VILNIUS UNIVERSITY

Algis Braun

SOCIOLINGUISTIC AND CROSSLINGUISTIC ASPECTS OF THE ACQUISITION OF ENGLISH BY LITHUANIAN UNIVERSITY STUDENTS

Doctoral dissertation
Humanities, philology (04H)

Vilnius, 2009

Research for this project was carried out from 2005–2009 at Vilnius University, Kaunas Faculty of Humanities.

Research supervisor:

Doc. Dr. Ala Likhachiova (Vilnius University, Humanities, Philology – 04H)

VILNIAUS UNIVERSITETAS

Algis Braun

SOCIOLINGVISTINIAI IR TARPKALBINIAI LIETUVIŲ STUDENTŲ ANGLŲ KALBOS ĮSISAVINIMO ASPEKTAI

Daktaro disertacija
Humanitariniai mokslai, filologija (04H)

Vilnius, 2009

Disertacija rengta 2005–2009 metais Vilniaus universiteto Kauno humanitariniame fakultete.

Mokslinis vadovas:

Doc. Dr. Ala Lichačiova (Vilniaus universitetas, humanitariniai mokslai, filologija – 04H)

CONTENTS

INTRODUCTION 9
1. MODELING SPEECH PRODUCTION 16
1.1 Levelt's Speaking Model 16
1.1.1 The conceptualizer 17
1.1.2 The lexicon and lexical items 19
1.1.2.1 The lexicon 20
1.1.2.2 Lexical items 22
1.1.3 The formulator and the articulator 28
1.1.3.1 Grammatical encoding and lemmas 28
1.1.3.2 Phonological encoding and lexemes 32
1.1.3.3 Articulation 34
1.1.4 The monitor 36
1.2 Multilingual Speech Production 39
1.2.1 Models of multilingualism 41
1.2.1.1 The revised hierarchical model 41
1.2.1.2 The inhibitory control model 44
1.2.1.3 The competition model 46
1.2.1.4 The dynamic model of multilingualism 50
1.2.2 Concepts 57
1.2.2.1 The common underlying conceptual base 58
1.2.2.2 Lexical concepts and cognitive models 61
1.3 Summary and Discussion 64
2. SOCIOLINGUISTIC ASPECTS OF THE ACQUISITION OF ENGLISH 66
2.1 Society, Culture, Language 66
2.2 Sociocultural Language Use Survey 71
2.2.1 Demographic data 72
2.2.2 Language history data 73
2.2.3 Cultural data 81
2.3 Summary and Discussion 98
3. CROSSLINGUISTIC ASPECTS OF ENGLISH SPEECH PRODUCTION 106
3.1 Types of Crosslinguistic Influence 106
3.2 CLI in Speech Production 110
3.2.1 Hesitation 111
3.2.2 Code-switching 114
3.2.3 Neologisms and foreignizing 122
3.2.4 Transfer 125
3.2.5 Phonology 130
3.2.6 Interactional strategies 134
3.2.7 Pragmatic failures 138
3.3 Summary and Discussion 140
4. CROSSLINGUISTIC ASPECTS OF ENGLISH WRITING PRODUCTION 146

4.1 Learner Constructions and Conceptual Blending 147
4.1.1 The acquisition of constructions 148
4.1.2 Grammatical blending 154
4.1.3 Crosslinguistic blending 160
4.1.4 Chunks, blending, and learner constructions 166
4.1.5 Analyzing the transfer mechanism: methodology 170
4.2 Crosslinguistic Blending in Writing 174
4.3 Summary and Discussion 196
CONCLUSIONS 199
REFERENCES 201
Appendix 1 214
Appendix 2 217
Appendix 3 222
Appendix 4 223

FIGURES

Figure 1. Levelt's model of speech production. 17
Figure 2. Incremental speech production. 18
Figure 3. The lexicon as a spreading activation network. 21
Figure 4. Structure of a lexical item. 23
Figure 5. Revised structure of a lexical item. 24
Figure 6. Intra- and crosslinguistic links for features of the lexical item RED. 26
Figure 7. Proposed structure of a lexical item. 27
Figure 8. Grammatical encoding. 28
Figure 9. Constituent assembly. 30
Figure 10. The revised hierarchical model. 42
Figure 11. The "re-revised" hierarchical model. 44
Figure 12. The inhibitory control model. 45
Figure 13. Evidence and analogy in the competition model. 47
Figure 14. Language system development in the DMM. 54
Figure 15. Effect of LME on LS development. 55
Figure 16. Development of a multilingual language system. 56
Figure 17. Proposed amendments to the DMM. 56
Figure 18. Development of a bilingual CUCB. 61
Figure 19. Facets, cognitive models, lexical concepts, and word forms. 63
Figure 20. Ethnic distribution in Lithuania and in the AnRK study program. . 72
Figure 21. Preferred direction of translation. 80
Figure 22. Cultural stereotypes among students of AnRK. 84
Figure 23. Cultural stereotypes as a function of nationality. 86
Figure 24. AnRK identification with Lithuanian culture. 90
Figure 25. AnRK identification with Russian culture. 91
Figure 26. AnRK identification with British/American culture. 92
Figure 27. Lithuanian scores with corresponding Russian and English. 93
Figure 28. Russian scores with corresponding Lithuanian and English. 94

Figure 29. Lithuanian sub-group scores sorted on Lithuanian culture. 94
Figure 30. Minority sub-group scores sorted on Lithuanian culture. 96
Figure 31. Lithuanian sub-group scores sorted on Russian culture. 97
Figure 32. Minority sub-group scores sorted on Russian culture. 97
Figure 33. Russian-only scores sorted on Russian culture. 98
Figure 34. Lexical acquisition proceeds from lexeme to lemma to concept. . 150
Figure 35. Incorrect access route when L2 lexeme links to L1 lemma. 151
Figure 36. Overt and implicit constructions at the lemma level. 153
Figure 37. Two equivalent representations of one conceptual blend. 157
Figure 38. Grammatical blending in the double object construction. 159
Figure 39. Translation as two independent blending operations. 161
Figure 40. Lemma information in the resultative construction. 162

CONCEPTUAL BLENDS

Blend 1. “I wiped my glasses clean.” 162
Blend 2. “*Švariai nusišluosčiau akinius.*” 163
Blend 3. “*Aš nušluosčiau mano akinius švarūs*.” 165
Blend 4. “not into her voice” 175
Blend 5. “we with our team” 177
Blend 6. “have and their bad side” 180
Blend 7. “almost didnt broke” 182
Blend 8. “weddings” 185
Blend 9. “the last our project” 188
Blend 10. “I even will not write” 190
Blend 11. “about spiritual or moral things I even will not write” 192
Blend 12. “*Įrodyti savo, kaip šokėjos, galimybes*” 194
Blend 13. “To prove my as dancers possibilities” 195

COMMONLY USED ABBREVIATIONS

CAIS	constantly available interacting system
CLI	crosslinguistic influence
CUCB	common underlying conceptual base
DMM	dynamic model of multilingualism
DST	dynamic systems theory
EFL	English as a Foreign Language
FP	filled pause
FS	false start
ICM	inhibitory control model
LME	language maintenance effort
LS	language system
NS	native speaker
NNS	non-native speaker
RHM	revised hierarchical model
RP	repeat
TOT	tip-of-the-tongue
UP	unfilled pause

STYLISTIC CONVENTIONS

RED	concept / cognitive model or facet
[RED]	lemma / lexical concept
“red”	lexeme / word form
RED	lexical item (lemma/lexeme combined)
red	example of usage

TRANSCRIPTION CONVENTIONS

(adapted from Chafe, 1998 and Ford, Fox & Thompson, 2003)

yes.	falling intonation
so,	rising intonation
m:	elongated sound
but-	unfinished/interrupted word or phrase
…(1.6)	timed pause (to nearest tenth of second)
[okay]	overlapping speech
@	laughter
#	inhalation
H	exhalation
↑that’s	high pitch
^	cleared throat
(likes)	uncertain transcription
(*)	indecipherable word or phrase

INTRODUCTION

This dissertation combines psycholinguistic, sociolinguistic and cognitive linguistic approaches to study the language production phenomenon of crosslinguistic influence (CLI). Modern psycholinguistics has taken an active interest in language production studies (Carroll, 2008), offering a particularly influential model of speech production (Levelt, 1989) that will be examined in detail. And yet, as many researchers have argued, language data alone are insufficient to paint an accurate portrait of a given speaker's language use. Whether mono- or multilingual, a speaker's system of language(s) is itself a sub-system of the language system of the speaker's family, community, culture, etc. As Cook (2003: 2) writes, "Since the first language and the other language or languages are in the same mind, they must form a language super-system at some level rather than be completely isolated systems." In an **ecological perspective** (Waugh, Fonseca-Greber, Vickers & Eröz, 2007: 120 ff.), language is viewed as a sub-system embedded in a complex ecological system that includes at least the following co-dependent facets: linguistic, cognitive, social, cultural, historical, ideological, and biological. In other words, "discourse influences and is influenced by all the other facets of its ecological setting." This means that, prior to attempting to analyze speech production, "the researcher needs to collect social and cultural/ethnographic information" in order to be able to interpret the results more accurately. At the same time, recent cognitive linguistic advances in the study of conceptual metaphor, conceptual blending, and construction grammar offer unique insights into the structure of the lexicon and even concepts themselves. Each of these fields, of course, has its own terminology, even when they are dealing with the same phenomena (e.g., concepts and the mental lexicon).

A growing number of researchers agree that monolingual speech production is inherently different from multilingual (Grosjean, 1988; de Bot, 1992; Oller, 1997; Kecskes & Papp, 2000; Fouser, 2001; Cook, 2003; Costa, 2005). In this dissertation, the term **multilingual** will be used to mean anyone

who has acquired at least basic communicative competence in at least one language other than his or her first language (L1). This is to avoid the terminological confusion surrounding such terms as "bilingual," which in Li Wei (2000) had even 37 different definitions. Moreover, following Tomlinson (2007), language acquisition will be distinguished from language development. Acquisition is understood as the process by which a learner acquires basic communicative competence in a foreign language (FL). Development is the process by which this ability is extended to a wide range of situations, which includes the development of cultural familiarity and pragmatic skills not normally focused on in language classrooms, eventually leading to a proficiency threshold (Kecskes & Papp, 2000) beyond which the FL becomes a second language (L2) that can be used fluently.

This dissertation will therefore examine the speech production of multilinguals whose English language system has been acquired, but is still developing. Such a system is dynamic (van Geert, 1994; Larsen-Freeman, 1997, 2007; de Bot, Lowie & Verspoor, 2007) and subject to unpredictable dynamic effects, here grouped under the general rubric of **crosslinguistic influence**, or **CLI** (Kellerman & Sharwood Smith, 1986; Cenoz, Hufeisen & Jessner, 2001). As a natural byproduct of human conceptual organization and cognitive processes, CLI is unavoidable, yet it is often interpreted by teachers and monolingual interlocutors as erroneous.

This area of research is of particular interest in Lithuania, where some form of multilingualism is the norm (Statistics Lithuania, 2008). As of December 2005, 71% of the population (2.46 million people) claimed command of at least one language besides their mother tongue. These languages include, in decreasing order, Russian, English, Lithuanian (for those whose L1 is not Lithuanian), Polish, German, and French, among others. Russian was the most important official foreign language for state business when Lithuania was a Soviet republic, but since regaining independence in 1990, Lithuanian has been the only officially recognized state language

(Hogan-Brun & Ramonienė, 2004; Lithuanian Constitution, 1992). Now that Lithuania has joined the European Union, English has become the most necessary foreign language for dealing with the EU. Not coincidentally, the Lithuanian-Russian-English speaker is the most common type of multilingual in Lithuania: according to Statistics Lithuania (2008), over 2 million multilinguals speak Russian, while almost 590,000 speak English. This type of multilingual is the focus of this dissertation.

In this multilingual environment Lithuania experiences **diglossia** (Ferguson, 1959; Fishman, 1967; Hudson, 1991), the situation when one language is considered more appropriate for certain situations than another. This has been the norm for centuries; as described in Hogan-Brun & Ramonienė (2005) and Grumadienė (2005), Lithuanian has co-existed with Polish, Russian, and/or German since the early 1500s. At present, Russian is still the most common second language. Newspapers and magazines are published in Russian, the state radio station broadcasts a Russian-language news segment (Lithuanian Radio, 2008), and it was only in July 2007 that the state television channel took its Russian news program off the air. In this diglossic situation, however, Lithuanian is the prestige language. Attitudes towards Russian speakers speaking Russian vary, naturally, but range from indifference to open hostility. As Hogan-Brun & Ramonienė (2005: 429) write, "it has become socially more prestigious to be Lithuanian than it had been in Soviet times."

English, of course, is not currently imposed on Lithuanians as Russian was prior to 1990. Nevertheless, it is taught as the primary foreign language in schools, often beginning in the second grade. Films are frequently shown in English with Lithuanian subtitles, and radio stations play English language songs. Thus great numbers of Lithuanians are now routinely exposed to English, at least in receptive settings. As Lithuania continues to develop political and economic ties with the European Union, of course, the numbers of Lithuanians using English for business and other international relations will

grow rapidly. This alludes to the status of **English as a lingua franca** (Kecskes & Papp, 2000; Prodromou, 2007), or the use of English in an international context between people whose L1 may or may not also be English. For all of these reasons, an investigation into the language systems of Lithuanian-Russian-English multilinguals is both appropriate and timely.

The object of the dissertation is the speech production process, as illustrated in a 25,000-word corpus of university students' written and spoken English discourse. The data provided by a 63-question sociolinguistic survey explicate students' sociocultural attitudes and patterns of language use.

The **aim** of the dissertation is to provide a detailed sociocultural and psycholinguistic description of a group of multilingual Lithuanian university students in order to better understand the principles underlying crosslinguistic influence.

The following **objectives** have been set in accordance with the main aim:

- to unify the terminology of psycholinguistic and cognitive linguistic approaches to language production;
- to survey the student group in detail, paying close attention to demographic, language history, and cultural factors;
- to construct a database of student online speech production under extreme time limits and to analyze this data for evidence of CLI;
- to construct a database of student offline written production with no time limits and to analyze this data for evidence of conceptual blending.

The hypotheses. This dissertation addresses the following hypotheses:

1. The linguistic attitudes of the multilingual subjects affect their communicative competence in FL.
2. The L1 linguistic and sociolinguistic competence of the multilingual subjects affects their FL competence.
3. Crosslinguistic influence takes the form of transfer of L1 skills into the FL.
4. The mechanism underlying the transfer of L1 skills into the FL is conceptual blending.

Scientific novelty of the research. Although many studies of the effects of CLI and/or transfer on English language production have been done, none have looked specifically at the effects of Lithuanian on English. In addition, the multidisciplinary approach undertaken here is of particular interest as it attempts to bring the most useful insights of three different yet related fields to bear on one highly focused synchronic study of language use.

The results of the study should be useful not only to the advancement of all three fields – psycholinguistics, sociolinguistics, and cognitive linguistics – but may also have practical applications for the development of teaching methods that are targeted to the specifics of the Lithuanian environment. Thus this study indirectly bears on the field of applied linguistics as well.

The conceptual blending approach developed in this dissertation is unique and has not been systematically studied elsewhere. It bears directly upon language acquisition and development studies, as well as on the study of translation. In the modern world, where politics and business both require the translation of sensitive documents, advertisements, and the like, a deeper understanding of the language- and culture-specificity of thought, speech, and writing is essential. The woeful inadequacy of machine translation engines is a case in point. The programmers of computerized translation systems may therefore also benefit from research of the type undertaken here.

Research methods. The dissertation was prepared by combining general descriptive and analytical research methods. This strategy is the most appropriate for both the representation of the theoretical background and for the explication of empirical data.

The **synthesis** of several theoretical paradigms – sociolinguistic, psycholinguistic, and cognitive linguistic – provides a focused empirical analysis. The descriptive aspect is primarily concerned with providing as accurate a description as possible of the nature or conditions of a given situation (in this case, the effects of CLI on students' use of English), with less focus on interpretation or judgment. For these reasons the ethnopragmatic,

ecological approach, which perceives "errors" as culturally and socially unavoidable, has been chosen.

Three analytical tools were selected as best achieving the goals of these methods. A survey was undertaken using a specially designed questionnaire, which was analyzed using elements of the quantitative method of analysis; student speeches were recorded using a digital voice recorder (producing .wav files) and transcribed by hand; a database of written learner constructions was also collected and transcribed in electronic form.

The collection of empirical data was based on a nine-step process developed in Waugh et al. (2007: 141) and suggested as an appropriate, objective, and trustworthy research method. It involves: 1) asking a research question, 2) surveying the potential database (e.g., multilinguals) to determine which would most fruitfully address the question, 3) creating initial hypotheses, 4) collection of corpora: suggested sources include audio- and/or videotaping, teacher observation notes, interviews, and questionnaires, 5) transcription and collection of data, 6) selection of "striking and clear" examples, 7) interpretation of the results, 8) comparison of interpretation with published literature, and 9) drawing conclusions.

Structure of the dissertation. The dissertation comprises this Introduction, four Parts, Conclusions, references, and appendices. Part 1, *Modeling Speech Production*, presents the monolingual psycholinguistic speech production model of Levelt (1989), which includes various processing components and a mental lexicon. This model is criticized from the point of view of multilingual research, and several models proposed within the multilingual framework are also discussed. Part 1 ends with an investigation of psycholinguistic and cognitive linguistic approaches to concepts.

Part 2, *Sociolinguistic Aspects of the Acquisition of English*, presents an overview of modern research into the effects of culture and society on language use. It presupposes an ethnopragmatic approach, which recognizes that speech patterns are formed through socially and culturally situated

experience. Language use is therefore seen as partially dependent on acculturation. The differences between inner and outer speech are also reviewed. Part 2 then turns to the survey, which investigates the students' demographic backgrounds, language learning histories, and attitudes towards Lithuanian, Russian, and English-speaking cultures.

Parts 3 and 4 turn to an analysis of English production, both spoken and written. It was decided to analyze the spoken production first, as this is the type of production the model presented in Part 1 was designed to describe. The spoken language data may thus be used to develop and complement the ideas presented in Part 1 regarding the online speech production of multilinguals. Writing, on the other hand, is usually an offline process and is therefore corrected by students as much as possible prior to handing in. Thus, where CLI can be found in the written corpus, it may be assumed that the students cannot correct it on their own; it stems not from online production constraints (e.g., speed, anxiety) but from the structure of the students' language systems.

Part 3, *Crosslinguistic Aspects of English Speech Production*, defines and describes seven types of CLI that can be found in the corpus of spoken data collected for this dissertation. Each is then discussed with examples from the corpus.

Part 4, *Crosslinguistic Aspects of English Writing Production*, defines learner constructions and places them within the wider cognitive linguistic context of construction grammar. A theoretical framework is developed which links the development of constructions to the psycholinguistic description of the lexicon presented in Part 1. Constructions are then shown to be grammatical blends. The process of blending is further developed and applied to the study of transfer, which is hereby understood to involve a process of crosslinguistic blending. This framework is finally applied to the analysis of eight learner constructions encountered in the written corpus.

The hypotheses are confirmed or disproved in the Conclusions.

1. MODELING SPEECH PRODUCTION

The best-known model of speech production is that of Levelt (1989). Despite working at the Max-Planck-Institut für Psycholinguistik in Germany, certainly a multilingual environment, Levelt explicitly tailors his model for monolingual production, perhaps because the book was published at MIT and aimed at an English-speaking audience. This decision has been criticized and modifications have been proposed (e.g., de Bot, 1992; Green, 1993; Poulisse, 1993; Kecskes & Papp, 2000). However, as this model has been influential and, moreover, remains the most complete speech production model to date, it will be discussed in some detail in section 1.1. Section 1.2 turns to ideas about multilingualism that have been developed into models of language acquisition, representation, and processing. It also discusses the language- and culture-dependence of concepts, which will be seen to be particularly important for explaining certain types of multilingual speech production phenomena. Section 1.3 summarizes the results.

1.1 Levelt's Speaking Model

The model as drawn by Levelt (1989: 9) is represented in Figure 1. As can be seen at a glance, the model is complex and consists of many sub-models. A detailed description of the model is beyond the limits of this dissertation (Levelt's book requires 500 pages to fully describe it), so the discussion here will necessarily be rather concise. Particular effort will be made to point out those aspects of the model which are most relevant to the understanding of multilingual production. It should also be noted that the model makes no attempt to describe the processes at work in speech comprehension; many other models already exist for this purpose. It is a model of **online** speech production; that is, it attempts to describe the process as it occurs in real time, during actual conversation[1]. Finally, it is important that the model does not make room for processes of change in the form of language

[1] As opposed to **offline** production, as when speakers can prepare utterances in advance and merely have to articulate them at the moment of speaking.

acquisition; it is, in de Bot's (1992: 421) term, a **steady-state** model, and thus does not provide any way of describing crosslinguistic influence or transfer without modification.

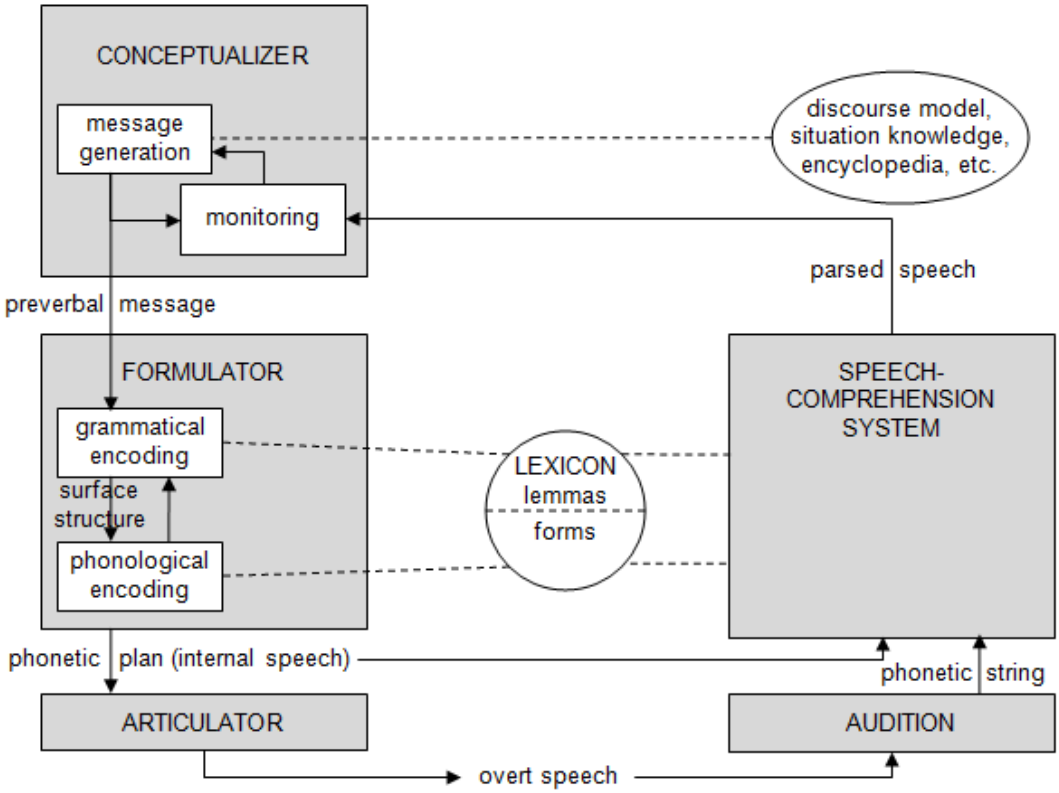


Figure 1. Levelt's model of speech production (Levelt, 1989: 9).

1.1.1 The conceptualizer

It is in the conceptualizer where a speaker's intentions are turned into **preverbal messages**. Messages are defined as "more or less complex conceptual structures that relate entities of different categories to one another" (Levelt, 1989: 79). The process by which messages are generated occurs in two stages: **macroplanning** and **microplanning**. In macroplanning, illocutionary intentions are turned into **speech-act intentions** (cf. Austin, 1962). This essentially involves deciding which mood (declarative, interrogative, or imperative) will be most instrumental in realizing the speaker's communicative goals. Macroplanning is assumed to be language-independent.

As soon as a speech-act intention has been produced, the process of microplanning can begin. Obviously, some messages will be composed of

more than one speech act, and it should therefore be noted that every step in this model of speech production is assumed to be **incremental** (cf. Kempen & Hoenkamp, 1982, 1987). Incremental production means that each component of the production system begins working as soon as it receives appropriate input from the component immediately above it in the chain of production; thus, the first speech-act intention (SA_1) to be produced is immediately sent to the microplanner, even as the macroplanner continues producing SA_2, SA_3, etc. This aspect of production is modeled in Figure 2, where it can be seen that the articulation of *John played* occurs in parallel with the formulation of *in Amsterdam* and the conceptualization of *last week*.

The process of microplanning turns a speech-act intention into a preverbal message that can be encoded by the formulator. The processes involved in microplanning are language-specific insofar as different languages have different requirements for preverbal messages. "In a language that has a tense system, for instance, it is obligatory to encode in the preverbal message the deictic and intrinsic temporal properties of a state or event" (Levelt, 1989: 157). Thus, in English, preverbal messages must be marked as PAST, PRESENT, etc., while in a language such as Malay (Levelt, 1989: 104) there is no tense system and thus no need for tense marking in the preverbal message.

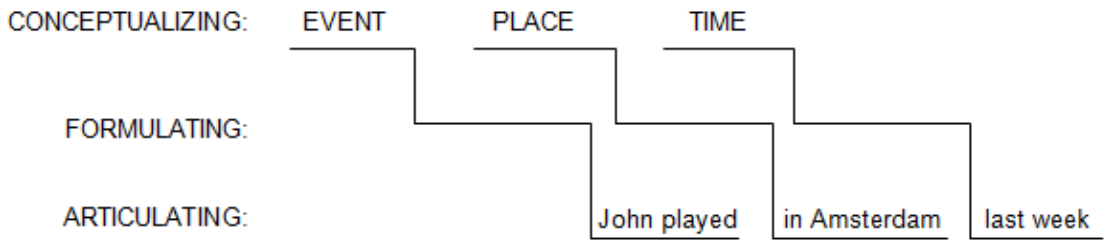


Figure 2. Incremental speech production (Levelt, 1989: 25).

Comments regarding multilingual conceptualization:

1. Although macroplanning is language-independent, microplanning is language-specific. As production is incremental, this is the first point in Levelt's model where multilingual production can differ from monolingual. Imagine that a multilingual speaker trying to generate a message in a foreign

language has difficulty microplanning SA_1, for example, due to unfamiliarity with the language-specific requirements of the FL. It is not necessarily the case that the macroplanner will stop producing SA_2, SA_3, etc. These later speech-act intentions will have to be stored in a **buffer** while the microplanner works on SA_1; this in turn deprives all processing components down the production line of some amount of the **attentional resources**[2] that they will also need, potentially resulting in such overt phenomena as hesitations and/or speech errors.

2. According to Levelt (1989: 157), the microplanning process is automated: "It is unlikely that these computations require special attentional effort." However, depending on a multilingual's level of proficiency in a given language, it may well be the case that conscious effort is indeed required, again with a corresponding drain on resources.

3. The conceptualizer is not only responsible for message generation. It is also the site of the monitor, which will be discussed in greater detail below. Here, however, an interesting point may be made: because the monitor is part of the same conceptual base which generates preverbal messages in the first place, it may be unable to register certain kinds of errors (of conceptualization) *as errors*. This may be especially true in the multilingual case, when a developing or less-proficient language system attempts to monitor itself.

1.1.2 The lexicon and lexical items

Although the preverbal message is sent to the formulator, before turning to its description it is necessary first to discuss the structure of the lexicon and lexical items. According to Levelt's model the formulator interacts closely with the lexicon, which (as shown in Figure 1) consists of lemmas and lexemes (or word forms). The grammatical encoder works with lemmas and the phonological encoder works with lexemes; for this reason it is impossible to discuss the formulation of a phonetic plan (the output of the formulator)

[2] This term, used in Levelt (1989), is similar to **working memory**. As noted in Skehan (1998), working memory is essential for online processing and the allocation of attention.

without first exploring the structure of the lexicon and individual lexical items. Of course, it is understood in Levelt's model that the lexicon is monolingual, although the model can easily accommodate a bilingual lexicon as well.

1.1.2.1 The lexicon

Although Levelt (1989: 183) does not describe the structure of the lexicon in much detail, he does indicate the existence of "item relations *within* and *between* entries" (emphasis in the original). Verb inflections are an example of within-entry relations, while the various members of a **semantic field** (cf. Miller, 1978) are related between themselves. The lexicon itself is assumed to be a **spreading activation network** (Bock & Levelt, 1994), wherein the various nodes representing lexical items are interconnected and in which activation of one node (e.g., lemma selection by the formulator) automatically raises the level of activation of all nodes to which it is connected in the net. "Lexical access in this model is represented by activation spreading from the conceptual level to the lemma level to the lexeme level" (Bock & Levelt, 1994: 952).

This type of network is closely related to the **subset hypothesis** proposed by Paradis (1987) to explain bilingual storage. According to this proposal, all lexical items are stored within a single system, with stronger connections formed between the elements of one language than between languages; this gives the appearance of two separate lexicons, as elements can easily be retrieved separately. However, at least in early stages of language acquisition, the elements of the L2 are strongly linked to those of the L1 through crosslinguistic links that, as proficiency grows, weaken or atrophy and are replaced by intralanguage links. Figure 3, based on Herwig (2001), attempts to illustrate this process. The gray nodes represent part of the L1 network (solid lines), while the open nodes are the developing L2 network (dashed lines). Dotted lines represent crosslinguistic links. As the system develops towards proficiency, the few original L2 lexical items (bolded on the right) lose many

of their crosslinguistic links, while at the same time forming intralanguage links with new L2 nodes.

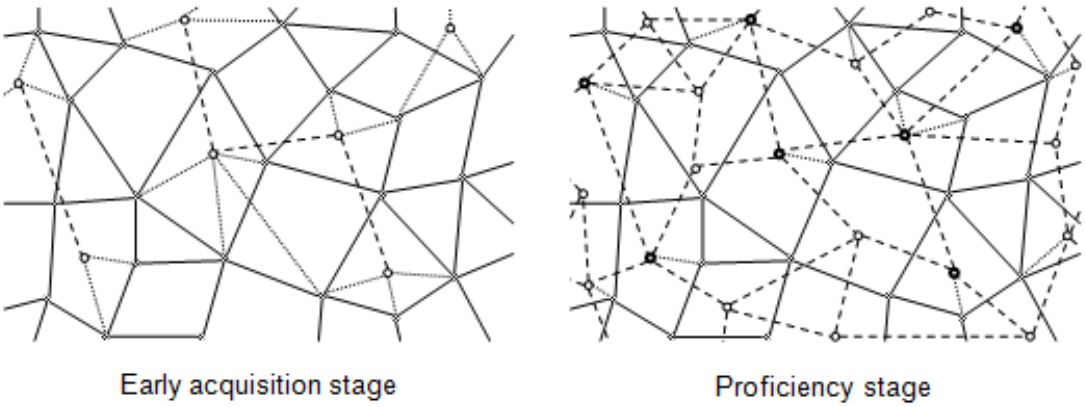


Figure 3. The lexicon as a spreading activation network.

A lexicon in the form of a spreading activation network is assumed by many researchers today (see, e.g., Dell, 1986; Herwig, 2001). It is important that activation spreads in the direction concept → lemma → lexeme, as it is during this process, even in a monolingual system, that difficulties can arise. One example of such difficulties is the **tip-of-the-tongue (TOT) phenomenon**, which can be explained as "a failure to access the lexeme from the lemma" (Bock & Levelt, 1994: 953). When this occurs, a speaker is able to remember a word's meaning, but its actual form is blocked. Indeed, speakers are often able to remember first letters, number of syllables, and even stress patterns for these "missing words" – but not the words themselves (Brown & McNeill, 1966; Jones & Langford, 1987; Levelt, 1989). Thus, it could be said that a lemma is something like a definition without a word attached.

Lemmas are usually accessed without difficulty, but sometimes they too can fail to be accessed or be accessed erroneously, resulting in three well-known classes of (monolingual) speech error: **word blends**, **substitutions**, and **exchanges**. Briefly, word blends[3] are two words fused into one; some common examples from Fromkin (1973) include *stougher*

[3] While it is extremely likely that word blends are caused by conceptual blending (Fauconnier & Turner, 2002), the two terms are meant to describe different phenomena. In this dissertation, therefore, when the term *blend* (or *blending*) is used alone, it is meant to refer to conceptual blending. Word blends will always be referred to as such.

(stiffer/tougher), *clear* (close/near), and *watch* (what/which). Substitutions occur when one word is erroneously uttered in another's place; Fromkin (1973) cites the example *Don't burn your toes* where the intended word was *fingers*. In an exchange, two words or phrases from different parts of the utterance are switched: *a hole full of floors* (instead of: *a floor full of holes*; Fromkin, 1973), for example. Such errors provide considerable support for the depiction of the lexicon as a spreading activation network.

Comments regarding the multilingual lexicon:

1. As noted above, the speech production model under discussion is a steady-state model and takes no account of the development of language systems. The language systems of multilingual speakers, however, are constantly changing in dynamic ways (see below). New words added to the lexicon create new connections and allow activation to spread in new directions, making possible transfer and crosslinguistic influence phenomena of all kinds.

2. The three speech errors discussed in this section are to be distinguished from the kinds of "errors" produced by language learners. As will be seen in Part 3, learner-produced utterances may not exhibit any of these errors, yet still be considered erroneous. The **learner constructions** (Waara, 2004) that will be studied in this dissertation are of this type, and can be understood to stem directly from the still-developing, dynamically-changing nature of their language systems and lexicons.

1.1.2.2 Lexical items

The lexicon as represented in Figure 3 is slightly misleading, as it implies that lexical items are unitary entities with no internal structure. This is far from the case. Unfortunately, researchers disagree on the best representation of this structure, so it will be the purpose of this section to review two of the more useful proposals and to offer a possible unification.

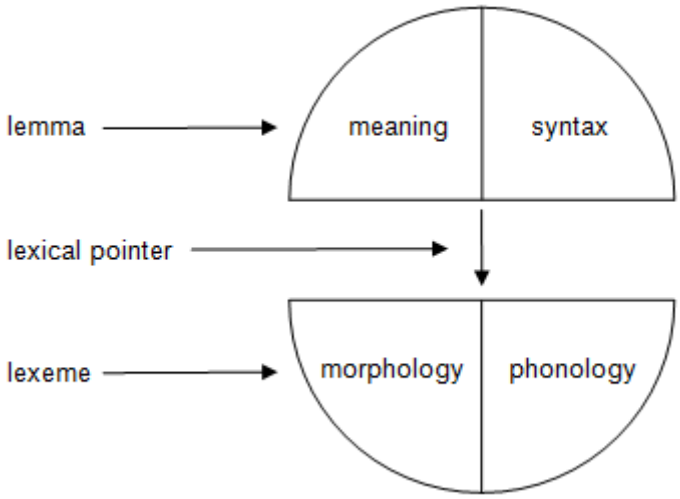


Figure 4. Structure of a lexical item (Levelt, 1989: 188).

Levelt's representation of the structure of a lexical item is reproduced in Figure 4. Here it can be seen that the lemma contains information relating to the meaning and syntactic qualities of an item, while the lexeme contains morphological and phonological information[4]. The two "halves" of the item are connected by a lexical pointer moving from lemma to lexeme. Thus, what are typically thought of as units, "words," are actually entities with an internal structure.

One can have a lexeme without a lemma: this happens every time one encounters a new word in a foreign language. The question "What does this word mean?" is essentially a request for a lemma to be supplied. One can also have a lemma without a lexeme: this happens in TOT states. This issue will be discussed in greater detail in Part 4.

Because the processes of grammatical and phonological encoding work with lemmas and lexemes respectively, Levelt goes so far as to say that the lexicon is in a sense partitioned into two kinds of store: "Lemmas can be said to be 'in the lemma lexicon,' and morpho-phonological forms to be 'in the form lexicon'" (Levelt, 1989: 187).[5] Herwig (2001: 119) offers essentially the

[4] Välimaa-Blum (2005: 242) has recently argued convincingly that it is "not possible to separate phonology from morphology."
[5] This notion is particularly important for the process of translation blending proposed in Part 4.

same opinion: "[T]he so-called 'mental lexicon' consists not in one but in two distinct network systems and the associations that hold between them."

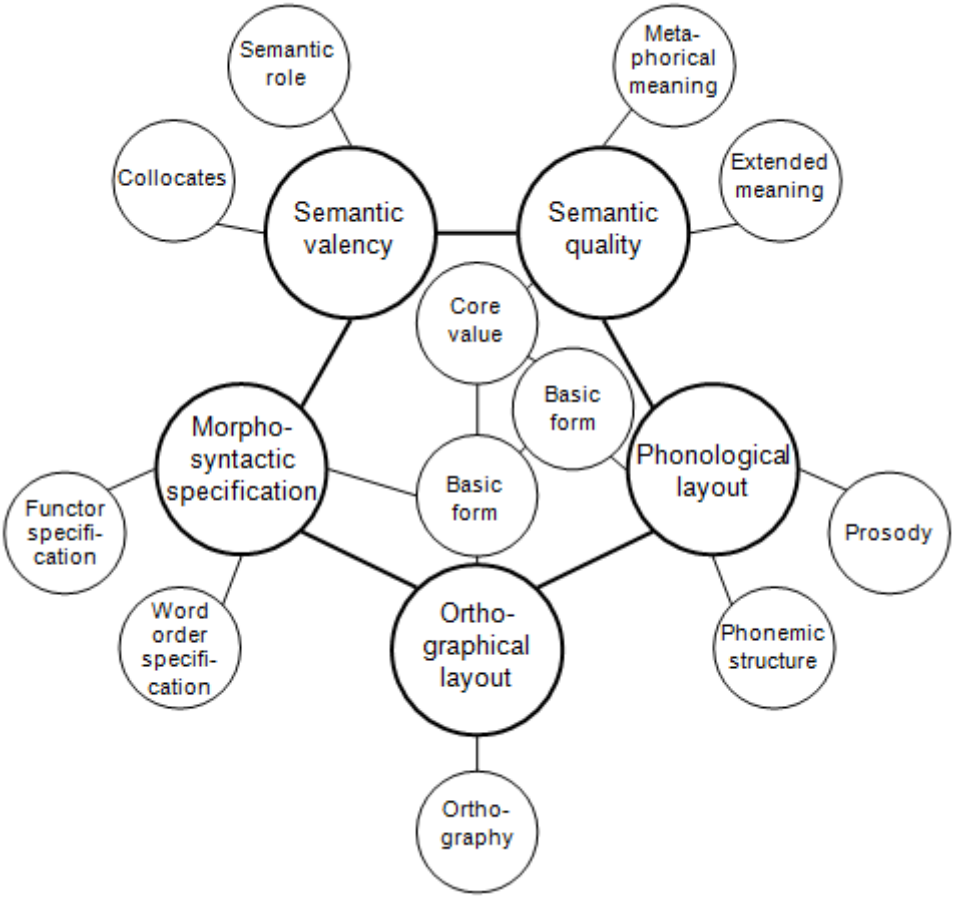


Figure 5. Revised structure of a lexical item (Herwig, 2001: 122).

Clearly such an interpretation does not sit well with the Subset Hypothesis as outlined above. It seems best to understand that the links between elements may be of several different kinds. As an example, consider Herwig's (2001) suggestion for the structure of a lexical item, reproduced, in slightly adapted form, as Figure 5. Here the distinction between lemma and lexeme has, unfortunately, been lost. In Herwig's model, semantic valency and quality form the **semantic pole** of a lexical item, while morphosyntax, orthography, and phonology make up the **formal pole**. Although this suggestion is supported by Langacker (1987) and other cognitive linguists (cf. Evans & Green, 2006), there is simply too much evidence supporting the lemma/lexeme distinction to justify ignoring it.

There are, however, some very useful ideas in Herwig's model. First, she includes the orthography of a word as part of its lexical entry. Herwig, whose paper describes her work with Irish students of Dutch, German, and Swedish, thus acknowledges the important role that the written forms of words play in the process of language acquisition and, indeed, in production[6]. Second, the different-sized circles in the drawing are meant to indicate that the five main attributes of a given word (large circles) can be modified (e.g., contextually) by adjustment of the values attached to them (small circles). This depiction is therefore better able to account for figurative uses of language than the rigid structure proposed by Levelt. Third, as she writes, "The interconnectivity of lexical items can then be marked transparently as associative links at various levels, both intra- and cross-linguistically" (Herwig, 2001: 123). In other words, some lexical items may be connected at the level of semantics, while others may be completely unrelated except through their written forms. Figure 6 is a highly simplified attempt to show how such links might be represented in a lexical network, assuming the structure proposed in Figure 5.

As mentioned above, intralanguage links can be of different kinds. Herwig's model makes it explicit what kinds they may be. In Figure 6, solid black lines indicate connections of semantic quality: the lexical concept [RED] is here shown connected to the lexical items BLUE and COLOR (the internal structure of nodes other than RED is not shown). BLUE is similarly linked to COLOR and also COLD. As RED generally describes objects that can have a hue[7], this semantic valency attribute is also linked to BLUE and COLOR, the dashed lines indicating this connection. BLUE and COLOR are similarly linked in this way, though BLUE and COLD are not: COLD's semantic valency relates to objects that can have a temperature, and while the two sets of objects no doubt overlap to a great extent, they are probably not identical.

[6] It seems possible that, given the use of specialized forms of language for specific purposes, there may be words that some people (almost) never actually speak aloud, but only write in the appropriate contexts; *orthography*, for instance.

[7] To paraphrase Chomsky (1965), ideas really are colorless, not being objects that can have a hue in the visible world, and it would make little sense to speak of (for example) *red ideas*.

Morphosyntactic links (dotted gray lines) connect RED with BLUE and COLD as all three are adjectives. Orthographical links (square black dots) connect the letters (printed and written) R, E, and D with READ and REED, as well as COLD with COLOR (C, O, and L being shared); while phonological links (dotted black curves) connect the sound /ɹɛd/ with the past tense (and also the sound /ɹiːd/ in REED with the present tense) of READ. Finally, solid gray lines represent possible crosslinguistic links that could form if an English speaker were to learn the Lithuanian equivalent of RED, RAUDONAS.

