers who see Russian as a means of constructing identities in opposition to the traditional cultural stereotypes.

What we cannot see from Aner's recording is what triggers the switch from one conversational language to another. Unlike Sevil's talk it does not correspond to a change in identity role. Though Aner discusses a number of topics, he does not seem to adopt a teacher role like Sevil or Aygun. The change in language could be motivated by the participation framework of the conversation, to build solidarity with Aner's interlocutor at the time, or it could be a result of the bilingual information state (his own as well as that evolving in the conversation) – it was simply the language that came to mind at the time.

7.5 Discussion

From these case studies a number of conclusions about code-switching/mixing in Azerbaijan can be drawn:

- It is normally, but not always, conditioned by an information state where both speaker and hearer are proficient in both languages.
- Different social identities call for different language use patterns including code-switching and/or mixing.
- A shift in language use patterns probably corresponds with a shift in conversational role and/or the identity being talked out by the speaker.
- Individual occurrences of code-switching/mixing often relate to the action structure of the conversation either as contextualization cues to emphasize the content of the spoken text or as textualization cues to organize the spoken text.

Table 7-32 lists some of the possible identities that can be enacted linguistically in Azerbaijan.

Table 7-32 Sample of identities present in case studies

Identity	**Language Norms**	**Information State**	**Speaker**
Educated Azerbaijani	Code-switching for larger conversational sections; avoid insertion and peripheral alternation	Bilingual speaker and hearer	Aner, Sevil, Aygun
Government contractor/worker	Speak Azerbaijani	Hearer/speaker proficient in Azerbaijani	Sevil
Modern, young Azerbaijani female	High levels of code-switching/mixing of all types within conversational turns	Bilingual speaker and hearer	Tarana
Mother	Speak Azerbaijani (some Russian insertion possible)	Hearer/speaker proficient in Azerbaijani	Sevil
Teacher	Speak language being taught (Russian/English)	Bilingual speaker, hearer with lower L2 proficiency	Sevil (Russian), Aygun (English)
Various family roles – wife, relative, sister, etc.	Depends on family	Depends on family	Sevil, Aygun, Tarana, Aner

Moreover, we see that subjects shift between roles and therefore linguistic norms differently. Tarana shows very little change in her linguistic behavior over her recording. This could be due to lack of opportunity – the one recording does not show her in a variety of contexts – but it could also be that she does not change her language to the extent that Sevil does. Sevil on the other hand, shows a great deal of fluidity in the way she moves from one set of role/identity/language norms to another. Though the results presented here present some of the ways that code-switching/mixing is used to negotiate language and identity in Azerbaijan, clearly more research is needed – both qualitative work involving a wider range of participants, and quantitative work to show how the society at large views the individual behaviors.

CHAPTER 8

RESULTS, IMPLICATIONS, AND FUTURE DIRECTIONS

8.1 Introduction

The history of Azerbaijan has produced a language situation where many people are proficient in both Azerbaijani and Russian, but with the collapse of the U.S.S.R. Russian is left with virtually no official role. It has been demoted from the inter-nationality language of a super power to being just another 'foreign' language. At the same time, Azerbaijani has received a great amount of governmental support. This shift in language planning policies can be seen in the data for this dissertation. Azerbaijani is the dominant language in all but one recording, and in the Staged recordings Russian speakers deferred to their Azerbaijani speaking interlocutors in almost every case. However, Russian's dis-prefered/foreign status combined with its wide distribution in the country make it useful to construct identities in opposition to Azerbaijan's acceptable cultural norms.

This chapter will provide a brief overview of the most salient results of this dissertation project and provide a framework for future research. These results include the forms that codeswitching/mixing takes (Muysken's typology as well as Meyers-Scotton's model), the pragmatic functions and motivations for code-switching, as well as implications for identity construction.

8.2 Summary of results

8.2.1 Code-switching/mixing and syntax

8.2.1.1 Muysken's typology

The code-switching/mixing in the data for this research project cover the full range of Muysken's Typology: alternation, insertion, and congruent lexicalization. Of these, insertion is the most common and congruent lexicalization least common.

While there are examples of larger constituents being inserted, the most common insertions in this data are single lexical items. Common nouns are inserted most frequently (sometimes morphologically integrated, sometimes not), followed by adverbs, pronouns, and adjectives. The direction of language transfer is overwhelmingly Russian to Azerbaijani – Russian elements are most often inserted into Azerbaijani CPs rather than the reverse. Another form of insertion is dummy word constructions using a Russian non-finite verbs with an Azerbaijani helping verb. This strategy allows subjects to use the Russian verbal lexis while maintaining Azerbaijani syntax. In the data for this project, the reverse never occurs – Azerbaijani non-finite verbs never occur with Russian helping verbs. Lastly larger constituents are sometimes inserted: noun phrases, prepositional phrases, and adjective phrases. Like single word insertions, most are Russian constituents inserted into Azerbaijani clauses.

In Muysken's conceptualization, alternational code-switching can take place in three ways: between conversational turns, between CPs in the same conversational turn, and between elements that are within the same clause. While alternation between conversational turns and speakers seems quite possible given the level of multi-

lingualism in Azerbaijan, it is nearly absent from the analysis of the Staged recordings. If there were any doubt which language was appropriate, the speakers simply agreed in the first turn and proceeded in the agreed upon language: most often Azerbaijani, occasionally Russian, and once English. There was very little negotiation, and no evident competition for linguistic dominance in terms of language choice. Thus, most of the patterns of alternation proposed by Auer (1995: 125-126) were absent. Given the current political attitude toward language in Azerbaijan, it is not surprising that Azerbaijani was dominant in most conversations, even when one speaker exhibited limited proficiency.

All subjects in the Home recordings corpora showed language alternation between CPs in the same conversational turn. Since all were proficient Russian speakers, this is not surprising, nor is this phenomenon particularly interesting from a syntactic perspective. Rather, such behavior can be interpreted as a manifestation of higher level discourse issues and/or identity construction.

Of more syntactic interest is alternation within CPs. While Muysken lists a number of ways that this can occur, the most common in this dataset is peripheral alternation. Here other language lexical items fill [Spec, CP] slots somewhat separated from the matrix structure of the CP. This occurs most often with adverbial elements, but is also possible with conjunctions. Like insertion, it is most common to append Russian peripheral elements to Azerbaijani CPs, though the reverse is also possible.

Chapter 3 concludes by predicting that congruent lexicalization between Azerbaijani and Russian should not happen, but it does seem to occur to a limited extent in the Home recordings data. While Russian often utilizes a zero copula in the present

tense, formal descriptions of Azerbaijani require an overt stative verb in all clauses. Analysis of 10 minute samples of the Home recordings data, with the addition of Tamam (a subject from the Staged recordings who does not speak Russian), demonstrate, however, that a zero copula must also be posited for Azerbaijani in informal contexts. With bilingual speakers this produces an area of ambiguity. When stative clauses occur with no overt copula, it is impossible to determine the matrix language. Thus stative CPs with lexical items from both languages and no overt verb were classified as congruent lexicalization.

8.2.1.2 Myers-Scotton's 4-M model

Myers-Scotton's 4-M model divides all morphemes into classes based on three binary oppositions

(2002: 73):

[+/- conceptually activated]

[+/- thematic role receiver/assigner]

[+/- looks outside its immediate maximal projection for information about its form]

This results in four classes of morphemes:

Content morphemes	Nouns	Must have thematic roles to overtly appear in a CP. *(Most commonly switched)*
Early system morphemes	Plural morphology	No thematic role. *(Rarely switched)*
	Adverbs	No thematic role. *(Commonly switched but not to the extent that nouns are)*

Bridge late system morphemes	Nominal case morphology	Integrates the object N into a larger constituent. *(Very rarely switched)*
Outside late system morphemes	Verbal agreement	Must look outside VP to subject for its morphological form. *(Should never be switched)*

The findings of this study agree with the predictions made by this model. Nouns are indeed the most commonly inserted items, followed at some distance by adverbs. There are no instances of a switch in language between a verb stem and its agreement morphology. Rather the dummy constructions discussed above allow speakers to use Russian verbs in Azerbaijani CPs. While there are cases of Russian nouns receiving Azerbaijani nominal morphology, this only occurs in Azerbaijani CPs, and with morphemes which can be interpreted as cliticized post-positions rather than nominal case. Thus, these data confirm Myers-Scotton's proposals for a new language pair – Azerbaijani and Russian.

8.2.2 Motivations for code-switching/mixing

The literature on code-switching/mixing details numerous reasons that these linguistic phenomena occur. The most basic motivation for code-switching/mixing is psycholinguistic. According to Myers-Scotton, code-switching always originates at the conceptual level – it occurs because it represents the speaker's intention. Thus all code-switching/mixing is psycholinguistically motivated. In Schiffrin's terminology this reflects the speaker's bilingual information state, a necessary condition for the switching/mixing. Along this vein are discussions of which language mode or modes are activated while a speaker is producing the utterance. According to Muysken (2000) both

languages are activated to some extent in bilingual speech. Thus, at a psycholinguistic level, an utterance in one language may trigger more items in that language. We see this briefly in Sevil's case study where a mixed stative CP ending in Russian precedes fully Russian CPs. It is possible that the Russian CPs were triggered by the preceding Russian lexical items. This, however does not seem to be very common in the Home recordings corpora. Rather, several subjects switched rapidly and frequently "in a way which minimized the salience of the switch points, and where the switches formed part of an overall discourse strategy to use both languages…" (Poplack 1988: 230).

This discourse strategy is conditioned by two things: a bilingual information state, and a participation framework where the speaker believes the switching/mixing behavior is appropriate (Schiffrin 1987). In this dissertation these issues enter most into the discussion of Sevil's recording. Though she uses relatively low levels of Russian language items, she does so only when the participation framework allows it given the conversation's information state. While it would be possible for a speaker to intentionally use a language knowing that the interlocutor did not understand it, we never see it in any of the data for this study.

Rather, at the pragmatic level, we see code-switching/mixing most often used as contextualization cues. In Sevil's conversation we see contextualization both for contrast and inherent meaning potential (Auer 1995). For example, in her second phone conversation she utters the adverb *наоборот* [naoborot] 'backwards' in Russian rather than Azerbaijani. Here it appears that the change in language itself (much like a rise in voice pitch or a change in tempo) is the contextualization adding emphasis to the adverb.

In other cases the language of the lexical items carries meaning (contextualization by inherent meaning potential). This may be the case when Sevil uses Russian insertion when speaking about medicine. For example:

Республиканск-ий-də	операционн-ı-da
republican-m.s.n/a-loc	operating.room-3.sg.possd-loc
[rɛspublikansk-iy]	[opɛatsion]
'Republic hospital'	'operating room'

8.2.3 Code-switching/mixing in identity construction

Finally, we see the subjects in this study using code-switching and code-mixing to talk out a range of social identities. While these language choices are conditioned by the proper combination of information state and participation framework, these speakers use them to tap into various social Discourses (Gee 1999) in order to control how their interlocutors perceive, classify, and/or interpret them as individuals. Table 7-32 lists a number of identities being constructed in the Home recordings data. While Sevil uses her language skills with great fluidity to talk out a range of socially approved identities for her age and level of education, Tarana uses a higher level of Russian admixture to talk out an image of modern, young femininity in contrast to the politically/socially approved norms. Thus, very similar linguistic resources can be used for opposing identity construction efforts.