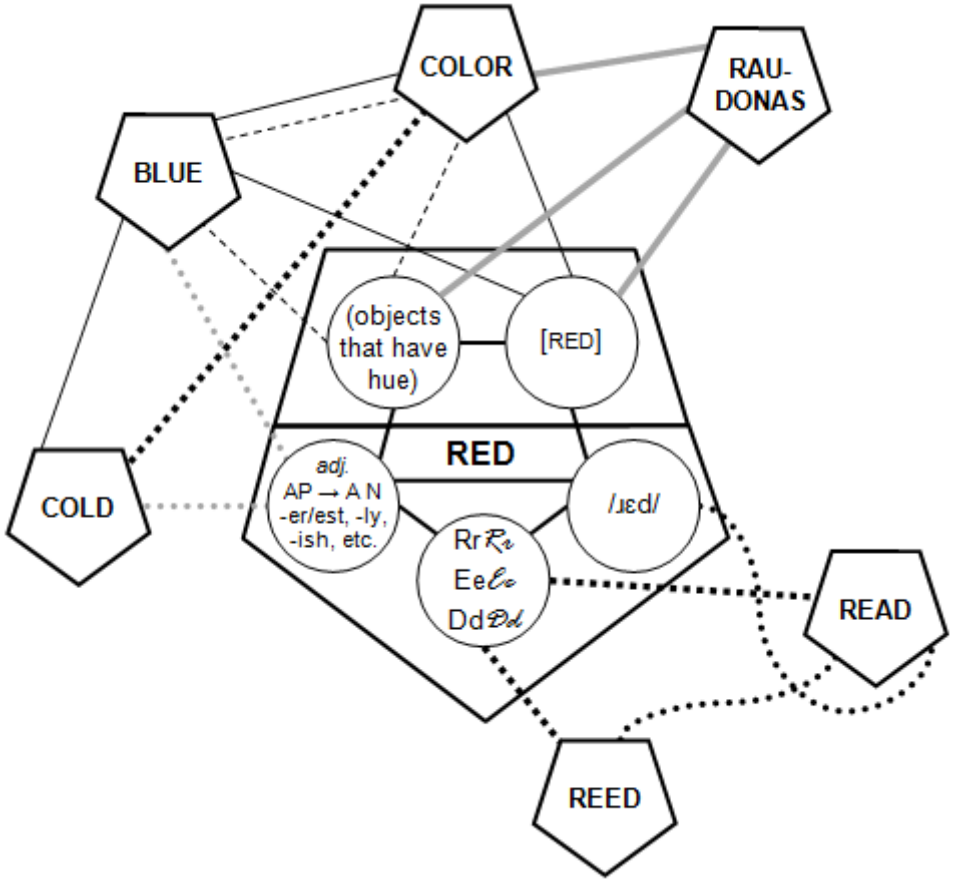


Figure 6. Intra- and crosslinguistic links for various features of the lexical item RED.

To summarize, while it is necessary for Levelt's speech production model to maintain the lemma/lexeme (and, therefore, the syntax/morphology) distinction, it is also necessary to include orthography into the lexeme. In this dissertation, then, when drawing and discussing the internal structure of lexical items, it will be assumed that they can be described as shown in Figure 7. The two circles in the lemma represent the semantic and syntactic information

necessary for grammatical encoding, while the three circles in the lexeme represent the morphological, phonological, and orthographic information necessary for phonological encoding and producing written language. All of this information is in the form of attribute-value sets (as shown in detail in Figure 5), each of which forms its own type of link within the overall spreading activation network of the lexicon.

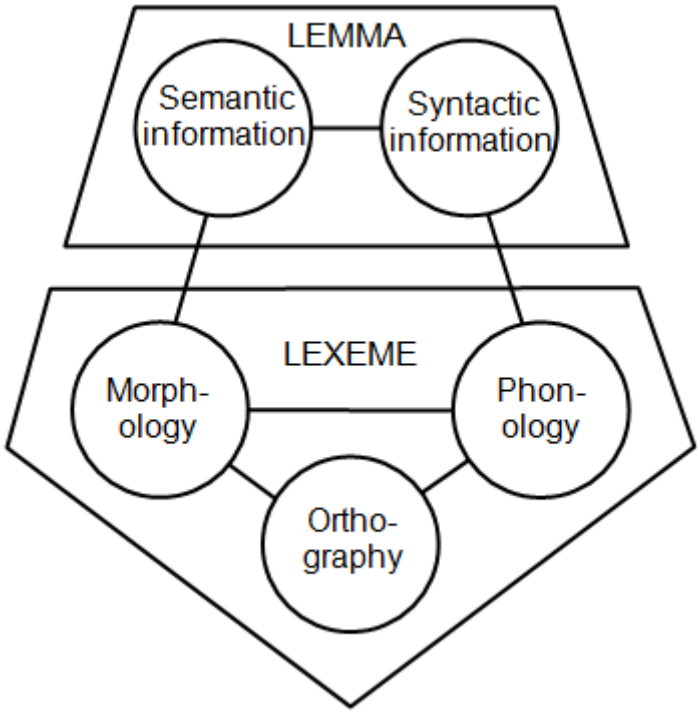


Figure 7. Proposed structure of a lexical item.

It is particularly important to remember that the lemma contains no overt form information. Although, for the purpose of exposition, it is more convenient to refer to lemmas by name, in the speech production process they can be (and sometimes are, as the TOT phenomenon shows) selected without their accompanying lexemes. In other words, lemmas are not words, but information about words, and cannot be produced unless accompanied by lexemes.

When referring to lemmas (without their lexemes) in this dissertation, small capitals in square brackets will be used, as in the example [RED]. When the referral is to a lexeme (without its lemma), it will be written in quotation

marks, e.g., “red.” When the whole lexical item is meant – lemma and lexeme both – the word will be written in capitals, as in RED.

1.1.3 The formulator and the articulator

In the formulator, preverbal messages are turned into a phonetic plan through two processes, grammatical encoding and phonological encoding. These will be discussed individually. This dissertation will not consider errors of articulation, so the articulator will not be discussed in its own subsection. Its activity will be briefly described after phonological encoding.

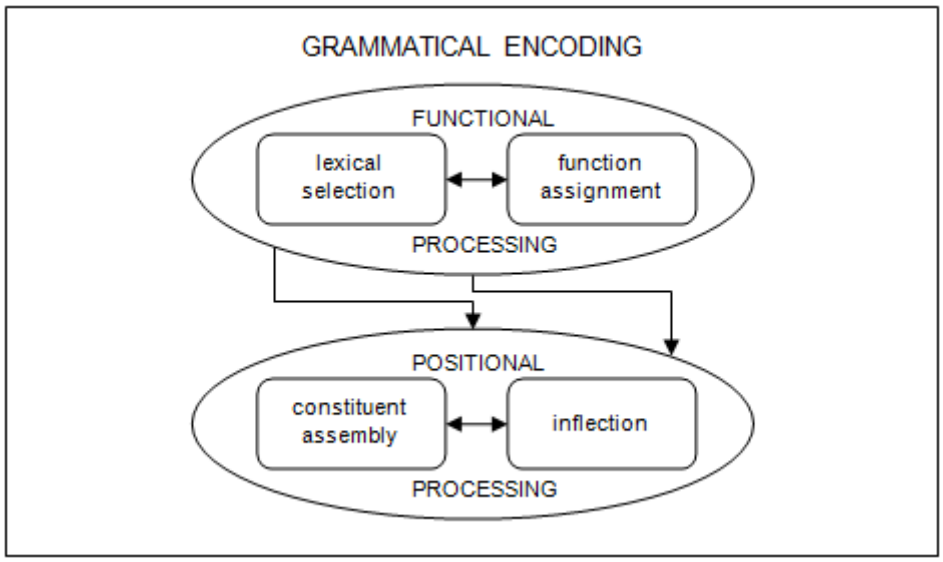


Figure 8. Grammatical encoding (Bock & Levelt, 1994: 946).

1.1.3.1 Grammatical encoding and lemmas

The grammatical encoding process was further developed in Bock & Levelt (1994), and is reproduced in Figure 8. This figure can be understood as an expanded version of the box labeled *grammatical encoding* in Figure 1. Because Bock & Levelt (1994) expands upon the ideas laid out in Levelt (1989), it will be the source for the discussion that follows.

The two stages of grammatical encoding are processed incrementally, as indicated by the staggered arrows connecting them. The first stage is **functional processing**. This stage consists of two subcomponents, **lexical selection** and **function assignment**, which do not occur sequentially, but work together (as indicated by the double-headed arrow connecting them). In lexical

selection, lemmas are accessed that are appropriate for conveying the meaning contained in the preverbal message (or part thereof) received from the conceptualizer. The speech error called substitution, discussed above, is an error of lexical selection. In function assignment, grammatical roles (subject, object, etc.) are determined, with, in languages where such is necessary, the assignment of appropriate case. An error in function assignment can lead to the speech error called exchange.

As each lemma is selected and assigned its appropriate role, it is sent on for positional processing. This stage, too, consists of two subcomponents: **constituent assembly** and **inflection**. Constituent assembly, as described in Bock & Levelt (1994: 947-8), "is the creation of a control hierarchy for phrasal constituents that manages the order of word production and captures dependencies among syntactic functions." As will be discussed in more detail in Part 4, this process is essentially the same as **grammatical blending** (Mandelblit, 1997; Fauconnier & Turner, 2002). The illustration provided in Bock & Levelt (1994) for this process (depicting the example sentence *She was handing him some broccoli*) is reproduced here as Figure 9. In this figure, the words *she*, *him*, *broccoli*, and *hand* are meant to be lemmas[8], and do not represent the lexeme forms of these words, which are accessed only in phonological encoding.

Finally, within each position of the hierarchy, a plan for inflection is created, verbs, for example, being further refined into stem/affix pairs, or number being assigned to nouns. Errors in the inflection process can lead to two characteristic speech errors: **stranding** and **shifts**. (Again, these are monolingual errors.) The exchange error seen above, *a hole full of floors* (Fromkin, 1973) is also an example of stranding: the plural morpheme, attached to the second noun in the phrase during inflection, stays there even when the lemmas [HOLE] and [FLOOR] exchange places (e.g., the error is not

[8] In this dissertation they would be written [SHE], [HE], [BROCCOLI], and [HAND]. There can be no lemma [HIM], as case is attached to the lemma [HE] during function assignment, but only realized as *him* during phonological encoding, which requires accessing the morphological information in the lexeme "he". Thus this figure is slightly misleading.

holes full of a floor). In a shift, conversely, a usually bound morpheme moves to a different position in the hierarchy (e.g., if the error were *a floor fulls of hole*).

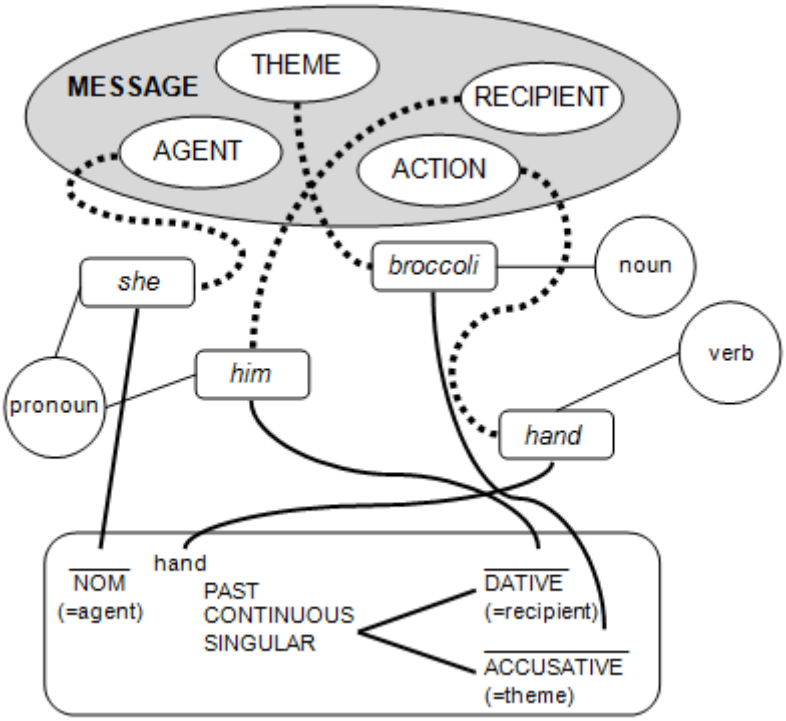


Figure 9. Constituent assembly (Bock & Levelt, 1994: 968).

Levelt (1989) emphasizes more than once the automatic nature of grammatical encoding. For example: "It is likely that this process is highly automatic and nonintentional. A speaker will not, for every message, consider which of various grammatical alternatives would be most effective in reaching some communicative goal" (Levelt, 1989: 282). Thus, in the steady-state, monolingual version of the model that he presents, grammatical encoding is seen as a process that does not require many, if any, attentional resources.

In grammatical encoding, what begins as a preverbal message is, through the four processes described above, turned into a surface structure that can be sent on to the phonological encoder. Clearly the grammatical encoder works very closely with the lexicon. Indeed, if the theory of **logogens** (Morton, 1979) is correct[9], it may be the case that lemmas are directly selected from the

[9] Logogen theory proposes that lexical items have threshhold levels of activation, and that higher-threshold items will be activated later than lower-threshold items. The threshold level

conceptualizer before being sent to the grammatical encoder for functional and positional processing.

Comments regarding multilingual grammatical encoding:

1. For monolinguals, lexical access is only rarely a problem: the TOT phenomenon occurs very rarely (outside psycholinguistic laboratories), and shifts do not occur with great regularity. Multilinguals, on the other hand, are frequently faced with the problem of trying to find an L2 equivalent for an L1 concept and coming up wanting. The problem of **lexical gap** occurs whenever there is no L2 lemma to match with the (preverbal) concepts chosen by the conceptualizer[10]. A related problem occurs when students learn vocabulary from textbooks in decontextualized lists. Under such conditions they are most likely to develop L2 lexemes without L2-appropriate lemmas. In other words, L2 lexemes are attached to L1 lemmas, creating hybrid lexical items that cannot be encoded properly. This process will be developed in detail in Part 4.

2. A native speaker of a (nearly) caseless language such as English may find function assignment particularly troublesome when attempting to speak a case-dependent language like Lithuanian. Similarly, due to the crosslinguistic links in their lexicons, multilinguals may be prone to certain abnormalities in role assignment, especially when the valency of (apparently) equivalent verbs differs in different languages. As an example, consider EN *say* (2 arguments) and LT *sakyti* (up to 3 arguments).

3. The structures created by constituent assembly, as will be argued in greater detail below, are **constructions**[11] (and, typically, grammatical blends). Such constructions are often language-specific, and multilinguals must therefore be careful to use the appropriate constructions for whichever

---

of an item depends on its frequency of activation, e.g., more frequently and/or recently accessed items have lower thresholds than rarer items. This theory also fits well with the idea of resonance in the competition model; see below.

[10] There are, in fact, two kinds of gaps: a true lexical gap occurs when an L1 concept does not exist in L2. A learner's personal lexical gap occurs when he or she simply has not yet learned the appropriate L2 word.

[11] For construction grammar, see Part 4.

language they are attempting to speak. This is particularly of concern when the languages in question differ in configuration, like English and Lithuanian. Because English is configurational (i.e., its grammatical functions are mostly expressed through word order, typically S–V–O), a speaker of Lithuanian (a nonconfigurational language in which grammatical functions are expressed by case endings that leave word order relatively unfixed) may use nonstandard word orders when speaking English. For a monolingual speaker, of course, this issue does not exist.

4. Depending on the speaker's level of fluency in a given language, the automatic nature of grammatical encoding may break down. It can hardly be disputed that grammatical encoding in a native language is automated to a large degree, but multilingual speakers, especially at earlier stages of acquisition, are likely to encounter encoding difficulties when they have not yet automated the language-specific grammatical roles, functions, control hierarchies, and inflections of the target language. Thus, such processes as assigning tense to verbs (inflection) may require conscious control (in the form of recalling irregular verb forms, choosing among the many English tenses, etc.), at some cost to the limited supply of attentional resources.

1.1.3.2 Phonological encoding and lexemes

The phonological encoder takes the surface structure generated by the grammatical encoder and produces a phonetic plan. The phonetic plan can be sent on to the articulator for production as voiced speech, sent to the speech-comprehension system (as **inner speech**[12]) for monitoring, or – this assertion is not directly made by Levelt (1989), but is probable, especially if lexemes contain orthographical information as asserted by Herwig (2001) – buffered until it can be produced in written form.

The phonological encoder works with lexemes. As mentioned above, activation spreads in the direction concept → lemma → lexeme. The morphological and phonological information in the lexeme is similarly

[12] This is not the "inner speech" of Vygotsky; see Part 2.

accessed through spreading activation. Here the direction is morpheme → syllable → phonemes, with the syllable being understood as the most important element in the creation of the phonetic plan. "[T]he phonological specifications of lexical items are retrieved and mapped onto a fluently pronounceable string of syllables" (Levelt, 1989: 361).

The analysis of crosslinguistic influence in Part 3 of this dissertation will not focus on phonological transfer, so it is not necessary to go into great detail in this overview of phonological encoding[13]. In brief, then: the morpheme level accesses word stems and morphological affixes, the syllable level breaks them (if necessary) into onset-nucleus-coda triples, and these are (in the presence of clusters) subdivided into individual phonemes. The phonological encoder also contains a prosody generator, responsible for the generation of rhythm and pitch contours in connected speech.

The types of (monolingual) speech errors associated with mismanaged phonological encoding are: **exchange**, **substitution**, **omission**, **addition**, and **shift**. Here, however, the elements involved are not words or word phrases, but syllables or parts thereof: onsets exchange with onsets, codas are omitted or shift position, etc.

Finally, the syllable string can be modified to accommodate assimilation: "The generation of word forms in context causes striking deviations from the words' citation forms" (Levelt, 1989: 411). Items can be merged and resyllabified in accordance with their phrase-structural relations. The phonetic plan that is sent to the articulator is based on just such restructured material.

Comments regarding multilingual phonological encoding:

1. It is well known that children acquire the ability to perceive language-specific phonetic contrasts in their native language within the first 12 months of life (Sebastián-Gallés & Bosch, 2005). Indeed, it appears that young children are capable of acquiring any phoneme belonging to any language,

[13] Greater detail can be found in Levelt (1989) and Gerken (1994).

provided it is a salient part of their language input, up to a **critical period** ending at about age 12, after which, according to Scovel (2001), biological and neurological constraints make native-like phonology essentially impossible. For multilinguals for whom successive languages are acquired late, phonological encoding is unavoidably hampered by their reliance on L1 phonology. While this usually does not lead to speech errors per se, it can be the cause of certain problems (more on this in Part 3).

2. English is particularly prone to restructurization (merging and resyllabification) of the kind discussed above. Multilinguals with a non-English L1 who are reluctant (or unable, due to a lack of familiarity) to restructure their phonetic plans in English-appropriate ways are likely to develop an "overprecise" speech style that sounds unnatural to native English speakers. Again, this cannot be considered an error, but it may lead to psychological discomfort and/or avoidance issues.

3. English orthography (unmentioned in Levelt's model) is famous for its opacity. Non-native speakers of English often encounter difficulties arising from an overreliance on the written forms of English words[14]. Silent letters and syllables are often pronounced, while syllable nuclei can be pronounced incorrectly (e.g., substituting short vowels for diphthongs in words like *basic*, pronounced /**bɑsɪk**/ instead of /**beɪsɪk**/). Errors such as these may or may not lead to miscommunication, depending on whether the mispronounced form is also an English word.

1.1.3.3 Articulation

The articulator is Levelt's (1989) term for the control mechanism that unifies the nearly 100 muscles, distributed in three anatomically distinct systems (respiratory, laryngeal, and supralaryngeal), for the production of speech sounds at an average rate of 15 per second. In spontaneous speech, the articulator only begins working when it receives a certain minimal input (in the form of a phonetic plan from the formulator). Based on a review of

[14] This supports the idea that lexemes and lemmas can be acquired separately; see Part 4.

experimental data with onset latencies in one and two-syllable words, Levelt (1989) theorizes that the minimum input required by the articulator is one phonetic word. Speech need not, of course, be spontaneous. The phonetic plan can be stored in an **articulatory buffer** for later retrieval. In this case it can be monitored prior to production and, if necessary, sent back for reformulation.

The main speech units dealt with by the formulator are syllables. Evidence indicates that syllables are organized in such a way as to minimize articulatory effort while maximizing perceptual distinctiveness (Lindblom, 1983). Similarly, the systems of the articulator can be adapted to account for environmental disturbances such as noise, food in the mouth, illness or tiredness, etc.

Levelt's model, as already noted, does not discuss written production. Writing is a **graphomotor** skill (Thorne, 2009) that uses completely different anatomical systems than speaking and accesses the orthographical, rather than phonological, information in the lexeme. Moreover, the rate of production of handwriting is much slower than speaking. It is unclear whether this means there is a separate "graphomotor buffer," or whether the same articulatory buffer can be used for both processes. Informal observation of children learning to write indicates that they often mouth words while writing them[15], an indication that the motor skills of writing, at least while they are being learned, piggyback onto the same articulatory system used for speaking. In any case, as writing takes significantly more time to produce, the buffered phonetic plan can be scrutinized, edited, and/or returned for reformulation by the monitor at leisure. For this reason among others, as will be seen in Part 4, written production contains significantly fewer learner constructions than spoken.

[15] Personal observation.

Comment regarding multilingual articulation:

When a learner begins to study a foreign language, he or she is likely to encounter speech sounds that do not exist in his or her native language. The English "th" sounds /**ð**/ and /**θ**/, for example, are not present in Lithuanian; speakers of Lithuanian as an L1 therefore often approximate them: /**v**/, /**z**/, or /**w**/ for /**ð**/ and /**f**/, /**t**/, or /**s**/ instead of /**θ**/[16]. These same speakers have no trouble comprehending the sounds /**ð**/ and /**θ**/, however, indicating that this is a problem of articulation only. They simply have not yet learned to coordinate the speaking apparatus in such a way as to produce these sounds. This does occasionally lead to miscommunication, as when the approximated sound creates another acceptable word (e.g., *first* instead of *thirst*).

1.1.4 The monitor

Once a phonetic plan has been produced by the formulator, it can immediately be monitored (in the form of inner speech) or articulated and monitored (via the speech-comprehension system) as overt speech. This **double perceptual loop** (Levelt, 1989: 469) is not, however, infallible. Errors do occasionally slip through, and in a significant number of cases are never corrected, even in monolingual speech; Nooteboom (1980) noted that 25% of phonological errors and even 47% of lexical errors in one corpus (Meringer, 1908) were not repaired by the speakers[17]. Thus, it is supposed that the monitor is dependent on the limited attentional resources available to the speaker, the level of which fluctuates throughout the speech production process. Evidence suggests that at phrase boundaries attention becomes available for evaluation of the current speech output (Levelt, 1989: 467). Moreover, monitoring is apparently context-sensitive: if the speaker is focusing on grammatical precision, the probability that lexical or phonological errors may slip past the monitor increases.

[16] Difficulties also occur for speakers of English as an L1 when dealing with the "soft" consonants of Lithuanian.
[17] In Levelt (1983), 46% of lexical errors similarly remained uncorrected.

Assuming that the monitor detects an error, a **repair** is initiated. Repairs are generally preceded by **editing expressions** (Hockett, 1967), the most common of which is *uh* (often also written as *er*). According to James (1972), *uh* generally signals that the speaker has forgotten something. Other editing expressions, classified by DuBois (1974), include *that is* (used to specify a referent: *He hit Mary – that is, Bill did*), *or rather* (used to specify a nuance in meaning: *I am trying to lease, or rather, sublease my apartment*), and *I mean* (used when the word is entirely incorrect: *I really like to – I mean – hate to get up in the morning*[18]). Repairs typically follow the **main interruption rule**: "Stop the flow of speech immediately upon detecting trouble" (Levelt, 1989: 478).

The editing expression *uh* forms a significant part of the spoken corpus collected for this dissertation – 10.3% of all words produced. According to Levelt (1989: 483), *uh* is "the only editing expression that is practically universal; it exists, with only minor phonetic variations, in many if not all languages." (The Lithuanian version of *uh* is usually pronounced /ɜ/, rendered in the corpus tapescripts as *eh*.) *Uh* is most frequently used when repairs are **covert**, a covert repair being one in which the erroneous element is never actually articulated. In Levelt's (1989: 483) example, *We start with a green – uh – green point*, the speaker seems to have been planning a word other than *point*, but realized its incorrectness in time to avoid uttering any part of it.

There are two categories of (monolingual) **overt** repair: error repairs and appropriateness repairs. In an error repair, the erroneous element is identified and replaced through the process of **lemma substitution**. In such a repair, the syntax of the original utterance is fully preserved; only the erroneous word is replaced. In an appropriateness repair, new syntactic structure may need to be introduced through the process of **reformulation**: the erroneous element is replaced, but the syntactic information in the lemma does not match with the

[18] Examples from Dubois (1974).

old surface structure, so the entire phrase is returned to the formulator. All three examples provided by DuBois (1974; see above) were overt repairs.

Finally, while monitoring is usually direct (through the double perceptual loop in Levelt's speaking model), it may be indirect, as well. Indirect monitoring relies on the interlocutor to mark incomprehension either overtly (e.g., by asking a question like *what do you mean?*) or nonverbally (e.g., by looking puzzled, startled, etc.).

Comments regarding multilingual monitoring:

1. Even monolingual speakers never correct a significant portion of the errors they produce; as discussed above, this is probably due to the allocation of limited attentional resources. But recall that for monolingual speakers, the processes of the formulator are assumed to be automatic and not under conscious control (thus, not resource-costly). For multilinguals, however, depending on their level of proficiency, the processes of grammatical and phonological encoding may also require conscious control, at a cost to attentional resources. This leaves even fewer resources available for monitoring, thus increasing the likelihood of uncaught errors.

2. Multilinguals at low levels of proficiency may produce errors because of a misunderstanding regarding the rules of the language they are speaking – these are usually referred to in the literature as **interlanguage** errors (Selinker, 1972; De Angelis & Selinker, 2001; Swan & Smith, 2001). Recall that the monitor is assumed to be part of the same conceptualizer that creates preverbal messages in the first place, and cannot be expected to repair utterances that seem accurate. To put it bluntly, you can't detect what you think is correct.

3. Even if errors are caught by a multilingual monitor, there are cases where speakers simply do not know how to correct them. Students of English are often at a loss when asked to provide particles for phrasal verbs, for example, frequently guessing random prepositions in hopes of stumbling

across the correct one[19]. Such "IDK[20] errors" may, in fact, be caught by the monitor; speakers may realize they are saying something incorrectly, but have no other choice than to ignore the error in hopes that they will be understood nonetheless. A similar effect can be noticed for errors of articulation (discussed above): students may sense that they are mispronouncing, e.g., /**ð**/ or /**θ**/, but simply have no alternative.

4. Another type of error unique to language learners is the "NMF[21] error" caused by inaccurate bilingual dictionaries. One electronic dictionary still popular in Lithuania, Alkonas (Piesarskas, 1998), contains a number of simply incorrect Lithuanian–English translations which, unsurprisingly, sometimes find their way into the spoken and, especially, written output of Lithuanian students of English. For example, the Lithuanian word *apuokas* "horned owl" is translated as *turkey* (LT *kalakutas*). Such errors are in no way the fault of the speakers (or writers) who produce them, and obviously cannot be expected to be caught by the monitor.

1.2 Multilingual Speech Production

Section 1.1 described those aspects of Levelt's (1989) speaking model that are most relevant to the issues raised in this dissertation. It also attempted to draw attention to features of the model which may operate somewhat differently for multilingual speakers. As pointed out by the Dutch bilingualism researcher de Bot (1992: 421), "It could be argued that because every unilingual speaker has the potential to become bilingual, the validity of a model can be tested by examining whether it is suitable for bilingualism. <…> [O]ne could argue that the basic model should be concerned with bilingualism, with an option to have a unilingual version." Considering the comments about multilinguals throughout section 1.1, this suggestion seems extremely valid.

[19] Boers & Lindstromberg (2006) suggest that such difficulties may be due to unfamiliarity with the metaphorical structure of English, and that students should be made aware of the motivated nature of the metaphorical extensions such verbs encode.
[20] "I-don't-know."
[21] "Not-my-fault."

Like de Bot, Kecskes & Papp (2000), the Hungarian researchers who proposed the idea of the common underlying conceptual base (see below), have noticed the need for revision of the Levelt model. Their influential critique (Kecskes & Papp, 2000: 40-47) is based on four main points, which will be briefly summarized here.

First, they note that most concepts are not independent of language, but language-specific. This is in agreement with the **ethnopragmatic** approach of Wierzbicka (e.g., 2006) and Goddard (2004, 2006), which proposes the existence of only about 60 "universal" concepts[22]. Green (1993) has also noted that multilinguals must often deal with the situation in which the L2 does not provide a lexical concept (i.e., lemma) which exists in the L1. In his opinion, this shows the language-dependence of concepts. Second, concepts are not only language-specific, but culture-specific as well, "not only in the process of conceptualization (i.e., message generation) but also in the blueprints of most concepts" (Kecskes & Papp, 2000: 41). This will be seen in the discussion of the common underlying conceptual base, below. Support for this claim may be found in Pavlenko (1996), where Russian-English bilinguals were compared with Russian and English monolinguals. Pavlenko found that bilingual cognition is dependent on concepts, the development of which is itself dependent on cultural exposure. Third, the weak version of the Sapir-Whorf theory (Sapir, 1921; Whorf, 1956), or **linguistic relativity**, is supported by research on multilingualism (cf. Gumperz & Levinson, 1996). Whether language actually shapes thought or, as Slobin (1996) believes, acts as a filter through which thought can be discussed, is moot. Pinker (1994), eloquent as his rebuttal of linguistic relativity may be, provides no multilingual evidence against it. Fourth, Kecskes & Papp (2000) emphasize that the two (or more) language channels of a multilingual form a **constantly available interacting system (CAIS)**. This can result in code-switching and other forms of

[22] All concepts other than these, it must be assumed, are more or less language- and culture-specific. Ethnopragmatics will be discussed in more detail in Part 2.

crosslinguistic influence, phenomena that a monolingual speaking model (and, perhaps, a monolingual speaker) is unprepared to account for.

This section, then, examines some of these other aspects of multilingual speech production, aspects overlooked by Levelt's model. First, four of the many models of bi- or multilingualism that have been developed will be reviewed. Because most of these models make reference to concepts and conceptualization, and because conceptual blending will be particularly important later in this dissertation, some theories of concepts and conceptual structure will then be presented.

1.2.1 Models of multilingualism

Many models of bi- or multilingualism have been developed, each examining slightly different aspects of the issue. Some focus on lexical representation or acquisition, others on code-switching or comprehension. Among these models are: the revised hierarchical model of Kroll & Stewart (1994), the inhibitory control model of Green (1986, 1998), the competition model of Bates & MacWhinney (1982) and MacWhinney (2005), the dynamic model of multilingualism of Herdina & Jessner (2002), the distributed features model of de Groot (1992), the matrix language frame model of Myers-Scotton (1993) and Myers-Scotton & Jake (1995, 2001), the bilingual model of lexical access (Grosjean, 1988), and acquisition by processing theory (Truscott & Sharwood Smith, 2004). Only the first four of these will be discussed in this dissertation, as they bear the most directly on the issues of speech production that are being discussed.

1.2.1.1 The revised hierarchical model

One of the most influential and best studied models is the revised hierarchical model (RHM) first proposed by Kroll & Stewart (1994). It has undergone considerable testing and its strengths and weaknesses are therefore well described in the literature. Kroll has continued to explicate the structure and consequences of her model in many recent publications (e.g., Kroll &

Tokowicz, 2005; Kroll & Dussias, 2004; Kroll & Tokowicz, 2001). The basic shape of the model is shown in Figure 10. As can be seen, it is assumed in the model that the L1 system is larger than that of the L2. Unfortunately, an explication of the internal structure of this system is lacking in any of the sources cited above. It is therefore unclear whether "L1" and "L2" in this model are meant to refer to the lexicon alone, or to the entire speech production and comprehension system (though the former possibility is certainly more likely).

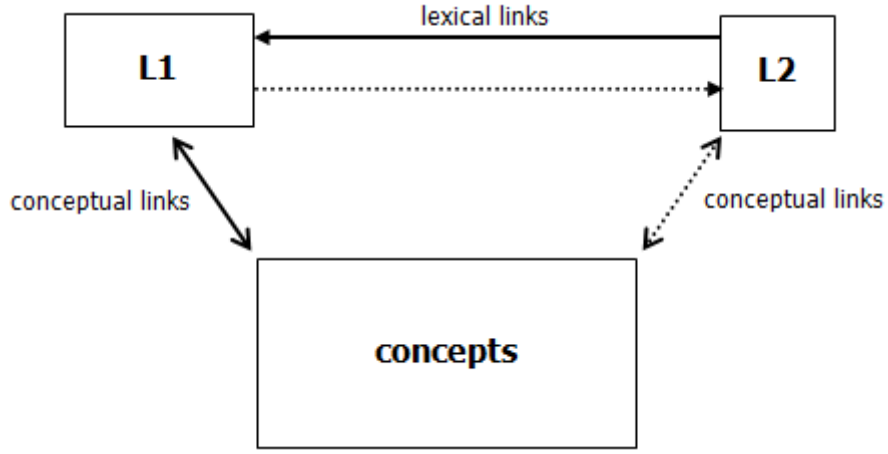


Figure 10. The revised hierarchical model (Kroll & Tokowicz, 2001: 51).

According to the RHM, both languages can access the underlying conceptual base via conceptual links that allow lexicalization of concepts in either language. Solid lines indicate stronger connections. Here, too, this model differs from the speech production model of Levelt (1989), where there are no direct links from the conceptualizer to the lexicon. However, as discussed above, the Levelt model could easily be made to accommodate them through the addition of logogens.

The lexical links that can be seen between the two languages are intentionally asymmetrical, as in early stages of language acquisition the L2 is heavily dependent on the L1 for access to the conceptual base. In other words, an L2 word must first be translated to L1 before it can be understood conceptually. The reverse is also true: a concept must first be lexicalized as L1, then translated to L2. This is in agreement with the language- and

cultural-specificity of concepts proposals of Kecskes & Papp (2000) and Green (1993). The weak direct connection between L2 and the conceptual store grows stronger with later proficiency. The weak connection in early stages leads to **translation latencies**, or slower response times, when translation is from L1 → L2 as opposed to L2 → L1.

Kroll and Dussias (2004: 178) write that the RHM "is fundamentally a model of the development of L2 proficiency." In Kroll & Tokowicz (2001) it is noted that, in early stages of L2 acquisition, lexical form is the most important cue to meaning comprehension. Some obstacles to the development of the direct connection between L2 and concepts are also mentioned, most of which stem from the problem of linking new words to old concepts. Because the lexicalization of concepts into L1 is highly practiced, there is strong competition from the L1 for any lexicalization (as per the competition model, below). The authors suggest that in immersion situations, the unique environmental cues associated with the L2 can help overcome this competition.

Heredia (1996; Heredia & Brown, 2004) has examined the RHM in detail, and suggests "re-revising" it. While he agrees with the asymmetrical nature of the connections depicted in the model, he notes that bilingual memory is in large part a function of which language is used more often. He describes cases of immigration in which the L2 comes to dominate the L1, a situation that in extreme cases can lead to attrition or even loss of the L1 (cf. Jessner, 2003; Herdina & Jessner, 2002; Scovel, 2001; 1998). Therefore Heredia suggests relabeling the model, as in Figure 11. Here, MDL stands for **more dominant language**, and LDL for **less dominant language**. This may more accurately reflect the dynamic, ever-changing bilingual language system. Heredia's revision is also compatible with the theoretical foundation of the dynamic model of multilingualism (see below), in which the process he describes is referred to as **transitional bilingualism** (Herdina & Jessner, 2002).

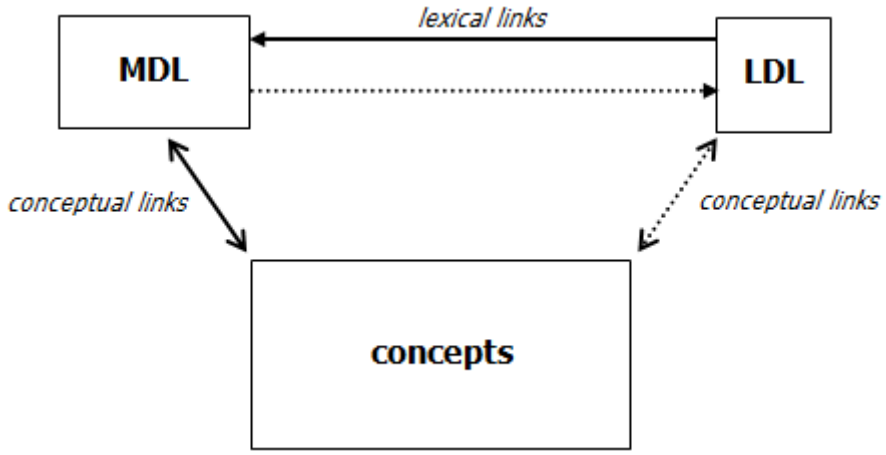


Figure 11. The "re-revised" hierarchical model (Heredia, 1996).

One of the major insights of the RHM is that the representation system of bilinguals changes with increased proficiency. Specifically, direct links develop between the L2 and the conceptual store, links whose absence in early stages is compensated for by translation. It is therefore suggested by many authors (e.g. Michael & Gollan, 2005; Kroll & Tokowicz, 2001; Wong Fillmore, 1991) that **immersion** is much more effective than classroom education in developing direct conceptual links to and from L2.

1.2.1.2 The inhibitory control model

The inhibitory control model (ICM), first proposed in Green (1986; revised 1998), is based on the idea that in order to produce a word in one language, a bilingual speaker must also inhibit or suppress the equivalent word(s) in his or her other language. This supports the notion of the constantly available interacting system (Kecskes & Papp, 2000) and suggests a mechanism by which language mixing and other forms of crosslinguistic influence can be reduced. It should, however, be noted that according to both Kecskes & Papp (2000) and Grosjean (2001), it is impossible ever to suppress a language system fully or otherwise "turn it off"; all language systems are always available to the speaker. Indeed, Roelofs (1998) proposes that during the multilingual speech production process, utterances are planned in parallel in both (or all) languages, with the unnecessary lemmas being suppressed as much as possible prior to articulation. This proposal was tested and evidence

supporting it was found by Dewaele (2001). Incidentally, this proposal explains why code-switches happen so fluently: the switch language utterance has already been prepared and does not need to generated from scratch.

The representation of the model provided in Figure 12 has been adapted from Green (1986). It has been simplified here to exclude the comprehension subsystem, as production is the main issue in this dissertation. The two main features of the model which differ from the RHM and Levelt's model are the **specifier** and **resource generator**. The specifier is a subsystem proposed by Green (1986) which determines both the language of production and the register, or speech style, to be used. This is accomplished through control instructions (diamond arrows) that make use of **tags**, or feature labels, assigned to each lexical item.

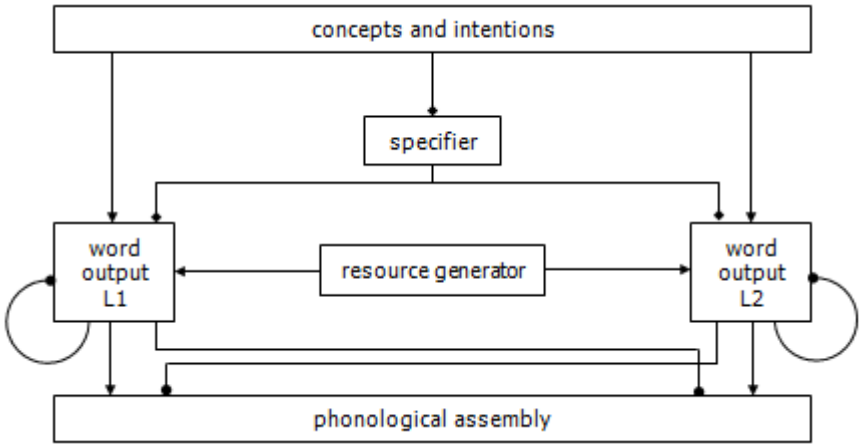


Figure 12. The inhibitory control model (Green, 1986).

The resource generator is postulated to account for **working memory** constraints. Attentional resources in this model can be either **excitatory** (pointed arrows) or **inhibitory** (round arrows), and it is assumed that they are produced by the generator at a fixed rate. Obviously the rate at which they are consumed varies, depending on the task, such that it is possible to overwhelm the generator. Although the model is drawn for bilinguals, Green (1986) discusses trilinguals and polyglots as well. He stresses the fact that with an increase in the number of languages comes an increase in problems of control and suppression.

The ICM assumes that both of a bilingual's languages are always active and that, in order for speech production to succeed, one of them must always be inhibited. Two types of inhibition can be seen: **internal** and **external**. The looped inhibitory arrows represent internal suppression within the language output subsystem; the crossed inhibitory arrows represent external suppression of one language by the other. In other words, when a bilingual is speaking L1 spontaneously, this externally suppresses the L2 before the stage of phonological assembly[23]. However, when bilinguals translate from L2 to L1, the L2 suppresses itself.

Kroll & Tokowicz (2001) note that the translation asymmetries seen in the RHM can also be explained by the ICM: as the L1 is presumably used more frequently and constantly in a higher resting state of activation, it is consequently harder to suppress than the L2: this suppression consumes more resources and leads to longer response times. This issue has been experimentally studied by Lee & Williams (2001), who affirm that lexical competition, both within and between languages, is resolved by inhibition.

Finally, this model may lead one to consider other factors which may require inhibition (and therefore consume attentional resources) during the speech production process, particularly distractions external to the process itself, such as background noise (here construed to be not only construction equipment or music, but also speech overheard by, but not aimed at, the speaker; when one is trying to produce an L2 utterance, overheard L1 input can be especially distracting) and emotional or other affect (e.g., overcoming shyness to speak before a group).

1.2.1.3 The competition model

This model, first proposed by Bates & MacWhinney (1982), is primarily a model of child language acquisition, though it has also been used to test various hypotheses about comprehension and production in second language

[23] It is unclear whether the equivalent in Levelt's model of the ICM's *phonological assembly* is phonological encoding or articulation, though the latter appears more likely.

acquisition as well (Kilborn, 1994). There are many aspects to the model, most of which do not bear directly on the issues of concern in this dissertation[24]. Excellent summaries are provided in MacWhinney (2004, 2005).

As the name implies, **competition** is at the core of this theory. The two competing forces in language acquisition are **analogy** and **evidence**. In order to produce novel utterances, children can draw analogies based on what they have learned of a language's grammar, phonology, etc. At the same time, through interaction with others, children gather evidence about new structures, rules, etc. When evidence contradicts analogy there is an internal competition for production. Children learning English as an L1 invariably come up with a rule that regularizes all past tense verbs to *-ed*: they produce *goed*, *writed*, *bringed*, etc. (Pinker, 1994). At the same time, they can hear adults saying *went*, *wrote*, *brought*, and so forth. Eventually, according to the model, analogy and evidence even out and result in a grammar that accepts a small core of irregular verbs and a large number of regular ones. To summarize, MacWhinney (2004) provides a general schema for competitive acquisition of meaning, reproduced as Figure 13.

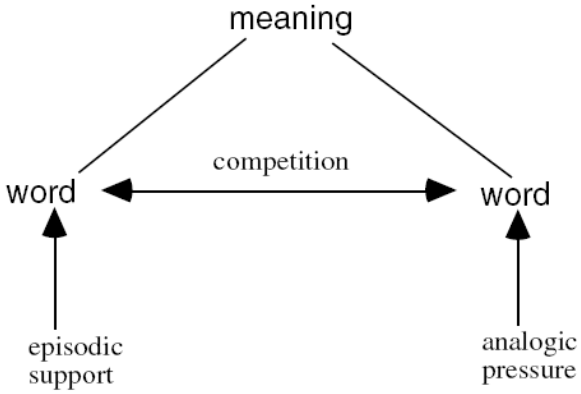


Figure 13. Evidence and analogy in the competition model (MacWhinney, 2004: 901).

Unlike Levelt (1989), MacWhinney (2005) specifically claims that the competition model can be applicable not only to monolingualism and child

[24] Indeed, MacWhinney (2005: 50) specifically states that his "is not to be interpreted as a processing model. Rather, it is a logical decomposition of the general problem of language learning…."

language acquisition, but also to second language acquisition and multilingualism in both children and adults. In multilinguals, competition comes to refer to the choices a speaker must make both within and between languages. According to MacWhinney (2005), competitive processes affect both speech production and comprehension. He uses the term **arenas** to describe those aspects of the processes affected by competition. The five competitive production arenas[25] are summarized in Table 1 and described briefly below.

Table 1. Five competitive arenas in speech production.

production	phonology	*articulatory planning*	*+ orthography (writing)*
	lexicon	*expressive lexicalization*	
	morphosyntax	*sentence planning*	
	conceptualization	*message formulation*	

Competition affects the four arenas of production in the following ways. In **articulatory planning** there is a competition between types of phonemes, between syllables for insertion into a rhythmic phrasal output pattern, and between overall intonation patterns. This can lead both to the characteristic speech errors described in Levelt's model, and also to the multilingual errors noted above. In **expressive lexicalization**, individual lexical items compete for the expression of communicative intent. In order to request a particular item a speaker must decide how best to refer to it: *that*, *that thing*, *it*, *the book*, *Cinderella*, etc. A multilingual speaker must choose not only among these possibilities, but also those available in his or her other language(s).

In terms of **sentence planning** (also called **morphosyntactic arrangement**), the competing elements are arguments which compete for attachment to predicate slots. In English, for example, one can say *I lied* (one argument), *I didn't say the truth* (two arguments), or *I told him a lie* (three

[25] Comprehension arenas will not be discussed in this dissertation.

arguments). Ungrammatical options, created on the basis of analogy, such as **I said him a lie* may also compete, until they are eliminated by the weight of evidence against them. In **message formulation** there is a competition between communicative goals. A speaker may begin asking a question but end up telling a story, or ask a different question than was originally intended. In addition, speakers must learn the (often socioculturally determined) rules for turn-taking and appropriacy in conversation, all of which may be determined through competitive processes. Finally, even **writing** is seen as a competitive process, as writers must choose among handwriting styles (printing or cursive) and letter/sound combinations[26]; multilingual writers may even need to choose between alphabets[27].

It may be helpful to emphasize the developmental nature of this process. The balance between evidence and analogy is in a constant state of flux. Every new utterance the learner hears can change the weight of evidence in favor of one or another analogically created rule. As time passes, more and more rules become overwhelmingly biased in favor of one particular interpretation. By adulthood, then, language learners have more or less determined all of the biases in their language systems. But even then, new evidence can cause them to reevaluate and update their grammars throughout the developmental process.

The term **resonance** (MacWhinney, 2005: 50, 60-63) is also important to understanding the competitive biases developed by language learners. It refers directly to neural networks of the sort assumed in the models (especially lexical) that have been discussed above. In essence, it is assumed that the more a set of reciprocal connections is used, the more resonant it becomes. Resonance can thus be understood as a lowered activation threshold (cf. logogen theory; Morton, 1979): more resonant items and connections (e.g., everyday vocabulary) will "come to mind" first, before more rarified forms (such as the formal, academic language suitable for scientific conferences). As

[26] E.g., the English phoneme [**f**], which can be written *f, ff, ph*, and even *gh*.
[27] One student (L1 = Russian), on a test of English syntax, wrote "я" instead of "I", presumably due to stress.

such, resonance explains the frequency effects so often noted in psycholinguistic experiments (e.g., Ellis, 2002; Jurafsky, 2003). Thus, transfer in the competition model is seen as a competitive process heavily influenced by resonance. An adult trying to learn a second language cannot avoid transfer from his first language in all areas, from articulation and audition to interpretation and pragmatics. As MacWhinney (2005: 55) puts it, "whatever can transfer will." The biases a learner develops in learning an L1 will for a time outcompete any differently biased rules from the L2 due to their overwhelming resonance. Interlanguage errors can also be explained by resonance: an established and practiced (incorrect) pattern (e.g., regularized irregular verbs) will be more resonant than the (correct) alternatives (e.g., irregular forms), meaning that it will outcompete them during the process of speech production.

1.2.1.4 The dynamic model of multilingualism

Recently, an entire issue of the journal *Bilingualism: Language and Cognition* (2007: volume 10, issue 1) was devoted to a discussion about **dynamic systems theory (DST)** among leading researchers in the field. The keynote article, by de Bot et al. (2007), details many of the most important issues associated with the view of language as a dynamic system. Before discussing these issues, however, it should be established just what is meant by dynamic systems. Van Geert (1994: 50) provides the following description of a system:

> A system... is more than just a collection of variables or observables we have isolated from the rest of the world. It is a system primarily because the variables mutually interact. That is, each variable affects all the other variables contained in the system, and thus also affects itself.... In this sense, a system is, by definition, a dynamic system....

This description applies not only to biological or physical systems such as organisms or the behavior of atoms, but also to languages. Even in the case of monolinguals, for whom multilingual issues such as crosslinguistic influence

do not apply, the single language that they know forms a system of mutually interacting variables as described by van Geert.

In relation to DST, the most important issues mentioned by de Bot et al. (2007) are these: a) dynamic systems exhibit **complete interconnectedness**; b) there are systems within systems in a **nested** sense, and every system both contains smaller sub-systems and is itself a sub-system of a larger system; c) systems settle into **attractor states**, which are by definition temporary but often stable over time; d) dynamic systems develop unpredictably and are sensitive to **initial conditions**.

Complete interconnectedness means that all variables interact, as expressed above by van Geert (1994). This mutual interaction among variables is a key quality of all dynamic systems. Viewed in this way, features of the (second, third, etc.) language learning process such as crosslinguistic influence and code-switching can no longer be seen as errors, but rather as natural (and unavoidable) feedback-like effects. Any change to a language system, e.g., the introduction of new vocabulary, cannot occur without consequences for the existing system. Similarly, each individual speaker's system will in large part depend on individual factors. Thus, no two students can be treated as equal because the language systems that they bring to the learning environment are completely unique.

By saying that systems are nested, de Bot et al. (2007) draw attention to the fact that any given sub-system of the language learning process (e.g., speech production) is itself composed of sub-sub-systems (as seen above in Levelt's model). Thus, traditional (non-dynamic) approaches to language acquisition, which attempt to study such sub- or sub-sub-systems in isolation and then draw conclusions about the larger system, are bound to be inadequate. Research into second language acquisition, according to Larsen-Freeman (2007; 1997), has been hindered by exactly these types of reductionist tendencies. She expresses concern with the prevailing assumption that "by studying influences on the process in a piecemeal fashion, and then

aggregating the findings, we would be able to explain the whole" (Larsen-Freeman, 2007: 35). In Larsen-Freeman's opinion, what is called for is a theory that recognizes the chaos and complexity of the acquisition process.

The concept of attractor states is taken from chaos theory (for an excellent introduction, see Gleick, 1987). In language acquisition terms, de Bot et al. (2007) provide examples such as **fossilization**: the incorrect, yet relatively stable, (interlanguage) structures produced by some learners are those learners' attempts to organize the chaos of a new language system into some sort of order. That some order would be imposed from any existing language system(s) seems highly likely: such imposition, of course, is referred to in this dissertation as crosslinguistic influence. Attractors are further discussed in van Geert (2007), who explains that such states are usually insensitive to small perturbations, but can be changed with a large enough disruption. Moreover, attractors are emergent: they spontaneously self-organize from the chaos of the new language system, and the same attractor state can be reached from different sets of initial conditions. Note also the similarity of this concept to that of resonance (from the competition model).

The sensitive dependence of dynamic systems on initial conditions is the fourth point mentioned by de Bot et al. (2007). Their discussion focuses on the problems caused by phonological misperceptions in the earliest stages of language acquisition. For an L1, such problems can be caused by, e.g., hearing disorders, ear infections, etc. However, such early childhood-based conditions are far beyond the capabilities of this dissertation, which examines crosslinguistic influence in trilingual university students. These students come to university with a long language learning history behind them, and the best that can be done is to attempt to determine at least some of the conditions present before any experimental data are collected (see Part 2).

Many language acquisition researchers may well be leery of the approach implicit in DST, for it certainly creates problems when discussing research

findings. Indeed, if all variables are interconnected, then the very idea of empirical research must be called into question. How can one conduct a controlled experiment without the ability to control variables? How can any sort of teaching materials or methodology be developed at all? And yet the fact remains that they are, and that they work. Students routinely learn languages regardless of the problems just discussed. If dynamic systems are so sensitive to initial conditions, how is it possible that the same attractor state can be reached from different starting conditions? As Ellis (2007: 23-24) writes, "There are regularities. They are not prewired. They are not learned by simple imitation and memorization.... [H]ow do these patterns emerge from the interaction of these forces integrated over the processing of each and every utterance and exemplar of language?"

The **dynamic model of multilingualism (DMM)** proposed by Herdina & Jessner (2002) is an admirable attempt to rise to the challenge. This model rejects outright the linear theory of language development, in which language learning is seen as a series of steps to be climbed in a particular order. Instead, the DMM assumes that language development is characterized by nonlinearity, reversibility (also known as attrition), stability (discussed above as attractor states), interdependence of variables, complexity, and changes of quality through individual stages of improvement and restructuring. These changes are manifested as phases of growth and retardation. The process of language development in the DMM is modeled as a sine curve, as shown in Figure 14. It is understood that this curve is an idealization, and that every individual will manifest a differently shaped learning curve. Language development is a process in which slow initial growth gradually accelerates, only to be slowed again and finally achieve a state of equilibrium. The causes of this slowing include limited learner resources (in terms of time and energy, attentional resources, motivation, etc.) and the language maintenance effort required to uphold the growing language system (LS in the figure).

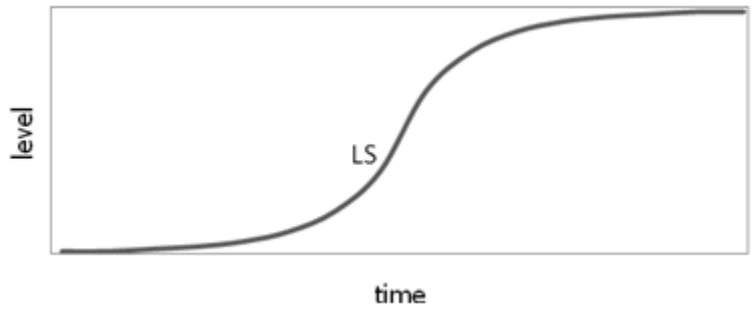


Figure 14. Language system development in the DMM (Herdina & Jessner, 2002: 100).

Language maintenance effort (LME) is composed of two factors: **language use** and **corroboration**. The language use factor describes the (rather commonsense) fact that simply using a language counteracts its attrition. The corroboration factor describes the fact that such activities as "looking up the spelling of a word or reflecting on the systematicity of certain grammatical aspects [or inquiring] about the appropriacy of punctuation" (Herdina & Jessner, 2002: 98) constitute a renewal and entrenching of different language sub-systems. LME creates an upper limit to the development of a language system because the effort required to maintain a system eventually exceeds the increase in actual language knowledge.

LME therefore poses a threat to the stability of any system comprising more than one language. Herdina & Jessner (2002: 103) propose the **principle of theoretical monolingualism**, which states that "monolingualism is the natural state of a speaker." This is because multilingual speakers are under constant (internal, competitive) pressure to optimize their linguistic systems, and the redundancy of lexical items, grammatical and syntactical structures, etc., required by multilingual systems is inefficient and non-optimal. Code-switching, seen above in other models, is viewed in the DMM as an indication of the erosion of one or all of the language sub-systems, which are taken to be in competition for LME. Thus, the DMM is also a model of transitional bilingualism, in which one language system is gradually replaced by another: recall Heredia's (1996) re-revised hierarchical model.

The impact of LME on the developing LS can be seen in Figure 15. As elsewhere in the DMM, the figure is an idealization. In early stages of learning, the amount of LME required to increase the level of knowledge of the system is comparatively high. For this reason, language learning is generally viewed as a challenging task. The amount of effort required to learn new vocabulary, grammar rules, phonetic features, etc., hardly seems worth the trouble in the early stages, as relative gain appears to be quite low. However, as time progresses, the level of knowledge can outstrip the amount of LME that the learner is able to invest, with a dual result: a) the learning curve flattens and stabilizes; b) the amount of LME required to maintain the given level of knowledge drops considerably (relative to knowledge). For this reason, many people who have taken a semester or two of a language in school or university can still, even decades later, recall a few phrases or the numbers from one to ten. The amount of effort required to maintain a very small level of knowledge is itself very small.

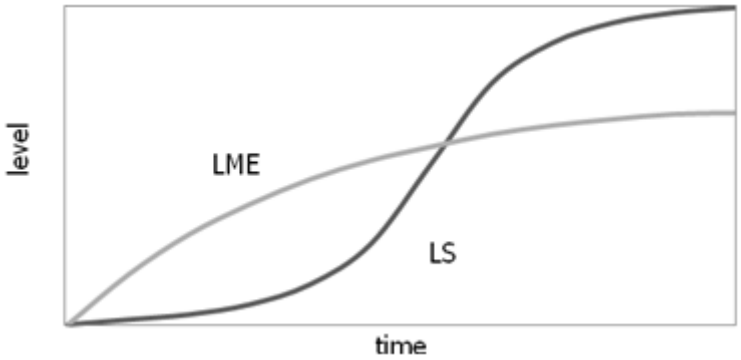


Figure 15. Effect of LME on LS development (Herdina & Jessner, 2002: 113).

Finally, the DMM (unlike any other model) attempts to describe the development of the multilingual language system (Figure 16). According to the model, the introduction of LS2 (and its subsequent drain on LME resources) flattens the curve of LS1. At the same time, LS2 is unable to develop to the same level as LS1 because a portion of the effort needed to acquire LS2 is already being used to maintain LS1. The two language systems settle into a stable relationship and are maintained in this way, with LS1 clearly dominant, until the introduction of LS3.

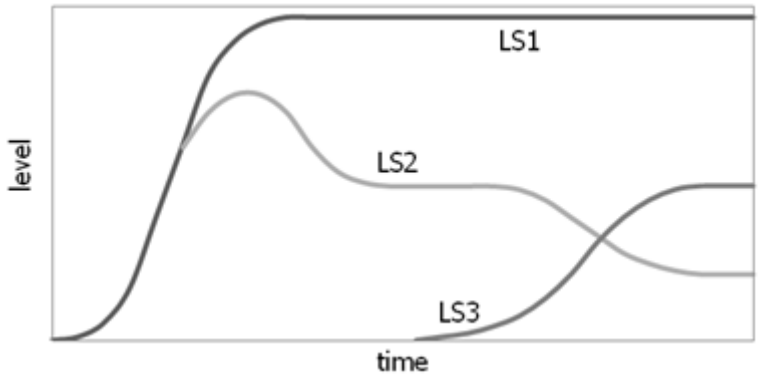


Figure 16. Development of a multilingual language system (Herdina & Jessner, 2002: 124).

At this point the model shows the impact of LS3 on LS2, but with no simultaneous impact on LS1. Whether it is possible in a dynamic system to introduce a new sub-system without impacting both previous sub-systems is not mentioned. However, DST clearly states that all variables are interconnected. Thus it appears that this part of the model is either inaccurate or not fully developed. Furthermore, having introduced LS3, the model predicts that this new system will grow at the expense of the older (yet still non-native) LS2, eventually replacing it. While this may indeed be the case in certain specific contexts, logically the intersection of LS2 and LS3 would be better understood as the beginning of a ripple effect in the graph, with the two systems alternating in dominance until a new equilibrium can be reached. Figure 17 is a proposal for how such an effect might look as a graph of the sort drawn in the DMM.

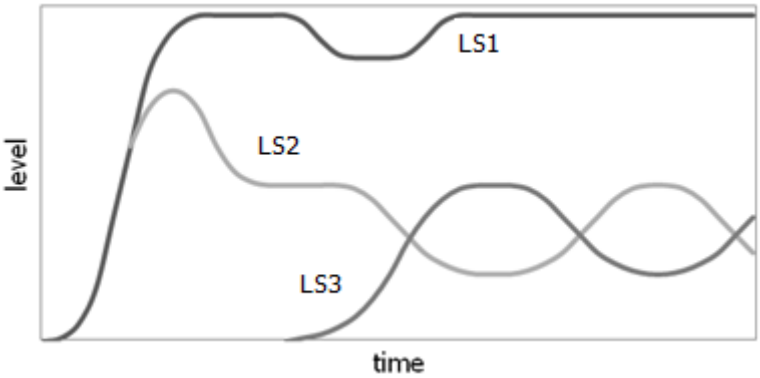


Figure 17. Proposed amendments to the DMM.

1.2.2 Concepts

Many of the models that have been examined thus far have made rather uncritical use of the term *concepts*. Levelt proposes a conceptualizer that turns concepts into preverbal messages; Kroll & Stewart (1994) link language systems to a conceptual base; in Green's (1986) inhibitory control model, speech production also begins with concepts and intentions. Even such groundbreaking works as Lakoff & Johnson (1980), which introduces the idea of **conceptual metaphors**, and Fauconnier & Turner (2002), which establishes conceptual blending as a quintessentially human cognitive process, fail to define the term *concept*. It is apparently taken for granted that readers know *a priori* what concepts are. At the other extreme is Fodor (1998), who claims that concepts have been misunderstood and misrepresented in the cognitive sciences, thus invalidating most research in these fields.

It is not the aim of this dissertation to contribute to what may be an essentially philosophical debate. It should, however, be noted that not all concepts are equivalent. According to **prototype theory** (Rosch, 1978; Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976), concepts are structured in categories that are more or less **inclusive**. Inclusiveness refers to the level of detail of a category, such that the category *animal* is less detailed than *dog*, which itself contains less detail than *chihuahua*. Indeed, someone who has never before encountered a particular breed of dog cannot include it as part of the *dog* category; thus, dog breeders and veterinarians will have a rather different conceptualization of dogs than people who take no interest in them. Moreover, the more abstract the category, the more likely it is to have **fuzzy boundaries** (Evans & Green, 2006). For example, consider again the category *animal*. Depending on their religious beliefs, some would be more likely than others to assign *human beings* to this category. Thus, concepts are also culture-specific, developing out of such cultural phenomena as religion,

traditions, and stereotypes[28]. As can be seen, even among monolingual English speakers conceptual structure is far from universal: it is, in fact, essentially personal, arising organically from a given individual's experiences, interests, and memories.

Because conceptual blending will be claimed to be the mechanism underlying transfer (in Part 4), some discussion of the structure and organization of the conceptual base is required. This section now turns to two theories about concepts that are compatible with the issues of multilingual speech production: the common underlying conceptual base of Kecskes & Papp (2000) and the theory of lexical concepts and cognitive models of Evans (2006).

1.2.2.1 The common underlying conceptual base

For monolinguals, there is no essential difference between concepts and words, because concepts can only be expressed in the words of a single language, so no alternative possibilities can be imagined or deemed necessary[29]. For multilinguals, however, lexicalizing concepts is often easier in one language or the other. Most multilinguals have found themselves at times unable to express in one language what sounds natural and proper in another. Meanings are "lost in translation" or, worse, "untranslatable." Students of English as a Foreign Language, for whom many L1 concepts are simply unavailable as English words, are bound to have difficulties with vocabulary, to say nothing of applying morphosyntactic structure to the words they do have.

According to Kecskes & Papp (2000), the **common underlying conceptual base (CUCB)** is a unique feature of multilinguals; monolinguals do not develop a CUCB and have no need for one. The CUCB is itself a dynamic system apart from the language systems of multilinguals, and its

[28] Lakoff (1987) has argued that social stereotypes are less personal, emerging from public discussion. Yet such discussion is itself inextricably tied to the culture in which it takes place.
[29] This is not entirely true. Neologisms are obviously created to lexicalize concepts that have no set form of expression.

nature and composition are constantly changing in reaction to such variables as the speaker's language learning history, current environment and communicative needs, and opinion as to which language(s) are prestige. Multilinguals develop a CUCB after reaching **threshold exposure** to a foreign language, and the issue of when such a threshold can be said to have been reached is discussed at length in Kecskes & Papp (2000). Although no definitive answer is given, the authors discuss the results of an experiment indicating that exposure to a foreign language (FL) of 4 hours per week or less cannot be expected to lead to the formation of a CUCB; in other words, only fairly intensive exposure – more than is typical for EFL classrooms – can cause students to develop a CUCB. Thus, the authors draw a distinction between the terms **FL (foreign language**: one which is studied primarily in an educational setting removed from its natural sociocultural context) and **L2 (second language**: one which is acquired within its natural sociocultural context, with or without the benefit of language courses), and suggest that only the latter situation can cause a CUCB to develop. As they write,

> L2 language production is *heavily influenced by the L1-dominated conceptual base* until the language learner reaches an advanced acculturation threshold. [This] depends not only on the development of L2 language proficiency but also the willingness of the speaker to acquire the new sociocultural frameworks (Kecskes & Papp, 2000: 108; emphasis added).

The CUCB is described as a container of concepts and knowledge. A very small number of concepts are thought to be language-neutral and, perhaps, universal; these are the same "close to 60" concepts referred to as semantic primes in the work of, for example, Wierzbicka (1998: 114) and Goddard (2004). The majority of concepts in the CUCB, however, are language- and culture-specific, a point acknowledged by Wierzbicka (1998) as well. This is because concepts enter the CUCB through one or the other language channel. Many concepts are unique to a given culture. For example, the English concept BASEBALL has no Lithuanian equivalent, and has been rendered as BEISBOLAS through borrowing. The Lithuanian concept GIRA,

conversely, has no English equivalent. To express this concept in English one must describe it using a long phrase (e.g., "a fizzy drink made of fermented rye bread and sugar").

According to Kecskes & Papp (2000), new information entering the CUCB through one language channel must be **neutralized** before becoming available through other channels. A neutralized concept is one which can be expressed equally well in more than one language. Neutralization is dependent on the nature of the concept and environmental pressure to neutralize it. An English concept such as MONEY[30] can be neutralized by a Lithuanian student simply by learning the Lithuanian equivalent PINIGAI. Complex and culture-specific concepts, however, such as EN YUPPIE or LT ŠIMTADIENIS, may require a great deal of cognitive effort to neutralize, effort that will not be spent unless absolutely necessary. Again, it is important to stress that the CUCB is a dynamic system whose features change in response to environmental variables: a Lithuanian-English bilingual might know the English concept YUPPIE but have no need to ever refer to it in Lithuanian; or he might interact with such people daily, in which case he might decide to neutralize it by thinking of an equivalent phrase (cf. Tildė, 2006: "*praturtėjęs ir mėgstantis puikuotis prabanga jaunuolis*").

Kecskes & Papp (2000) provide no illustration of the CUCB. Figure 18 is therefore an attempt to model the CUCB in two stages of development. In an early stage of acquisition, the CUCB (and, therefore, the conceptualizer as well) is dominated by L1-specific concepts. A small number of L2-specific concepts may have been acquired, and some concepts will have been neutralized as well. The dotted lines are meant to show that the region of neutralized concepts is not in some way separate from the rest of the CUCB; concepts can flow into it from both sides as required. In a stage of high

[30] Throughout this dissertation, concepts will be marked in small capitals, in boxes. This convention is meant to distinguish the *concept* of MONEY from the English *word* MONEY, from its lemma [MONEY], and from its lexeme "money," all of which represent only one possible lexicalization of the concept (other possibilities include, e.g., DOUGH, MOOLAH, BREAD, etc.).

bilingual proficiency, the CUCB is dominated by neutralized concepts which the speaker can express in either language with relative ease or fluency; however, some amount of language- and culture-specific concepts will always remain unique to the L1 and the L2.

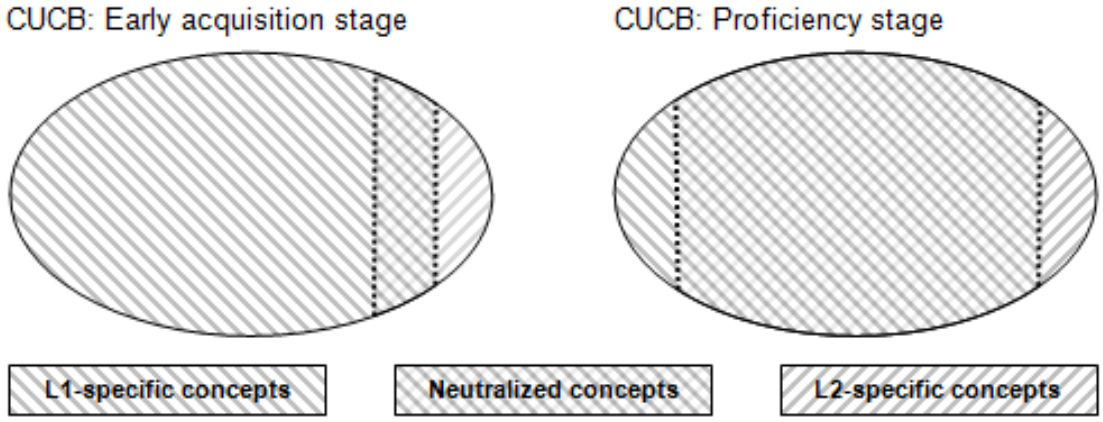


Figure 18. Development of a bilingual CUCB.

For multilinguals, then, an additional level of "personalization" of concepts must be considered. Concepts literally change through neutralization. A monolingual English speaker's concept of RED clearly differs from a monolingual Lithuanian speaker's concept of RAUDONAS, but both of these differ from the neutralized RED/RAUDONAS concept of a bilingual English-Lithuanian speaker.

1.2.2.2 Lexical concepts and cognitive models

Evans' (2006) theory of **lexical concepts** and **cognitive models** is primarily a theory of meaning construction (therefore, of comprehension). However, the process as described can presumably work in reverse for the purpose of speech production. This theory, developed by a British cognitive linguist with no reference to the psycholinguistic ideas discussed above, is also a theory of monolingualism. How this theory may apply to multilinguals will therefore be discussed, as well.

According to Evans, the idea that words have **sense-units** independent of their context in **situated language use** is a fallacy: "The precise semantic contribution of any word is a function of the utterance context in which it is

embedded and, moreover, the sorts of (conceptual) knowledge these lexical entities provide access to…. In other words, words don't have 'meanings' in and of themselves" (Evans, 2006: 492).

The theory, as its name implies, rests primarily on the explication of lexical concepts and cognitive models. Lexical concepts are defined as "stored linguistic knowledge units" (Evans, 2006: 496). These may be words, bound morphemes, idiomatic phrases, and even **implicit forms** such as grammatical constructions. It is particularly important that lexical concepts are both language- and culture-specific (Evans, 2006: 502, 509). Moreover, while lexical concepts are inherently form-specific (that is, are tied to a specific lexical form), forms are not lexical concept-specific. In other words, the same form can activate different lexical concepts[31] (e.g., "bread" as noted above, may access both [FOOD] and [MONEY]). Thus, lexical concepts are roughly equivalent to lemmas.

Cognitive models are defined as "conceptual knowledge structures which constitute the semantic potential that lexical concepts provide access to" (Evans, 2006: 496). These structures are presumed to be non-linguistic in nature, and thus are roughly equivalent to concepts. Cognitive models consist of **facets**[32] and the **relations** that hold between them. As such, they are **holistic** in the sense that they are richer and more detailed than the sum of the lexical concepts they are accessed by, just as concepts are more than the words (lemma/lexeme pairs) that name them. Finally, cognitive models are dynamic, being constantly updated via ongoing experience, yet forming temporarily stable cognitive structures (cf. attractor states). Figure 19, adapted from Evans (2006: 520) shows these different levels for the lexical item BOOK. The arrow linking the cognitive models (or concepts) BOOK and READING represents the relation that holds between them (namely, that books are read by *readers*).

[31] This is, of course, another way of saying that forms may be polysemous.

[32] Facets are also called secondary cognitive models (Evans, 2006: 513), especially when they themselves link to "deeper" facets.

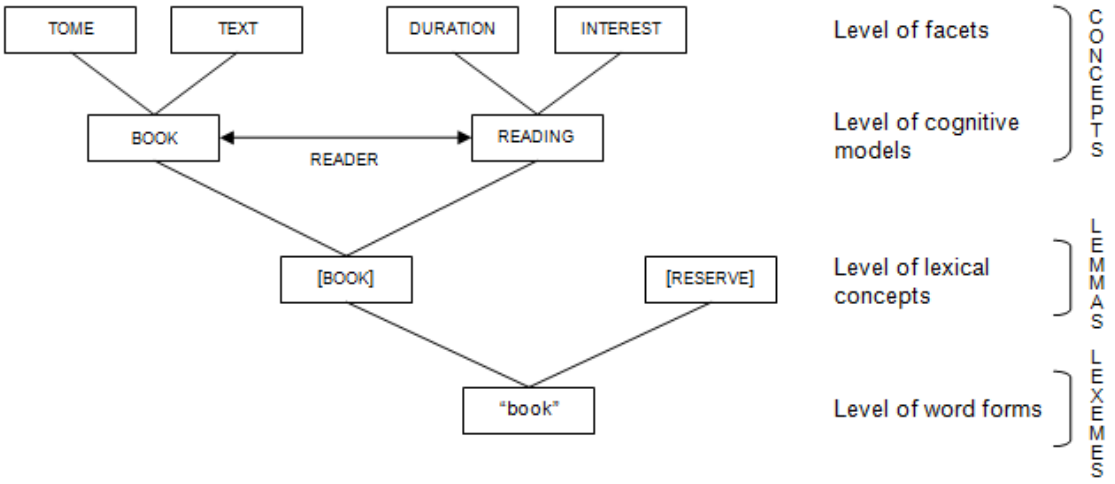


Figure 19. Facets, cognitive models, lexical concepts, and word forms.

In terms of meaning-construction, Evans (2006) proposes that lexical concepts establish **access routes** through the system of cognitive models. This explains their dependence on situated language use. Looking at the (simplified) structure shown in Figure 19, consider the following examples (adapted from Evans, 2006):

(a) This book is interesting.
(b) This book is 300 pages long.

In (a), the form (lexeme) "book" accesses the lexical concept (lemma) [BOOK], which, in turn, accesses the cognitive model (concept) READING and, through it, the facet INTEREST. In (b), on the other hand, the same form and lexical concept access a different cognitive model, BOOK, which picks out the TOME facet.

Other than mentioning that lexical concepts are language-specific (see above), Evans (2006) does not specifically discuss the multilingual situation. He does, however, note that meaning-construction is essentially cooperative: speakers encode meaning which listeners must infer from the situated forms (both overt and implicit) produced. A multilingual speaker who uses forms and constructions that are unfamiliar to his or her listener therefore runs the risk of being misunderstood. Such a situation can easily arise when multilinguals try to neutralize language-specific L1 concepts into an FL. If the result is nonstandard, listeners unfamiliar with the speaker's L1 will either assume that

the speaker has spoken in error, or that he or she is being intentionally difficult or even rude.

1.3 Summary and Discussion

This Part examined the production of speech by both mono- and multilinguals. Levelt's (1989) model of online monolingual speech production was introduced. In this model, intentions arise in a conceptualizer which creates preverbal messages through the processes of macro- and microplanning. These messages are sent to a formulator for grammatical and phonological encoding, processes which work closely with the lexicon. Grammatical encoding works with lemmas to create a surface structure, which is then phonologically encoded with lexeme information to create a phonetic plan. This plan is sent to the articulator for production as overt speech. A monitor (located in the conceptualizer) can check the phonetic plan prior to articulation, as inner speech, or (through the speech-comprehension system) check the spoken output and interrupt it if necessary. Interruption usually takes the form of an editing expression followed by a repair of the erroneous element or phrase.

It was noted throughout this section that the model as applied to multilinguals is underdeveloped. Thus, the second section turned to a review of what is known about multilingual speech production. Of the many models of multilingualism, four were chosen to discuss: the revised hierarchical model (Kroll & Stewart, 1994), the inhibitory control model (Green, 1986), the competition model (MacWhinney, 2005), and the dynamic model of multilingualism (Herdina & Jessner, 2002). Each highlight slightly different aspects of the issues facing multilinguals. The section then turned to a closer examination of concepts, utilizing the insights of the common underlying conceptual base (Kecskes & Papp, 2000) and the theory of lexical concepts and cognitive models (Evans, 2006). It was seen that for multilinguals, the conceptual base is a dynamic, constantly changing system, and that the lexicalization of concepts is far from straightforward.

The online production of speech is an extremely complex, interacting set of processes that must be performed at very high speeds. According to de Bot (1992) and Dąbrowska (2004), the average rate of informal conversation is 150 words per minute, with turns typically starting within 500 milliseconds of the end of the prior turn (and often actually overlapping). This means that speakers must plan their turns while still listening to their interlocutors, surely a drain on attentional resources. The resources of multilingual speakers are further drained by the need to inhibit currently unneeded language systems, to choose among lexical items, to grammatically and phonologically encode less resonant structures and phonemes, and to control articulatory organs in less familiar ways[33]. That some amount of error would occur in such a remarkably intricate process is only to be expected.

The "errors" that multilinguals produce, however, are different from those produced by monolinguals, especially when their knowledge of the foreign language(s) is mostly theoretical (i.e., has developed in classroom environments rather than in a naturalistic L2 setting). The demands of high-speed processing and the language- and culture-specificity of the concepts to be expressed are likely to result in an increased probability of abnormality as compared to a monolingual native-speaker standard. These issues will be developed in Part 2, where it is shown that sociolinguistic and ethnopragmatic factors also have a role to play in language acquisition and production.

[33] However, Bialystok (2005) has argued that multilinguals are better at inhibition than monolinguals, having in a sense trained themselves at it.

2. SOCIOLINGUISTIC ASPECTS OF THE ACQUISITION OF ENGLISH

Speech production, as described in Part 1, is an inherently personal process. Until a speaker articulates his or her utterance, there is no way for an interlocutor to know what it might be. Many of the stages of speech production described in Levelt's (1989) model have been inferred mostly from speech error data, but are not available for direct study[34]. Moreover, the concepts with which speakers conceptualize are personal as well, being developed out of personal experience and culture. However, there is good reason to believe that speech – and indeed, language itself – is also a social process that cannot be properly understood without some reference to the sociocultural context in which conversation takes place.

Section 2.1 discusses the interface of society, culture, and language and its effects on both overt (or outer) and inner speech. Section 2.2 describes a sociocultural language use survey given to the students whose speech and writing will be examined in Parts 3 and 4. The results are discussed in section 2.3.

2.1 Society, Culture, Language

According to Kasper & Blum-Kulka (1993: 3), **interlanguage pragmatics** research focuses on "the study of non-native speakers' use and acquisition of linguistic action patterns in a second language." In particular, research in this area has looked at what have been called **pragmatic failures** (Cenoz, 2003): not speech errors *per se*, for in these uses of language all lexical and grammatical elements are more or less correct, but nevertheless instances of deviation from native-speaker norms. A non-native speaker (NNS) who says, to a native English speaker, *Tell me the time*, would be considered to have violated some native speaker (NS) norm of politeness (cf. Searle, 1975), tact (cf. Leech, 1983) or face (cf. Brown & Levinson, 1978). This is because

[34] Articulation, of course, can be directly observed, as can some forms of monitoring. In addition, researchers such as Dechert (1987) and Herwig (2001) use an interesting technique of observing lexical selection during processing via think-aloud protocols.

English **contextualization cues** (Gumperz, 1982) require more indirect requests like *Can you tell me the time?* (Cenoz, 2003; Thornburg & Panther, 1997; Clark & Schunk, 1980). In other words, because *Tell me the time* is not erroneous, an English NS is likely to assume that the NNS said it intentionally, and will therefore interpret it as if it had been produced by another English NS. Thus, NNSs can unintentionally give the impression of rudeness simply by following conversational norms (and/or utilizing contextualization cues) that are entirely appropriate in their own sociocultural environments.

The **ethnopragmatic** research of Wierzbicka (2006) and Goddard (2004, 2006) has developed the idea of **cultural scripts** (Goddard & Wierzbicka, 2004) to confront what they see as an Anglocentric trend in interlanguage pragmatics. This trend is best described by Wierzbicka (1985: 145): "[researchers] take it for granted that what seems to hold for the speakers of English must hold for 'people generally.'" A cultural script is defined as "a statement – framed largely or entirely within the non-ethnocentric metalanguage of semantic primes[35] – of some particular attitude, evaluation, or assumption which is hypothesized to be widely known and shared among people of a given speech community" (Goddard, 2006: 5). In ethnopragmatics, cultural scripts are used to highlight the cultural specificity of the concepts, attitudes, and values underlying preferred speech patterns such as those discussed above (Goddard, 2004).

A **speech community** is defined (in Hymes, 1986: 36) as "a community sharing rules for the conduct and interpretation of speech, and rules for the interpretation of at least one linguistic variety." In other words, a speech community shares norms for **communication**. This logically implies that in different speech communities there are different rules of communication. These rules are not personal, but social. As Karaliūnas (1997: 95) writes, "Komunikacija yra tokia interakcija, kurios struktūra nepriklauso nuo atskirų

[35] For these, see Wierzbicka (1998).

jos dalyvių elgsenos. <...> Individas neišranda savo komunikacijos sistemos: atėjęs į visuomenę, jis gauna ją gatavą ir funkcionuojančią.[36]"

This social language can be called **outer speech**, and it can be opposed to **inner speech**, but here it must be noted that this inner speech is not that described by Levelt (1989). In his model, "inner speech" refers to a prepared utterance which has not yet been articulated: in other words, a phonetic plan which may be monitored for errors. However, a very different interpretation of inner speech exists in a tradition going back to Vygotsky, a tradition which takes the very process of learning to speak to be a social phenomenon (Ehrich, 2006). In this tradition, inner speech refers essentially to the process of thinking itself, or perhaps, in Levelt's terms, to conceptualizing. According to Vygotsky (1986), children develop inner speech in stages. First there is social speech, adult-directed and used to achieve specific purposes. Later comes egocentric speech, as when children "think out loud." Finally this turns into inner speech. Inner speech is qualitatively different from outer speech (Vygotsky, 1986; Ehrich, 2006; Karaliūnas, 1997). According to Tomlinson & Avila (2007: 66-68), among the characteristics of inner speech[37] are the following: it is fast, elliptical, implicit, vague, partial, simple, novel, relevant, and idiosyncratic. It would be incoherent to anyone else, but to the "inner speaker" it makes perfect sense.

A child's development, according to Vygotsky (1978: 57), is heavily dependent on the sociocultural environment, as "[e]very function in the child's cultural development appears twice: first, on the social level, and later, on the individual level." Thus, "Vygotsky saw the child's development as being the internalization of outside influences or social co-ordinations" (Perret-Clermont, Carugati & Oates, 2004: 310). It can therefore be said that not only speech and language, but the very process of thought itself – indeed, a person's whole concept of self – is at the most basic level both culture- and language-specific.

[36] "Communication is a kind of interaction whose structure does not depend on the behavior of its participants. <…> An individual does not invent his communicative system: he receives it upon entering society, complete and already functioning." (aut. trans.)
[37] Also referred to as the "inner voice."

Lakoff & Johnson (1980: 57) write: "[A]ll experience is cultural through and through… we experience our 'world' in such a way that our culture is already present in the very experience itself." Similarly, Karaliūnas (1997: 329) writes: "Ir apskritai mūsų mokami ir vartojami žodžiai yra kultūros, kurios dalis mes patys esame, bei viso mūsų gyvenimo išraiška. <...> [I]ndividas internalizuoja visuomenės kultūrą ir tampa kalbinės bendruomenės nariu[38]." Kecskes & Papp (2000: 107) agree: "Culture… serves as a catalyst for cognitive growth. It is therefore crucial to focus on the cultural environment – with its values, beliefs, world views, and presuppositions – in which multilingual development occurs…."