8.3 Wider implications

In addition to showing us how a number of individuals utilize Azerbaijani and Russian in conversation, the data for this study give us hints at the language situation of Azerbaijan in general. While providing us with a typology of code-switching/mixing Muysken (2000) discusses a number of sociolinguistic situations and the types of

switching/mixing that can be expected in them. He claims that insertion should occur: in colonial settings and among recent migrant communities, and when there is considerable asymmetry in speakers' proficiency in the two languages. A shift in language dominance should be accompanied by a shift in the direction of insertion. As discussed in chapters 5 and 6 code-switching/mixing in the data for this study is dominated by the insertion of Russian lexical items into Azerbaijani clauses. Given the dramatic switch in language dominance since the collapse of the U.S.S.R. in 1991, we can predict that the opposite was true before: that before 1991 code-switching/mixing in Azerbaijan may have been characterized by the insertion of Azerbaijani lexical items into Russian CPs. The enrollment decline in Russian language elementary and secondary schools as well as in Russian sectors in Azerbaijani universities, has resulted in a drop in the number of Azerbaijanis who are highly proficient in Russian. Several Russian teachers complained in interviews that students nowadays "just do not speak Russian very well." Thus while Azerbaijani proficiency may be on the rise, there is an increasing asymmetry in different speakers proficiency in Russian. There has been some difficulty in classifying the former Republics of the U.S.S.R. since Soviet rhetoric classified them as voluntary members of the union, but history shows the unification to have been less than voluntary in many cases. Therefore should we classify Azerbaijan as post-colonial (Laitin 1998), a developing country (Garibova 1999), or an ancient nation with an ancient literary tradition (Dövlət Dilinin Təkmilləşdirilməsi 2001)? While all may be true to an extent, the types of code-switching present in the data for this project seem typical of a colonial (or now post-colonial) situation.

As outlined in section 1.3.3 the Azerbaijani language has undergone significant status, acquisition, and corpus planning since 1991. The data for this project provide some reflection on the effectiveness of these legislatively based language planning efforts. Sevil uses almost no Russian items to talk out her role as "governmental contractor" while in Soviet times, this role would probably call for almost exclusive Russian speech. Many of Sevil's phone conversations and most of the Staged conversational data show that Azerbaijani is now the default language for public conversation. For Sevil's generation (at least) Russian is still a resource to be used in private, and she makes focused efforts to ensure that her daughter also has this resource. As Tarana and Aygun show, Russian language and mixture can also be used to construct identities that are may be politically incorrect. So, while the government's language planning efforts have been quite effective, they are still subject to resistance for the purpose of constructing alternative identities, particularly of femininity. Thus, we see in chapter 6 that there can be wide variation in language choice, especially among younger citizens, in Azerbaijan. Shahin speaks almost all Russian in his Home recording, but shows in his Staged conversations that he is also proficient in Azerbaijani. Returning to a previous study of language behavior in Azerbaijan, an interview subject for Zuercher 2004 said:

> In the 1970s and 80s we used to see Heydar Aliyev [then holding various prominent positions within the Communist Party hierarchy] on television speaking Russian, and we all wanted to speak Russian too. Now, we see his son,

> Ilham, on television speaking English with foreign diplomats and we all want to speak English too (Zuercher 2004: 100).

Clearly the attitude of the Azerbaijani government influences the languages that Azerbaijani citizens want to learn and use, but with the current administration walking a political tight-rope between U.S./European influence and Russian influence, there is no way to guess what the future will hold. A study of the roles that English plays in Azerbaijani society could well shed some light on how the 'common man' sees this.

8.4 Strengths and weaknesses of this project

The present study has much to contribute both to our understanding of Azerbaijan itself and code-switching/mixing. Though a recent edition of the International Journal of the Sociology of Language (July 2009) has contributed much to our understanding of sociolinguistic issues in Azerbaijan, clearly all of the ground has not been covered. One of the gaps in this literature is analysis of naturally occurring language data. While there are a number of studies with a theoretical orientation (Garibova 2009, Mammadov 2009), and others that provide a thorough analysis of reported language use (Clifton 2009, Zuercher 2009), this appears to be the first study to use transcriptions of authentic, contextualized language behaviors for analysis, providing a good counterpoint to previous research. While the results may not be generalizable to the whole population of Azerbaijan, the analyses of these data do provide a solid description of code-switching and code-mixing as they occur in Azerbaijan today. Lastly, by analyzing a heretofore unstudied language pair, Azerbaijani and Russian, this study confirms the hypotheses of both Pieter Muysken and Carol Myers-Scotton.

Like all research, this study also has its weak points. Though part of the analysis focuses on a sequential analysis of the conversational data, this is hindered by practical limitations on the research methodology. Though recording only one side of the conversations made the research more feasible given the regulatory environment in which it took place, it prevents true Conversational Analysis. In addition, the small sample size prevents more detailed quantitative analysis, thereby making it impossible to posit general claims regarding the wider speech community. This study focused on a detailed analysis of an obtainable set of data which enabled a solid description of the types of code-switching/mixing in Azerbaijan, provided some general information on how subjects use these linguistic phenomena in conversation, and allowed us to make projections of how cs/cm can be used to navigate the complex and changing palette of potential social identities available in Azerbaijan.

8.5 Future research directions

Rather than being a final word on language issues in Azerbaijan, the goal of this study was to lay a solid foundation for future research. While it gives indications of how code-switching/mixing is used in conversation, further studies are needed. Given the not so distant history of social repression under various Soviet regimes combined with current governmental support for Azerbaijani language use, subjects may have "strategic reasons to misrepresent their language use patterns" (Laitin 1998: 369). Therefore any survey based research on language in Azerbaijan must be balanced by qualitative research into actual language production by individuals involved in the study. One way to obtain quantitative data on how Azerbaijani society interprets the use of code-

switching/mixing behavior, would be to construct a matched guize study (Lambert et al. 1960; Lambert 1967) based on the analysis in this dissertation. The typology and analysis presented here could be used to construct a series of staged speech samples which use code-switching in different ways and to different extents. These samples could then be played to a large number of Azerbaijani subjects from a variety of demographic backgrounds to see how they rank the speaker on a number of axes: friendliness, professionalism, level of education, etc. By demonstrating how the society at large interprets code-switching/mixing behavior it would allow us to extrapolate speakers' intentions in producing similar utterances. As well, qualitative studies are needed that take on a true Conversational Analytic methodology. This would be particularly useful if retrospective interviews were conducted shortly after the data collection to study what the speakers thought they intended to do through their code-switching/mixing actions.

8.6 Language, code-switching/mixing, and identity construction in Azerbaijan: Macro/micro orientations

As mentioned in chapter 2, McKay and Hornberger (1996) provide a useful way to organize sociolinguistic research along two axes: macro to micro levels of social analysis, and macro to micro levels of linguistic analysis. The various types of code-switching/mixing listed by Muysken (2000) represent a wide range of linguistic choices and habits born, in the Azerbaijani case, from a rapidly changing socio/political language situation. In all likelihood the differences between the percentage of CPs without overt verbs or with Russian-like movement do not represent conscious language choices, rather it appears that we are seeing evidence of the two languages interacting with each other in each participant's bilingual information state. This and the analysis of variation between

subjects in this study would be classified as micro-linguistic with macro-social implications (the sample size is too small in this study to be generalizable). In addition, we see the influence of governmental language planning and policy on the language choices of individuals as members of the society at large – the macro-linguistic/macro-social quadrant in Table 2-1. By proposing ways in which subjects in this study use code-switching/mixing to construct or talk out various social identities, this project also has implications for the macro-linguistic/macro-social quadrant. However, the strength of this study is in its detailed linguistic analysis of a small group of subjects – micro-linguistic/micro-social. The author hopes that by building a solid foundation of linguistic analysis, utilizing both quantitative and qualitative methodologies, he will pave the way to future research that enhances our understanding of Azerbaijan, the former Soviet space in general, as well as code-switching/code-mixing.

APPENDIX A

SURVEYS, INFORMATION SHEETS, AND PERMISSION FORMS USED IN GATHERING RECORDED DATA
ENGLISH VERSIONS

Information Sheet

My name is Kenneth Zuercher and I am a graduate student in linguistics at The University of Texas at Arlington. As part of the research for my dissertation, I am studying they ways in which language is used by Azerbaijanis living in Baku. I am asking you to participate in my research by tape recording a conversation. While this study is focused on the *forms* of language, I am also interested in your views on some language related issues – please see the list of Conversation Topics. If you would like, you can use these to get you started talking. I would like the recording to be 20 to 30 minutes long, though longer is fine.

If you choose to participate in this study, I will supply you with a tape recorder, two small microphones, and a cassette tape. I ask you to do the following:

1. Complete the questionnaire.
2. Choose a quiet place to record the conversation.
3. Attach one of the microphones to your shirt or collar about 15 centimeters from your mouth. Your conversation partner should do the same with the other.
4. Press the Record button on the tape recorder, say your first name, have your conversation partner say their first name, and then talk informally about some of the attached Conversation Topics, or whatever else you would like to discuss. If the tape runs out, and you wish to continue, turn the tape over and press Record.
5. If the conversation includes embarrassing topics or other personal information that you do not want me to hear, please play through the tape and record over those sections.

After the study is complete, only one copy of this recording will be kept. It will be securely stored in a locked facility in the Department of Linguistics and TESOL, Hammond Hall, The University of Texas at Arlington. The only people who will be able to listen to it will be me (the principal investigator), Dr. Laurel Stvan (my supervisor for this project), and Dr. David Silva (chair of the Department of Linguistics and TESOL).

Given the nature of this study I foresee no risks or discomfort for you, but you are free to stop the conversation and/or withdraw from the study at any time. Should you complete the recording and choose to let me use it in my research, I will be sure to maintain your confidentiality at all times. No names will be included in any reports or in my final dissertation.

If you have any questions about the research, please feel free to contact either me or my faculty supervisor:

Kenneth Zuercher: zuercher@uta.edu
Dr. Laurel Stvan: stvan@uta.edu

For answers to questions you might have regarding your rights as a research subject for a project sponsored by The University of Texas at Arlington, please contact the Office of Research at: kvalsin@uta.edu.

Thank you for taking the time to consider my request.