Language learning, however, is not typically equated with culture learning. Although some language teachers make efforts to introduce their students to, e.g., "French culture," this is often in the form of reading French literature or eating French food[39]. Such experiences, however, are not likely to impact students' inner voices. Only long-term personal exposure to the sociocultural environment in which a given language is spoken could begin to lead to the development of what could be called multicultural multilingualism[40], though even this is in no way guaranteed and depends in large part on the willingness of the speaker to, so to speak, "make room for" the L2 culture. As Adamson (1988) notes, NNSs who live in an L2 environment often remain unwilling to accept L2 values and beliefs – L2 culture – regardless of length of residence or level of L2 proficiency. Similarly, Acton & Walker de Felix (1986) found that L2 production is usually based on the L1 socioculture until an advanced **acculturation stage** is reached. Kecskes & Papp (2000: 108) comment that reaching this stage "depends not only on the

[38] "And in general the words we know and use are an expression of both the culture of which we ourselves are a part, and of our lives. <…> An individual internalizes the culture of society [the socioculture] and becomes a member of the speech community." (aut. trans.)
[39] Personal observation; but see also von Knorring (2007) on French perceptions of British culture and vice versa, and Scovel (2001: 28-29), which describes one teacher's attempt to "give her [Japanese] students a genuine [American] acculturation experience" by having them eat apples with peanut butter.
[40] This is similar to the idea of multicompetence (Cook, 1991, 2003).

development of L2 language proficiency but also the willingness of the speaker to acquire the new sociocultural frameworks and make them a functional part of [the] CUCB."

Like language, then, culture is also both social and personal. It is after all quite possible that a student of English may not want to reach a high level of fluency or to "erase" his or her accent; such steps perhaps seeming to be a betrayal of his or her native culture. Thus, in addition to transfer of phonology, grammar, and syntax, it should come as no surprise to find transfer of pragmatic, socially- or culturally-based patterns as well (Cenoz, 2003; Fouser, 2001). Even when such transfer is unintentional, it is also very difficult for the monitor to catch: recall that monitoring is both context-sensitive and conceptual in nature. The likelihood that the kinds of pragmatic failures discussed above will occur, therefore, is especially high for students learning languages in FL environments. Under such conditions, there are no real possibilities to internalize the cultural standards of the FL.

The Lithuanian and Russian students of EFL whose work is examined in this dissertation come from sociocultural backgrounds very different from those of native English speakers. The CUCB that such a speaker brings to the English acquisition process is therefore dominated by Lithuanian- and Russian-specific linguistic and conceptual structures (cognitive models) that are likely to affect every stage of English speech production, from conceptualization to articulation and monitoring. The CLI data and learner constructions analyzed in Parts 3 and 4 are not to be understood as "errors," however. Following the ethnopragmatic approach, English native-speaker norms are not taken to be prescriptive or even applicable to these students. Such norms are only of interest insofar as they can affect how certain exemplars of English produced by these students might be interpreted by speakers with native English sociocultural backgrounds.

The students examined in this dissertation are, for the most part, Lithuanian-Russian bilinguals[41] with a long history of EFL learning. Because most of them have only academic experience with English, however, their English language systems are still developing. Most of their experience with English comes from classrooms; they are, therefore, highly dependent on L1 and/or L2 when speaking English. According to Kecskes & Papp (2000: 117), "Until multicompetence occurs, a typical language learner will *think in the L1*, following previously established patterns, norms, and sequences of activation" (emphasis added). As just discussed, this suggests that these students' inner voices are most likely to speak Lithuanian or Russian, not English, regardless of which language they are using for outer speech.

In terms of traditional EFL teaching, most of them place as Intermediate to Upper Intermediate level students on tests of English grammar. However, the English channel of their language systems is underdeveloped as compared to the Lithuanian and Russian channels. At the same time, through their participation in the university's study program, they are constantly receiving new input through all three language channels[42]. These students are therefore engaged in a great deal of conceptual development and restructuring at any given time. One consequence of such mental activity may be a high proportion of CLI in production, as students adapt their language systems to the constant influx of new information, the shifting communicative needs of the academic environment, exposure to new sociocultural information, and the necessity to neutralize both familiar and novel concepts.

2.2 Sociocultural Language Use Survey

This section presents the results of a sociocultural language use survey of the students of the English and Russian Languages study program (hereafter referred to as AnRK, after its university code) at the Vilnius University Kaunas Faculty of Humanities. For the reasons outlined in section 2.1, it was

[41] Or Russian-Lithuanian.

[42] Lithuanian, Russian, and English are all used as languages of instruction (in different subjects, usually, though some language mixing may occur within subjects and even lessons).

considered essential to establish a preliminary picture of their linguistic background and language systems. 51 students returned the questionnaire: 23 first-year students (out of 24 = 96%), 13 second-year students (out of 14 = 93%), 10 third-year students (out of 17 = 59%), and 5 fourth-year students (out of 18 = 28%). Because of the poor response rate of the fourth-year students, responses were gathered from only 70% (51 out of 73) of all of the students in the program. However, as fourth-year students are not included in the sample studied in this dissertation, 84% of all potential subjects responded. The questionnaire is reproduced in Appendix 1.

2.2.1 Demographic data

Within Lithuania as a whole, Russians form only 6.31% of the population (Statistics Lithuania, 2008). Slightly outnumbered by Poles (6.74%), they are the third largest ethnic group in the country, with Lithuanians forming the vast majority (83.45%). This proportion is not maintained within the AnRK program, however. According to the results of Question 10, which asked them to self-report their nationality, only 53% of the students are Lithuanian. Fully 31% are Russian, with Poles taking a distant third place (10%). The remaining 6% are Ukrainian and Armenian. These differences are illustrated in Figure 20.

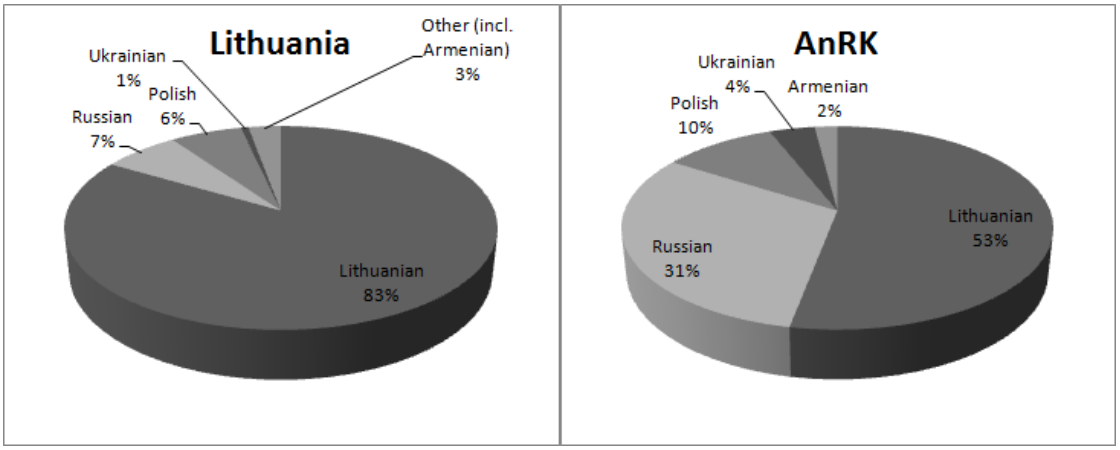


Figure 20. Ethnic distribution in Lithuania and in the AnRK study program.

Similar results are to be found in regards to the students' native language (L1). 49% of the group speak Lithuanian as their native language, 35% speak Russian, 6% speak Polish, and the remaining 10% are bilingual from birth:

four students speak both Russian and Lithuanian, and the other speaks Ukrainian and Bulgarian. Interestingly, of the four Lithuanian/Russian native speakers, three list their nationality as Lithuanian, and only one as Russian. It may be recalled that Hogan-Brun & Ramonienė (2005) also found a lack of congruence between nationality and native language.

Almost all of the students were born in Lithuania (90%), and the remainder were born in either Russia, the Ukraine, or Turkmenistan. Students' ages are generally what is to be expected, with 96% falling between 18 and 23 years of age. Two students stand out as considerably older than the norm: a 28-year-old first-year student, and a 31-year-old third-year student.

2.2.2 Language history data

Students' familiarity with the languages in question (English, Russian, and Lithuanian) was also investigated in the survey. 67% have been studying Lithuanian since the first grade in school, and all but one respondent began learning it in one of the primary school grades. English is most commonly introduced in the fourth grade (41%), but many schools begin English instruction earlier, such that even 78% of all respondents began to study English in one of the primary school grades, alongside Lithuanian. The story with Russian is slightly different because many students come to this study program from one of the Russian schools (schools where Russian is the primary language of instruction[43]). Thus, 31% began learning Russian in the first grade, while another 37% began learning it in the sixth grade. The remaining 32% are rather randomly distributed among grades 2, 5, 7, 9, 11, and even 12.

Taking the majority scores as a basis for describing the "average subject" of this study, they would lead to the following conclusions: such a student has been formally studying Lithuanian for at least 8 years, English for at least 8

[43] According to Hogan-Brun & Ramonienė (2004), there were 128 such schools in Lithuania in the 2000–2001 school year. In Kaunas, where this survey took place, at present there is only one Russian school remaining (Čubajevaitė, 2009).

years, and Russian for at least 6 years, and is a native speaker of either Lithuanian (49%), Russian (35%), or both (8%; total, 92%).

Questions 2 and 3 of the questionnaire were intended to help determine which languages are interesting to AnRK students. Question 2 asked them to list the languages they could "speak or write in," a purposely vague formulation designed to elicit as many languages as possible. Similarly, Question 3 asked for even more languages, those of which students might know only "some words or phrases." Table 2 collects the responses to these two questions and totals them, to find a "total interest" value for each language. Looking at this column of the table, it can be seen that the table clearly subdivides into three groups: a) the languages of the study program, English, Russian, and Lithuanian, with 100% student interest; b) languages with moderate student interest (11–25 students, or 22–49% interest); and c) those with only weak or sporadic student interest (≤6 students, or ≤12% interest). The list in group b) can be further reduced by excluding Latin, which many students added because a course in elementary Latin is included as part of the requirements of the AnRK study program. This leaves a rather short list: Spanish, German, Polish, Italian, and French. Three of these languages are related Romance languages, and, with the exception of Polish[44], all are major languages of Western Europe.

Students were also asked to name their "favorite" language, and to provide a reason for their choice. Leading the list of answers by a large margin is Russian, with 41% of the students giving it their vote. At some distance behind is English, with 29% of the vote. The remaining 30% of students are distributed among a variety of answers. They are: Russian and English (3 students), Polish (2), Italian (2), and, with one vote each, Spanish, Arabic, French, German/French, Hebrew, and Bulgarian. Significantly, only two students (4%) claim Lithuanian as their favorite language. The subordination

[44] The presence of Polish in this list is unsurprising, given the close historical ties between Poland and Lithuania and the typological similarity between Polish and Russian. Many Russians who "don't speak Polish" claim to be able to understand it anyway, simply through their knowledge of Russian.

of the Lithuanian language and culture to that of Russian and English is a general trend among AnRK students which will be discussed in detail below.

Table 2. Language interests of AnRK students.

Q2: "speak or write"	Q2 Total	Q3: "words or phrases"	Q3 Total	Total
English	51			**51**
Russian	51			**51**
Lithuanian	51			**51**
Spanish	2	Spanish	23	**25**
German	2	German	23	**25**
Polish	11	Polish	12	**23**
Italian	1	Italian	17	**18**
		Latin	15	**15**
		French	11	**11**
		Greek	6	**6**
Ukrainian	3	Ukrainian	1	**4**
		Esperanto	4	**4**
Armenian	1	Armenian	1	**2**
		Latvian	2	**2**
		Arabic	2	**2**
		Turkish	2	**2**
		Hebrew	2	**2**
Belarussian	1			**1**
Bulgarian	1			**1**
		Norwegian	1	**1**
		Japanese	1	**1**
		Dutch	1	**1**
		Turkmenian	1	**1**
		Moldavian	1	**1**

As mentioned, many (though not all) students also provided reasons for their choices. These reasons are collected in Table 3, where answers have been edited and grouped together. Many of the reasons given for preferring Russian are related to its culture, literature, and the richness of its vocabulary. This is in contrast to the reasons given for preferring English, which often focus on its status as a world language. For both languages, "easiness" is preferential, though how students define this quality is unclear. Similarly, many students prefer a language for being "beautiful" or "nice," also rather vaguely defined adjectives. Two students prefer English because it is "popular." Whether this means that it is spoken by many people in the world, or that it is a popular choice among school/university students, however, is unknown.

Table 3. Students' reasons for "favorite" language status (number of students).

Russian	it's easier (7), it's my native language (5), it's beautiful (3), I like Russian culture (3), I can express myself (3), good literature (2), it's a "rich" language (2), good vocabulary (2), I know it better (1), it's interesting to learn (1), I like to speak it (1), I'm fond of its grammar, lexis and literature (1)
English	it sounds nice/beautiful (7), it's easier (4), all the world knows it (2), it's "popular" (2), I understand it (2), it's interesting to learn (2), I like it (1), I love how you can express your thoughts (1), I know it better (1), you can use it everywhere (1)
Lithuanian	I speak it fluently and don't look for words (1), I like everything related to Lithuania (1)
Other languages	*Polish*: it's my native language (1), it's funny (1) *Italian*: it sounds nice/beautiful (2), I like Italian culture (1) *Spanish*: I only listen to Spanish songs (1) *Arabic*: it's hard to learn and very interesting (1) *Hebrew*: it's similar to Russian (1) *German and French*: they sound sexy (1) *Bulgarian*: I like everything about it (1)

According to Kellerman (1983), one of the most significant factors affecting L1→L2 influence in a learner is the learner's **perceived distance** (rather than the actual typological distance) between the two languages. In other words, if a learner perceives a language as very different from his own, he or she will be more likely to have difficulties learning it than one who perceives it as similar. Question 5 attempted to address this issue by asking students to decide which pair of languages (LT/RU, RU/EN, or LT/EN) were "the most similar." The results were strongly in favor of similarity between Lithuanian and Russian, with 63% of the students choosing this pair. A small number of students (16%) felt that Lithuanian and English were most similar, while only 6% chose Russian and English. Three students didn't respond. AnRK students, then, can on the whole be expected to do better learning Russian than English. Whether this theoretically defined tendency is reflected in real terms, e.g. course marks, has yet to be investigated. Moreover, it could be hypothesized that AnRK students should, on average, experience the most pronounced CLI when they are speaking English, while neither Russian nor Lithuanian native speakers should have much difficulty inhibiting CLI when speaking, respectively, Lithuanian and Russian.

Questions 15 and 16 asked students to indicate which language was the "best" for them in school, and which is "best" now. The adjective was purposely left undefined and placed in quotation marks on the questionnaire, to allow students room to interpret it in whatever way they felt appropriate: best marks, or favorite, or most interesting, etc. The results for the "in school" condition were as follows: Russian, 47%; English, 27%; Lithuanian, 22%; Polish and German, 2% each. For the "now" condition, the results were: Russian, 39%; English, 31%; Lithuanian, 27%; Polish, 2%; one student did not respond. The fact that Russian would drop in relation to English could have been expected; AnRK students consistently complain that their Russian courses are "very hard" compared with their English courses. It is interesting, however, that more students report Lithuanian as their best language now, too. In school, Lithuanian was also taught as a language course, whereas now it is only a language of instruction for non-specialist subjects such as psychology or philosophy. This may have caused some students to acquire a new appreciation for it as a language of communication.

The questions grouped together under number 6 in the questionnaire address the wider sociocultural background against which the students' specialized instruction takes place. Regardless of the language of instruction in classrooms, the fact is that AnRK students spend vastly more time outside the classroom, immersed in the local environment. As discussed in section 2.1, however, students' immersion in the local socioculture is a key factor in success of acquisition. Thus, Question 6 examined students' preferential use of language in various social (non-academic) situations. These questions were also designed to examine those situations where the use of language would be more productive, as opposed to receptive.

It was thought that the most common answer to these questions would be Lithuanian, because of the demographic data discussed in section 2.2.1, except perhaps for question 6.4 (about the Internet), which would probably be English. The results are collated in Table 4. Here, the results are expressed as

the total number of students who gave each answer, rather than in percentiles, as not every student answered every question[45]. As can be seen, the hypothesis in regards to these related questions was correct: Lithuanian is indeed the most common answer for every question except 6.4. In fact, the table slightly discredits Lithuanian's prominence in these areas, as many of the answers grouped in the "Other" slots are of the "English and Lithuanian" or "all three" variety, where Lithuanian is being used together with some other language. These results show that, no matter what a student's native language may be, he or she is immersed, in a wide variety of situations, in a context that is predominantly Lithuanian/Russian. The overwhelming presence of these two languages in the answers to these questions means that they are highly resonant. It therefore stands to reason that English, for all of these students, should be the most difficult language to feel fluent in, as it is rarely used outside the classroom or the Internet.

Question 7 was also related to Question 6, in that it addresses the sociocultural issue, but its focus is more on personal pastimes and receptive uses of language. Here the answers showed more variation, including several answers along the lines of "it doesn't matter," "it depends on my mood," and even, "all the languages of our world." In response to question 7, 21 students prefer reading only in Lithuanian, and 12 only in Russian. The remaining 18 respondents give various mixed-language answers with no clear preference for any one language. Looking, however, at the total number of students who read in a given language, these numbers appear: Lithuanian, 30; Russian, 22; English, 16; Polish, 2. To be clear, in these combined results, a student who answers "Lithuanian and Russian" is counted twice, once for the Lithuanian score and once for the Russian score. In response to 7.1, 23 students prefer listening to music only in English, while only two prefer exclusively Russian music. Not a single student listed Lithuanian music as the only choice. Combining the results as before, we find the following numbers of students

[45] For example, question 6.3, about jobs, was answered only by the 25 students who work.

who listen to music in a given language: English, 46; Russian, 22; Lithuanian, 8; other languages, 6. Finally, question 7.2 asked about watching films or TV. 11 students prefer watching films only in Russian, and 8 only in English. The combined results are: Russian, 38; English, 31; Lithuanian, 19. One student answered that she watches films in their "original language," but added that she uses Lithuanian subtitles. For this reason her answer was added to the Lithuanian total.

Table 4. Sociocultural language use.

Q. No.	Summary	Responses		Q. No.	Summary	Responses	
6	language used at home	Lithuanian Russian LT/RU Polish Other	25 19 3 3 1	6.4	language used on the Internet	English Lithuanian EN/LT Russian Other	18 11 10 5 7
6.1	language used with relatives	Lithuanian Russian LT/RU Polish Other	22 17 7 3 2	6.5	language used on summer holidays	Lithuanian Russian English LT/RU Other	20 17 5 3 5
6.2	language used with friends	Lithuanian Russian LT/RU Other	22 10 15 4	6.6	language "used" when thinking	Lithuanian Russian Polish English	29 19 3 2
6.3	language used at work	Lithuanian Russian English Other	20 1 1 3	6.7	language "used" when dreaming	Lithuanian Russian Polish Other	28 15 3 3

In terms of students' proficiency in English, at least, these answers may seem to be heartening: 31% read, 90% listen to music, and 61% watch films in English. However, it should be remembered that these are receptive skills. Most students listen to music while doing other things, such as cooking or preparing Russian grammar homework, so the overall benefit of this resonance-enhancing activity is likely to be quite small. In addition, several students have admitted that the only reading they do in English is what they are required to do for homework. Many of them do such reading both sporadically and under less-than-ideal conditions (e.g., reading 100 pages of a book the night before a quiz, falling asleep in between chapters, and not reading any

more until the next quiz[46]). They are thus unlikely to recall or retain much of what they have read even the next morning, to say nothing of any long-term benefits.

Question 17 was written bearing in mind the results of research done on the RHM (described in section 1.2) for translation latencies. Theoretically, L1→L2 translation should be more difficult and take longer than L2→L1. However, because the AnRK students are all multilingual, the situation is more complicated. It was expected that students would generally prefer translating from English into either Lithuanian or Russian, as these are the majority L1s, and that translation from either of these languages into English would be seen as more difficult. Moreover, L2↔L3 translation should be seen as the most difficult; more difficult than either L1↔L2 or L1↔L3. The results of the survey (in total number of students) are presented in Figure 21, where it can be seen that this is indeed the case. RU→EN translation is the least preferred, and RU→LT is most preferred. This is to be expected, as Russian and English are L2 and L3 for the majority of the group.

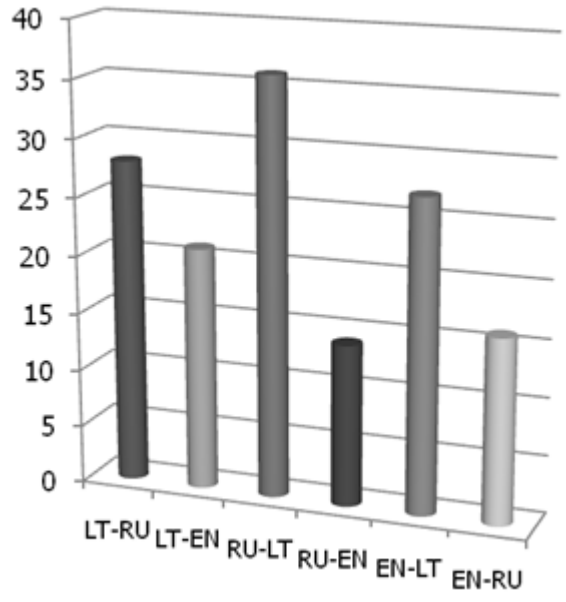


Figure 21. Preferred direction of translation.

[46] Personal communication with students.

The last question to be discussed in this section is Question 18, in which students were asked to label themselves as either mono-, bi-, tri-, or multilingual. As expected, a solid majority (71%) claim to be tri- or multilingual. It was somewhat surprising, however, to find that not only do fully 18% claim to be only bilingual, but 4% even label themselves monolingual. It is unclear whether to interpret this last result as pessimism, modesty, or sarcasm on the students' part. The remaining 7% either did not answer, or claimed not to know: a fair answer, indeed, considering that linguists themselves have a difficult time agreeing on the definitions of these terms.

2.2.3 Cultural data

In this section, the results of those questions aimed at determining students' cultural preferences are discussed. Question 8, for example, asked students to decide which language's jokes are the funniest. Clearly the ability to understand and appreciate the humor of jokes doesn't only depend on which language they are in, so this question is really aimed at seeing if there is a preference among AnRK students for the cultural humor of either Lithuania, Russia, or England/America. As it turns out, the preference is clear: 37 students chose Russian jokes, as opposed to 12 for Lithuanian and only 3 for English. It can be assumed that these students fail to appreciate English humor. A very large number of English jokes are based on puns: a particularly low form of humor, yet one which requires a particularly high awareness of English homonymy. Moreover, it can be assumed that most students translate from English to Lithuanian when trying to comprehend novel utterances. However, as discussed in Mandelblit (1997; see also Braun, 2008), translating is a complicated form of conceptual blending[47] (Coulson, 2001; Fauconnier & Turner, 2002). This blending is made even more complicated by the utterances' status as jokes, the punch lines of which are frequently polysemous. Thus, the low score for English jokes is easily understood. Harder to

[47] This issue will be discussed further in Part 4.

understand is the low score for Lithuanian jokes. Surely no students have any trouble with comprehension[48]. Presumably this is yet another manifestation of the general tendency among AnRK students to elevate Russian culture above Lithuanian.

Question 9, about having an accent, was intended as a demographic or, perhaps, language history question. It has been included in this section because of a tendency that appeared when the answers were collated. It was expected that nearly every student would answer "yes" to the question about having an accent when speaking English, as it was known from personal experience that no student in the program speaks English without an accent. In spite of this, a whole 20% of the group believes that their English is accent-free. Furthermore, of those who agreed that they have an accent, a significant number qualified their answers with expressions such as, "I think so," "I guess," "a little," and "sometimes." For this reason the discussion is included in this section. AnRK students are, by education, trilingual speakers of Lithuanian, Russian, and English. Socioculturally, however, there is not a single native English speaker among them. Until becoming AnRK students, many of them had never been taught by a native speaker of English. They simply have no idea whether their Lithuanian/Russian accents are strong or weak when speaking English. As for Russian and Lithuanian, it can be assumed that their experiences with these languages as "living" features of the environment (recall Waugh et al.'s (2007) ecological perspective) makes them more accurate judges. 15 students give themselves an accent when speaking Lithuanian, and 26 students claim an accent when speaking Russian. There is a correlation here with the students' native languages: 18 students listed Russian as their native language (of which 15 have an accent when speaking Lithuanian), and 28 students listed Lithuanian or Polish as their native language (of which 26 have an accent when speaking Russian).

[48] Again, the issue here is more about cultural stereotypes than jokes themselves. A joke can be stereotypically "English" yet be translated into Lithuanian; such a joke would probably be considered "a Lithuanian joke" in the vague formulation of the survey.

In Question 19, students were asked to write a few adjectives describing the "stereotypical style of interaction" of Lithuanians, Russians, and British/Americans, and then to choose one style which they felt was "closest" to them. The results were fascinating, and deserve a closer look.

Most of the 115 total adjectives or adjective-like phrases were only written by one or two students, but others appeared much more frequently. The most common adjective, used to describe all three nationalities, was *friendly*. It was used a total of 39 times. Overall, it was used to describe Russians 19 times, English speakers 13 times, and Lithuanians 7 times. Even with this most positive adjective, Lithuanians are still evaluated least favorably.

The adjectives the students wrote for each nationality were divided into three categories: positive, neutral, and negative[49]. Table 5 compiles all of the adjectives in each category for each language in decreasing order, down to a frequency of 3. Adjectives listed by only one or two students are excluded from this table[50]. Appendix 2 includes charts containing all of the responses listed for every nationality, subdivided by students' year of study, with totals.

Table 5. Cultural stereotypes by language and polarity.

Lithuanian	*Positive:* friendly, funny *Neutral:* reserved *Negative:* cold, shy, boring, rude, unfriendly, closed, jealous, pessimistic
Russian	*Positive:* friendly, communicative, funny, generous, kind *Neutral:* (none) *Negative:* alcoholics
English-speaking	*Positive:* friendly, funny, communicative, happy, polite, helpful, smiling, well-mannered *Neutral:* emotional *Negative:* (none)

One difference should be apparent almost immediately: Russians and English speakers are characterized by a variety of positive adjectives, whereas only two are used to describe Lithuanians. At the same time, Lithuanians are characterized by a wide range of negative adjectives, which are extremely

[49] For more information about how and why this division was performed, see Appendix 2.

[50] It was felt that while two students might arrive at the same adjective by chance, the odds that three or more students would do so were extremely small. Thus, these data represent actual stereotypes held by the students surveyed.

infrequent when describing Russians and/or English speakers. This trend remains pronounced even when all the one- and two-vote adjectives are included into the count. For Lithuanians, 60% of all the adjectives used are negative, with another 10% neutral, leaving only 30% positive adjectives. For Russians, on the other hand, a full 72% of all adjectives are positive, with 9% neutral and only 19% negative. The results for English speakers are even more upbeat, with 81% positive adjectives, 7% neutral, and 12% negative. These figures are shown for comparison as pie charts in Figure 22.

Figure 22. Cultural stereotypes among students of AnRK.

The results are even more interesting, however, when reanalyzed based on the students' native language. For the reanalysis, students were divided into two groups: L1 = Lithuanian, and L1 = Russian, Polish, and Other. While lumping Polish and Russian speakers together in the same group may not necessarily be the most accurate representation of ethnic allegiances, it is nonetheless true that both Polish and Russian native speakers represent minorities within the Lithuanian linguaculture, and can therefore be opposed to the majority Lithuanian native speakers.

As can be seen in Figure 23, column 1, both Lithuanians and Minorities are extremely critical of Lithuanian culture. Lithuanians give themselves 50% negative adjectives, while Minorities give them 74%. That a minority group such as Russians should be critical of the dominant ethnic group is by no means surprising. What is surprising is that the dominant group appears to agree with the minority. The only positive adjective receiving 3 or more mentions by the Lithuanian sub-group is, of course, *friendly*. No neutral adjectives receive any consensus. Two negative adjectives receive votes of 3 or

more in this group: *cold* and *shy*. The Lithuanian sub-group, then, has an easier time agreeing on negative adjectives to describe itself, than it does on positive. The Minority sub-group, interestingly, does not agree that Lithuanians are friendly. Indeed, that adjective was chosen by only one Minority person to describe Lithuanians. Moreover, while individual Minorities agree that Lithuanians can be *broad-minded*, *funny*, *kind*, *sociable*, and even *self-confident*, no positive adjectives received more than two mentions. The neutral adjective *reserved* was mentioned by exactly three Minorities. Negative adjectives, however, received much more consensus: *boring* (4 votes), *cold* (4 votes), *rude* (3 votes), *shy* (4 votes), and *unfriendly* (4 votes). Thus, even ignoring the other negative adjectives mentioned by the Minority group, both groups agree that Lithuanians in general are *cold* and *shy*. This self-image, it should be noted, cannot be expected to positively impact the long-term prospects of any given Lithuanian in the study program.

Now, looking at column 2, it can be seen that both Lithuanians and Minorities are extremely positive about Russian culture. Lithuanians give Russians 69% positive adjectives, while Minorities give them 76%. (Coincidentally or not, the Minority group itself is 76% Russian.) Within the Minority group, three positive adjectives reached consensus: *friendly*, *communicative*, and *kind*. And, as might be expected from the results shown in Figure 23, no neutral or negative adjective was given more than two votes. Moreover, of the individual negative adjectives that did appear, they were often qualified: "some of them *unfriendly*," "sometimes *strange*," etc. Thus the Minority sub-group, in stark contrast to the Lithuanian, only agrees on positive adjectives, and apparently has misgivings about even mentioning the negative ones. At the same time, the Lithuanian sub-group is rather generous with its praise of Russian culture, even giving it two of the same positive adjectives: *friendly* and *communicative*. The third positive adjective to receive more than two votes is *funny*; recall that Russian humor was highly evaluated, as well. No neutral adjectives reach consensus among the Lithuanians, and the only

negative adjective to do so is *alcoholic*. As the stereotype of Russians as heavy drinkers is known throughout the world, this is hardly surprising or necessarily reflective of real familiarity with Russian culture. To summarize, both groups are very positive about Russian culture, even agreeing on the adjectives used to describe it, while neither group is consistently negative. Thus, any Russians in the study program would seem to be at a real psychological advantage compared to Lithuanians, as they are held in high regard by both major ethnicities in the group. Such an atmosphere can be expected to lead, over time, to greater feelings of self-confidence among Russian students. Whether this translates to better average marks has not been studied.

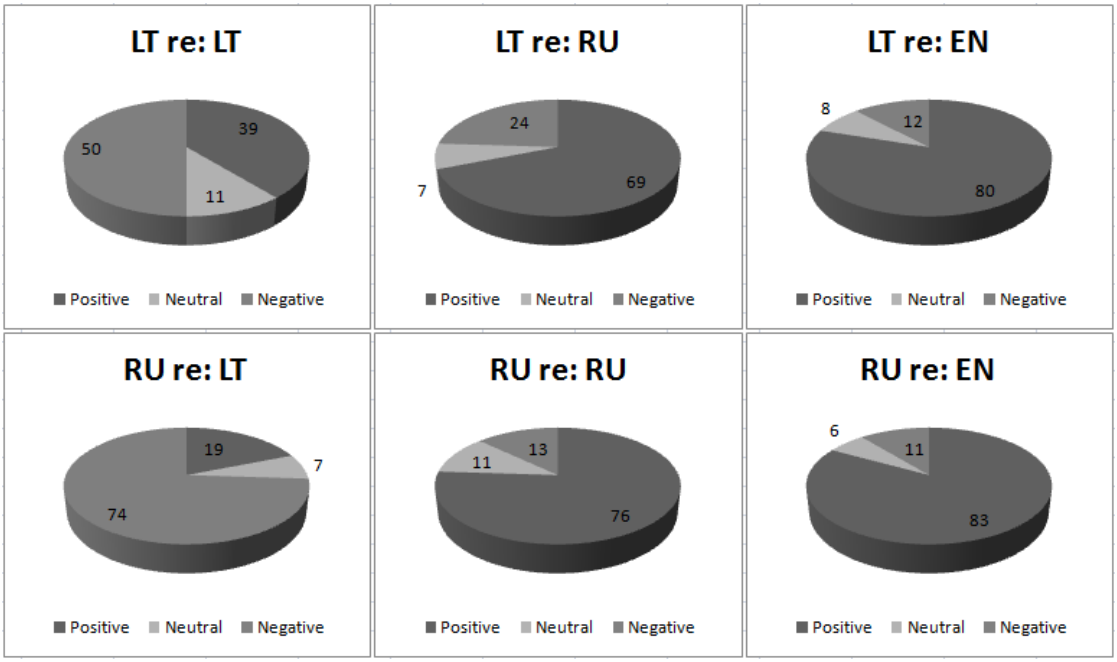


Figure 23. Cultural stereotypes as a function of nationality.

Column 3 shows the results regarding English-speaking cultures. It should be stressed that American and British cultures are not very familiar to these students. While some have first-hand experience with these cultures, it is generally as tourists or summer job-seekers, and such students are extremely few overall. The majority have only theoretical knowledge of these cultures. Thus, it may be unsurprising to find that both Lithuanians and Minorities give English speakers a large number of positive adjectives (more, even, than they

give themselves): *friendly*, *funny*, *happy*, *helpful*, *polite*, *well-mannered*[51]. Interestingly, the Lithuanian sub-group, which ascribes to itself the negative adjective *cold*, gives English speakers the neutral adjective *emotional* (3 votes). Clearly, this adjective can be considered either positive or negative, depending on the context, and one wonders whether there isn't a negative opposition implied: perhaps, to the Lithuanian group, coldness and non-emotionality are, in fact, positive qualities? Unfortunately, students were not asked to comment on their choices of adjectives, and none volunteered, so this possibility has not been investigated further. The Minority sub-group doesn't agree on any neutral adjectives to describe English speakers, and neither group reaches consensus on any negative adjectives. British and American culture, then, are highly idealized among both groups of students. It would, of course, be nice if English speakers all possessed the charming positive qualities ascribed to them above, but in the real world this is far from the case. These adjectives simply show how unfamiliar students are with the cultures whose languages they are studying.

Having written their adjectives, students were asked to decide which style – Lithuanian, Russian, or British/American – felt "closest" to them. Once again, the results are telling, even though only 62% of the group answered the question. Of those 62%, 17 students identify with Russian culture, followed by 11 students who identify with British/American culture. Only 4 students out of the entire AnRK group claim that Lithuanian culture is closest to them. And this, despite the fact that they use the Lithuanian language in nearly every imaginable social situation. To be clear, students state that they identify more closely with a completely foreign culture, one with which most of them have no firsthand experience, than they do with the culture of the country they were born and live in and whose language they use every day.

[51] In this discussion, only those adjectives mentioned three or more times within the sub-groups are listed. In Table 5, adjectives such as *smiling* appear because they were mentioned more than twice by both groups combined.

These results can be profitably contrasted with the answers students provided to the last section of the questionnaire, titled "Cultural scales." This section gave students the opportunity to agree or disagree with various statements about Lithuanian, Russian, and English-speaking cultures on a 5-point Likert-type scale. Questions 20–23 and 25–26 were related, as can be seen in this excerpt, which provides only the Lithuanian culture questions. (Again, the complete questionnaire can be found in Appendix 1.)

20. Lithuanian culture is a part of my life.	1.....2.....3.....4.....5
21. I feel like a Lithuanian.	1.....2.....3.....4.....5
22. Lithuanians would think I'm Lithuanian.	1.....2.....3.....4.....5
23. Most of my friends are Lithuanian.	1.....2.....3.....4.....5
25. I want to speak Lithuanian fluently.	1.....2.....3.....4.....5
26. I'm comfortable speaking to native Lithuanians.	1.....2.....3.....4.....5

The questions were intended to see how closely students identify with the three cultures whose languages they study. It was assumed that a student who answers mostly 4 ("partially agree") or 5 ("strongly agree") to these statements, regardless of his or her nationality or native language, values Lithuanian culture and the Lithuanian language more highly than one who answers mostly 2 ("partially disagree") or 1 ("strongly disagree"). Such a student feels comfortable in Lithuania, identifies with and is an active part of its ethno- and linguaculture, and – most importantly for the purposes of this dissertation – will be more likely to experience CLI from Lithuanian when speaking Russian or English than vice versa. Similarly, students who provide high numbers for the same questions about Russian culture are more likely to entertain feelings of minority pride and/or Russian nationalism (in a weak sense[52]), and perhaps to see themselves in opposition to the majority Lithuanian students in the group. In terms of CLI, of course, such students can be expected to show more influence from Russian when speaking Lithuanian

[52] Strong Russian nationalism in a negative, racist sense does not seem to be a problem for this group, perhaps because many Russian AnRK students feel somewhat "provincial" compared to Russians in Russia. One Russian student described visiting Moscow, where she was unpleasantly surprised to find that she was treated as a foreigner because she spoke Russian with an accent.

or English than vice versa. It was not expected that any students would answer these questions with 4s or 5s when referring to British/American culture; the questions were mostly included for symmetry. On the other hand, there do occasionally appear certain anglophilic students who seem to value American or British culture above all others, and it was of interest to see whether any such students were included in this group. The questions in this section, then, attempt to look beyond the students' stereotypes, which may be surface-level responses (as the "alcoholic Russians" example indicates) rather than deeply held beliefs; it was hoped that the questions would reveal more about individual students' real attitudes towards the cultures in question than those grouped under Question 19.

The first observation to make in regards to these "cultural scales" is that they clearly demonstrate that Lithuanians are not nearly as self-critical as their lists of stereotypical adjectives seemed to indicate above. Neither are they as fond of Russian culture as it seemed. In order to fully appreciate the results provided by these scales, the analysis has again been performed twice: first, looking at the answers provided by the entire AnRK group as a whole; and then again by examining Lithuanian as opposed to Minority sub-group answers. The results are best appreciated when drawn as scatter plots with trendlines, as will be seen in the illustrations below.

The discussion begins by focusing first on the answers given by the entire AnRK group regarding each culture in general. Figure 24 plots all of the results about Lithuanian culture in ascending order. Each point on the graph represents the average of the answers to all six questions listed above. As can clearly be seen, a few answers fall below 3 ("maybe"), indicating students who feel distant from Lithuania and its culture. However, the majority of the answers (33 out of 51, or 65%) are 4 or higher; the mean is 4.02, standard deviation 0.86. These results are unsurprising, considering that almost all of the students have lived in Lithuania for most of their lives. Thus, in spite of the negative stereotypes regarding Lithuanians held by AnRK students, a solid majority of

them nevertheless feel close to Lithuanian culture and see themselves as an integral part of it. This answer is particularly interesting, as it may be remembered that Lithuanians form only 53% of the group. Thus, even some Russian students, whose stereotypes of Lithuanians were seen to be highly negative in the discussion above, identify themselves rather strongly with Lithuanian culture. Exactly how far this identification goes will be examined more closely below.

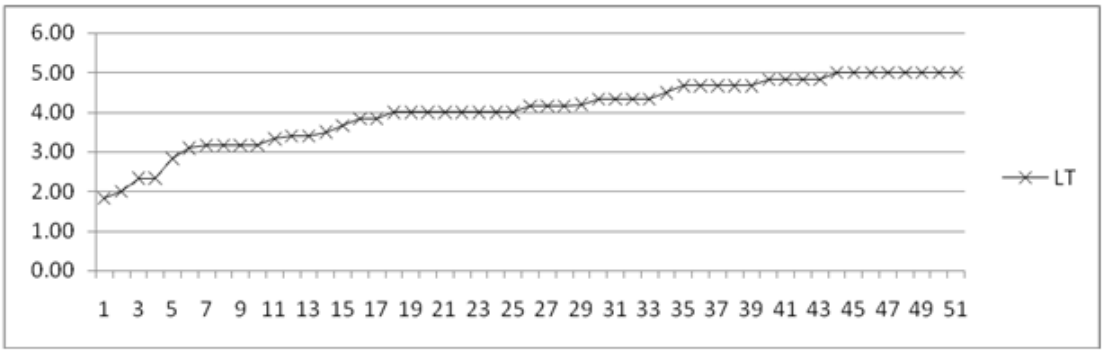


Figure 24. AnRK identification with Lithuanian culture.

In Figure 25 are plotted the results regarding Russian culture, also in ascending order. Here it can be seen that a smaller proportion – only 26 out of 51, or 51% – award Russian culture scores of 4 or higher. For Russian culture, the mean score is 3.78 (standard deviation 0.87). These, then, are the students who, regardless of their nationality, feel comfortable in a Russian cultural setting. It may here be remembered that Russians form only 31% of the group, which means that a large number of Lithuanians and/or Poles identify with Russian culture. This, of course, could have been predicted from the adjectives they provided in Question 19. Of those adjectives, 72% of those applied to Russians were positive, indicating that many non-Russian students hold favorable stereotypes regarding Russian culture. However, significantly fewer (only 51%) actually feel as though they belong to that culture.

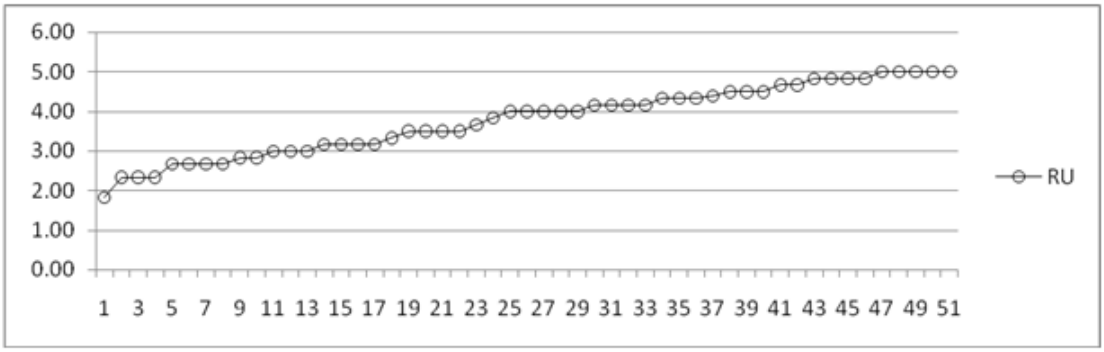


Figure 25. AnRK identification with Russian culture.

Although AnRK students were very positive in their stereotypical descriptions of British and American culture, they are not willing (or, more likely, not able) to fully commit themselves to those cultures while living in Lithuania. As can be seen in Figure 26, not a single student's average response to the six questions about English-speaking cultures reached 4 ("partially agree"). Indeed, the mean response was only 2.56 (standard deviation 0.5) – clearly not the answers of a group of anglophiles. That Lithuanian and Russian students living in Lithuania do not feel as though they belong to British or American culture is, of course, not a significant result – only the opposite would be noteworthy. However, it does corroborate the contention that the extremely positive adjectives given above were the result of students' unfamiliarity with and mostly theoretical knowledge of these cultures. As these cultural scales clearly demonstrate, no matter how *friendly*, *happy*, or *helpful* English speakers are thought to be, they are also clearly recognized by the students as *other* and *different*. Unfortunately, the questionnaire did not include any questions about students' experience(s) abroad, so it is impossible to know whether the higher English evaluations came from those students who have lived in English-speaking countries. It would be interesting to devise a follow-up questionnaire focusing on questions, such as this one, raised by the results of the survey under discussion.

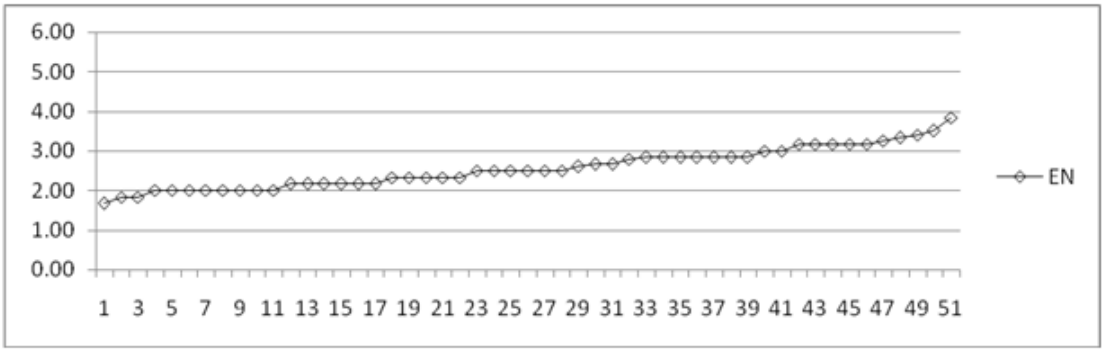


Figure 26. AnRK identification with British/American culture.

Chi-square, a statistical significance measure, was also computed for the above information. This procedure tests the strength of the relationship between/among variables, and does not indicate cause or effect (Hatch & Lazaraton, 1991). For the answers to the six cultural scale questions, answers were weighted and totaled. The expected result was 918 (this would be the weighted total were all 51 students to give an answer of 3 for all 6 questions: 51 x (3 x 6) = 918). The observed results were as follows: Lithuanian culture, 1214; Russian culture, 1148; English-speaking cultures, 763. The test therefore showed that the observed scores are statistically highly significant: $\chi^2_{(2)} = 179.23$, $p < .001$.

Now consider again the information represented in Figures 24 and 25. In Figure 24, the answers to the questions about Lithuanian culture were averaged and plotted for all 51 respondents. They were also arranged in order from least to most agreement with the statements given in the questionnaire. In Figure 25, the same was done for the questions about Russian culture. The two charts therefore represent two different arrangements of the same data. In Figure 27, on the other hand, the Lithuanian answers have again been arranged in ascending order. In addition, the Russian answers have also been added to the chart. As can be seen, the Russian answers, when plotted as a function of the Lithuanian answers, not only do not increase, but actually decrease rather sharply. The solid "RU trend" line noticeably drops from 4.6 to just under 3, while the Lithuanian results climb from 1.8 to 5. For comparison, the results for English-speaking cultures have also been plotted, such that the dotted "EN

trend" line can be seen to rise slightly, from 2.3 to 2.8, over the same set of answers. These results suggest that the Lithuanian and Russian cultures are indeed seen in some form of opposition within the AnRK group, even if that opposition is generally not expressed. At the same time, both groups feel more or less equally indifferent towards British/American culture.

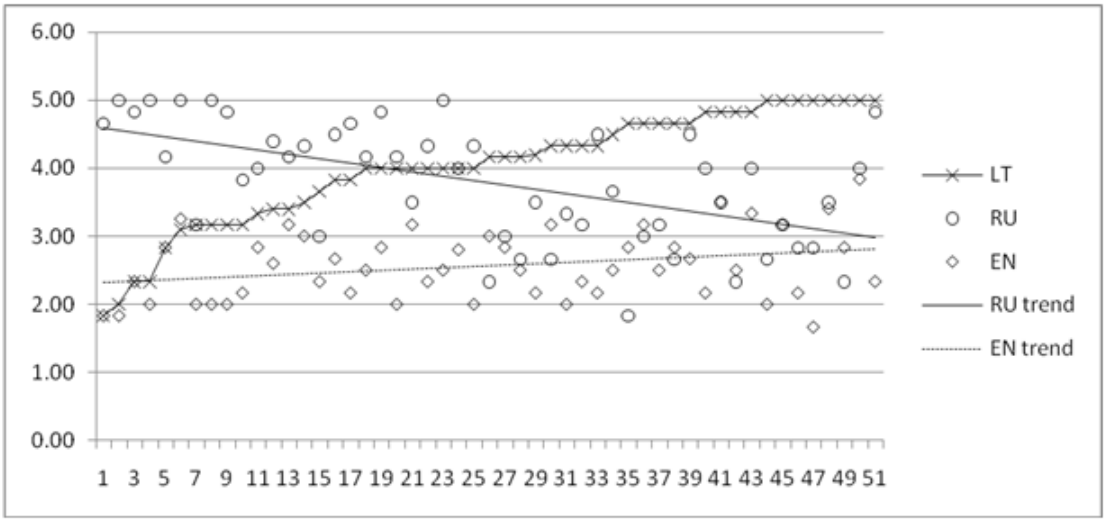


Figure 27. Lithuanian scores with corresponding Russian and English scores superimposed.

Figure 28 is the opposite of Figure 27: here, the Russian scores from Figure 25 have again been arranged in ascending order, with the corresponding Lithuanian scores superimposed. And, as could only be expected, this time the Lithuanian trend line drops noticeably, from 4.8 to 3.2. Thus, the closer a given student identifies with Russian culture, the farther that same student is likely to feel from Lithuanian culture – and vice versa. As before, the English-speaking results have also been calculated, and the trend line can be seen to drop very slightly, from 2.8 to 2.4, again indicating the students' general unfamiliarity with and distance from British/American culture.

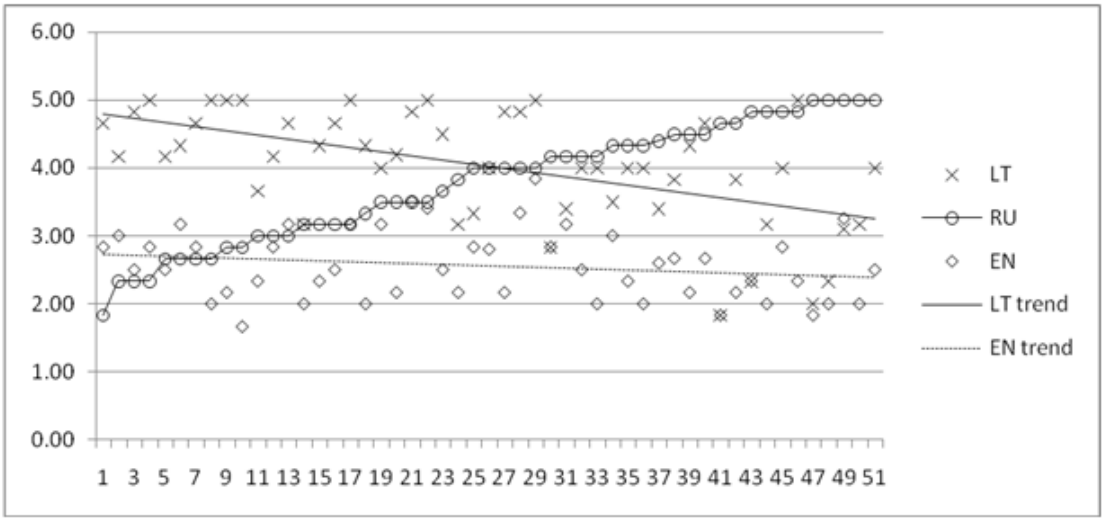


Figure 28. Russian scores with corresponding Lithuanian and English scores superimposed.

The charts in Figures 24–28 plot the answers for the entire AnRK group as a whole. However, as the results more and more clearly indicate, the group is far from homogeneous. For this reason the charts were recalculated based on students' stated nationalities, as above in the discussion of adjectives. Figures 29–33, then, present the data subdivided into the group of Lithuanians as opposed to the group of Russians, Poles, and other nationalities.

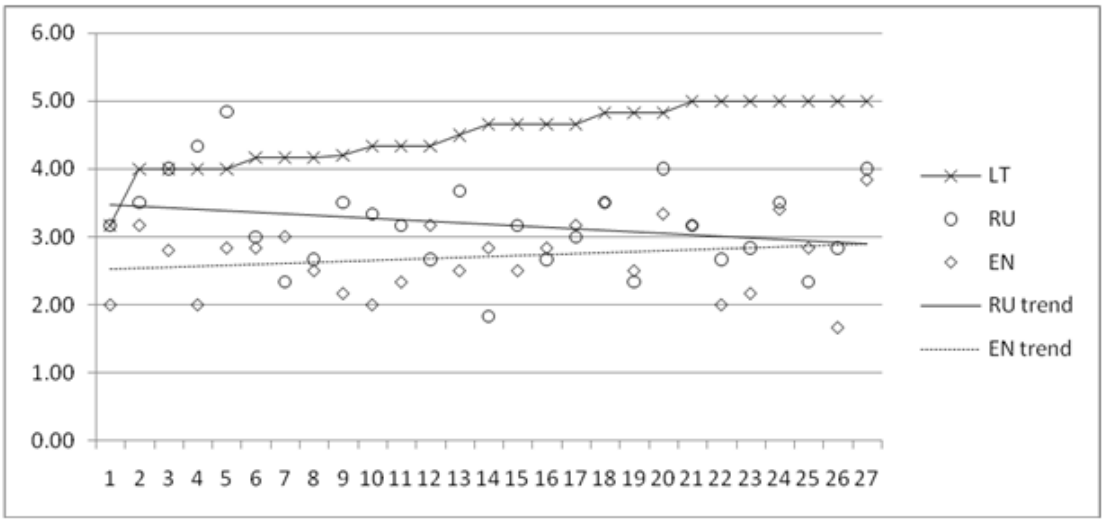


Figure 29. Lithuanian sub-group scores sorted on Lithuanian culture.

Figure 29 shows the answers of the Lithuanian sub-group, sorted in ascending order according to their answers to the questions about Lithuanian culture. As can be seen immediately, this sub-group clearly and definitively

identifies with Lithuanian culture: even 96% of the answers score at 4 or higher. Recall that, in the entire AnRK group, such answers formed only 65% of the total – a large number, to be sure, but not as overwhelming as within the Lithuanian sub-group. Meanwhile, the trend line for the Russian answers does not drop as sharply as it does over the group as a whole – it moves from 3.4 to 2.9, a distance of 0.5, as opposed to a drop of 1.6 points over the AnRK group as a whole – but this is due to its much lower starting point. Both lines end at approximately 2.9, but the "RU trend" line in Figure 29 begins 1.2 points lower than that in Figure 27. In general, the angle of the line is much less significant than its average height, which, except for the very first score, is well below that of the Lithuanian line. The Lithuanian sub-group, then, identifies much more clearly with Lithuanian culture than with Russian, and the closer they feel to Lithuanian culture, the farther they do from Russian culture.

Figure 30 plots the answers of the Minority sub-group, sorted in ascending order according to their answers to the questions about Lithuanian culture, just as in Figure 29. The answers, as can clearly be seen at a glance, cover a much wider range than in the Lithuanian sub-group, from less than 2 to 5. Moreover, it can be seen that answers of 4 or higher form a much smaller proportion of this group: only 33% of Minority students feel so close to Lithuanian culture. At the same time, the entire Russian trendline stays high above the 4 mark (and, mostly, the Lithuanian line), dropping only very slightly from 4.6 to 4.2. Interestingly, the line does drop, indicating that even among Minority students, those who identify most strongly with Lithuanian culture feel least close to Russian culture.

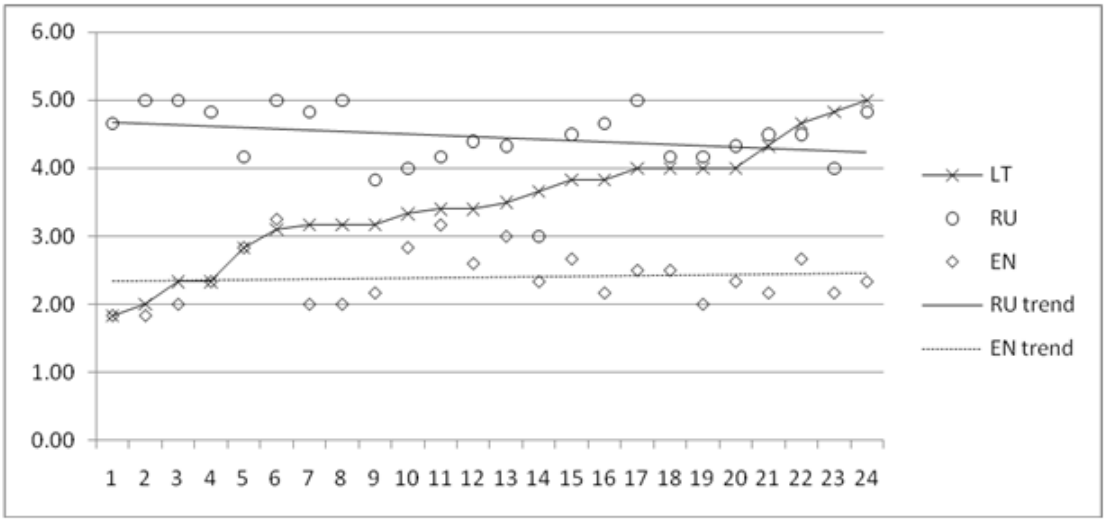


Figure 30. Minority sub-group scores sorted on Lithuanian culture.

In both Figures 29 and 30, the English results were also plotted (dotted trendlines). In both groups, these lines are noticeably lower than any others, again demonstrating students' essential distance from English-speaking cultures. Note that in Figure 29, the English and Russian trendlines end at the same point. Apparently, for those students who identify most strongly with Lithuanian culture, neither Russian nor English is preferred; any foreign culture at all is dispreferred about equally.

The last three charts in this section show the sub-groups' results when sorted in ascending order according to the answers to the questions about Russian culture. The same patterns that have applied to all the charts thus far examined apply here, as well. In Figure 31, the answers of the Lithuanian sub-group are shown. Only 5 students give Russian culture a score of 4 or higher (a mere 19% of the sub-group); indeed, the mean score for Russian culture in this sub-group is only 3.19. Lithuanian culture, on the other hand, receives very high scores: the "LT trend" line moves down only slightly, from 4.7 to 4.3.

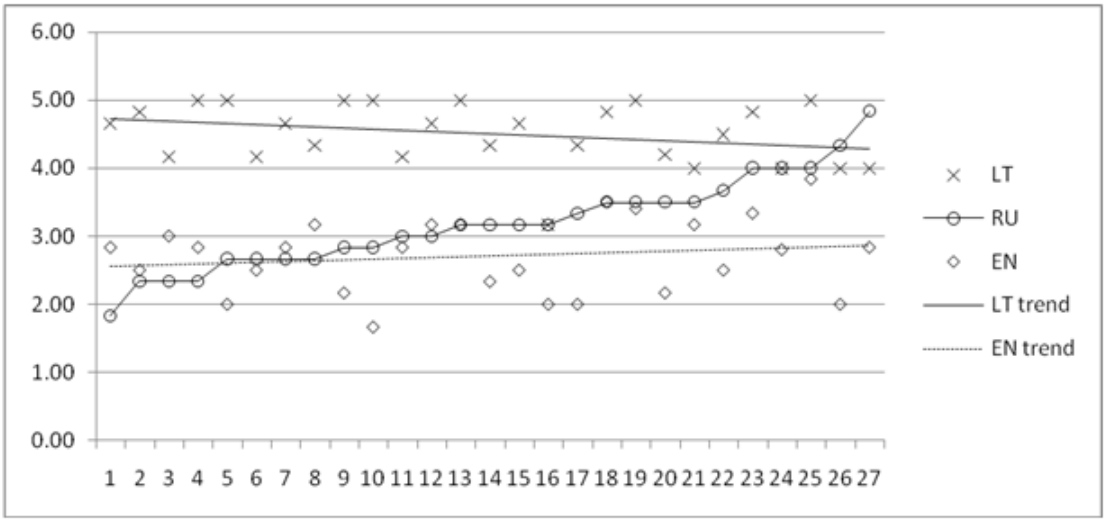


Figure 31. Lithuanian sub-group scores sorted on Russian culture.

Figure 32 shows the answers of the Minority sub-group. Here only two students gave Russian culture a score less than 4; 92% of the answers are 4 or more, with a mean of 4.45. And, as could only be expected, Lithuanian culture remains much lower, with its trend line dropping from 3.9 to just over 3.

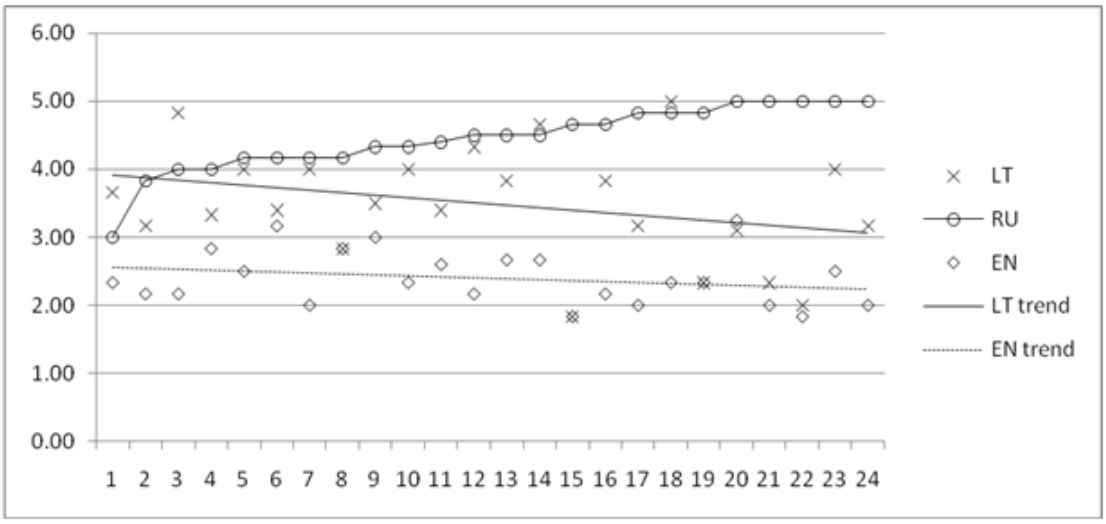


Figure 32. Minority sub-group scores sorted on Russian culture.

Finally, for comparison, the answers of just the Russian students within the Minority sub-group were extracted. These are plotted, again in ascending order, in Figure 33. This time, every answer about Russian culture is 4 or higher, and the Lithuanian line drops sharply indeed, from 4.2 to 2.8, a distance of 1.4 points. This drop is much more pronounced than the corresponding drop

for Russian culture among the Lithuanian sub-group when the answers were sorted on Lithuanian culture (see Figure 29). There, the drop was from 3.4 to 2.9, or only 0.6 points.

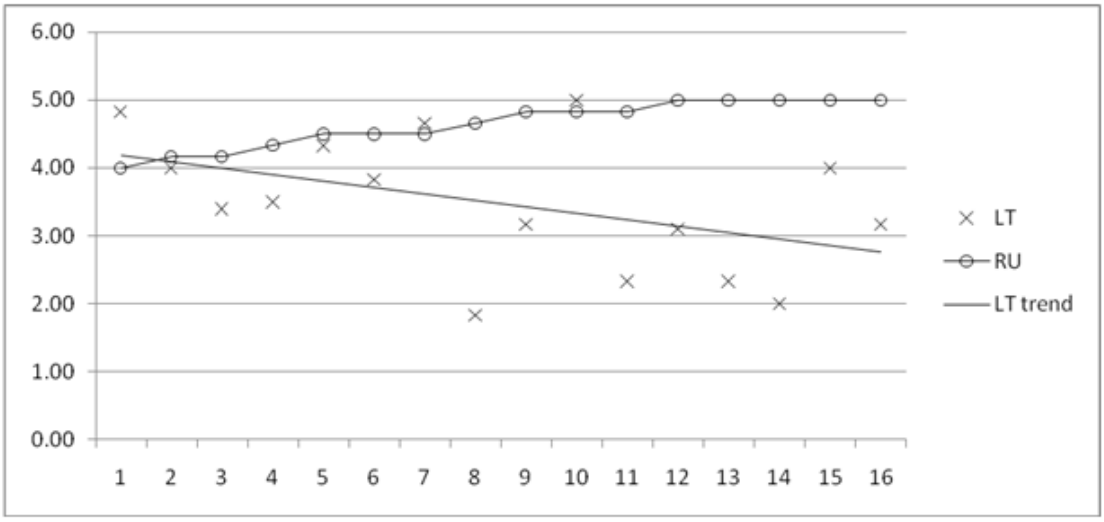


Figure 33. Russian-only scores sorted on Russian culture.

To summarize, it can be seen that as Russian identification with Russian culture grows, the distance from Lithuanian culture increases quickly. At the same time, however, Russians in general feel closer to Lithuanian culture than Lithuanians do to Russian culture. The Russian culture trendline in Figure 29 begins 0.8 points lower than the Lithuanian culture trendline in Figure 33 (though, again, both end almost at the same point). Thus, the sociocultural environment can be seen to affect the students' answers to these questions, too. The majority of Lithuanians do not feel especially close to Russian culture, as could only be expected considering the ethnic distribution of the country as a whole. Russians, however, having grown up as a small minority in the general Lithuanian socioculture, cannot help identifying with it to some degree.

2.3 Summary and Discussion

The results of this survey show that the students under investigation are deeply affected by their sociocultural environment. Whatever their nationality, they are situated in a world of stereotypes and allegiances that extend far beyond the often extremely confined bounds of the circle of acquaintances

with whom a language can be spoken. This environment permeates their relationships with family and friends from an early age.

Before entering the AnRK program, students grow up in various, more or less ethnically diverse, parts of Lithuania, where the state language is Lithuanian and attitudes towards minorities are not always favorable. As members of the prestige culture and native speakers of the prestige language, Lithuanian students can be expected to enter the group feeling confident and sure of their social position. Russian students, on the other hand, even if they grow up in Russian-dominated enclaves, have spent their lives with the knowledge of their difference from the majority culture. They might entertain feelings of weak nationalism and ethnic pride, but they can also be expected to enter the group feeling less secure about their social position.

However, the story changes rather quickly once the AnRK study program begins. Within the program, Russians remain a minority, but a significantly more populous one, numbering more than 30% of the total group. This fact alone could account for a number of the phenomena seen in the survey. Russians will quickly recognize the changed ethnic distribution and, perhaps, band together (especially if they have lived in Russian-majority enclaves as children). They may also feel a certain lowering of the usual social restrictions on Russians in the wider culture, as here they are not only more numerous, but the Russian language (unlike Lithuanian) is one of the two major languages of study. As Russian is their own native language, some aspects of Russian linguistics (e.g. grammar, lexis) should be easier for them than for the Lithuanian-native contingent, another factor contributing to Russian self-confidence. At the same time, Lithuanians cannot help but notice the ethnic redistribution as well. Their language and culture are not studied as they were in school, except accidentally (as when Lithuanian is the language of instruction); instead, they must learn a great deal about the Russian language and culture. Their lower fluency in Russian and non-native familiarity with the culture may also put them at an academic disadvantage when compared to the

Russian natives. All of this could have the unintended consequence of causing Lithuanians to feel socially inhibited or marginalized within the group, despite their still-dominant numbers.

Nearly half of the group listed Russian as their favorite language, as opposed to only two students who chose Lithuanian. The main reasons for this had to do with the rich vocabulary and literature of Russian. These are rather academic or linguistic reasons which support the theory advanced above, namely, that due to its status as one of the two languages studied in the AnRK program, Russian becomes a prestige language within this group. Similarly, when asked which style of interaction feels closest to them, the majority of respondents choose Russian, with only four choosing Lithuanian. This also supports the idea that, within the AnRK group, and in clear opposition to the larger Lithuanian socioculture, the prestige language and culture are Russian. This suggestion is further supported by the students' choices of adjectives to describe members of the three cultures. There, Russians are given a large number of positive adjectives by both Minority and Lithuanian students. Lithuanians, in contrast, are negatively evaluated by both Minority and Lithuanian students themselves. Thus, all ethnic groups rate Russians positively and Lithuanians negatively. Students are, of course, unaware of the results of the present survey, yet this situation cannot remain unnoticed, at least subconsciously, within the group. Such an atmosphere may serve to reinforce Russian feelings of self-confidence while, at the same time, causing Lithuanians to experience a kind of culture shock. Such a feeling of cultural displacement without ever having left one's home country may have unforeseen consequences for these students. Such consequences may range from feelings of inhibition (recall that *shy* and *cold* were the two most frequent adjectives given to describe Lithuanians) to lowered expectations and demotivation, at least when the language of study is Russian[53].

[53] Recently one Lithuanian student dropped out of the AnRK study program, only to re-enter the university the next year in the English Philology (AnF) program. AnF students take no Russian language courses.

Where English is concerned, all students are more or less equally optimistic. The adjectives given to describe English speakers are not only positive but almost idealized. This most likely reflects the students' lack of personal experience with English-speaking cultures. Foreign EFL textbooks used in Lithuania, such as the formerly popular *Headway* series and others based on that model, do not generally portray native English speakers in a negative light, and students whose experience of such people is limited to that presented in such textbooks may very well assign them descriptions like *happy*, *helpful*, and *polite*. These stereotypes will be held regardless of nationality, as all students are equally underexposed to living examples of British or American culture. This may also explain such phenomena as students' unwillingness to recognize their own accents when speaking English, and also the fact that English is perceived as more distant typologically from both Lithuanian and Russian than those two languages are from each other.

Another factor that probably contributes to the lack of identification with "English-speaking" cultures is the vagueness of this term. The debates about English as a world language (e.g., Kachru, 1985, 1991; Scovel, 2001; Alptekin, 2002; Prodromou, 2007) that no longer has one single culturally- or geographically-defined core point to the heart of the problem. If "English culture" is indefinable, then there is no single English culture for students to identify with. If there is no concrete model, students will have a difficult time acquiring anything approaching stereotypical NS norms. It should come as no surprise that students' English accents, when they are not influenced by Lithuanian or Russian, show characteristics of both Standard British and General American.

What limited experience students have of English culture outside the classroom tends to come from books, films, music, and the Internet – all of which mix British, American, and other "Englishes" as a matter of routine. Except for the Internet, none of these areas are interactive in a way that would require real online production of English. Even the Internet is mostly used by

students as a kind of encyclopedia. (The survey did not actually ask students what they use the Internet for, but informal discussions have revealed that very few of them engage in any real-time chatting with native speakers of English. Its main purpose seems to be as a source of ideas for homework and other assignments.) Recalling the DMM concept of language maintenance effort (LME), it can be seen that students do invest some effort into their English receptive skills, yet there are reasons to doubt the effectiveness of even this amount of practice. Reading is often done hurriedly, under pressure of deadlines, late at night, while ill, and so forth; in other words, under non-ideal conditions. Similarly, music is mostly used as background noise, playing while students are engaged in other activities, so one may question whether students actually "listen" to music at all. Meanwhile, as was shown in Table 4, the languages students use for production are overwhelmingly Lithuanian and Russian. This can easily be sensed by attending any English lecture. Many students bring Lithuanian newspapers to leaf through or read, others work on homework assignments for Russian courses they are taking, and almost all *private* student-student interaction is in either Lithuanian or Russian[54]. Thus, despite the efforts of teachers who attempt to maintain monolingual English communication in the classroom, AnRK students seem unwilling to inhibit Lithuanian and/or Russian to any significant degree.

This brings up the issue of **language mode** (Grosjean, 1997, 2001; Dewaele, 2001), defined in Grosjean (2001: 3) as the "state of activation of the bilingual's languages and language processing mechanisms at a given point in time." Although he writes specifically about bilinguals, the language mode framework can in principle be extended to cover any number of languages. According to this idea, a multilingual can be seen as being in either a fully monolingual mode (although this rarely occurs in practice), in which case all

[54] Private interaction describes those instances when students talk to each other, usually in whispers, for personal reasons. Such conversations are not meant to be overheard by others. Private can be opposed to *public* interaction, which is usually at the request of the teacher and spoken loudly, clearly, for all to hear, and always in English (except for code-switching; see Part 3).

other languages are inhibited as much as possible; a fully multilingual mode, when all languages are highly active; or anywhere on the continuum in between. The level of activation changes constantly in response to the conversational situation and is sensitive to the perceived mono- or multilingualism of any interlocutor(s). It is also a habitual pattern which can be more or less well established in different speech communities. For these students, a mixed mode is natural in all social situations but for the classroom – so it is only natural if they continue to use it there, too.

The language they choose to switch into for private interaction depends on both their nationality and that of the interlocutor. It is often (though certainly not always) the case that students of the same nationality will sit together, though this tendency seems to decrease throughout the years of study. (This is an informal observation and has not been studied.) In any case, the cultural scales questions clearly show that students are well aware of nationality, that they identify more closely with that of their own culture, and, of course, that no students identify with English-speaking cultures. Actually, as fully 65% of the group show a strong allegiance to Lithuanian culture, it could be said that even many minority students recognize their position as part of the Lithuanian socioculture. Be that as it may, the fact that Russian and Lithuanian scores on these scales seem to rise and fall in inverse proportion (see Figures 27–33) does indicate that all students are aware of and affected by their nationality.

It is recognized, of course, that the charts mask individual differences among these students, working as they do from average trends. For example, even in the Russian-only sub-group (see Figure 33), three students actually identify more strongly with Lithuanian culture than with Russian. However, the purpose of this survey was not to provide case-studies of any individual students, but to examine general trends within the entire group.

In a group such as this, crosslinguistic influence is unavoidable. As the results of this survey indicate, the greatest amount of CLI will be felt whenever

these students attempt to speak or write in English. Not only is English the language with the greatest perceived distance from their L1, but it is also unavoidably underrepresented in the environment. Students have little practice using English productively and maintain a bilingual language mode even in English language classrooms. They can thus be expected to rely heavily on their more resonant language(s) for conceptualizing and formulating utterances, which must then be translated from Lithuanian or Russian to English. As will be seen in Part 4, at least for the Lithuanian students, this is indeed the case. Students need to make conscious efforts to increase their exposure to (especially productive uses of) English outside the classroom, in order to increase English resonance, gain practice inhibiting Lithuanian and Russian CLI, and acquire more natural syntactical and lexical patterns.

Finally, the status of Russian as a prestige language within the group could be studied more closely. As can be seen at a glance in Figure 23, one could easily expect group dynamics based on nationality to develop. In fact, upon entering the program, students are traditionally given a Russian language test, the results of which are used to sort them into two sub-groups[55]. The test does not sort students on the basis of nationality, as both sub-groups end up composed of both Lithuanian and minority students. However, it is the case that the group which scores higher on the test usually contains a higher proportion of Russians than the group which scores lower (and has a higher proportion of Lithuanians). These groups study some subjects separately, though others (especially elective courses from outside the AnRK program) are studied together. As it turns out, there does seem to be a certain amount of rivalry or opposition between these two sub-groups[56]. That such dynamics are not more openly visible to date may be a result of the unusually positive evaluation of Russians across the entire group, coupled with the low evaluation

[55] This is a purely administrative procedure, and is done for scheduling purposes that have nothing to do with students' nationality or academic aptitude.

[56] One student in the Lithuanian-majority sub-group, while giving a speech on an unrelated topic, referred to the other group as "not so nice," and was immediately seconded by two students in the audience.

of Lithuanians. In general, to judge by the answers given on the survey, Lithuanian loses out to Russian in nearly all areas. It may thus be assumed that Russians should feel more self-confident in the AnRK study program, and may even tend to do better in the long run as a result. Thus far, however, no attempt has been made to study whether either group is academically more or less successful than the other.

In regards to the types of EFL pragmatic failures discussed above, then, three observations seem to be of particular importance. First, these students' speech community is inherently Lithuanian/Russian. Lithuanian and Russian pragmatic norms can therefore be expected to transfer to English as a matter of course. Second, students' inner voices are Lithuanian or Russian, too. As noted by Kecskes & Papp (2000), these are the languages they think in, even when speaking English, and this cannot but lead to CLI and transfer at many different levels. Third, because they have studied English as an FL in classroom environments, they have had little or no exposure to genuine English-speaking culture(s). None of these students can be expected to have reached anything approaching an English acculturation threshold, and, as seen in all of the charts above, English cultures therefore remain foreign to them. Pragmatically, this can be expected to result in large numbers of utterances that do not congrue with English NS norms.

3. CROSSLINGUISTIC ASPECTS OF ENGLISH SPEECH PRODUCTION

Multilingual speech production in anything other than the native language is fundamentally different from monolingual speech production. The basic process, as outlined in Levelt's (1989) model, remains the same; however, as seen in Part 1, multilingual production can be affected at all levels, from conceptualization to monitoring, by factors that monolinguals simply do not encounter. The multilingual lexicon is also structured differently from that of monolinguals. Models of multilingualism take into account such issues as translation latencies, inhibition, competition and resonance, and the dynamic interaction of variables, none of which affect monolingual production.

Moreover, as was seen in Part 2, multilinguals studying languages in non-immersion settings (e.g., as foreign languages) are likely to produce pragmatic failures, even if all other aspects of production are correct. This is due to the language and cultural specificity of their inner speech processes. As seen in the responses to the sociocultural language use survey, the students of AnRK are underexposed to English cultural norms.

This Part analyzes the English speech production of those AnRK students whose L1 is Lithuanian, on the assumption that Lithuanian linguistic and sociocultural norms will unavoidably affect their production of English. This type of affect is what has been called CLI throughout this dissertation. Section 3.1 reviews the seven different types of CLI that can be found in the spoken corpus. In section 3.2, examples of each type are presented and discussed. Conclusions are drawn in section 3.3.

3.1 Types of Crosslinguistic Influence

The term crosslinguistic influence (CLI) is meant to be a catch-all term covering as many language interaction phenomena as possible. It covers both transfer and interference, as well as such phenomena as code-switching and borrowing. It can also be used to cover such non-predictable dynamic effects as transitional bilingualism and language attrition, avoidance and maintenance

techniques (Herdina & Jessner, 2002), choice of language mode (Grosjean, 2001; 1997), fossilization, and the effects of the L2 on the L1 (cf. Cook, 2003; Kecskes & Papp, 2000).

In analyzing the spoken corpus, seven distinct types of CLI were found. Each will now be briefly defined and discussed.

The first and easily most prevalent type of CLI is **hesitation**, also called **pausing**. This phenomenon is by no means exclusive to multilinguals. It has long been known that "[a]n average of about 40–50 percent of utterance time is occupied by pauses. <…> [P]ausing is as much part of the act of speaking as the vocal utterance of words itself" (Goldman-Eisler, 1964: 98-99). Pauses may be **filled** or **unfilled**: filled pauses are most typically expressed as *uh*, as discussed above (Levelt, 1989), while unfilled pauses are silent. According to Maclay & Osgood (1959), hesitation can actually be of four kinds: filled and unfilled pauses[57], **repeats**, and **false starts**. A repeat is any superfluous repetition of a word, while a false start is any abortive beginning of a sentence or phrase. All of these, but especially filled pauses, are believed to represent the speaker's attempts to keep control of the speaking turn (Carroll, 2008). Moreover, speakers typically alternate between **hesitant phases** and **fluent phases** in speaking turns (Henderson, Goldman-Eisler, & Skarbek, 1966; Beattie, 1983). Goldman-Eisler (1961) suggests that pausing may be related to both emotional factors and cognitive activity; Mahl (1956) investigates hesitation in the context of anxiety. According to Dewaele (2001, 1996), hesitations in multilinguals are especially common before and after lexical gaps, and are indicative of cognitive activity. In fact, all of these factors are undoubtedly present when multilinguals speak an FL, especially (as in the student corpus) when in the context of a marked in-class assignment before a group of peers. Moreover, as the processes of grammatical and phonological encoding are not automated for multilinguals as they are for monolinguals, pausing as a manifestation of cognitive activity is to be expected.

[57] In Quinting (1971), the median length of unfilled pauses was .87 seconds.

The second type of CLI is **code-switching** (Poplack, 2000; Myers-Scotton & Jake, 2001; Myers-Scotton, 2005). Different language researchers define this term in different ways. In this dissertation it is used to mean any insertion of a non-English word, phrase, or sentence into the stream of English speech. Monolinguals do this, too, usually as a stylistic device – witness the many Latin phrases adopted for (usually abbreviated) use in academic writing. Words or phrases that have been **borrowed** from other languages into English (e.g., *burrito*, *spaghetti*, and *croissant*) are also a form of code-switching, of course. However, monolinguals would not be able to switch completely into production of, e.g., Latin, Spanish, Italian, or French.

The third type of CLI is the creation of **neologisms** and/or use of **foreignizing** (Cenoz, 2001; Poulisse, 1990). By neologism is meant any nonce creation that is not already a standard English word, usually through incorrect derivation. Foreignizing is defined as "the use of an L1 (or L*n*) word with phonological and morphological adaptation" (Poulisse, 1990: 111) to the language being spoken. English monolinguals create neologisms too, of course, but the new words are grammatically correct with respect to the rules of English, thus distinguishing them from the kinds of neologisms being discussed here. Monolinguals rarely, if ever, foreignize.

Transfer is the fourth type of CLI found in the corpus. As will be argued extensively in Part 4, transfer is a kind of conceptual blending in which L1 concepts and lemmas are blended with L2 lexemes. When monolinguals conceptually blend, the results are acceptable and comprehensible. When multilinguals conceptually blend, the results are often **learner constructions** that do not meet NS norms (Waara, 2004).

The fifth type of CLI manifests in students' phonology. It goes without saying that these students speak English with an accent. Monolingual native speakers, naturally, do not. Most of the time students' accents are no hindrance to communication, but occasionally certain peculiarities of pronunciation can

lead to difficulty in comprehension; such examples will be identified and discussed.

The sixth type of CLI is **interactional strategies**, defined in Cenoz (2001: 12) as "direct or indirect appeals to the interlocutor in order to get help;" as such, the term includes Hammarberg's (2001) CLI category of **elicitations**. Under the rubric of interactional strategies are also included those cases where students appeal to the teacher or the audience for reasons other than to get help (i.e., to make a joke) which are nevertheless outside the defined topic of conversation. This Part will also discuss those cases where speakers assume an understanding on the part of the audience that may or may not be there. Monolinguals employ these strategies as well, of course; anyone experiencing a TOT phenomenon usually asks others, if they are available, to supply the sought-after word. Levelt (1989) also notes that speakers exploit interlocutors' cooperativeness on the assumption that if they haven't been understood, a question will be asked. Multilingual speakers do this as well; however, the linguistic clues they provide for the interpretation of meaning can be insufficient or confusing.

The seventh type of CLI, described above in Part 2, is pragmatic failure. As discussed in Cenoz (2003), this phenomenon is most noticeable in questions, particularly requests. Blum-Kulka, House & Kasper (1989) have identified five elements that English NSs use to modify requests with the purpose of reducing their impositive effect. They are: **alerters**, **request strategies**, **downgraders** (which may be either syntactic or lexical), and **mitigating supportives**. Alerters are attention-drawing elements like names, titles, offensive or endearing terms, etc. Request strategies in English are conventionally indirect, including such statements as *I'd like to*, *how about*, and *could I*, but may also include the use of the imperative (Cenoz, 2003). Syntactic downgraders modify the request through the use of past tense (e.g., *I was wondering*), conditional clauses, etc., while lexical downgraders modify it with words or phrases like *please*, *I'm afraid*, and *you know*. Finally,

mitigating supportives are used to prepare the listener (e.g., *I'd like to ask*), justify the reason for the request, and/or promise a reward for fulfilling the request. Monolingual English speakers usually include at least some of these elements, and leaving them out intentionally is considered impositive and face-threatening (Cenoz, 2003). For this reason the pragmatic failure of multilinguals to include them can unintentionally give the impression of rudeness, and is therefore potentially more likely to cause miscommunication than any other form of CLI.

Finally, it should be mentioned that the seven categories of CLI just discussed are not mutually exclusive. Code-switching, for example, can occur in the middle of almost any other type of CLI, as well. For this reason, some examples will appear more than once, under different headings.

3.2 CLI in Speech Production

This section examines data from the spoken corpus collected for this dissertation. The corpus was collected in the spring of 2008. Students were asked to speak for up to five minutes about simple topics (e.g., "Why I (don't) have a pet," "Sleeping," "My dream house," etc.) and then answer any questions posed by the audience. They did not know their topics in advance, drawing them randomly from a pile, and were thus required to begin speaking with no preparation at all. Moreover, they were made aware of the fact that they were being recorded and that their recordings would be analyzed in this dissertation. This was intentional for two reasons: first, so as not to record anyone without their informed consent; second, it was hoped that the high-stress situation thus created would elicit production at the opposite extreme from writing (which will be analyzed in Part 4). Indeed, in the spoken corpus can be found the kinds of unusual uses of language and CLI phenomena that appear when students are most pressured and have least time to think through the language options available to them. Over the course of the semester 36 students were recorded, and, when their speeches were transcribed, the resulting corpus totaled approximately 20,000 words. Of these,

half (18 speeches) were selected for analysis on the basis of the students' self-reported nationality and native language. Again, only those students for whom L1 = Lithuanian were analyzed. Some information relating to the spoken corpus is presented in Table 6.

Table 6. Summary of the AnRK spoken corpus.