Demographic Questionnaire

Your Name: _____ Age: _____

What is your nationality? Azerbaijani Russian Other: _____

What is your father's nationality? Azerbaijani Russian Other: _____

What is your mother's nationality? Azerbaijani Russian Other: _____

What language did you speak first as a child? Azerbaijani Russian Other: _____

What language(s) did your family use at home when you were a child? Azerbaijani Russian Other: _____

What was the language of instruction at your primary school? Azerbaijani Russian Other: _____

At your secondary school? Azerbaijani Russian Other: _____

At your university? Azerbaijani Russian Other: _____

For each of the following statements, please mark whether you totally disagree, disagree somewhat, are unsure, somewhat agree, or totally agree. Please mark only one box:

	Totally disagree	Disagree somewhat	Unsure	Somewhat agree	Totally agree
Everyone in Azerbaijan should speak Azerbaijani.					
Everyone in Azerbaijan should speak Russian.					
Azerbaijani is sufficient for discussing any subject.					
Some subjects are easier to discuss in Russian.					
I like to hear Azerbaijani people speaking in Russian.					
I do not like to hear Azerbaijani people speaking in Russian.					

For each of the following statements, please mark whether you totally disagree, disagree somewhat, are unsure, somewhat agree, or totally agree. Please mark only one box:

	Totally disagree	Disagree somewhat	Unsure	Somewhat agree	Totally agree
I generally like Russians.					
I like Russians for close friends.					
I like Russians as neighbors.					
I like to work with Russians.					
I resemble Russians.					
I would like Russians as relatives.					
I avoid contact with Russians.					

Conversation Topics

When you are asked to talk for a tape recorder, it can be difficult to think of things to say. Here are some questions about language in Azerbaijan that I think are interesting. If you would like, you can use them to help you get started:

Language and Education

Did you go to Azerbaijani or Russian primary and secondary schools? Why did your parents choose those schools?
How was school language different from the language you spoke at home?
What was the language situation at school? How did primary school, secondary school, and university differ?

Home Language versus Public Language

Do you speak the same language(s) at home with your family as you do in public in Azerbaijan?
How the languages spoken in private and public situations different in Azerbaijan?
Why do you think people speak more Russian in Baku than in other parts of Azerbaijan?

Language and Gender

Do women/girls speak more Russian than men/boys? Why?
Do you think that fathers or mothers speak more Russian with their children? Why?
Some people say that young women speak more Russian than other people in Azerbaijan. Why would that be true?

APPENDIX B

ALPHABETS USED IN AZERBAIJAN

Azeri Latin	Azeri Cyrillic	IPA	Example
Aa	Аа	a	ata [ata] “father”
Bb	Бб	b	baba [baba] “grandfather”
Cc	Ҷҷ	dʒ	cəm [dʒæm] “jam”
Çç	Чч	tʃ	çay [tʃaj] “tea”
Dd	Дд	d	daş [daʃ] “stone”
Ee	Ее	ɛ	etmək [etmæk] “to do”
Əə	Əə	æ	əl [æl] “hand”
Ff	Фф	f	fındık [fɯndɯk] “hazlenut”
Gg	Ҝк	g	getmək [gælmæk] “to go”
Ğğ	Ғғ	ʁ	ağ [aʁ] “white”
Hh	Һh	h	hə [hæ] “yes/yea”
Xx	Хх	χ	xan [χan] “king”
Iı	Ыы	ɯ	bığ [bɯʁ] “moustache”
İi	Ии	i	iman [iman] “faith”
Jj	Жж	ʒ	jaket [ʒakɛt] “jacket”
Kk	Кк	k	kabab [kabab] “shish kebab”
Qq	Гг	q	qaçmaq [qatʃmaq] “to run”
Ll	Лл	l	lalə [lalæ] “poppy”
Mm	Мм	m	mən [mæn] “I”
Nn	Нн	n	neft [nɛft] “oil”
Oo	Оо	o	od [od] “fire”
Öö	Өө	ø	ölmək [ølmæk] “to die”
Pp	Пп	p	papaq [papaq] “hat”
Rr	Рр	r	rəqs [ræqs] “dance”
Ss	Сс	s	səhər [sæhær] “morning”
Şş	Шш	ʃ	şəhər [ʃæhær] “city”
Tt	Тт	t	top [top] “ball”
Uu	Уу	u	un [un] “flour”
Üü	Үү	y	üz [yz] “face”
Vv	Вв	v	vaxt [vaχt] “time”
Yy	Jj	j	yağ [jaʁ] “fat/butter”
Zz	Зз	z	zəhər [zæhær] “poison”

Russian	IPA	Example
Аа	a	Америка [amɛrika] “America”
Бб	b	белый [bɛlɯj] “white”
Вв	v	вечер [vʲɛtʃɛr] “evening”
Гг	g	газ [gaz] “natural gas”
Дд	d	да [da] “yes”
Ее	ʲɛ	есть [ʲɛst] “to be”
Ёё	ʲo	её [ʲɛʲo] “hers”
Жж	ʒ	жить [ʒit] “to live”
Зз	z	завод [zavod] “factory”
Ии	i	икра [ikra] “caviar”
Кк	k	карта [karta] “map”
Лл	l	лампа [lampa] “lamp”
Мм	m	мама [mama] “mother”
Нн	n	ночь [notʃ] “night”
Оо	o	обед [obɛd] “lunch”

Russian	IPA	Example
Пп	p	папа [papa] “father”
Рр	r	работа [rabota] “work”
Сс	s	сейчас [sɛjtʃas] “now”
Тт	t	так [tak] “so, thus”
Уу	u	уже [uʒɛ] “already”
Фф	f	фарш [farʃ] “chopped meat”
Хх	x	хлеб [xlʲɛb] “bread”
Цц	ts	центр [tsɛntr] “downtown”
Чч	tʃ	чай [tʃaj] “tea”
Шш	ʃ	школа [ʃkola] “school”
Щщ	ʃtʃ	щи [ʃtʃi] “cabbage soup”
Ээ	ɛ	это [ɛto] “this is”
Юю	ju	юбка [jubka] “skirt”
Яя	ja	я [ja] “I”
ь	palatalizes preceding consonant	
ъ	makes preceding consonant fortis	