Total number of speeches	**36**	Total words, all participants	**20,688**
Total length (hours)	**3:05:29**	Total words, speakers only	**15,789**
Longest speech (minutes)	**10:02**	Longest speech (words)	**896**
Shortest speech (minutes)	**2:10**	Shortest speech (words)	**180**
Average length (minutes)	**5:09**	Average number words per speech	**439**

3.2.1 Hesitation

As discussed above, Maclay & Osgood (1959) identified four types of hesitation: filled pauses (FP), unfilled pauses (UP), repeats (RP) and false starts (FS). Example (1), a 34-second extract from one speech, exhibits all four types of hesitation at once.

(1) and **um** …(1.5) I **um-** for me it was **uh** too hard to understand how people **uh** can **uh** ^ ^ live **uh** with **uh** ancie- **uh** with **uh** ancient condi- **uh** …(1.5) conditions **uh** how they **uh uh** what they **uh** eat how they **uh** what they *what what* they **m m:** …(2.1) how they mind or: what they do [I303[58]]

In this extract there are 17 FPs[59] (marked in bold), three UPs of 5.1 seconds' total duration, one RP (italic), and six FSs (underlined). The FSs make up a total of 9 more words or interrupted word fragments. Thus in (1) FPs account for 17 of 57 words, or almost 30% of the passage. Yet if the RP and FSs are included, the total grows to 28 of 57 words, or 49%. Put another way, almost

[58] Each example is followed by its coded number in the spoken corpus.

[59] Throat clearing, marked in this dissertation with the symbol ^, was not added to the total number of FP. It does, however, often seem to be another form of FP, as do some forms of laughter (marked with @ in the examples). Throat clearing is not mentioned in the hesitation literature, nor is laughter treated as hesitation, hence their exclusion from the counts in this dissertation.

every other word in this extract is some form of hesitation, not counting UPs at all.

This example is rather extreme, of course, and most student speech is less hesitant than this. It is not the aim of this section to analyze student hesitations in any great detail, but rather simply to demonstrate that they are common and can be viewed as evidence of CLI. Frequency was therefore calculated only for FPs in the speech production of the Lithuanian sub-group, and is presented in Table 7.

Table 7. Filled pauses (FP) in speeches of Lithuanian sub-group.

Total number of speeches	**18**	Most FP in one speech**	**62/321 = 19.3%**
Total number of FP	**874**	Least FP in one speech	**11/297 = 3.7%**
Most FP in one speech*	**64**		
Least FP in one speech*	**11**	Most frequent FP***	**157/62 = 2.5s**
Average number of FP per speech*	**37**	Least frequent FP	**187/12 = 15.6s**

* These numbers total only FPs produced by the designated speakers; audience contributions were not included in the totals.
** (FP/words) = percentage of all words that are FPs
*** (seconds/FP) = average number of seconds between FPs

Now consider the extract in (1) with the hesitations removed. Example (2) leaves only those words which directly contribute to furthering the thought being expressed:

(2) for me it was too hard to understand how people can live with ancient conditions, what they eat, how they mind or what they do

Example (2) is less than half the length of (1) and significantly easier to comprehend; however, it is still marred by serious errors of syntax, grammar, vocabulary, and even pronunciation (e.g., the word *ancient*, which in this passage was pronounced /ˈ**ænsənt**/). All of this indicates that for this student, at least, the effort of producing English speech is nearly insurmountable and comes at great cognitive cost.

Similar examples can be found throughout the corpus. In example (3), UPs and RPs again indicate use of attentional resources during online processing.

(3) …(1.8) and uh for us they are the …(3.6) m: m the: the the the: lovest uh …(0.8) uh people in the world [I213]

A pause of 3.6 seconds is extremely long. Were such a pause to occur during normal conversation, it would be a signal to other participants that the speaker was relinquishing control of the conversational turn. In the formal classroom speech scenario, however, she is able to continue her thought after the pause because the audience is unwilling to interrupt, perhaps for fear of reprisal from the teacher. After such a lengthy pause one might expect a well formed utterance to emerge; however, the speaker follows it up with six essentially meaningless words (of which four are RPs), only to finally produce a nontransparent neologism, *lovest*. This word is itself followed immediately by more hesitation. Notice also the large number of RPs of the word *the*. In Maclay & Osgood (1959) it was found that RPs typically involve function words, a finding borne up by the examples seen so far. That study focused on monolinguals, however. In this multilingual corpus, it appears that FPs themselves are also repeated: *m* immediately, and *uh* after a short UP.

Hesitations such as these are clear indications of CLI. In both examples, it appears that the students are searching their lexicons for appropriate English words to express the L1 concepts they wish to convey. Interestingly, in both cases they end up making lexical errors (*how they mind*, probably a nonce-translation of LT *kaip jie mąsto* "how they think"; and *lovest*, an attempt to express LT *mieliausi* "the dearest"). Despite some significant stalling, neither student is able to arrive at an acceptable English formulation of her thoughts. This indicates that online production is still extremely resource-costly for these students, and supports the contention in Part 2 that students continue to think in their L1 while producing FL utterances.

In writing production, only the final product is seen – example (2) might approximate what the student would have produced had she been given the time needed to organize her thoughts. In speech production, conversely, the

processes of conceptualization, grammatical or phonological encoding, lemma or lexeme access, and so forth can be glimpsed through hesitations such as those seen in (1) and (3). In (3), for example, the passage *m: m the: the the the: lovest* clearly indicates trouble with lexicalization: for this student a personal lexical gap exists, which she is unable to fill with any words from her lexicon. The LT word *mieliausi* "the dearest" is derived from *meilė* "love," so the student quite logically attempts to derive an English word by the same reasoning. The result is, of course, unacceptable by English norms, but again, it is only accidentally an "error." It should rather be viewed as a particularly creative solution to a communicative problem, one that, of course, only a multilingual experiencing CLI from her native language could possibly invent.

3.2.2 Code-switching

When faced with a lexical gap during online production, multilinguals must quickly decide how to deal with it. One option, as seen in examples (1) and (3), is to engage in a prolonged search of the lexicon. A second, more expedient option is to simply insert the L1 word that expresses the concept which cannot be found in the FL, and continue speaking. This is both time- and attentional resource-efficient, and has the benefit of leaving the speech processors free to continue their incremental planning without buffering.

The code-switches found in the corpus can be divided into four common types: 1) **interjections**; 2) **unneutralized lexical gaps**[60] which students fill with an L1 word – this is done mostly as a resource-efficient time-saving device, though in one case (the example of *skola*; see discussion below) their teacher may be to blame for it; 3) requests for translation; and 4) **asides**, here used to mean L1 comments usually directed at the speaker from the audience.

The first category is interjections, as demonstrated in examples (4)–(6):

[60] These are not to be confused with borrowings, which are accepted into the English NS speech community.

(4) R: *Oi*. What I hated most this year at KHF.
O: Yesterday. Filming.
R: *Ne*:, it was uh beautiful boys. [I209]

(5) um: …(1.1) uh *ai* and the maybe the most uh thing that I really hate uh is eight o'clock lectures [I308]

(6) R: H should people go to Mars.
T: Yeah.
C: *Ei*, not- not this topic. I have- topic on my- [I407]

Other interjections found in the corpus, but not shown in these examples, are *nu*, *fu*, *a*, and *oho*.

In example (4) student R has just drawn her topic at random. Her use of the interjection *oi* indicates that she may have been hoping for a different topic. She reads the topic aloud and immediately an audience member suggests a possible situation with which to begin. However, R does not accept the suggestion, negating it with the Lithuanian *ne* and then explaining why the suggestion was inappropriate.

Example (5) begins with two FPs separated by a UP, showing that the speaker is in the middle of active online processing. Her use of *ai* may indicate the sudden appearance of a solution to whatever linguistic problem she was trying to solve; indeed, the next FP does not appear until 2.75 seconds (six words) later.

In example (6), student R reads her topic aloud, but it turns out to be similar to the topic of student C's prepared presentation, which C intends to present after the speeches are finished. C therefore jumps in quickly to halt the speech before it has properly begun, requesting that R be allowed to choose a different topic for her speech.

Why do students switch to L1 for these interjections? There appear to be three main reasons. First, there may be no English equivalents for certain Lithuanian interjections, especially *oi* and *nu*. In English it is possible to say *oy*, which is an accepted borrowing from Yiddish, but the use of this term is

restricted unless the speaker is actually Jewish; and the interjection *nu* can be used in so many different contexts to mean so many different things that it cannot possibly be identified with a single English equivalent. Alternatively, students may not be familiar with the English equivalents. The English equivalent of *ai* in the context of example (5) is probably *oh*; it is impossible to know whether the speaker was familiar with that usage of *oh*, but the possibility remains that she was not[61]. The most likely reason of all may be resonance: interjections are almost never studied and practiced in language classrooms, and therefore may never be consciously learned by students. Students thus continue to use their L1 for interjections long after reaching rather high levels of proficiency in the production of those standard grammar and vocabulary items that are taught in classrooms, as Lithuanian interjections are vastly more resonant for them than English ones.

The second category of code-switching is unneutralized lexical gaps. Here is an example:

(7) my friends uh uh …(2.8) *ištraukia iš namų* @ [I211]

In this case the speaker seems to make an attempt (evidenced by the RP) to invest the effort necessary to recall an English phrase that would be equivalent to *ištraukia iš namų* "get (me) out of the house," but fails. Either she is unable to find one or has not yet learned it. She does not, however, try to elicit a solution from the teacher or audience, but simply code-switches. Interestingly, she acknowledges (through laughter) the fact that it is a code-switch. (Compare examples (4)–(6), in which the L1 interjections are unacknowledged in any way, suggesting that they may be produced without conscious awareness by the speaker.)

Similar examples can be seen in (8) and (9):

[61] Students often use *oh* in phrases like *oh my God* or *oh, no*, but may be less familiar with its use as a stand-alone interjection.

(8) So when you will be rich and middle-age woman you will choose to- …(1.3) to ***pasamdyti*** maid? @ [I304]

(9) But some dogs are- like ***pitbuliai*** are very angry and I don't see a point uh how- why to keep them. [I305]

Example (8) follows the same pattern as (7). Notice that in both examples, the speakers pause significantly before the code-switch. This indicates that both attempted a cursory search of their English lexicons for the appropriate word or phrase but, finding them missing, proceeded with the L1 switches seen here. Similarly, both speakers follow their switches with laughter. In (9), however, the switch is essentially unacknowledged by the speaker. It is true that the switch in (9) follows a break in the train of thought at the word *are*, but this FS is probably not due to an attempt to recall the English word *pit bull*. It seems as though the student began the sentence with the vague formulation *some dogs*, then decided to include an example of which kinds of dogs she had in mind; that this example appeared as a code-switch can be attributed therefore to its appearance in the wider context as an afterthought.

Special mention should be made of the Lithuanian term *skola*, seen in examples (10) and (11):

(10) U: You maybe hate her just because she gives you that ***skola***.
R: [No, no, I haven't ***sko***-]
E: [What ***skola***? I have ***skola***! @]
R: I haven't ***skola***, I just didn't pass my exam for her. [I209]

(11) I don't like uh that I have ***skola***, but I think I will correct this mistake [I211]

In these examples students are using the term *skola* "a debt" in its academic meaning, to refer to the fact that they have one (or more) incomplete marks for subjects taken the previous semester. Unfortunately, it seems likely that their teacher is at least partially responsible for this code-switch. The term had come up during one lesson, and while the students granted that *debt* as an equivalent

was unacceptable, they were unable to come to agreement on an appropriate alternative. Their teacher suggested the term *an incomplete,* but admitted that he uses the code-switch *skola* himself. As he was the one recording their speeches, it is likely that the students therefore considered *skola* to have been officially "borrowed" into the speech community of his classroom.

There is one other kind of code-switch found in this category of unneutralized lexical gaps. Consider example (12):

(12) I was watching TV, uh the program was *Stilius* [I310]

The names of television shows and magazines pose a special problem when speaking a foreign language. American television shows broadcast on Lithuanian channels are translated and dubbed, as are their names: *Friends* becomes *Draugai*, *The Simpsons* becomes *Simpsonai*, *E.R.* becomes *Ligoninės priimamasis*, etc. Yet students are unwilling to reverse the process when speaking English. *Stilius* should, by the same logic, become *Style*. That it does not seems to indicate that the speaker is afraid of being misunderstood. Perhaps she is worried that if she translates the name, her listeners might not understand whether the television show is Lithuanian or foreign.

The third category of code-switching found in the spoken corpus is requests for translation. The following words and phrases were all requested, either from the teacher or from the audience: *priedai* "additives," *santykiai* "relationship," *paskirstyti* "share," *persikų* "peach," *tarybiniai* "Soviet," *iš pirmo karto* "on the first try," *aplinka* "environment," *sutrikimai* "dysfunctions," *gaila* "feel sorry for," *vilkšunis* "wolfhound," *persų* "Persian," *kilmės* "breed," *kančia* "suffering," *grikiai* "buckwheat," and *zodiako ženklas* "Zodiac sign." Although the majority of these requests were indicated only through intonation (typically expressed as either a rising tone or a whisper), they are occasionally framed as requests, sometimes in Lithuanian, and sometimes in English, as in example (13):

(13) N: she was eating only ten uh big uh spoons s:- for soup, m of *grikiai.*
T: M, m-hm.
N: Uh how it?
T: Buckwheat. [I310]

Because the majority of request-for-translation code-switches are unframed, students are prepared to offer translations with very little prompting – even when no request is actually made. Consider example (14):

(14) L: I keep two *papūga*s.
U: Parrots.
L: At dorm I keep. But I like cats. [I212]

Not only does student L not request the translation offered by U, she ignores it and continues her train of thought as if U hadn't spoken at all. Sometimes no code-switch is even required for a student to suggest a translation. It is almost as if some members of the audience are attempting to anticipate potential code-switches, or are perhaps interpreting hesitations as evidence of lexicon-searching. This can happen on the basis of very little evidence, as in example (15):

(15) G: Whatever. ^ It is not so so-
T: Important. @
G: important. I guess. [I306]

Here, student G is interrupted by T's suggestion, even though there is only one RP to suggest that G was in need of help. It does seem rather unlikely that G would have been unable to produce this particular word without prompting. Note also that T laughs after her suggestion, as if aware of the fact that it could be interpreted as presumptuous.

The fourth and final type of code-switching, asides, can also be distracting to speakers, especially in the context of a recorded classroom

assignment. In example (16), student Y seems momentarily unable to understand a question posed in Lithuanian:

(16) Y: Curtains. Also flowers on the- on by the window.
E: *Tai o ko jisai geresnis?*
Y: What? [I208]

In this case it may be that the task itself has caused Y, the primary speaker, to push herself towards a monolingual English mode, in which L1 is inhibited as much as possible. For this reason E's L1 question ("So why is it better?"), which may not obviously relate to what she was just saying, requires more time to process than usual, and Y buys this time with a request for repetition. While some asides can thus be seen to have a disrupting effect on communication, others are easily accepted, as example (17) shows:

(17) O: it was very good weather, it was s: shiny, sunny @
R: *Naujus žodžius randa*.
O: Yeah, ^ but I like always was the first [I203]

In this case speaker O acknowledges R's aside ("She's finding new words") with a simple *yeah* and then continues her story where she left off. Note also that in both (16) and (17) the switches are full sentences. This suggests that the students who produced them, being members of the audience, feel in some way sanctioned to use L1 more freely than the speakers; after all, it is the speakers who will receive marks for their speeches, not the audience. This frees the audience members to relax back into a fully bilingual mode in which English and Lithuanian are interchangeable. This may also be a manifestation of the phenomenon of private speech, discussed in Part 2. The audience members, not being "on the spot," direct their asides to the speakers (or, sometimes, each other) in L1, exactly as they would during a lecture by the teacher. The only difference is that, in a lecture situation, these private asides can be whispered,

and thus remain truly private. During a student's speech, the asides have to be loud enough to hear – thus turning them into public code-switches.

Code-switching, then, is especially common in the spoken production of multilingual FL students. This feature of the spoken corpus points to the suggestion that the monolingual language mode is nearly impossible to establish in the FL classroom. Certain linguistic features, such as interjections, seem to escape students' conscious awareness, appearing in their L1 forms without any surrounding hesitations. Unneutralized lexical gaps, on the other hand, are often preceded by pauses and followed by laughter, sure signs that the speakers recognize them as code-switches but, for various reasons, are unable or unwilling to invest the resources required to produce them in English. One such reason could be habit: as discussed in Part 2, code-switching is frequent in the predominantly Lithuanian/Russian sociocultural environment they come from[62]. This naturally translates to code-switching in the classroom, especially when (as in this group) it is known that the teacher is multilingual and/or tolerant of the practice.

When students have attentional resources to spare, they often pause after such switches or include them within the frame (either L1 or English) of a request. Their listeners seem eager to provide information, perhaps gaining some self-confidence from the knowledge that they are able to recall a word another student cannot; however, this eagerness occasionally overextends to the point of suggesting translations even when they are not requested. Finally, audience members themselves feel free to direct questions or comments at the speaker in L1, which can sometimes be disruptive. It may be assumed that within the context of the impromptu speech assignment that elicited this corpus, the audience members felt more secure in switching to L1 than the speakers did. Speakers use switches only as a last resort.

[62] And possibly in the English language classrooms they studied in at school. One former student is now an English teacher, and, in her own words, "teaches English in Lithuanian." In other words, her students "learn English grammar" by discussing the rules in Lithuanian.

3.2.3 Neologisms and foreignizing

In this section the term neologism is used to mean any English-like word or phrase which the student essentially invents on the spot. These words can be taken as examples of the state of the students' English interlanguage, as they show evidence of a certain (albeit shaky) grasp of the principles of word derivation in English. By foreignisms are meant those cases where an L1 word has been adapted to English by means of phonological and/or morphological change. Foreignizing is thus a particular kind of code-switching, one in which the switched words are adapted to the rules of English, while neologisms are created mostly from the students' English knowledge. Thus foreignisms are based mostly on the L1, while neologisms are based mostly on the FL.

In Cenoz (2001), foreignisms were found to be rare; less than 25% of her sampled Basque children used them in their speech. In this corpus, they are even less common: only 2 students (14%) produced them at all, and even then only once each. As such, they represent 2 words of a total of 15,789, or 0.01% of the total. However, despite their low frequency – or perhaps because of it – they are interesting examples of CLI, as they can only be created by the interaction of the two languages. Examples (18) and (19) reproduce both instances:

(18) Because eh they look like skelets or something @ else, [I206]

(19) I keep two *papūga*s. [I212]

There are some words, particularly proper names and the category sometimes called "international words," which appear remarkably similar in Lithuanian and English: consider, for example, *telefonas ~ telephone*, *cigaretė ~ cigarette*, and *robotas ~ robot*. In all of these cases, the main morphological difference is the addition, in Lithuanian, of a masculine or feminine suffix to the root. In example (18), then, the student has assumed that an English

equivalent to the Lithuanian *skeletas* can be formed by dropping the masculine morpheme. In (18) the foreignism *skelet* is then given an English plural morpheme, and the lack of hesitations immediately surrounding the word indicate that the speaker is convinced of its appropriacy.

In (19), a different student has solved a lexical gap by simply attaching an English plural morpheme to an L1 word. The pronunciation is English-appropriate: /**pəˈpu:gʌz**/, rather than the Lithuanian plural form, *papūgos* /**pʌˈpu:gɒs**/. Two phonological adaptations indicate that the word has been borrowed for this use: /**ə**/ rather than /**ʌ**/ in the first syllable, and /**ʌz**/ rather than /**ɒs**/ in the final syllable. Interestingly, as seen above in example (14), the student herself seems unflustered by this borrowing, and does not acknowledge her classmates' English translation.

Examples (20) and (21) show students' faulty use of English derivational suffixes in the creation of neologisms:

(20) I think that it is n- not fair and uh it means hu- these people are- not humanic [I305]

(21) In a- in a forest and wilderning [I208]

In these examples, students have attempted to derive the words *humane* and *wilderness*. Here the words are not foreignisms, but neologisms: these are nonce inventions based on partial knowledge of English. The suffixes *-ic* and *-ing* are relatively common, at least in comparison to those they have replaced (*-e* and *-ness* respectively). These examples demonstrate that some English derivational patterns are beginning to establish themselves in these students' developing language systems, but have yet to become reliable templates for the creation of new words.

In the next example, an English word is reformulated as a different part of speech:

(22) I like this newspaper, uh and I really uh read it uh when I get it and uh uh uh I don't know it's not quite uh large or long, it's uh- there ^ a lot of ^ ^ interesting events but they are uh uh written very- how to say, uh- economically? @ Uh little uh <u>smalls</u> maybe sometimes uh um m tales [I407]

Here the topic is the free daily newspaper *15 Minutes*, and the student is apparently trying to express her appreciation for the brevity of the articles. Two forces may be at work in the creation of the neologism *smalls*. First, students are surely aware that many English words can be used as different parts of speech with no morphological change in the root: *record*, for example, can be both a noun and a verb; *present* can be a noun, a verb, and an adjective. No rules seem able to explain why this is true of some words but not others, and students simply have to learn them on a case-by-case basis. At the same time, the Lithuanian word *mažiukas* "a small (one)" is a countable noun. Working together, these two facts can easily explain how the speaker of (22) could derive the countable noun *smalls*.

Both foreignisms and neologisms are clear examples of CLI. Such derivations would be difficult, if not impossible, to imagine coming from monolingual speakers. Moreover, while the neologisms can be attributed to the developing state of the speakers' interlanguage, the foreignisms are necessarily language-specific, being based purely on L1. In other words, any speaker with any L1 could also be found creating the neologisms in examples (20)–(22), as they are based in part on the very rules of English derivation that they flaunt. However, the foreignisms seen in examples (18) and (19) come directly from Lithuanian. Although *skeleton* in Russian is скелет and in French *squelette*, in both of these languages the stress is on the second syllable; in Lithuanian, it is on the first, /ˈ**skɛlɛtʌs**/. The foreignism *skelets* was pronounced /ˈ**skɛlɛts**/. This is not to say that Russian-speakers or French-speakers couldn't produce something like *skelets* /**skɛˈlɛts**/. However, speakers of Spanish, for whom *skeleton* is *esqueleto* /**ɛskɛˈleto**/, would be less likely to do so. Thus the production of foreignisms is highly dependent on the L1.

3.2.4 Transfer

Transfer is a well known phenomenon in which elements of the L1 can be found in L2 production; it can happen at all linguistic levels and processing stages, from pragmatics down to phonology and from conceptualization down to articulation, as suggested by MacWhinney's (2005: 55) slogan, "whatever can transfer will." Kilborn (1994), working within the perspective of the competition model, identifies four types of transfer, which he terms **forward transfer**, **backward transfer**, **amalgamation**, and **differentiation**. As he suggests, these probably represent a trajectory in the process of acquiring an L2. In forward transfer, L1 strategies are applied to the L2, often inappropriately; these will be the main focus of the discussion here and in Part 4. In backward transfer, L2 strategies may be found to affect the L1: such processes have been studied in depth in, e.g., Kecskes & Papp (2000), Cenoz et al. (2001), and Cook (2003). In amalgamation, a set of strategies neither exclusively L1- or L2-based is created from both language systems; this suggestion seems borne up by the idea of neutralized concepts in the CUCB, as developed in Kecskes & Papp (2000). Finally, differentiation leads to the use of different language-specific strategies with different language systems, e.g., L1 strategies for L1 and L2 strategies for L2. This situation, according to more recent research (especially Cook, 2003, Grosjean, 2001, and Herdina & Jessner, 2002) seems to be extremely unlikely in practice, due to the dynamic nature of the language systems themselves. Differentiation probably represents an idealized "multilingual" system from the point of view of monolingual theorists who find such systems more parsimonious.

For an idea of the nature of forward transfer, consider examples (23) and (24):

(23) we just uh <u>with friend</u> m were reading some uh magazine [I310]

(24) G: I said that it isn't a good idea, yeah, because children have damage.
R: But they will have damage and the other way, too- [I213]

Examples such as these will be called learner constructions after Waara (2004: 53), where it is stated that "[a] learner construction is a construction, i.e., a meaning and syntax correspondence, but which is used in a slightly unconventional manner. Although usage does not result in a communication breakdown between participants, it deviates in some way." According to Waara (2004: 69), such constructions "reflect elements of transfer, blending, and overgeneralizations of the developing L2 system." As such, they are clearly the result of learners' unfamiliarity with the language- and culture-specific patterns and structures of the FL they are trying to speak. In Part 4 it will be argued that conceptual blending can, in fact, account for a great deal of the transfer data found in student production, whether spoken or written. This section will simply show that under the pressure of online production, transfer occurs as a form of CLI, and that its main purposes are a) to provide structural frameworks on which to hang L2 lexical items, and b) to fill lexical gaps. Transfer of phonology will be discussed in the next section, and pragmatic transfer will be taken up after that.

Returning to example (23), the construction *we with friend* is this student's way of expressing the common Lithuanian construction *mes su drauge* "my friend and I." As such it is a clear example of forward transfer. It should be emphasized that while the result is non-standard in English, it is nonetheless a solution to the essential problem of communication, namely, the need to use words to express concepts. As seen in Part 1, concepts are layered knowledge structures that are linked via lemmas to lexemes in a large spreading activation network. As discussed in Evans (2006), lemmas (or lexical concepts) can be overt (in the case of lexical items) or implicit (in the case of constructions). Thus, the pseudo-English expression *we with friend* demonstrates transfer of the implicit L1 (Lithuanian) construction *mes su*

drauge. This structure is simply filled in with FL lexemes. The same is true of the use of *and* in *they will have damage and the other way* (instead of: *they will also have damage the other way* or *they will have damage the other way, too*) in example (24). In this case, the L1 (Lithuanian) use of *ir* "and" in situations where English prefers *too* or *also* reflects forward transfer as well. In (23), the use of the FPs *uh* and *m* indicates that the speaker may be unsure of the accuracy of her learner construction; in (24), however, there are no FPs at all, which suggests that this speaker does not see her construction as erroneous in any way. Neither "error" is corrected. As discussed in Part 1, this is probably because these students do not see them as errors; recall that the monitor is itself part of the same conceptualizer (CUCB) for which the L1 constructions being transferred are natural and, more importantly, resonant means of expressing concepts.

Now consider example (25):

(25) I more like sit at home uh but I don't uh sometimes have some opportunity [I211]

The underlined clause is difficult to understand in English, because the status of *like* is in doubt: it is not immediately apparent whether it is being used as a preposition or a verb. In Lithuanian there is no such problem: *Man labiau patinka sėdėti namie*. A better expression of this thought in English could be achieved with *I'd rather sit at home* or *I prefer sitting at home*[63]. In any case, by transferring an L1 construction into English and filling it with English words, the speaker has produced a learner construction which makes her utterance difficult (though not impossible) to understand.

The construction in (26), on the other hand, might require more effort to decipher than an English NS interlocutor would be willing to invest:

[63] *Sit*, of course, should probably be replaced with *stay*.

(26) No, eh usually before steli- stelerization, @ uh people give one time eh cat children. And then. [I305]

The discussion was about the question of neutering cats, and the student wanted to say that cats are usually given the opportunity to have children once before being neutered. Leaving aside her mispronunciation of the word *sterilization*, consider how the underlined phrase could sound in Lithuanian: *žmonės duoda vieną kartą katei vaikus*. In Lithuanian, of course, the indirect object *katei* is in the dative case while the direct object *vaikus* is accusative. In English, with its lack of cases, word order and prepositions must do the work of the Lithuanian cases. Moreover, the verb *give* is inappropriate in this case, as the cat itself produces the children in question; it can, as mentioned above, be given the opportunity to have them, but a more appropriate lexical choice here would be the verb *let*. The English version should thus be expressed as *people let the cat have children one time*. Again, by transferring the Lithuanian model and filling it with English words, the student has created a sentence that may be prohibitively difficult to understand for anyone unfamiliar with Lithuanian word-order patterns.

Example (27) shows one more common transfer of structure:

(27) And uh it's very good diet. When you eat for- yes, for five months, from the September till the January. And very helped. [I310]

The use of *very* follows the Lithuanian pattern precisely: *Ir labai padėjo*. Here, a correct English expression with *very* would be *And it helped very much*; or, to follow the Lithuanian pattern more closely, *very* could be replaced by *really*, as in *And it really helped*. (27) is not a particularly serious error; while it is a mismatch, it is quite understandable. Interestingly, the word *it* is also missing from the utterance, where it should be in the subject position. This reflects the forward transfer of another L1 implicit structure, namely, the fact that Lithuanian is a null-subject language whose verb morphology is complex and allows subjects to be dropped. In English, of course, *it* needs to be expressed

overtly. Yet even with these "errors" present, the example is understandable and unlikely to cause miscommunication. This in itself is problematic for students of English as an FL. If everyone who hears incorrect formulations like these nevertheless understands them and responds to their semantic content, without reacting to their erroneous syntax, the speakers may never learn that they have said anything unusual.

It was stated above that transfer is also used to fill lexical gaps, as the next two examples demonstrate:

(28) you uh are trying to catch uh it uh while driving a trolleybus or bus [I407]

(29) Oh, it- it depends on the- on the- …(1.6) human who is keeping it. [I305]

In (28), the speaker uses *driving* to mean *riding*. In Lithuanian the verb *važiuoti* is polysemous, and can mean both *drive* (as in *važiavau namo* "I drove home" or *važiuojam į Vilnių* "let's drive to Vilnius") and *ride* (as in *atvažiuoju į mokyklą troleibusu* "I ride to school by trolleybus"). In English, of course, the two words have different meanings, and the result in (28) is miscommunication: the student unintentionally implies that she is a bus driver. In (29) a different student translates the Lithuanian *žmogus* "person; human being" as simply *human* (instead of: *person*), perhaps to distinguish human beings from the animals they keep (the topic of her speech). This translation appears frequently in AnRK learner constructions. The student herself seems to be aware that the choice may be inappropriate, as evidenced by the various hesitations preceding the word, but having produced it continues with her speech. She is uncorrected by the audience, again suggesting that some of these constructions may establish themselves as a side-effect of the FL environment. Students and their NNS teachers simply do not recognize them as erroneous – which indeed they are not according to the sociocultural and linguistic norms

of the L1. Only as compared to monolingual English NS norms do they become problematic.

3.2.5 Phonology

According to Scovel (2001), while non-native speakers of English may, with time, become remarkably proficient and accurate in terms of grammar, vocabulary, etc., they are never able to fully rid themselves of an accent when speaking. This section therefore takes it as given that many or even most English words are pronounced with a Lithuanian accent by these students. The discussion will focus on some of those instances where pronunciation and/or intonation become particularly troublesome for these students, to the point of hindering communication or even stopping the flow of conversation.

Before turning to such examples, however, consider one example of phonological transfer:

(30) they uh place is only- not at home, and only in the nature [I405]

In (30) the word *and* is pronounced /**ɒnd**/. It is not the case that the speaker always mispronounces this word; in her speech *and* is produced eleven times in all, seven of which are pronounced /**ænd**/, three more are pronounced /**ənd**/, and only this one is mispronounced. Now imagine if this were a Lithuanian sentence: *jų vieta yra tik- ne namie, o tik gamtoje*. The Lithuanian conjunction *o* "but, while, and" is pronounced /**ɒ**/ - the very pronunciation used for *and* in (30). In this example it seems as though the student momentarily transferred the Lithuanian pronunciation of *o* with the English word *and* to produce an error in pronunciation. That it is also an error of grammar – it should be *but*, not *and* – is also of note. As discussed in Part 1, the monitor is context-sensitive, meaning that it cannot monitor production for every aspect at once. In this example the student may be focusing mostly on macro- and microplanning. There is only one FP, suggesting that her primary concern in

this passage is not with lexicalization or pronunciation, but structure or content.

In (31), the English phoneme /**ð**/ seems to be interchangeable in this speaker's FL system with /**w**/. Neither phoneme, it should be noted, is part of the Lithuanian phonological system, so these troubles are essentially FL-specific.

(31) So I didn't say **that** they shouldn't divorce, I think **that** they- if you don't have to divorce, yeah, but **they** w- ha- uh **they** have to look for some solution to be together, but if it is impossible, so they should divorce. [I213]

In this example, bolded words are pronounced correctly (with /**ð**/), while underlined words have /**ð**/ replaced by /**w**/. Moreover, the diphthong /**eɪ**/ in two cases is replaced, once with the long vowel /**i:**/ and once with the short vowel /**ɪ**/, such that the three underlined words, in order, are pronounced /**wi:**/, /**weɪ**/, and /**wɪ**/. It is impossible to explain such apparently random phonological behavior, but it can safely be said that a monolingual English speaker would not exhibit such difficulties. Unfortunately, the effect of all this is potential miscommunication, as in two places it sounds as if she is saying *we* rather than *they*: in the first clause, *So I didn't say that "we" shouldn't divorce*, and in the last clause, *so "we" should divorce*. In context, where the pronoun is pronounced both correctly and incorrectly, it is unlikely to cause confusion; but were this student to produce only one utterance in which *they* is pronounced /**wi:**/, awkward misunderstanding could result.

In the next example, the speaker becomes unsure of her choice (or, perhaps, pronunciation) of a word and corrects herself; unfortunately, her "correction" is actually incorrect:

(32) they often have a green- green color outside [I208]

The underlined word is pronounced /**gɹɪn**/. The only explanation seems to be that under the pressure of online production, this student's monitor has become oversensitized. It is often the case that Russian speakers lengthen English vowels unnecessarily, e.g. pronouncing *live* as /**li:v**/ and *good* as /**gu:d**/. Although the speaker of (32) is Lithuanian, she studies in a speech community in which the Russian language plays a major role, as discussed in Part 2. It is thus possible that in this example she was momentarily unsure whether *green* should contain a long or short vowel, and corrected to short as a way of "playing it safe." Unfortunately, her effort is wasted as, of course, the correct pronunciation is /**gɹi:n**/.

A number of students while giving these speeches had difficulty pronouncing words that were new to them, and as a result the corpus contains several examples of a student's attempts to internalize the pronunciation of a newly learned word. In example (33), the student is speaking on the topic "Why I (don't) like dieting."

(33) N: Why I don't like- <u>dieting</u>? What does it mean?
T: Like or don't like. Uh dieting.
N: …(9.4) I don't like **dieting** because uh I think that is- it is unnecessary for me… <…> so I don't like uh <u>dieting</u>, because my sister was on the **diet**, and uh seen- she uh- a month she l- lose uh about uh ten kilos, she was uh eh eating some special **diet**… <…> I think that <u>dieting</u> is uh necessary for that people which uh have m m problems [I206]

As she reads the topic she comes across the word *dieting* but does not recognize it: she pronounces it /**di:tɪŋg**/. The teacher is very careful to provide a clear model for the problem word, enunciating slowly: /**daɪətɪŋ**/. Student N understands its spoken form, as she begins to speak on the topic (and not about something else), and even pronounces the word (marked in bold) correctly the first time. However, both uses of the word later in the speech (underlined) are incorrect, emerging as /**di:ɪtɪŋg**/. This is especially interesting as the word *diet*, used twice in the speech (bold), is pronounced correctly: /**daɪət**/. Apparently,

given the resource-demanding task of organizing and presenting her speech, student N is unable to analyze the new word *dieting* into its constituent parts. In effect, she treats it as a completely new word, unrelated to any others in her lexicon, and is thus unable to extend the pronunciation of the known word *diet* to the just-learned derivative *dieting*.

It was seen above that in these speeches, interjections (*ei*, *oi*, *nu*, etc.) are almost exclusively code-switches from L1. A related phenomenon may be the pronunciation of letter names. All learners of English have at some point memorized the English alphabet song and should therefore be able to use it to help them spell words out loud or name individual letters. This seems to be avoided, however. Students prefer to rely on L1 names for letters[64]. Perhaps recalling the alphabet song in the middle of online production is too troublesome and resource-costly; alternatively, it may seem somewhat childish, especially to young adult learners such as these. Examples (34) and (35) show how students pronounce letter names:

(34) With additives like with uh E and some numbers, and they are not very good. [I202]

(35) Okay, my topic is what I like most this year at KHF. [I211]

In (34) the student is talking about food additives, often marked on ingredient lists with code numbers beginning with the letter E. The student pronounces this as a Lithuanian letter, /**æ**/, instead of producing the English-appropriate /**i:**/. There are no letter names in English that sound anything like /**æ**/ – the only word that does is often written *aah!* and is an interjection expressing fright, pain, or surprise. Thus in (34) CLI leads to complete miscommunication. In (35), the letters KHF stand for Kaunas Faculty of Humanities, the Vilnius University faculty where these students study. The correct pronunciation would

[64] According to Likhachiova (personal communication), this is true also of Russian students who went to Lithuanian schools (i.e., where the language of instruction was Lithuanian): they spell Russian words using Lithuanian letter names.

be /**keɪ eɪʧ ɛf**/. What this student says is /**kə hɑʃ fʊ**/ – again, using the typical L1 alphabet pronunciation. Needless to say, this would be incomprehensible to anyone unfamiliar with the Lithuanian alphabet. These examples, together with the interjection data seen above, suggest that some types of production processes, such as interjecting, spelling out acronyms, and possibly counting[65], are so highly resonant that they overcome any attempts by FL lexemes (where phonological information is housed) to compete with them.

3.2.6 Interactional strategies

Interactional strategies rarely appear outside of a spoken corpus because they assume the presence of an audience with whom the speaker is interacting; and while certain written media exist that mirror speech to a small degree (e.g., letters, in which an imagined recipient is addressed more or less directly throughout), only in spoken interaction can a speaker receive immediate feedback on his or her interactional strategies. Such strategies in and of themselves are not erroneous, and may well be used by monolinguals, too, depending on the situation. However, as used by the students in this corpus, they can offer clues about how CLI affects online production.

Two kinds of interactional strategies related to CLI are present in the corpus. The first is elicitation or requesting help. In this corpus such appeals to the teacher or audience occur with and without code-switching; this section discusses those cases in which production remains entirely in the FL. Moreover, elicitation of unknown vocabulary is not the only purpose of this strategy: students also use it to elicit the complicity of their fellow students in reducing the difficulty of the speaking task. The other form of interactional strategy is called in this dissertation **assumed understanding**; in such cases, speakers forge ahead with grammatically or lexically inappropriate utterances, assuming (or, perhaps, hoping) that they have been understood. Such assumptions may be stated overtly or demonstrated in context.

Example (36) was recently seen in example (33):

[65] Informal observation.

(36) N: Why I don't like- dieting? What does it mean?
T: Like or don't like. Uh dieting.
N: …(9.4) I don't like dieting because… [I206]

Here the speaker, confronted with a word she does not recognize, immediately asks for clarification. Interestingly, she does not receive it. The teacher assumes that the problem is one of pronunciation, as seen above, and repeats the word correctly, rather than going into an explanation of what it means to diet. Student N then proceeds with her topic, proving that the teacher's assumption was correct. In this case, the student's problem seems to stem from orthography: she imagines the written word (seen on the topic card) with an unfamiliar pronunciation and does not recognize it; but when given the correct pronunciation, she recognizes the word and understands the topic.

As mentioned above, students also attempt to elicit the complicity of their fellow students in easing the speaking task. Examples (37) and (38) show how students can avoid being asked any questions at the end of a speech:

(37) T: No questions for Betty?
B: No! [I203]

(38) T: Uh any questions for her?
I: No questions please. [I204]

In both cases, not surprisingly, the students were not asked any questions, despite the teacher's explicit instructions to the audience (prior to the recording) to do so.

The next examples demonstrate assumed understanding on the part of the speakers. In these examples, syntactic and lexical nontransparency make the speakers' utterances particularly difficult to decipher. In (39) the speaker is describing a situation in which she was at a bar with two gay men:

(39) they even just- I don't know, they even just touched each- each other for, I don't know, as a girls, you know, then uh taking- by the hand or going or something, normal girls, but they even just touch each other. [I403]

In this description, false starts, repeats, and the insertion of the (essentially meaningless) phrases *I don't know* and *you know* fragment the syntax and make the logical structure very difficult to follow. This is coupled with the use of vague vocabulary (e.g., *going* and *something*). The key phrase, repeated three times, seems to be *they even just touched each other*, though it is not at all clear what exactly this means. Despite these drawbacks, no one in the audience admits to being unable to follow the train of thought. It thus seems as though speakers are justified in assuming that the audience understands them; even if their utterances are incomprehensible, they are only rarely questioned about them.

In example (40), seen above in (16), the speaker describes a location using a nontransparent prepositional phrase:

(40) Also flowers on the- on by the window. [I208]

By the window can, with some effort, be understood to mean *on the sill*. However, even in Lithuanian there is a large difference, both conceptually and grammatically, between *prie langо* "by/near the window" and *ant palangės* "on the sill". This phrase follows an FS, indicating that the speaker began with the intention of saying *on the sill* but quickly realized that she could not recall or did not know the word *sill*. She therefore has to switch to the alternative *by the window*, having already said *on*, leaving it up to her listeners to understand what she meant.

In example (41), student L is explicitly led to believe that the audience understands her meaning:

(41) L: Um I think that um if you a- just to become vegetarian only for- for your health so uh you can fall back really, just um-
V: M-hm.
L: Just fall back,
V: M-hm.
<…>
T: Uh by fall back you mean: start eating meat [again]
L: [Yes.] [I310]

Here student V's repeated *m-hm* is taken by the speaker to indicate that she has been understood. Indeed, perhaps she has, for the term *fall back* is most likely a nonce-translation from the Lithuanian *atkristi* "relapse." This term is usually used to talk about someone becoming ill again after a short recovery, and in this sense L's construction is a colorful and creative description of returning to a meat diet after being a vegetarian – as if eating meat were a disease or an addiction. In any case, V seems to understand her intent, being a Lithuanian speaker herself. The teacher, an English NS, is less sure, and asks for confirmation.

Example (42) is the one case when a speaker's assumed understanding is explicitly shown to be false:

(42) J: I know a lot of uh uh people who play @ different games computer in such m **good age** @ uh so I think that m a lot of people: are addicted to their computers.
T: Okay. Uh, questions for Jenny?
L: Um, uh: who are you talking about when you speak about adults who play games in computer?
J: Um.
L: Who are you talking about? You said that you know some people, about- who plays-
J: *A*!
L: games.
J: Uh, my neighbors. @ Yeah! [I505]

The speaker uses the phrase *good age*, marked with a sarcastic intonational contour, to indicate something about the people she knows. Student L seems to assume that what J means is that these people are adults, and that in J's opinion

computer games are only for children. L is not convinced of this, however, and asks for confirmation. Speaker J seems not to understand the question. L then reformulates her question without the key word *adults*, apparently deciding that her initial assumption was wrong. Unfortunately, J gives a particularly uninformative answer, as it is still not at all clear whether the neighbors are adults or not. L then decides not to pursue this line of questioning.