APPENDIX C

СОЛНЕЧНЫЙ КРУГ ‘THE CIRCLE OF THE SUN’
FULL LYRICS WITH ENGLISH TRANSLATION

Солнечный круг [solnɛtʃnɯj krug]
(May There Always be Sunshine 2009)

English translation

Солнечный круг, небо вокруг
Это рисунок мальчишки
Нарисовал он на листке
И подписал в уголке

The circle of the sun, around the sky
This is a child's drawing
He drew it on a sheet of paper
And signed it in the corner

Пусть всегда будет солнце
Пусть всегда будет небо
Пусть всегда будет мама
Пусть всегда буду я
(x2)

Let there always be the sun
Let there always be the sky
Let there always be mama
Let there always be me

Милый мой друг, добрый мой друг
Людям так хочется мира
И в тридцать пять сердце опять
Не устаёт повторять

Dear my friend, good friend
People want peace
And in the thirty-five heart again
Do not get tired to repeat

Пусть всегда будет солнце
Пусть всегда будет небо
Пусть всегда будет мама
Пусть всегда буду я
(x2)

Let there always be the sun
Let there always be the sky
Let there always be mama
Let there always be me

Тише солдат, слышишь солдат
Люди пугаются взрывов
Тысячи глаз в небо глядят
Губы упрямо твердят

Hush soldiers, listen soldiers
People fear explosions
Thousands of eyes face the sky
Lips stubbornly committed

Пусть всегда будет солнце
Пусть всегда будет небо
Пусть всегда будет мама
Пусть всегда буду я
(x2)

Let there always be the sun
Let there always be the sky
Let there always be mama
Let there always be me

Против беды, против войны
Станем за наших мальчишек
Солнце навек, счастье навек
Так повелел человек

Against evil, against war
Will for our children
Sun forever, happiness forever
The people demanded it

Пусть всегда будет солнце
Пусть всегда будет небо
Пусть всегда будет мама
Пусть всегда буду я
(x2)

Let there always be the sun
Let there always be the sky
Let there always be mama
Let there always be me

REFERENCES

Akhundov, Agamousa (ed.). 1997. *Азәрбајҹан дилинин изаһлы лүғәти, I-III ҹәлбләр* (*The explanatory dictionary of the Azerbaijani language, vols I-III*). Baku: Chiraq.

Altstadt, Audrey L. 1992. *The Azerbaijani Turks: Power and identity under Russian rule*. Stanford, California: Hoover Institution Press, Stanford University.

Auer, Peter. 1984. *Bilingual conversation*. Philadelphia: John Benjamins.

Auer, Peter. 1992. Introduction: John Gumperz' approach to contextualization. In Peter Auer & Aldo di Luzo (eds.), *The contextualization of language*, 1-38. Philadelphia: John Benjamins.

Auer, Peter. 1995. The pragmatics of code-switching: A sequential approach. In Lesley Milroy & Pieter Muysken (eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, 115-135. Cambridge: Cambridge University Press.

Bani-Shoraka, Helena. 2005. *Language choice & code-switching in the Azerbaijani community in Tehran: A conversation analytic approach to bilingual practices*. Uppsala, Sweden: Uppsala Universitet.

Bani-Shoraka, Helena. 2008. Challenging social hierarchy: Playing with oppositional identities in family talk. *Multilingua* 27(1-2). 13-35.

Bani-Shoraka, Helena. 2009. Cross-generational bilingual strategies among Azerbaijanis in Tehran. *International Journal of the Sociology of Language* 198. 105–127.

Blom, Jan-Peter., John J. Gumperz. 1972. Social meaning in linguistic structure. In John Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics: The ethnography of communication*, 407-434. New York: Holt, Rinehart, and Wilson.

Bokamba, Eyamba. G. 1989. Are there syntactic constraints on code-mixing? *World Englishes* 8. 277-292.

Borras, Frank Marshall & Reginald Frank Christian. 1984. *Russian syntax: Aspects of modern Russian syntax and vocabulary*. Oxford: Oxford University Press.

Chan, Alice. 2004. Syntactic transfer: Evidence from the interlanguage of Hong Kong Chinese ESL learners. *The Modern Language Journal* 88(1). 56-74.

Chan, Brian Hok-Shing. 2004. Beyond "contextualization": Code-switching as a "textualization cue". *Journal of Language and Social Psychology*. 23(1). 7-27.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Code of Federal Regulations. 2005. Part 46: Protection of human subjects. Retrieved August 13, 2009 from http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm.

Crandall, Christian S. 1991. Multiple stigma and AIDS: Medical stigma and attitudes toward homosexuals and IV-drug users in AIDS-related stigmatization. *Journal of Community and Applied Psychology* 1. 165-172.

Department of Health, Education, and Welfare. 1979. The Belmont report. Retrieved August 13, 2009 from

http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.htm.

Dövlət Dilinin Təkmilləşdirilməsi haqqında Azərbaycan Respublikası Prezidentinin Fərmanı #506 (2001). Retrieved April 23, 2008 from http://www.e-qanun.az/viewdoc.aspx?id=3568&type=1&state=1&df=2.

Ferguson, Charles. 1959. Diglossia. *Word* 15. 325-340.