Thus there may be two reasons why students develop assumed understanding. First, in many cases they are justified, as they share the same sociocultural background as their peers. As seen in Part 2, these students rarely have any opportunity to interact with English NSs; most of the time when they use English they are interacting with other NNSs who, like themselves, think in Lithuanian and are thus in a better position to understand the kinds of non-standard English constructions that they produce. Second, even if they are not understood, they are rarely questioned by their peers. This may be out of a sense of solidarity, as in those cases when students explicitly ask not to be questioned. Similarly, it might be socially proscribed to draw the teacher's attention to another student's incorrect language by asking questions about it. A third possibility is that an audience member may not want to admit that he or she has not understood something, under the assumption that if no one else is asking questions, he or she may be the only one who did not understand.

3.2.7 Pragmatic failures

As discussed above, the impositive effect of requesting is often mitigated by English NSs using such pragmatic elements as alerters, request strategies, syntactic and lexical downgraders, and mitigating supportives. In this corpus requests usually take the form of requests for information: as seen above, students ask for translations of L1 words (to fill lexical gaps), and sometimes ask each other for clarification or favors. This section looks at some examples of student requests from the point of view of their pragmatic failure to include many or most of the mitigating elements. It may be stated at the outset that in the majority of cases, none of the five elements were used at all. Most requests

are unmodified demands and, as such, would be considered especially rude by English NS interlocutors unfamiliar with Lithuanian sociocultural norms.

Consider example (43):

(43) L: she was eating only ten uh big uh spoons s: for soup, m of *grikiai*.
T: M, m-hm.
L: Uh how it?
T: Buckwheat. [I310]

In this example the student does not know the English equivalent of *grikiai*, and, perhaps, intends her code-switch to speak for itself as a kind of question. The teacher, however, simply murmurs to indicate that he has understood the switch. Student L must therefore ask him for a translation. Her question is completely unadorned with any English-appropriate requesting elements (e.g., *I wanted to ask*, *Could you tell me*, etc.). Such questions would not be inappropriate in Lithuanian, however, as the next example shows:

(44) R: Um: you- …(2.0) *kančia kaip*?
T: M?
R: *Kančia*.
T: *Ai*, it's uh- it suffers. [I305]

Here the student, trying to recall the word *suffer*, engages in a prolonged UP ending in a code-switch. The question *kančia kaip*? is literally the two words "suffering how," having the meaning of "how do you say 'suffering' in English." As can be seen, the teacher is unprepared for this question. The student takes the teacher's hesitation as a sign that he didn't hear the word *kančia* (her code-switched question is whispered, perhaps to distinguish it from the main flow of the speech itself), though it is also possible that, being an English NS, he is unprepared for such a pragmatically direct request.

Table 8 summarizes the data relating to the use of mitigators in requests by students in the spoken corpus. There were 34 requests in all, of which three

can be counted in two categories. While in theory any given request can be modified by more than one of the five elements discussed above, it is most often the case that requests are completely unadorned, as seen already in the examples above. The one alerter, *Ei*, was coupled with an imperative to produce a "request" (directed at the teacher) that would be considered especially rude by NS English standards. Indeed, even five requests were actually framed as imperatives, though one was coupled with a lexical downgrader in the form of *please*. In general, the students of AnRK, as might be expected, transfer L1-based norms in framing FL requests, which makes them appear direct to the point of rudeness from the point of view of English norms. Of all of the requests produced, only a handful are appropriate.

Table 8. Requesting behavior in the spoken corpus.

	Total (34)	Example(s)
unmodified	**24**	*Hm*?; *What*?; *So questions*?; *And- minutes*?; *What does it mean*?; *In America there is such, yeah*?
alerters	**1**	*Ei, not- not this topic.*
request strategies	**6**	*No*!; *Ask*!; *Are you recording already*?
syntactic downgraders	**2**	*Can I choose other one?*; *I can-?*
lexical downgraders	**2**	*No questions please*; *I hope that there aren't 'Why I liked Animal Dreams'*
mitigating supportives	**2**	*I don't know how in English*; *I can tell about my one drastic uh diet*

3.3 Summary and Discussion

Part 3 has analyzed seven categories of students' spoken production. It was shown that these samples of students' spoken production contain evidence of CLI. Moreover, the majority of these effects appear to be specific to spoken language: they simply do not appear in writing. This is because writing is an offline process. Between the words of a sentence written on the page, hours or days may have elapsed; there is no way to know. If speakers pause between the words of a sentence, however, they will lose their audience in minutes or, most likely, seconds: the speeds at which language is processed and spoken are

remarkable, and in everyday conversation the typical delay between turns in a conversation is about half a second (de Bot, 1992; Bull & Aylett, 1998; Dąbrowska, 2004). This puts great demands on the online processing systems. When the language of production is an FL, CLI appears in many forms. The seven kinds of CLI examined in this section were: hesitations, code-switching, neologisms and foreignizing, transfer, phonology, interactional strategies, and pragmatic failures.

There are many types of hesitations, from the prototypical *uh* to unfilled pauses, false starts, repeats, throat clearing, and so forth. All of these are ways of essentially buying time while the language production mechanisms work at a frantic pace. Hesitations probably occur no less frequently during the process of composition in writing than when speaking, but they are simply edited out of the final draft. Similarly, a student can pause when writing, refer to a dictionary, insert an appropriate English translation, and continue, thus masking the extent to which he or she is given to code-switching when speaking. In the same way, a neologism or foreignism can be double-checked against a dictionary before being included in a written draft. None of this is possible with speech.

When a lexical gap cannot be filled from current knowledge, speakers resort to a number of techniques in order to solve the problem. Perhaps the least cognitively demanding is to simply code-switch, filling the gap with a known word from another language (usually the L1). Of course, this behavior presupposes that the interlocutor will understand the switch. In the context under which these speeches were recorded, no such question could arise. Moreover, it should not be forgotten that AnRK students come from an environment tolerant of code-switching, and the habitual multilingual mode is very likely to be brought to the classroom. Code-switches seem to be of four major types: interjections, unneutralized lexical gaps, requests for help, and asides. Interjections appear to be unmonitored and emerge without the conscious awareness of the speaker. Unneutralized lexical gaps and requests

for help are two ways of using L1 switches to fill gaps, either by inserting an L1 word or by using it to ask for an equivalent FL word. Asides are usually directed at the speaker from the audience, who feel as though the requirement to maintain a monolingual FL mode does not apply to them.

Foreignizing is another technique for filling lexical gaps in which an L1 word is adapted, morphologically and/or phonologically, to suit the FL context. Occasionally this may happen from the belief that the foreignism is an actual English word. At other times, "false friends" – words which sound particularly alike in the two languages, yet have very different meanings – may be to blame. In addition, students sometimes create neologisms by attempting to derive new parts of speech, for example, by turning adjectives into nouns, or by attaching the wrong derivational suffixes.

Transfer can also be used to avoid lexical gaps, e.g. by filling them with translated L1-appropriate items. These are not always standard in the FL, however, as demonstrated by examples such as *driving* (instead of: *riding*) and *human* (instead of: *person*). Transfer also helps students structure sentences, by providing a grammatical framework upon which to hang FL lexical items.

Another example of the burden placed on the language processing system by speech are the phonological errors seen in the corpus. For example, the phoneme [**ð**], which does not exist in Lithuanian, appears to be produced only with conscious attention; when such attention wanders, the same phoneme can be mispronounced, even within the same sentence. Moreover, is it clear that processing itself inhibits acquisition, at least in the short term: in most cases where students were told new words during the course of speaking, their pronunciation of those words changed throughout the speech, indicating that, while the general shape of the word can be acquired, its internal structure requires more in-depth processing that must be performed offline.

Because language processing is so resource-costly, students have developed techniques to reduce the cognitive burden; these are referred to as interactional strategies. Students use interactional strategies to explicitly elicit

help, either from the teacher or the audience, in explaining unfamiliar vocabulary, confirming pronunciation and/or lexical gap fills, and even by asking for assistance in easing the speaking task (e.g., by requesting that no questions be asked). The other interactional strategy is assumed understanding, wherein a speaker simply assumes that his or her fragmentary clues are enough for the listeners to reconstruct the meaning he or she wishes to convey. Such assumptions may be unwarranted.

Students' requests were also examined for evidence of pragmatic failure. Although a few students made efforts to include imposition mitigating elements like downgraders or supportives, these were often coupled with pragmatically inappropriate request strategies such as the use of imperatives. The overall effect of students' requests is therefore quite rude by English NS norms.

The examples analyzed in this section clearly show the effects of CLI in students' spoken English production. While it is true that monolingual speakers hesitate, they rarely do so to the extent found in the extracts examined above. Moreover, a monolingual speaker, by definition, cannot code-switch or use foreignisms, has no other language system to transfer from, and has no difficulties with phonology. Finally, monolingual speakers have no sociocultural norms affecting their language production other than those of the language they speak, so their interactional strategies and pragmatic requesting behavior are not barriers to communication.

In Part 1, a number of comments were provided that pointed out potential trouble spots for the monolingual speech production model offered by Levelt when applied to multilinguals. The data in the spoken corpus allow several of those ideas to be further developed. For example, it can clearly be seen that many of the processes claimed by Levelt to be automatic (e.g., microplanning, grammatical encoding, articulation and monitoring) are, in fact, areas under conscious control, at a cost to attentional resources.

More importantly, the data analyzed in Part 3 point to two major factors influencing CLI in multilingual student production: lexical gaps and L1 sociocultural norms. The first four types of CLI – hesitation, code-switching, neologisms and foreignizing, and transfer – were all found to be reactions and responses to lexical gaps. When students are unable to lexicalize concepts, they pause, repeat, start anew, switch to L1, incorrectly derive new FL words from L1 or FL roots, and transfer structures into the FL. Taken together, these processes – along with the phonological inconsistencies, interactional strategies and pragmatic failures also reviewed above – clearly demonstrate the strong influence of the L1 on the students' developing FL system. Students follow L1 patterns at most levels, from phonology to pragmatics. This is due to their basic unfamiliarity with FL sociocultural norms. Despite years of studying English as an FL, students' inner voices remain Lithuanian, and this impacts their production of English in more or less drastic ways. This also proves that for these students, the developing English language system has not yet passed the proficiency threshold beyond which it becomes an L2.

Finally, it should again be stressed that the linguistic behavior analyzed here is not considered to be unequivocally erroneous. English is a world language, and native-speaker norms are only one standard of judgment. A more realistic criterion is simply whether or not learners are able to use the linguistic resources at their disposal to effectively communicate in the target language. By this definition, even code-switching, given an environment in which all participants understand the switch language, is not erroneous, but simply pragmatic. The students of AnRK communicate quite well, considering the sociocultural background from which they come and their total lack of experience living in a monolingual environment. Such an environment – one in which no Lithuanian-specific elements could be expected to be understood by any interlocutors – would, of course, force a monolingual language mode, encourage a greater focus on accuracy and less reliance on code-switching, stimulate the use of an English inner voice, and, perhaps, push the learner to

acculturate. All of this would not only assist in the development of language skills more acceptable to English NSs, but also push the language system to develop beyond the L2 proficiency threshold. As none of the students of AnRK live in such an environment, however, their current linguistic habits are more than sufficient to enable communication in English with other NNSs.

4. CROSSLINGUISTIC ASPECTS OF ENGLISH WRITING PRODUCTION

It has been seen that multilingual speech production differs from that of monolinguals, that multilinguals are deeply affected by their more resonant sociocultural context until reaching an acculturation threshold, that this threshold is nearly impossible to attain when the language of study is an FL taught in a predominantly L1 environment, and that all of this affects the speech production process as CLI, which can take many different forms. It was also shown that CLI in speech can be seen as a response to two major factors: lexical gaps and sociocultural norms (see Part 3).

Many of the CLI phenomena that appear in speech production, however, are usually hidden in writing[66]. Students may hesitate when writing – indeed, they may spend hours or days between sentences or even words – but this is, of course, invisible to a reader, who sees only the finished product. Code-switching also occurs, but, as students may refer to dictionaries, this is also much less common[67]. Phonology obviously cannot be a problem, though spelling certainly is. Handwritten assignments contain frequent errors in spelling, and even typewritten assignments are often not checked for spelling errors. However, as spelling errors are intra-lexemic and almost always part of the developing FL intralanguage system, they will not be considered here. As discussed in Part 3, interactional strategies are also not normally appropriate when writing, as writing is an offline process and the audience is not usually present during its composition. This leaves neologisms and foreignizing, transfer, and pragmatic failure as the most likely types of CLI to appear in writing. Part 4 will argue that all of these phenomena are, in fact, different manifestations of the general human ability to blend concepts in innovative and linguistically useful ways.

[66] This dissertation examines only writing that was assigned as homework; students thus had ample time to refer to sources such as dictionaries or the Internet to fill any lexical gaps. No impromptu writing assignments (e.g., in-class tests) were considered.

[67] Writing samples were collected from 23 L1 = Lithuanian students. Not one of them code-switched in writing. Of the 18 L1 = Lithuanian students whose speeches were analyzed in Part 3, ten (or 56%) code-switched at least once during a speech.

The purpose of Part 4 is to justify, present and explore a methodology for the analysis of multilingual FL learner constructions as conceptual blending. Section 4.1 provides the necessary theoretical background and shows that transfer can work at the level of concepts, lemmas, and lexemes, or various combinations of the three. It also establishes the methodology used to perform the analysis in section 4.2, where eight examples of learner constructions are shown to be conceptual blends. Conceptual integration networks are drawn for each example. In section 4.3 the results are discussed and situated within the wider context of the socio- and crosslinguistic data analyzed in Parts 1–3.

4.1 Learner Constructions and Conceptual Blending

This dissertation assumes the cognitive linguistic principles of **construction grammar** (Goldberg, 1995, 1998; Taylor, 1998; Kay & Fillmore, 1999; Fillmore, Kay & O'Connor, 2003; Evans, 2006; Evans & Green, 2006). Goldberg (1998: 205) defines a **construction** as "a form-function pair, such that some aspect of the form or some aspect of the function is not strictly predictable from [its] component parts." According to Taylor (1998), this definition allows that idioms, established non-idiomatic phrases (or chunks; see below), and even single-morpheme lexical items are constructions. Fillmore et al. (2003: 243) confirm that "constructions may specify, not only syntactic, but also lexical, semantic, and pragmatic information… lexical items, being mentionable in syntactic constructions, may be viewed, in many cases at least, as constructions themselves." This is important, as it allows that constructions are **recursive**: a construction may itself contain constructions which are made up of more constructions, etc.

Many constructions are implicit in the sense that they have no overt form. Goldberg (1995) has extensively described the **caused-motion construction**, as exemplified, e.g., by the sentence *John sneezed the napkin off the table*. In this sentence the subject, *John*, performs an action, *sneezing*, which in turn causes the object, *the napkin*, to move *off the table*. This is a **productive** construction in that it is the pattern of phrases, rather than any inherent

property of the lexical items, that is responsible for the caused-motion effect. Consider a few sentences constructed on the same model: *The audience laughed the hypnotist off the stage; The boss shouted me out of her office; They drove the cows into the corral*. As an example of a construction containing another construction, consider a caused-motion sentence containing an idiom: *I nagged that couch potato into the gym.* Here the overt construction (idiom) *couch potato* is included as the object of the implicit caused-motion construction.

It is another important property of constructions that they are to some extent language-specific. Consider the **resultative construction** (Goldberg, 1995), an extension of the caused-motion construction that can be exemplified by sentences like *He drank himself to death* and *I wiped my glasses clean*. These sentences cannot be translated word-for-word into Lithuanian: **Jis gėrė save į mirtį; *Aš nušluosčiau mano akinius švarūs*. Resultative clauses in Lithuanian must be expressed using Lithuanian-specific constructions: *Jis gėrė, kol galutinai nusigėrė ir numirė* and *Aš švariai nusišluosčiau (savo) akinius*, perhaps.

The sub-sections that follow examine various properties of constructions as they apply to the acquisition of lexical items (now understood to mean words, phrases, and constructions themselves) and their use in conceptual (and, especially, grammatical) blends.

4.1.1 The acquisition of constructions

This section proposes that lexical acquisition, especially in FL (rather than L2) settings as noted by Kecskes & Papp (2000), proceeds in the direction lexeme → lemma → concept, and that this process (which can take years to complete) can lead to transfer effects, including neologisms, foreignisms, pragmatic failures and structural mismatches that can be very disruptive to the overall goal of communication.

Students of foreign languages typically meet new words in classrooms or textbooks; in school, these most often take the form of lists of more or less

related words[68]. New words are therefore learned in the form of lexemes: written and/or spoken forms without any higher-level (e.g., lemma or conceptual) information attached. These FL "vocabulary items," as demonstrated in the revised hierarchical model (Kroll & Stewart, 1994), are for this reason typically linked by learners to L1 lemmas (which are themselves inextricably linked to L1 concepts, or cognitive models, that are language- and culture-specific). As students become more familiar with a given lexeme, they begin to develop an L2 lemma which will contain some (but certainly not all[69]) of the L2-specific semantic information pertaining to the word, together with the syntactic requirements for using the item in question. This L2 lemma will slowly replace the original L1 lemma[70], but will still remain linked to the original L1 concept. Only after much experience with situated L2 uses of a given lexical item can learners begin to develop an L2-appropriate concept to replace the L1 concept, which can take years. This process in schematic form is shown in Figure 34.

[68] This means that a learner's most prototypical first encounter with a new FL word is as a grapheme.

[69] See Part 1 for a discussion of the internal structure of lexical items.

[70] Bearing in mind that lexical items can be viewed as nodes in a spreading activation network and that the connections between nodes are resonant (meaning that more resonant connections fire faster than less resonant ones, that is, more resonant connections have lower activation thresholds), the process of replacing a resonant L1 lemma with a newly created, unresonant L2 lemma requires much practice and use of that item. This was expressed in the dynamic model of multilingualism as language maintenance effort (LME).

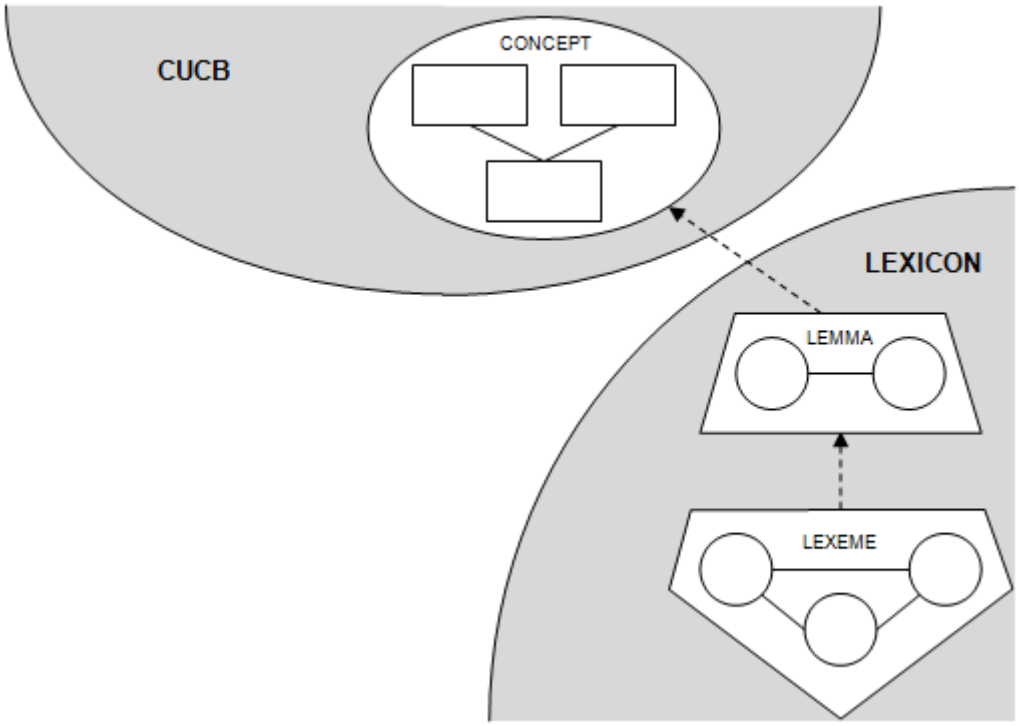


Figure 34. Lexical acquisition proceeds from lexeme to lemma to concept.

As a concrete example, consider the English phrase *white bread*. To an English speaker this phrase usually describes a foodstuff made from bleached wheat flour, white in color, which is baked in loaves and may be served sliced. A native speaker of English studying Lithuanian will at some point encounter the lexemes “balta” and “duona;” from these, he or she may decide that the Lithuanian phrase *balta duona* is an acceptable translation of *white bread*. This would be incorrect. *Balta duona* is made from a mix of wheat and rye flour, and is much darker and denser than *white bread*. A more accurate translation of *white bread* would, in fact, be *batonas* – itself probably a borrowing from the Russian батон. Thus what the English speaker intends by *balta duona* and what the Lithuanian listener understands upon hearing it are two entirely different things.

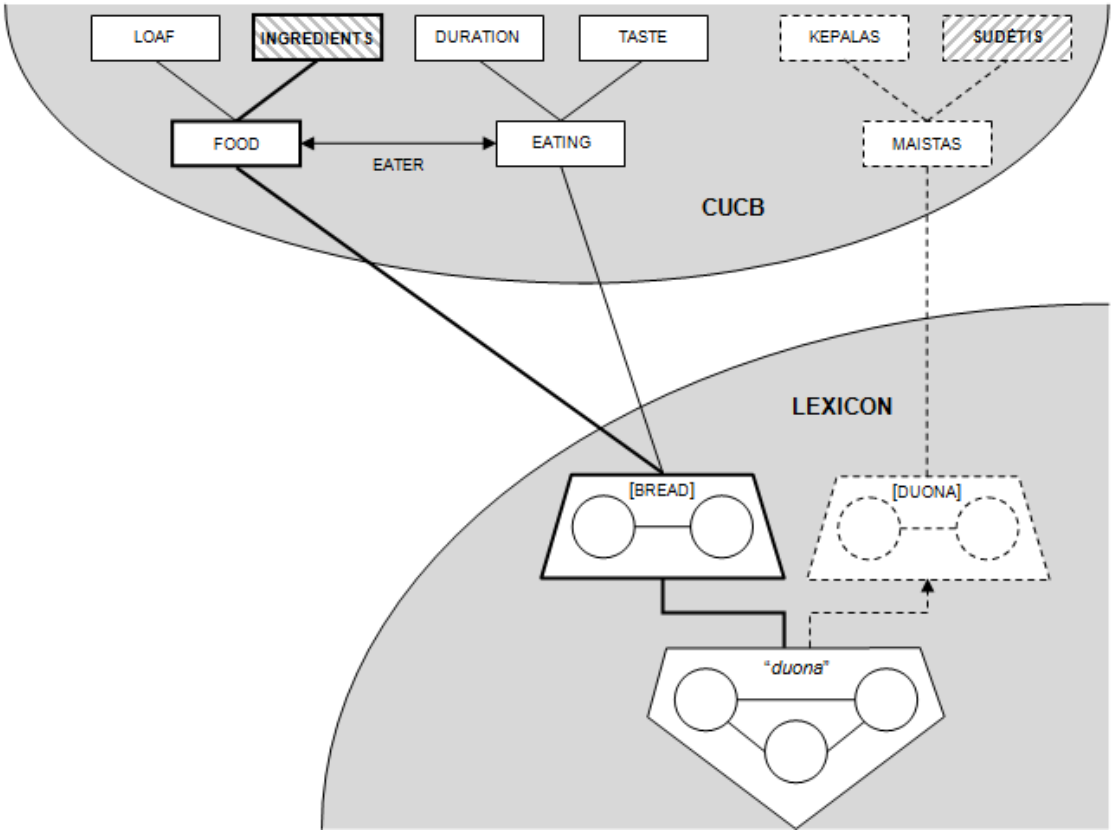


Figure 35. Incorrect access route established when L2 lexeme links to L1 lemma.

Figure 35 illustrates this process. Here the Lithuanian (FL) lexeme "duona" has been linked to the English (L1) lemma [BREAD]. This lemma is itself linked to an English conceptual knowledge structure in which the cognitive models FOOD and EATING are linked by the *eater* relation. Each cognitive model has facets attached. The intended access route to be established by the phrase *balta duona* is marked with darker lines, as the essential feature that distinguishes *white bread* from other types is the use of bleached wheat flour, an ingredient. This INGREDIENTS facet is shaded (cf. Figure 18) to indicate its language-specific nature.

The "error" arises from the fact that the ingredients of *balta duona* are not the same as those of *white bread*. The lexeme "duona" should be linked to a Lithuanian [DUONA] lemma (lexical concept), here indicated with dashed lines, which would access a Lithuanian-specific conceptual knowledge structure different from the English structure. In this example the most salient difference arises in the facet SUDĖTIS, which is also language-specific. As the

hypothetical English speaker becomes more familiar with Lithuanian bread, he or she may come to realize that the words DUONA and BREAD are not equivalent in all their features. This would be the beginning of the development of a Lithuanian-specific concept to match with the Lithuanian lexeme. Or, as Kecskes & Papp (2000) would have it, this would be the bigging of the concept‘s neutralization.

An English NS who creates the phrase *balta duona* in Lithuanian (with the intended meaning of *batonas*) is thus creating a learner construction which, by definition, is nearly appropriate but slightly off. This process arises organically from the organization of the lexicon and the structural properties of individual lexical items, the relation of the lexicon to the CUCB, and the processing mechanisms at work in the production of speech.

If lexical items are overt constructions, as argued above, then this process should be true of implicit constructions, too. Indeed it is, but it works at the level of lemmas rather than lexemes. This is because implicit constructions have no lexemes. Recall the discussion of grammatical encoding in Part 1. There the model of Bock & Levelt (1994) was presented using the example *She was handing him some broccoli*. This sentence exemplifies the **double object construction**, also discussed in Goldberg (1995). In the discussion in Part 1 it was shown that the lemmas representing the various lexical items are bound through the process of constituent assembly into a control hierarchy that determines word order. According to construction grammar, however, this control hierarchy is itself a specialized lemma (an implicit construction) which contains **slots** and semantic/syntactic instructions for how they can be filled with the lemmas of the overt lexical items ([SHE], [HAND], [HE], and [BROCCOLI] in this case). The [DOUBLE OBJECT CONSTRUCTION] lemma, having no lexeme of its own, simply acts as a semantic/syntactic frame for the overt lemma/lexeme pairs that are to be articulated.

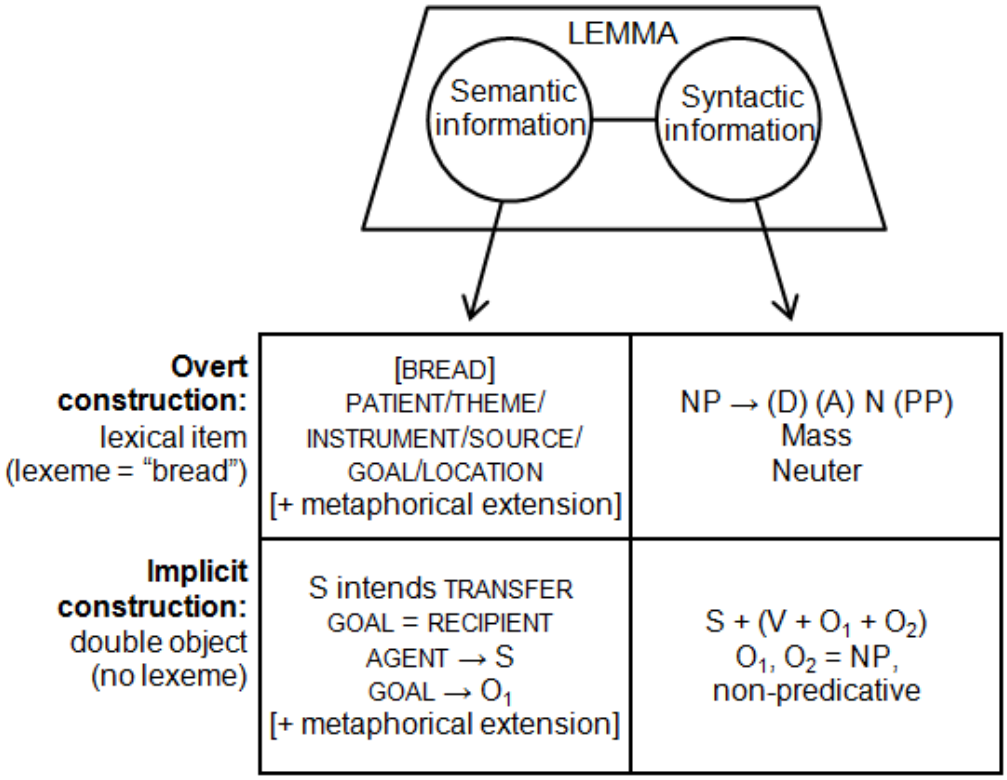


Figure 36. Overt and implicit constructions at the lemma (lexical concept) level.

In Figure 36, the lemma information for an overt construction (the lexical item *bread*) and an implicit construction (the double object construction) are compared. The information summarized for the double object construction is taken from Goldberg (1995) and Evans & Green (2006). The aim of this section is not to analyze any one construction in great detail, however, but rather to establish that implicit constructions, like other lexical items, have lemmas which attach to conceptual knowledge structures in the CUCB, and that the acquisition of constructions may also proceed in a piecemeal fashion. It is, moreover, hampered by the fact that implicit constructions have no overt (lexeme) forms, that is, are not available to be learned as vocabulary items in lists. For this reason they are usually a) explicitly pointed out by teachers and learned in the form of "grammar rules," or b) eventually abstracted, more or less accurately, from experience with situated exemplars by learners themselves. Process a) is in some ways faster, but process b) is more likely to

result in accurate usage[71], though it will require much more time and LME. It should therefore be stressed that implicit constructions, like lexical items, are acquired in the L1 through experience with their situated use in a sociocultural context. As Goldberg (1998: 208) writes, "what children learn when they learn the syntactic patterns of simple sentences is the particular way certain basic scenarios of human experience are paired with forms in their language." This is to say that constructions, like lexical items, are language- and culture-specific.

4.1.2 Grammatical blending

For most monolinguals, the distinction between concepts and words is blurred, and there is little or no psychological difference between them (Oller, 1997). For multilinguals, on the other hand, lexicalization is an essential problem of communication. This was seen especially in Part 3, where four types of CLI were attributed to the problem of lexical gaps. This section will examine grammatical encoding from the point of view of conceptual blending theory, which proposes that blending is the process by which the various conceptual knowledge structures (concepts and facets), lexical concepts (lemmas), and word forms (lexemes) are combined. These combinations may be useful and productive if done appropriately, that is, following language- and culture-specific principles. However, if they are done inappropriately, miscommunication in the form of CLI and learner constructions results.

According to Fauconnier & Turner (1994; 1996; 2002), conceptual blending is a process in which two or more **mental spaces** (Fauconnier, 1994; 1997) are selectively combined to yield a **blended space** that contains new, emergent meaning not present in either (or any) of the **input spaces**. In mental spaces theory, any utterance invites the listener to create a mental space (typically drawn as a circle) which contains elements representing the topic under discussion. As discourse unfolds, new spaces are added through the use

[71] It is no coincidence that process b) is the same as that used by children in acquiring their first language: "we might view the constructional semantics as emerging from an abstraction over the particular semantics of learned instances with particular verbs. <…> [T]he child is categorizing learned instances into more abstract patterns, and is associating a semantic category with a particular formal pattern" (Goldberg, 1998: 209).

of **space builders** (expressions such as *in this book, if, last year*, and *John thinks* are all space builders that cause new mental spaces to appear), linked to the already existing mental spaces through connectors that establish counterpart elements in the new spaces (Fauconnier, 1994; Evans & Green, 2006). This theory has been extended to describe conceptual blending, also called conceptual integration, in which different elements from different mental spaces are mapped onto each other in a third, blended space.

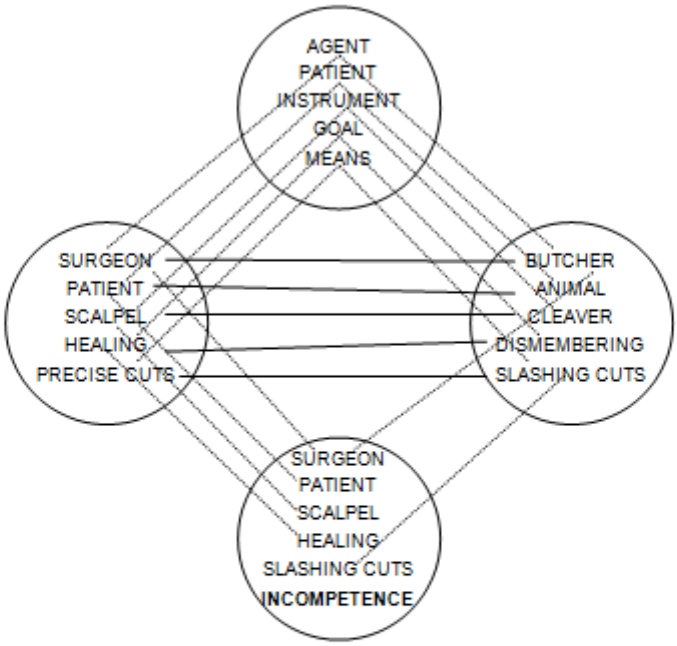
Consider the metaphoric expression *That surgeon is a butcher*, discussed in both Evans & Green (2006) and Coulson (2006). This expression manages to convey meaning that is not inherent in either of the input spaces: indeed, assuming prototypical representations of a surgeon and a butcher, both would be understood to be competent professionals. Yet the meaning of *That surgeon is a butcher* is, of course, that the surgeon is incompetent and should not be allowed to practice. This meaning arises through the nearly simultaneous creation of four mental input spaces. One input space represents the surgeon and another the butcher; again, these spaces contain only the most prototypical, default representations of the professions in question. It is not the point of the expression to imagine either profession in great detail, so extremely general characteristics are sufficient. The elements in these spaces are connected to a third, **generic space** which contains only the most general structural information, usually syntactic/semantic. In other words, the input spaces are created using lexeme information (morphological, phonological, and even orthographical), as guided by lemma information in the generic space.

In the next stage, **mapping** occurs between the input spaces (the two professions) in order to determine possible points of comparison (Fauconnier & Turner, 2002: 41). By finding a relationship such as Disanalogy, a **vital relation** (Fauconnier & Turner, 2002: 92) is established between the two spaces that enables them to be **compressed** into the final, blended space. This process of compression often results in the formation of a new vital relation altogether. In the example under discussion, the Disanalogy relation is chosen

because the expression focuses the listener's attention on what is different between the two professions. However, at the same time the listener is also asked to imagine a new type of profession altogether, one that would be able to embody properties of both a surgeon and a butcher. No such profession exists, of course, so this imaginary employee will automatically be one-of-a-kind. Thus, the Disanalogy relation becomes compressed into one of Uniqueness, as the meaning of the expression involves a single, unique doctor who contains elements of both inputs. In the blended space, certain features of surgeons and certain other features of butchers are **selectively projected** to create the image of a reckless surgeon who, rather than making precise incisions to minimize scarring and recovery time, slashes away at his patient-victim as if dismembering an animal for consumption. Such a surgeon would clearly be considered incompetent (if not insane), and therefore the idea of *incompetence* is said to be **emergent** in the final blend. Neither input space contains any such notion; prototypical surgeons and prototypical butchers are automatically competent in their respective fields. Thus *incompetence* is an emergent property of the blend itself.

There are two opinions in the literature as to how conceptual blends should be represented. Fauconnier & Turner (2002) favor using an array of circles called a **conceptual integration network**, as circles have traditionally been used in the analysis of mental spaces (cf. Fauconnier, 1994; 1997). Coulson (2006), however, favors presenting the data in tables (e.g., Coulson, 1996; 2001; 2006). Both methods are illustrated for the conceptual blend *That surgeon is a butcher* in Figure 37. This blend was chosen because it is discussed in both Evans & Green (2006), where it is presented using circles, and Coulson (2006), where it is presented in a table. Following Coulson (2006), tables will be used for the analyses in this dissertation.

Conceptual blend: "*That surgeon is a butcher.*"



Generic Space	Surgeon Input Space	Butcher Input Space	Blended Space
AGENT	Surgeon (s)	Butcher (b)	Surgeon (s/b)
PATIENT	Patient (p)	Animal (a)	Patient (p/a)
INSTRUMENT	Scalpel (k)	Cleaver (c)	Scalpel (k/c)
GOAL	heal patient	dismember animal	heal patient
MEANS	precise cuts	slashing cuts	slashing cuts
			Emergent inference: incompetent (s/b)

Figure 37. Two different, yet equivalent, representations of the same conceptual blend. Adapted from Evans & Green (2006: 406) and Coulson (2006: 193).

Conceptual blending is a powerful theoretical tool which can describe not only metaphorical expressions, but also such diverse human creations as clocks, money, complex numbers (Fauconnier & Turner, 2002), and cartoons (Turner, 2006). Humor is often based on blending (Evans & Green, 2006; Coulson, 2001; Braun, 2008), as is irony (Coulson, 1999). For this dissertation, **grammatical blends** (Mandelblit, 1997; Fauconnier & Turner, 2002) are of special importance. According to conceptual blending theory, many commonplace grammatical structures (e.g., constructions) are conceptual blends that have become stable and accepted as standard responses to communicative needs. Such structures include the caused-motion construction and French causative clauses (Fauconnier & Turner, 1996), the morphology of

Hebrew verbs (Mandelblit, 1997), privative adjectives (e.g., *a stone lion*, from Coulson & Fauconnier, 1999), resultative constructions and single-word blends such as *Chunnel* (Fauconnier & Turner, 2002). As Boers & Lindstromberg (2006: 313) put it, "grammatical constructions... reflect language users' choices of construal, that is, their choices of *how to conceptualize* a given situation or event" (emphasis added).

Such grammatical blends exist in every language. It is therefore important to stress that the lexicalizing of concepts is more than a simple matter of attaching labels to mental images. Many concepts can only be expressed through conceptual blends, which are often language-specific. Hebrew verb morphology, for example, allows causative morphemes to attach directly to verb stems, forming causative blends (Mandelblit, 1997) in a way that could never happen in English. Thus a large part of lexicalization is, in fact, conceptual blending. It is not only concepts that are language-specific, but also the blends used to express them. The process of neutralization (Kecskes & Papp, 2000) may well be understood as a process of acquiring appropriate language-specific blends for concepts learned in another language. Moreover, one of the threshold requirements for the development of a CUCB may be the acquisition of a sufficient quantity of FL grammatical blends, without which lexicalizing L1 concepts in the FL is resource-costly and difficult to achieve during high-speed online processing.

Now consider again the double object construction, as exemplified by *She was handing him some broccoli*. This same construction can be analyzed as a grammatical blend. Indeed, the figure provided by Bock & Levelt (1994) to illustrate the process of constituent assembly (see Figure 9) is remarkably similar to a conceptual integration network. For comparison, Figure 38 shows how this process would look as a conceptual integration network.

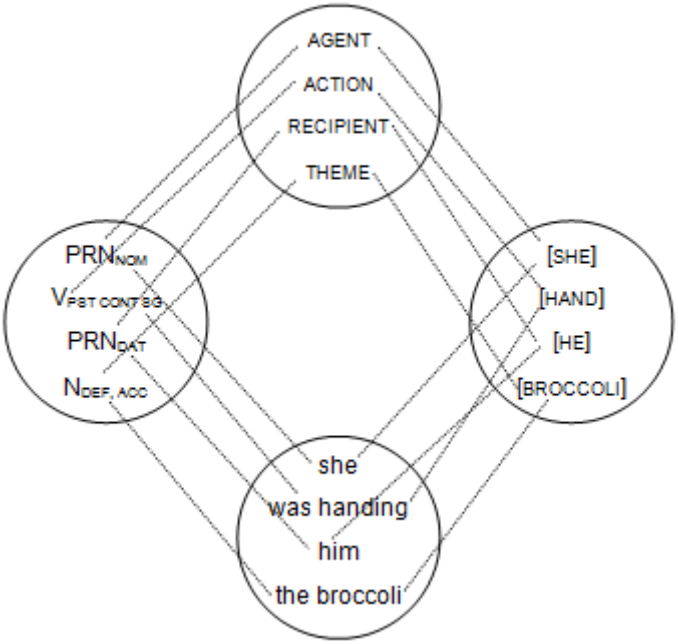


Figure 38. Grammatical blending in the double object construction.

It is unfortunately characteristic of the conceptual blends drawn by blending theorists that they are unspecific as to the exact level at which the blending actually takes place. Fauconnier & Turner (2002) demonstrate that blending occurs between and among:

- lexical items: *word blends, compound nouns, phrases, idioms*;
- morphemes: *prefixes, roots, suffixes;*
- constructions: *resultative, caused-motion, double object*, among others;
- conceptual representations: *counterfactual statements, lies, complex numbers*;
- visual representations: *advertisements, photographs, logos*;
- realia: *watches, altimeters, money*.

Thus the term *conceptual* blending is either too vague or too specific, depending on one's theoretical orientation.

In fact, the double object construction prompts for a blend whose four mental spaces can be quite precisely defined. In Figure 38, the generic space at the top is structured by the [DOUBLE OBJECT CONSTRUCTION] lemma detailed in Figure 36. The input space on the right actually contains four more lemmas.

This is unavoidable: the double object construction is implicit and requires overt lemmas to fill its slots. Indeed, it is the purpose of implicit constructions to help arrange overt lemmas in productive patterns that can best convey the meaning the speaker wishes to express. The input space on the left contains constituent assembly instructions coming from the formulator: grammatical blending, after all, occurs as part of the process of speech production that Levelt's model was designed to account for. Thus the blended space at the bottom represents the surface structure that will be sent on for phonological encoding and eventual articulation. Grammatical blending thus offers a way to illustrate the process of grammatical encoding and, in particular, provides a mechanism that explains how the formulator interacts with the lemmas in the lexicon.