Fasold, Ralph. 1990. *Sociolinguistics of Language*. Cambridge, MA: Basil Blackwell.

Fishman, Joshua A. 1972. *The sociology of language - An interdisciplinary social science approach to language in society*. Rowley, MA: Newbury House Publishers.

Gal, Susan. 1979. *Language shift: social determinants of linguistic change in bilingual Austria*. New York: Academic Press.

Garibova, Jala. 2009. Language policy in post-Soviet Azerbaijan: political aspects. *International Journal of the Sociology of Language* 198. 7–32.

Gee, James Paul. 1999. *An introduction to discourse analysis: Theory and method*. New York: Routledge.

Goldsmith, John A. 1990. *Autosegmental and metrical phonology*. Cambridge, MA: Blackwell.

Goodwin, Charles & John Heritage. 1990. Conversation analysis. *Annual Review of Anthropology* 19. 283-307.

Grice, Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics vol. 3: Speech acts*, 41-58. New York: Academic Press.

Gumperz. John J. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.

Haarmann, Harald. 1992. Measures to increase the importance of Russian within and

outside the Soviet Union - a case of covert language-spread policy (A historical outline). *International Journal of the Sociology of Language* 95. 109-129.

Haarmann, Harald. 1995. Multilingualism and ideology: The historical experiment of Soviet language politics. *European Journal of Intercultural Studies* 5(3). 6-17.

Haegeman, Liliane. 1994. *Introduction to government and binding*, 2nd edn. Cambridge, MA: Blackwell.

Hatch, Evelyn & Anne Lazaraton. 1991. *The research manual: design and statistics for applied linguists*. Boston, MA: Heinle & Heinle Publishers.

Hatcher, Lynley. 2009. Script change in Azerbaijan: Acts of identity. *International Journal of the Sociology of Language* 192. 105-116.

Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language* 26(2). 210-231.

Haugen, Einar. 1959. Planning for a standard language in modern Norway. *Anthropological Linguistics* 1(3). 8-21.

Haugen, Einar. 1983. The implementation of corpus planning: Theory and practice. In Juan Cobarrubias & Joshua Fishman (eds.), *Progress in language planning: International perspectives*, 269-290. Berlin: Mouton.

Heyat, Farideh, 2002. *Azeri women in transition: Women in Soviet and post-Soviet Azerbaijan*. Central Asia Research Forum Series. New York, NY: RoutledgeCurzon.

Hornberger, Nancy. 2006. Frameworks and models in language policy and planning. In Thomas Ricento (ed.), *An introduction to language policy: Theory and method*, 24-41. Malden, MA: Blackwell Publishing.

Hymes, Dell. 1972. Models of the interaction of language and social life. In John Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics: The ethnography of communication*. New York: Holdt, Rinehart, and Winston, 35-71.

Kulick, Don. 1992. *Language shift and cultural reproduction : Socialization, self, and syncretism in a Papua New Guinean village*. New York: Cambridge University Press.

Labov, William. 1972a. *Language in the inner city: Studies in the Black English vernacular*. Philadelphia: University of Pennsylvania Press.

Labov, William. 1972b. *Sociolinguistic Patterns*. Philadelphia: University of Philadelphia Press.

Laitin, David D. 1998. *Identity in formation: The Russian-speaking populations in the near abroad*. Ithaca, NY: Cornell University Press.

Lambert, Wallace. 1967. A social psychology of bilingualism. *Journal of Social Issues* 23(2). 91-109.

Lambert, Wallace, R. Hodgson, R Gardner, & S Fillenbaum. 1960. Evaluative reactions to spoken language. *Journal of Abnormal and Social Psychology* 60. 44-51.

Landau, Jacob M & Barbara Kellner-Henkele. 2001. *Politics of language in the ex-Soviet Muslim states: Azerbayjan, Uzbekistan, Kazakhstan, Kyrgyzstan, Turkmenistan and Tajikistan*. Ann Arbor: The University of Michigan Press.

Le Page, Robert Brock. 1968. Problems of description in multilingual communities. *Transactions of the Philological Society*, 189-212. Oxford: Blackwell.

Leed, Richard L. 1986. *5000 Russian words with all their inflected forms: A Russian-*

English dictionary. Bloomington, IL: Slavica Publishers.

Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Longacre, Robert E. 1996. *The grammar of discourse,* 2nd edn. New York: Plenum.

Mammadov, Azad. 2009. The issue of plurilingualism and language policy in Azerbaijan. *International Journal of the Sociology of Language* 198. 65-74.

Maschler, Yael. 1998. On the transition from code-switching to a mixed code. In Peter Auer (ed.), *Code-switching in conversation: Language, interaction and identity*, 125-149. New York: Routledge.

May There Always be Sunshine. 2009. Retrieved March 3, 2009 from http://en.wikipedia.org/wiki/May_There_Always_Be_Sunshine.

McClure, Erica, & Malcolm McClure. Macro- and micro-sociolinguistic dimensions of code-switching in Vingard. In Monica Heller (ed.), *Codeswitching: anthropological and sociolinguistic perspectives*, 25-51. Berlin: Walter de Gruyter & Co.

McDonald, Maryon. 1989. *"We are not French!": Language, culture, and identity in Brittany*. New York: Routledge.

McKay, Sandra & Nancy Hornberger. 1996. Preface. In Sandra McKay & Nancy Hornberger (eds.), *Sociolinguistics and language teaching*, ix-xi. New York: Cambridge University Press.

Musayev, Oruc (ed.). 1996. *Азәрбајчанча Ингилисчә лүғәт*, 1st edn. (*Azerbaijani-English Dictionary*). Baku: Azerbaijan State Institute of Languages.

Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.

Myers-Scotton, Carol. 1993. *Dueling languages: Grammatical structure in codeswitching*. Oxford: Clarendon Press.

Myers-Scotton, Carol. 1995. A lexically based model of code-switching. In Lesley Milroy & Pieter Muysken (eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, 233-256. Cambridge: Cambridge University Press.

Myers-Scotton, Carol. 2002. *Contact linguistics: bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.

Nartey, Jonas. 1982. Code-switching, interference or faddism? Language use among educated Ghanaians. *Anthropological Linguistics* 24(2). 183-192.

Nivens, Richard J. 2002. *Borrowing versus code-switching in West Tarangan (Indonesia)*. Dallas, Texas: SIL International.

Nunnally, Jum C. & Ira H. Bernstein. 1994. *Psychometric theory,* 3rd edn. New York: McGraw-Hill.

Ochs, Elinor. 19979. Transcription as theory. In Elinor Ochs & Bambi Schieffelin (eds.), *Developmental pragmatics*, 43-72. New York: Academic Press.

Ochs, Elinor. 1992. Indexing gender. In Alessandro Duranti & Charles Goodwin (eds.), *Rethinking context: Language as an interactive phenomenon*, 335-358. Cambridge: Cambridge University Press.

Patrick, George Z. 1989. *Roots of the Russian language: An elementary guide to*

wordbuilding. Lincolnwood, IL: Passport Books.

Piller, Ingrid. 2002. *Bilingual couples talk: The discursive construction of hybridity*. Philadelphia: John Benjamins.

Poplack, Shana. 1980. 'Sometimes I'll start a sentence in English y terminal Espanol': Towards a typology of code-switching. *Linguistics* 18. 581-618.

Poplack, Shana. 1988. Contrasting patterns of code-switching in two communities. In Monica Heller (ed.), *Codeswitching: Anthropological and sociolinguistic perspectives*, 215-244. Berlin: Walter de Gruyter & Co.

Poplack, Shana & Marjory Meechan. 1995. Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse. In Lesley Milroy & Pieter Muysken (eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, 199-232. Cambridge: Cambridge University Press.

Robinson, John P., Phillip R. Shaver. & Lawrence S. Wrightsman. 1999. *Measures of political attitudes: Volume 2 of measures of social psychological attitudes*. New York: Academic Press.

Sacks, Harvey, Emanuel Schegloff & Gail Jefferson. 1974. A Simplest systematics for the organization of turn-taking for conversation. *Language* 50(4). 696-735.

Sankoff, David & Shana Poplack. 1981. A formal grammar for code-switching. *Papers in Linguistics: International Journal of Human Communication* 14(1). 3-45.

Saussure, Ferdinand de. 1959. *Course in general linguistics*. New York: McGraw-Hill.

Schachter, Jacquelyn & William Rutherford. Discourse function and language transfer.

Working Papers on Bilingualism 19. 2-11.

Schiffrin, Deborah. 1987. *Discourse markers*. New York: Cambridge University Press.

Schiffrin, Deborah. 1994. *Approaches to discourse*. Cambridge, MA: Blackwell.

Searle, John. 1969. *Speech Acts*. Cambridge: Cambridge University Press.

Serra, Celia Oesch. 1998. Discourse connectives in bilingual conversation: The case of an emerging Italian-French mixed code. In Peter Auer (ed.), *Code-Switching in conversation: Language, interaction and identity*, 101-122. New York: Routledge.

Stroud, Christopher. 1998. Perspectives on cultural variability of discourse and some implications for code-switching. In Peter Auer (ed.), Code-Switching in *conversation: Language, interaction and identity*, 321-348. New York: Routledge.

Swietochowski, Tadeusz. 1995. *Russia and Azerbaijan: A borderland in transition*. New York: Columbia University Press.

Talbot, Mary M. 1998. *Language and gender: An introduction*. Malden, MA: Polity Press.

Trudgill, Peter. 1974. *The social stratification of English in Norwich*. New York: Cambridge University Press.

Veliyev, Arif (general supervisor). 2002. *Demographic indicators of Azerbaijan*. Baku: State Statistical Committee of Azerbaijan Republic.

Wade, Terence. 2000. *A comprehensive Russian grammar*. Malden, MA: Blackwell Publishers.

Waal, Thomas de. 2003. *Black garden: Armenia and Azerbaijan through peace and war*. New York: New York University Press.

Woods, Anthony, Paul Fletcher & Arthur Hughes. 1986. *Statistics in language studies*. New York: Cambridge University Press.

Zuercher, Kenneth. 2004. *Language choice in Azerbaijan: A Soviet legacy*. Arlington, TX: UT Arlington MA Thesis.

Zuercher, Kenneth. 2009. The linguistic construction of gender in Azerbaijan: Sociodemographic factors influencing Russian language choice. *International Journal of the Sociology of Language* 198. 47–64.

BIOGRAPHICAL INFORMATION

In defending this dissertation, Kenneth Zuercher is earning his third degree from The University of Texas at Arlington: a Bachelor's in German with a minor in Linguistics, a Master's in Linguistics with a Graduate Certificate in TESOL, and now a Ph.D. in Linguistics. Along the way he has worked in the Airline industry, done humanitarian work in Azerbaijan, and most recently worked as an Academic Advisor giving him the opportunity to assist many students in achieving their academic goals while allowing him to teach a number of courses. His linguistic research interests center on sociolinguistics, with a focus on language contact phenomena.