4.1.3 Crosslinguistic blending

Mandelblit (1997) is the only blending theorist who discusses blending and translation. As her interpretation differs slightly from that proposed in this dissertation, it bears some discussion. Figure 39 is a reproduction of Mandelblit's (1997) scheme for translation-blending. As her caption states, blending is seen as two independent processes, taking place first in the source language and then in the target language. In fact, the process as it occurs in multilinguals speaking an FL is much more interconnected. Mandelblit's scheme is an idealized version of the way translation should work, or the way successful translations are done; in practice, however, the process is often quite different. After the source language blend is created, this blend itself becomes an input (or, as shall be seen, several inputs) in a new blend of both languages, which blends the source language utterance with whatever knowledge the speaker has of the target language structures (usually lexemes) appropriate for this communicative event. That knowledge may be partial or complete, depending on the translator's knowledge of the target language. Finally, a new blend of source language concepts expressed as target language words and structures is created. Note that this model leaves much more room for error

than does the model illustrated in Figure 39. As shall be seen below, students' English utterances do contain such inappropriate two-language blends.

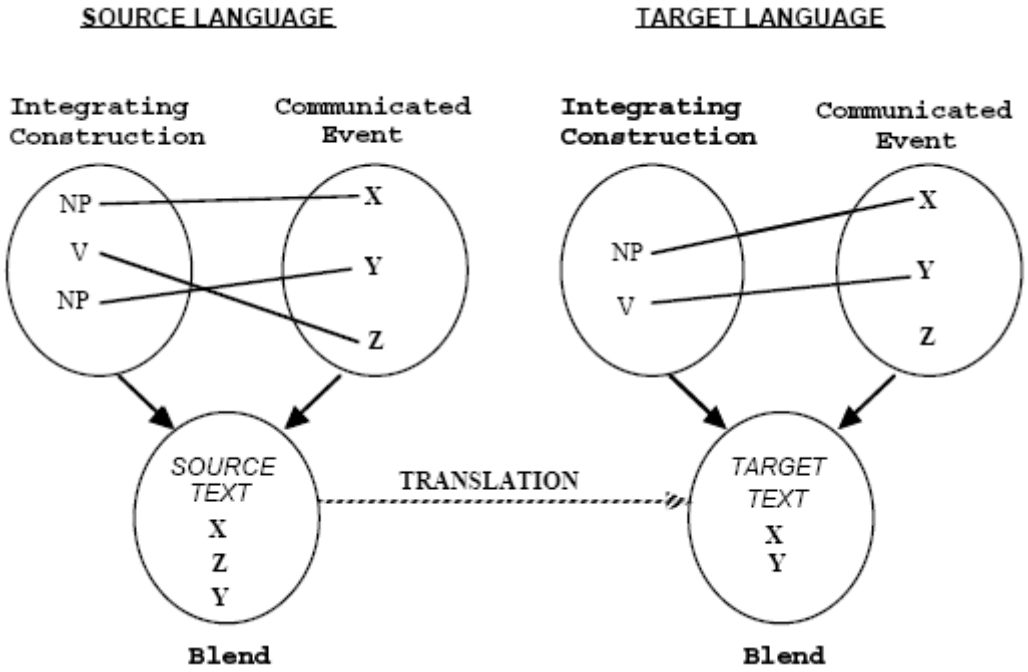


Figure 39. "Translation is the outcome of two independent blending operations." Reproduced from Mandelblit (1997: 194).

To illustrate the model proposed in this dissertation for what will now be called **crosslinguistic blending**, consider again the English-specific resultative construction as exemplified by the sentence *I wiped my glasses clean*. The resultative construction, like the double object construction, is implicit and exists only as a lemma: it has no overt form (lexeme), and acts as a generic space to structure the overt lemmas that will be used in the utterance. The semantic and syntactic information contained in the [RESULTATIVE CONSTRUCTION] lemma (cf. Goldberg, 1995; Evans & Green, 2006) is summarized in Figure 40.

Lemma: [RESULTATIVE CONSTRUCTION]	
semantic information	**syntactic information**
S → AGENT O → PATIENT V encodes CAUSATION A = BINARY A ≠ DEVERBAL [– metaphorical extension]	S + (V + O + $OBL_{PP/AP}$)

Figure 40. Lemma information in the resultative construction.

In English, this construction is used to generate resultative clauses that focus the listener's attention on the end state of an action. This is done via grammatical blending. Blend 1 illustrates the English-specific blend for the sentence *I wiped my glasses clean*.

Generic Space	Input Space 1	Input Space 2	Blended Space
AGENT ACTION PATIENT RESULT	PRN_{NOM} V_{PST} $N_{DEF, ACC}$ A_{BIN}	[I] [WIPE] [GLASSES] [CLEAN]	I wiped my glasses clean **Emergent inference:** CLEANLINESS

Blend 1. "I wiped my glasses clean."

The generic space is structured by the information in the [RESULTATIVE CONSTRUCTION] lemma. Four roles are required: a volitional AGENT, a verb that encodes CAUSATION as the immediate result of its ACTION, a PATIENT whose state undergoes a change as a result of the ACTION, and the RESULT itself, expressed as either a prepositional or adjectival phrase. If the phrase is adjectival, the adjective must be either binary or able to be interpreted as such (that is, as the endpoint of a scale). The first input space contains instructions from the formulator for constituent assembly. This includes part-of-speech information as well as case and inflection (see Part 1). The second input space contains the lemmas chosen according to the instructions contained in the preverbal message sent to the formulator from the conceptualizer. All of this is grammatically blended to produce the surface structure seen in the final, blended space, where: the pronoun lemma [I] is in the nominative case; the

verb lemma [WIPE] is in the past tense; the object noun lemma [GLASSES] is in the accusative case and has been made definite with the use of a determiner, *my*; and the binary adjective lemma [CLEAN] has been attached. All of this gives rise to the emergent concept of CLEANLINESS, which is, of course, the result of the action and the intention of the utterance.

As discussed above, there is no single Lithuanian construction that does the work of the English resultative construction. English may be more construction-laden than Lithuanian because, being a configurational language, word order rules are necessary cues for meaning construction. The nonconfigurational nature of Lithuanian means that word order can be pragmatic rather than grammatical. To focus attention on a result in Lithuanian, one can simply move it to the front of the sentence. Alternatively, one can exploit Lithuanian verb morphology, which allows verbs to have resultative forms through the use of prefixes. As Paulauskienė (1994: 275) writes, "priešdėliai parodo veiksmo rezultatinę baigtį.... <...> Grynoji rezultatinė veiksmo baigtis susiformuoja tuo atveju, kai priešdėlis parodo, jog pasiektas būtent tas rezultatas, į kurį buvo orientuotas pamatinio veiksmažodžio veiksmas[72]." The Lithuanian equivalent of *I wiped my glasses clean*, as discussed above, is probably *Švariai nusišluosčiau akinius*. Blend 2 illustrates this Lithuanian-specific grammatical blend.

Generic Space	Input Space 1	Input Space 2	Blended Space
RESULT ACTION AGENT PATIENT	Adv V RES REFL PST 1SG PRN NOM N PL, ACC	[ŠVARIAI] [ŠLUOSTYTI] [AŠ] [AKINIAI]	*Švariai nusišluosčiau akinius* **Emergent inference:** CLEANLINESS

Blend 2. "*Švariai nusišluosčiau akinius.*"

Again the blend begins from a generic space in which thematic roles are brought together to structure the final utterance. The order, however, has been

[72] "[P]refixes show the resultative end of the action…. <…> The true resultative end of an action is formed when the prefix shows the achievement of exactly that result towards which the action of the main verb was oriented." (aut. trans.)

changed to reflect Lithuanian free-word-order norms. The RESULT comes first, expressed as the adverb lemma [ŠVARIAI]. The ACTION and the AGENT are fused in the verb lemma [ŠLUOSTYTI] (Lithuanian being a null-subject language), which is also given the resultative prefix *nu-*, the reflexive particle *-si-*, and the first person singular past ending *-iau*[73]. Finally, the PATIENT role calls for the plural noun lemma [AKINIAI] to be given accusative case. This is combined in the blended space to produce an utterance that is appropriate according to Lithuanian linguistic and sociocultural norms. When this blend is complete, the formulator will have readied the surface structure *Švariai nusišluosčiau akinius* for phonological encoding.

Thus far the discussion has followed Mandelblit's (1997: 176) model closely; as she writes: "the translation process first requires a conscious process of 'de-integration' (or 'un-packing') of the source sentence into its conceptual and linguistic input structures, and then a 're-blending' operation of these structures into the target language's grammatical constructions." Indeed, the Lithuanian translation did not begin with the final English utterance, but rather with the structural information contained in the generic space. For less proficient FL speakers, however, it is precisely the fact that utterances can be **unblended** that leads to error. It has already been established that, due to the piecemeal process of lexical acquisition, learners almost cannot avoid linking FL lexemes to L1 lemmas in early stages of acquisition. They thus have no FL lemmas (or at least only partial ones) to guide the **reblending** of L1 constructions as FL constructions. They must therefore work with what they have, namely lexemes. Blend 3 illustrates what might happen if an English NS were to attempt to express the English-specific resultative construction *I wiped*

[73] To be clear, at this point the verb has no overt morphology added, as that occurs during phonological encoding when the morphological information in the lexeme is accessed. Its lemma (which is itself not an overt form, as TOT studies have shown) is simply tagged with the appropriate inflectional information as part of the grammatical encoding/blending process.

my glasses clean using only Lithuanian lexemes. The result, as already discussed above, might well be *Aš nušluosčiau mano akinius švarūs*[74].

Generic Space	English Input Spaces	Lithuanian Input Spaces		Blended Space
CLEANLINESS	*I*	"aš"	NOM	*Aš nušluosčiau mano akinius švarūs* **Emergent inference:** (mismatch/error)
	wiped	"šluostyti"	RES PST 1SG	
	my	"aš"	GEN	
	glasses	"akiniai"	PL, ACC	
	clean	"švarus"	PL	

Blend 3. "*Aš nušluosčiau mano akinius švarūs*."

This blend is radically different from the others. It begins in the generic space with the concept of CLEANLINESS that was emergent in the blended space of the original English resultative construction. This concept is linked to the English words that expressed it in the source utterance[75]. They are written here in italics, rather than as lemmas or lexemes, to represent the fact that they are treated in the English input spaces as finished products: that is, they are no longer lemmas, but words that have been grammatically and phonologically encoded and are ready to be articulated. These are matched with Lithuanian lexemes (in quotation marks) which contain morphological information (here made explicit in its own column). Thus *I* is expressed as nominative *Aš*, *wiped* becomes *nušluosčiau*, *my* becomes genitive *mano*, *akiniai* are declined as accusative *akinius*, and *clean* is assigned number and emerges as *švarūs*. The final result is unlikely to give rise to the desired emergent meaning. Depending on the willingness of a Lithuanian listener to "play along," this sentence will either be interpreted as a mismatch or an outright error.

[74] This actually gives the speaker the benefit of the doubt. There is no reason to assume that he or she would know to use the resultative prefix, how to assign the correct case to the object, nor that the object and its adjective must agree in number and gender (which, in English, they don't). It is therefore possible that even less appropriate utterances like *Aš šluosčiau mano akiniai švarios*, etc., could also be produced.

[75] Although the "prototypical" conceptual blend is made up of four spaces, there is, in fact, theoretically no upper limit to the number of spaces that can compose a blend (cf. Fauconnier & Turner, 2002, on megablends). In other words, blends are also recursive, and can take other blends as inputs.

Fauconnier & Turner (2002: 383) also comment on this interpersonal nature of blending. In the idealized monolingual situations they describe, grammatical blending "delivers slightly new expressions that, however novel, are intelligible precisely because they are for the most part strongly anchored to existing constructions. When we hear an expression, we try to construct an integration network…. <…> We do as much blending as we need to do to make sense of the utterance…." A Lithuanian NS who is highly proficient in English and, therefore, very familiar with the English resultative construction, would probably be able to understand the unfortunate utterance in Blend 3 by reblending it as an English utterance. This would, however, require more work than is typically necessary for comprehension. A Lithuanian NS who is only weakly familiar with English, however, might not be able to work out the appropriate meaning of the incorrect construction at all. The result would probably be interpreted as an error. The most significant error is that the FL learner construction *Aš nušluosčiau mano akinius švarūs* does not obviously lead to the emergent concept CLEANLINESS, thereby defeating its own purpose.

4.1.4 Chunks, blending, and learner constructions

An English NS who has expressed the English resultative construction *I wiped my glasses clean* in Lithuanian as *Aš nušluosčiau mano akinius švarūs* cannot be said to have neutralized the construction, as neutralization implies the effective and natural-sounding expression of an L*x* concept in an L*y*. Kecskes & Papp (2000) do not specifically discuss the neutralization of constructions such as those discussed here. Their discussion focuses on the types of concepts that can be expressed as words, e.g., BASEBALL or GIRA. Nonetheless, implicit constructions are clearly an essential part of the lexicon. Their status as lemmas means that they must also be linked to conceptual knowledge structures. Moreover, it is known that idioms and non-idiomatic phrases are also included in the lexicon. This section now looks more closely at such phrases, commonly known as **chunks**.

Much research in the field of second language acquisition during the 20th century was conducted within the Chomskyan paradigm of universal grammar, which postulates that humans are born predisposed to acquire a language, that language acquisition is mainly a matter of setting a finite number of principles and parameters, and that the lexicon – redundant, full of synonyms and semantically opaque idioms and polysemous entries – has little, if anything, to do with the formation of grammar. Recent research, however, has shown otherwise. Tomasello (2006; 2000), for example, describes child language acquisition as *item-based*, meaning that the structure of children's early utterances derives from concrete words and phrases rather than from any innate syntactic schemas. Children learn imitatively, picking up linguistic expressions from the environment. In the terms developed above, it could be said that children begin by memorizing grammatical blends. Over time, these expressions are analyzed, categorized, and schematized, and from them a grammar develops. In the competition model (MacWhinney, 2005: 55) the same process is termed chunking. By storing ever-larger chunks as lexical items, learners, whether children or adults, can not only speed up production, but also process and "induce the grammar from implicit generalization across stored chunks."

The idea of chunking stems from Miller's (1956) research into short-term memory capacity, and it is worth noting that human memory makes use of chunks regardless of the domain of activity: sequences of numbers can be chunked to form a single unit, as can sequences of actions. In a sense, all hierarchically organized systems can be said to be chunked, from the physical (atoms form chunks, called molecules, which in turn chunk into elements, etc.) and biological (cells form chunks, called organs, which in turn chunk into bodily systems) to the social (the behaviors of societies are determined by the behaviors of smaller chunked units of those societies, such as labor unions, political parties, etc.). Language, of course, is no exception. "Chunking,"

according to Ellis (2003: 78), "is a basic associative learning process which can occur in and between all representational systems."

Dąbrowska (2004) provides a detailed and useful discussion of such prefabricated lexical units, or **prefabs**[76], and the role they play in language processing and language acquisition. First and most importantly, chunks, like words and implicit constructions, are pairings of form and meaning; unlike many words, but like constructions, they also have internal grammatical structure and are often partly unspecified. This means that they may contain slots, like constructions. The average size of a chunk is between two and three words[77]. For this reason, chunks often need to be combined with other chunks or single words in order to form complete utterances. This is done by all speakers as a matter of routine, and allows extremely fast processing as chunks do not need to be composed (grammatically blended) online. Thus, chunks are a processing shortcut used by both speakers and listeners. That they are common is attested to by Erman & Warren (2000), who examined idiomatic word combinations in the spoken London–Lund corpus (idioms automatically being chunks). They found that nearly 60% of all of the almost 500,000 words in the corpus were parts of chunks. As these chunks did not include non-idiomatic expressions such as *How are you?*, *What are you doing?*, etc., it is quite likely that 60% is an underestimate of the total proportion of chunks in average discourse.

The combining of chunks, of course, is a form of grammatical blending. Adult native speakers are proficient at this process from long practice, but children's attempts to combine chunks do result in errors. Consider the examples below, taken from Clark (1974):

(a) I don't know where's Emma gone.
(b) I want I eat apple.

[76] Although *prefabs* is clearly meant to distinguish lexical chunks from other varieties, the more common term *chunk* will be used in the discussion that follows.
[77] Idioms, of course, are long chunks which even the most avid supporters of UG admit are memorized in the lexicon, as no compositional rules can explain their meaning.

No conceptual integration network is needed to see that in (a) the child has combined the chunk *I don't know* with the chunk *where's* _____ *gone* and the name *Emma*. Similarly, (b) can easily be formed by inserting the chunk *I eat* _____ into the slot in the chunk *I want* _____; as such, (b) represents a child's attempt to create an acceptable blend, *I want to* _____ *a(n)* _____, from scratch. There is no essential difference between child language errors such as these and the errors of learners of an FL of any age. In both cases, speakers with an incomplete repertoire of memorized chunks and/or lexical items are doing what they can to convey meaning. Their blends may not be as smooth and polished as the "official" blends that are "approved" by the language, but they often get the job done.

For bilingual children, the process is even more complicated, as they additionally have to keep language-specific chunks separate. A Lithuanian-English bilingual child, for example, must learn that Lithuanian *kauliukai* are English *dice*. It is not usually the case, however, that the parents of such children explicitly teach them translations of this sort. Much more commonly, children will learn such words as they need them from situational contexts, i.e., when playing board games with dice. As Kecskes & Papp (2000) point out, language enters the CUCB through one language channel and must then be neutralized. Thus a child who first learns *kauliukai* when playing a game with his Lithuanian mother, and then refers to them as *little bones* when playing the same game with his English-speaking father at a later date, is probably attempting to neutralize the term by creating a crosslinguistic blend.

Developing a CUCB and neutralizing concepts, then, are a matter of acquiring a threshold level of grammatical blends and lexical items, whether for bilingual children or older learners of foreign languages. Learners fall back on L1 chunks as a scaffolding method prior to developing an independent FL channel in the CUCB. Although a pre-threshold learner has begun to acquire some words and phrases of the FL, he or she still conceptualizes in the L1. Language acquisition is thus mainly a matter of automating more and more FL

chunks of ever-increasing size. The students whose work is examined below are still in this acquisition process. Their errors will therefore be analyzed as erroneous grammatical blends of Lithuanian concepts with English words, essentially the reverse of the process demonstrated above in Blend 3. The results of inappropriate blending are learner constructions.

4.1.5 Analyzing the transfer mechanism: methodology

It is the aim of Part 4 to examine the learner constructions produced by AnRK students when writing in English as evidence of CLI, especially transfer. Although the analyses look at the influence of Lithuanian as an L1 on English as an FL, it is not to be understood that these errors are predictable or bound to occur in any particular order. Level of proficiency also plays a key role in production accuracy, as do individual learner differences and students' attitudes towards the culture of the language being studied (see Part 2). Lithuanian learners of English very often have preconceived notions about certain English grammar constructions that are absent in Lithuanian (e.g., the present perfect tense and articles). Whether or not the expectation of difficulty leads to increased CLI is not within the scope of this dissertation. The learner constructions produced by the students in this study are not simply predictable from the presence or absence of a structure in one or the other language. In fact, they are entirely dependent on students' individual language learning histories and the amount of English-specific lexical items (be they words, chunks, or implicit constructions) they have acquired.

In order to analyze erroneous crosslinguistic blends, the following seven-step process was developed and implemented: 1. Isolate learner constructions; 2. Determine the appropriacy of such constructions in Lithuanian; 3. Work out the underlying generic space information for the Lithuanian blend; 4. Isolate any emergent inferences in the original Lithuanian utterance; 5. Decide how best in English to express the generic space concepts to achieve the desired emergent inference; 6. Construct a conceptual integration network to determine whether the error resulted from an

inappropriate grammatical blend; 7. Test for exceptionality by asking a control group of Lithuanian-Russian-English multilinguals to translate sentences that were designed to prompt for the incorrect blends chosen for analysis. Each step will now be discussed in more detail.

Step 1. As students turned in various written assignments for evaluation, these were checked for unusual uses of English. In many cases the learner constructions seem to literally "jump off the page," as they either prompt for the construction of unusual blends or simply seem to make no sense. Large numbers of these constructions were collected. Those which seemed most representative of the difficulties caused by crosslinguistic blending were chosen for discussion.

Step 2. Lithuanian and English are languages with some very obvious differences in structure, from the use of cases in Lithuanian to the many verb tenses of English. Word-for-word translation between the two languages is therefore generally not possible, as it might be between, for example, Spanish and French. Even an extremely simple question like *Kiek tau metų?* cannot be translated one word at a time: *How many for you years?* is almost incomprehensible. The opposite, of course, is also true: *How old are you?* should not be rendered as *Koks senas esi tu?* It is a useful rule of thumb[78] that any English sentence which can be translated word-for-word into correct Lithuanian is probably itself erroneous. Thus, if an erroneous English sentence is correct in Lithuanian, this can be taken as evidence that the error may be due to improper crosslinguistic blending.

It must be acknowledged that Lithuanian learners of English also make errors which cannot be classified as improper blends. Such errors include those involving articles, phrasal verbs, and verb tenses, among others. For example, students frequently produce verb tense errors such as *I am riding a bus to university every day*. As there are no continuous verb tenses in Lithuanian, this is clearly an interlanguage error resulting from the student's unfamiliarity with

[78] Though clearly not true in all cases.

the English system itself, rather than from the interaction of the English and Lithuanian systems. Such errors are referred to as **intraference** in Scovel (2001: 51), and defined as "the confusion a language learner experiences when confronting conflicting patterns *within* the structure of a newly acquired language" (emphasis in the original). Note that such errors, if translated word-for-word into Lithuanian, result in Lithuanian errors as well: *Aš esu važinėjantis autobusu į universitetą kiekvieną dieną*. Such errors were therefore excluded from the analysis.

Step 3. In order to determine where the crosslinguistic blending process went awry, it is first necessary to determine the principles according to which the original L1 blend was constructed. This can be accomplished by establishing the generic space information for the L1 sentence. Recall Blend 3, above. There, the generic space represented only the final stage of the original English blend, namely, the concept of CLEANLINESS. It is this concept that the hypothetical speaker was attempting to express in Lithuanian. The problem with the Lithuanian utterance is its lack of emergent inference; such inferences are not present in the generic space to begin with, but arise in blended spaces as part of the blending process. However, before an emergent inference can be found, it is first essential to know what concept(s) the blend was attempting to express.

Step 4. In addition to establishing generic space information, it is also necessary to establish the emergent inference of the original blend. The student who translates *I wiped my glasses clean* wants both the [RESULTATIVE CONSTRUCTION] lemma structure and the CLEANLINESS emergent inference to be present in the Lithuanian translation. The two are merged in the original, and thus an L1 concept or chunk cannot be said to be neutralized until an appropriate FL expression has been found that can similarly merge conceptual structure and inference. If no such expression exists, then the original utterance may be impossible to translate without some loss of meaning. For this reason,

establishing the emergent inference of the underlying Lithuanian blend is also necessary.

Step 5. Having discovered the generic space structure(s) or concept(s) and emergent inference(s) underlying the learner construction, it is then useful to examine how such concepts and inferences are typically lexicalized in English. To continue the example under discussion, the [RESULTATIVE CONSTRUCTION] lemma is not the only way to talk about the results of actions in English. Other ways to draw attention to the CLEANLINESS of one's glasses could include *Look how clean I got my glasses*, *My glasses were sure dirty*, or even simply *I cleaned my glasses*. All of these expressions more or less directly focus the listener's attention on the glasses' current state of CLEANLINESS.

Step 6. A conceptual integration network such as the three Blends above can be drawn to show how the English learner construction is derived from the original, language-specific Lithuanian blend.

Step 7. With any learner construction, it is possible that it can be accounted for by peculiarities in the individual learner's developing language system, i.e., that it is exceptional and unrepresentative of L1 → FL CLI. At the same time, when students were writing the assignments in which the analyzed learner constructions were found, their attention may have been directed towards meaning rather than form, that is, they may have been thinking less about accuracy of grammar and more about expressing their communicative intentions.

For each of the eight examples, therefore, Lithuanian sentences were constructed based on the results of Step 2. These were given to a control group of 20 multilingual speakers of Lithuanian, Russian, and English, also university students, but not those whose writing is analyzed here. They were asked to translate the sentences into English using any resources at their disposal. The translations were collected after one week[79]. Thus, this control

[79] The sentences, with instructions, can be found in Appendix 3.

group had both ample time to work and did not need to concentrate on meaning, but could focus entirely on accuracy of form. As will be seen, the learner constructions collected in the original written corpus were also produced by this control group of translators, proving that the analyzed constructions are not exceptional, but representative of those created by Lithuanian students of English as an FL.

4.2 Crosslinguistic Blending in Writing

This section examines eight learner constructions. The data for the analysis of written production were collected from 23 students of the English and Russian Languages study program at Vilnius University, Kaunas Faculty of Humanities[80]. Students were given written, marked assignments, and did not know that their work would be used for this purpose. These writing tasks were assigned as homework, and students had up to a week to prepare them. Unlike speaking assignments such as the one analyzed in Part 3, written assignments do not pressure students with time limits or other performance stressors (cf. Skehan, 1998). Moreover, students can refer to dictionaries, grammar books, the Internet, and other resources at their leisure in the preparation of such assignments. In other words, a written assignment represents the upper limit of a student's potential: what he or she can produce under ideal conditions.

Example (1) shows what happens when a student attempts to neutralize a language-specific idiomatic expression.

(1) I gripped her hand so tightly, that's why she screamed <u>not into her voice</u>.

The underlined learner construction would be entirely incomprehensible without some knowledge of Lithuanian, as it looks nothing like any typical English expressions. Translated into Lithuanian, however, it is perfectly acceptable: *ji suklykė nesavu balsu*. This Lithuanian chunk, *nesavu balsu*, is a

[80] Learner constructions were, in fact, collected from 45 students, but again, as in Part 3, only those students for whom L1 = Lithuanian were considered for the analysis in this dissertation.

colorful metaphoric expression which is clearly the result of grammatical blending. It is a chunk, a phrasal lexical item which results from a blending operation whose generic space codes for EMPHASIS; as such, the expression is equivalent to other emphatics, especially adverbials like *garsiai* "loudly", *šaižiai* "shrilly", etc. Adverbials often encode manner, and *nesavu balsu* is no exception; it does so, however, in the form of the emergent inference that attaches to the expression as a result of its being a grammatical blend. *Nesavu balsu* implies that the person in question was either terrified or in extreme pain, a meaning that would not automatically attach to the single-lexeme adverbials mentioned above. This lively Lithuanian chunk, then, goes beyond the simple concept of EMPHASIS to provide extra information about both manner and cause. In order to perform such a blend in English, the sentence would need to be radically rewritten. One English idiom that could be used here is *she screamed loud enough to wake the dead*, a chunk which similarly combines the concept of EMPHASIS with both manner and cause inferences. Alternatively, one could resort to a simile of some sort: *she screamed like a* _____, where any appropriate non-human source of noise could be inserted into the slot: *banshee*, *siren*, etc. Finally, it would also be possible to say *she screamed in a voice that was not her own*, perhaps the most faithful translation but one which sounds rather formal and wordy.

<table>
<tr><th>Generic Space</th><th colspan="2">Lithuanian Input Spaces</th><th>English Input Spaces</th><th>Blended Space</th></tr>
<tr><td rowspan="4">EMPHASIS</td><td rowspan="2">nesavu</td><td>"ne"</td><td>NOT</td><td rowspan="4">not into her voice

Emergent inference:
(none)</td></tr>
<tr><td>"savo"</td><td>SHE_{POSS}</td></tr>
<tr><td rowspan="2">balsu</td><td>"balsas"</td><td>VOICE</td></tr>
<tr><td>INST</td><td>INTO</td></tr>
</table>

Blend 4. "not into her voice"

Blend 4 illustrates the crosslinguistic blending operations that this student engaged in. The generic space is structured by the EMPHASIS concept and linked primarily to the Lithuanian-specific chunk *nesavu balsu*. This chunk is unblended into two separate words, each of which are further unblended into

lexemes and even one morpheme, the instrumental case. These are then linked to "equivalent" lexical items in the English space. That they are lexical items and not only lexemes is attested to by, first, the use of *her* in the finished blend (the production of this word requires accessing morphological information in the lexeme of SHE), and, second, by the correct structuring of the prepositional phrase *into her voice* (instead of: *her voice into*, which would follow the Lithuanian order more precisely). This demonstrates that, although this particular crosslinguistic blend is a failure, the student is nonetheless using English principles in its construction. She is actively trying to convey meaning using what knowledge she has of her FL.

The sentence given to the control group to translate was *Taip smarkiai suspaudžiau jos ranką, kad ji suklykė nesavu balsu*. The majority of the translations either avoided the word *nesavu* entirely[81], or translated it using adjectives: *foreign, unrecognisible*[82]*, ill-at-ease,* and *high*. Out of the 20 students in the control group, four (20%) attempted to translate *nesavu*: two wrote *she screamed not her own voice,* one wrote *she screamed not in her own voice,* and one wrote *she has cried not the voice*. For at least one-fifth of this group, this language-specific Lithuanian grammatical blend has not yet been neutralized.

The error exemplified in (2) is one of the most common caused by Lithuanian → English CLI[83]. In this case, both grammatical blending and cultural-pragmatic factors may play a role in causing this error and contribute to its prevalence.

(2) We with our team made desition to try to dance something like hiphop

[81] One student simply wrote a question mark.
[82] Student writings are presented exactly as they were written, and have not been edited for spelling or punctuation.
[83] It was also seen in Part 3 in the analysis of transfer, example (23).

This erroneous *we with* _____ pattern appears time and again, and clearly reflects the correct Lithuanian chunk *mes su* NP_{INST} (in this case, *komanda*). This chunk usually appears as the subject of the sentence; as such, it is an implicit construction for any first-person plural AGENT. In this case it is understood that the writer includes herself as one of the actors in the decision-making event. Such a construction could be lexicalized in one of three ways in English: *we*, *our team*, or *the team and I*. Similarly, in Lithuanian it would not be incorrect to say simply *mes* or *mūsų komanda*. The student's choice of *mes su komanda* reflects the pragmatic, emergent inference this grammatical blend conveys: by explicitly mentioning both the speaker and the other actors, the roles of both are highlighted such that the student's membership and active participation in the team's decision-making process are foregrounded. The English implicit construction that structures for first person plural AGENT with the emergent inference of group membership is *the team and I*.

Generic Space	Lithuanian Input Spaces	English Input Spaces	Blended Space
[AGENT 1 PL]	*mes*	WE	we with our team **Emergent inference:** (mismatch)
	su	WITH	
	[SAVO]	WE$_{POSS}$	
	komanda	TEAM	

Blend 5. "we with our team"

In Blend 5, the student has unblended the Lithuanian original into its constituent parts and translated each individually. This *we with NP* learner construction, as mentioned above, is quite common, and one reason for its prevalence might be the cultural-pragmatic associations of the constructions. Students are often aware of the English-appropriate version _____ *and I*, yet continue to produce the Lithuanian-influenced error seen in (2). Lithuanian being a nonconfigurational language with pragmatic word order, Lithuanian speakers are likely to be more sensitive to the nuances of word order than English speakers. To reverse the process, imagine an English speaker saying *komanda ir aš nusprendėme*. No Lithuanian speaker would accept such a

construction as natural-sounding. Indeed, by moving the pronoun *I* to the end of the phrase, the English construction implies a slight effacement of self which may sound particularly foreign to Lithuanian ears. Students could therefore be tempted to cling (at a conceptual level) to their Lithuanian-framed construction.

The presence of *our* in the final English blend, which may be absent from the Lithuanian original, can be explained at the lemma level. In Lithuanian it would be permissible to include the pronoun *savo* before *komanda*, though this is not a requirement. This has been indicated in the conceptual integration network by marking [SAVO] as a lemma, rather than an overt word form. In English, however, *we with team* would clearly be incorrect: an adjective or determiner is required to explain which team is being discussed. Thus, in addition to unblending the expression *mes su komanda* into three separate words, the student has also included the lemma [SAVO] into the final blend, showing that while she has developed some sensitivity to English norms, she has not yet developed enough self-confidence in her English skills to drop her reliance on Lithuanian models for crosslinguistic blends.

The control group translated a slightly modified version of this sentence: *Mes su draugais nusprendėme pabandyti sušokti hiphopą*. The original writing assignment contained many references to the team referenced in (2), whereas it was felt that for a context-free translation exercise a more common word like *draugai* "friends" would be more appropriate. Of the 20 students in the control group, eight (40%) created similar learner constructions: four wrote *we with friends*, and two wrote *I with my friends*; the other two wrote *we with our friends* and *we and our friends*.

The sentence in example (3) demonstrates another very common learner construction, produced by students at all ages and levels of acquisition[84].

(3) international marriages have and their bad side that one of the couple has to leave from his native country

[84] This, too, was seen in Part 3, example (24).

Here the focus of the discussion will be on this usage of *and* by Lithuanian speakers. The clause containing the learner construction could be expressed in accurate Lithuanian as *tarptautinės santuokos turi ir blogąją pusę*. In this Lithuanian construction, the word *ir* "and, also, too" is compressed from the larger construction (chunk) *ne tik* _____, *bet ir* _____ "not only _____, but also _____".

In Lithuanian, *ir* therefore ceases to mean only *and*, but takes on the emergent meaning of *also* or *too*. In English, however, *and* can only mean *and*, and the equivalent expression *not only* _____, *but also* _____ requires the use of a new lexeme, *also*.

In (3), the construction *ne tik* _____, *bet ir* _____ is only implied; the reader is expected to infer it from the context, and to mentally fill in the missing elements. These elements can be assumed to be parallel to those overtly mentioned; e.g., *blogąją pusę* is a noun phrase, so another noun phrase is likely to be implied. The Lithuanian construction could, for example, be expanded into *tarptautinės santuokos turi ne tik gerąją, bet ir blogąją pusę*. This implied contrast is the emergent meaning of the Lithuanian grammatical blend, and the author of (3) was hoping to convey the same implied contrast in English.

In this crosslinguistic blend, the Lithuanian *ir* lexicalizes the grammatical concept of third person plural THEME, even though the noun phrase explicitly mentioned in (3) is singular. The presence of *ir* sets up an internal (conceptual) conflict which can only be resolved by appealing to schematic knowledge about, e.g., the fact that objects are construed as having two sides, and that if an abstract noun is claimed to have a bad side, it must also have a good one, etc[85]. Unfortunately, when speaking English, *and* cannot be used in this way in this construction; either the formal construction *not only* _____, *but also* _____ must be used, or the informal *too*, as in *international marriages have a bad*

[85] The expression *blogoji pusė* "bad side" is itself a chunk, and should be translated as one; see below.

side, too. Notice that the expression with *too* mirrors the proposed Lithuanian construction: the other object of the verb *have* that is being contrasted with *a bad side* is only implied, and it is up to the reader to infer it from the context. Thus, the *too* construction contains the same emergent meaning that the author of (3) intended, but failed to convey, and it could be suggested as the best translation equivalent. However, (3) is a mismatch because the author has simply translated *ir* as *and*, thereby losing the emergent meaning.

Generic Space	Lithuanian Input Spaces	English Input Spaces	Blended Space
[THEME 3 PL]	*turi*	HAVE3PL	have and their bad side **Emergent inference: (mismatch)**
	ir	AND	
	[SAVO]	THEYPOSS	
	blogąją	BAD	
	pusę	SIDE	

Blend 6. "have and their bad side"

As an illustration, Blend 6 shows the unblending and translating process for (3). The generic space [THEME 3 PL] applies to the whole verb phrase. As above, the Lithuanian implicit [SAVO] lemma appears as an overt form in the crosslinguistic blend. That *blogąją pusę* has been separated into two individual inputs is attested to by the control group data. There, four students (out of the 20 who translated the sentence *Tarptautinės santuokos turi ir blogąją pusę – vienam iš poros teks gyventi toli nuo gimtosios šalies*) used English chunks or words instead of the phrase *bad side*: *flip-side* (chunk), *drawbacks* (compound word), *disadvantige* (word). At the same time, two students (10%) in the control group used *and* to create the same learner construction that was seen in (3). Finally, as proof that *ir* prompts for the *ne tik* _____, *bet ir* _____ construction, it should be noted that one student in the control group wrote *international marriages have positive and negative part*, even though nothing in the sentence they were given to translate mentions a positive or good side.

Example (4) presents a pair of sentences containing another relatively frequent error caused by Lithuanian-English CLI:

(4) a. She collapse so painfully on the stairs, she <u>almost didnt broke</u> her leg.
b. A pedestrian was walking on the street and he <u>almost wasn't hit</u> by a car.

In this pair, the Lithuanian-appropriate grammatical blend *almost* + V_{NEG} would cause confusion for English NSs, who would spend some time working out whether the event in question actually occurred or not. In English, with its strong rules of negation, (4a) means that she broke her leg, and (4b) means that he was hit by the car. In Lithuanian, however, these expressions have the opposite meaning. They are also meant to be used with the negative particle: *Ji taip skaudžiai nukrito ant laiptų, kad <u>vos nenusilaužė</u> kojąs*; *Pėstysis ėjo gatve ir jo <u>vos nepartrenkė</u> mašina*. The closest equivalent English expressions change the polarity of the verbs, as in *almost broke* or *was almost hit*, to arrive at the correct interpretation that in (4a) her leg did not break, and in (4b) the pedestrian was not hit.

Before proceeding, it will be useful to examine the Lithuanian word *vos* in more detail. When combined with a negative verb, as in the translations of (4a-b), it seems to be equivalent to the English *almost*. However, when combined with a positive verb, it has a different meaning. Consider the expression *Vos pradėjau rašyti, kai suskambėjo telefonas*. In this sentence *vos pradėjau rašyti* cannot be translated as *I almost started writing*, because that would imply that no writing had yet been accomplished before the telephone rang. This is clearly not the case. A more appropriate English equivalent would be *I had just started writing*. Thus, the Lithuanian *vos* and the English *almost* are not always synonyms, as the meaning of *vos* is partially unspecified and depends on the polarity of the verb that follows it.

Thus, the Lithuanian expression *vos* + a negative verb is a construction which could be glossed as [BEVEIK] "almost", while *vos* + a positive verb glosses as [TIK KĄ] "just". When the Lithuanian [TIK KĄ] construction is used with *vos*, the degree to which the event occurred is relatively small. If *I had just started writing*, this implies that a sentence or two had been written, no

more. When the Lithuanian [BEVEIK] construction is used with *vos*, the event did not occur, but the degree to which it failed to happen is similarly small – in (4b), for example, the car can be imagined as avoiding the pedestrian by only a few centimeters.

Generic Space	Lithuanian Input Spaces	English Input Spaces	Blended Space
[BEVEIK]	*vos*	ALMOST	almost didnt broke
	ne-	DO$_{PST}$	
		NOT	**Emergent inference:** (mismatch)
	nusilaužė	BREAK$_{PST}$	

Blend 7. "almost didnt broke"

Blend 7 shows the process by which the author of (4a) arrived at her mismatch. This example is especially interesting because it also contains an error in morphology, namely, the tense of the verb *break*. The generic space prompts for the Lithuanian-specific [BEVEIK] construction *vos* + V_{NEG} which is unblended into three Lithuanian input spaces. There are three rather than two because the student has been especially diligent about unblending the Lithuanian phrase. First, *vos* has been separated from its following verb and linked to the English lexical item ALMOST. Then the negative particle has also been separated from the verb. One explanation for this choice could be that the student's developing knowledge of English (in which the negative particle is always separated) is influencing the crosslinguistic blend. Finally, the verb *nusilaužė* has been faithfully translated in its past tense form, despite the fact that in English the past tense marker attaches to the negative particle in the form of an auxiliary verb, leaving the main verb in the infinitive. Such "double past" constructions are typical of the overgeneralizations found in child language acquisition (cf. Pinker, 1994), and it is therefore interesting to note that students of English as an FL are also subject to producing such learner constructions. The emergent meaning of this crosslinguistic blend is a mismatch, as it is possible to understand what the student intended to say, albeit with more effort than should be required for fluent conversation.

The control sentence for translation was *Ji nugriuvo ant laiptų ir vos nenusilaužė kojos*. Four students (20% of the control group) used some form of negation in the verb phrase: *nearly has not broken her foot*, *barely haven't broken her leg*, *hardly didn't break her leg*, and even *hardly unbroke her leg*. However, five more students (an additional 25%) used adverbial phrases that also imply that the leg was actually broken, e.g., *barely broke*, *hardly broken*, and *narrow broke*. Thus, nearly half of all the control subjects were unable to appropriately translate this Lithuanian-specific construction.

Another troublesome area for students is the language-specific differences in the countability of nouns, as exemplified by the two learner constructions in (5):

(5) a. me, my sister and mum went to Germany to my cousins weddings
b. A good new is that our government seems to fight with corruption

In (5a), a singular countable noun has been made plural, which in principle is not erroneous; however, the wider context reveals that in this situation, the choice of plural is inappropriate. In (5b), an uncountable plural noun has been made singular; as there is no singular English noun *new*, this word can only be interpreted as an adjective (as in, *a good new film*, etc.), causing confusion and possibly incomprehension on the part of any English NS readers.

NUMBER is a grammatical feature that is lexicalized differently in different languages. In English, for example, the following nouns are countable, which requires that individual exemplars be singular: *shirt*, *door*, *gate*, *year*, *ladder*, *funeral*, and *Christmas*. In Lithuanian, however, they are never singular without a change in meaning: *marškiniai*, *durys*, *vartai*, *metai*, *kopėčios*, *laidotuvės*, and *Kalėdos*. Similarly, the singular English *wedding* is represented in Lithuanian as *vestuvės*, a plural noun. Interestingly, both Lithuanian and English lexicalize the concept NEWS in the plural (LT *žinios*). However, in Lithuanian it is possible to use this word in its singular form, *žinia*

"a piece of news; a message," and the author of (5b) apparently believes that the same is possible in English.

By blending the PLURAL morpheme with the concept of a ceremony uniting two people in matrimony, Lithuanian imbues the word *vestuvės* with emergent meaning not found in the English equivalent. For one thing, it is very common for Lithuanians to marry twice in one day: once in a church, and then again in a civil office. Thus the Lithuanian concept includes much more than the marriage ceremony, as the plural marker seems to indicate that all of the events that occur throughout the entire day (or, often, days) during which the young couple are married and celebrate with family and friends are to be considered part of the overall concept VESTUVĖS. In English, a *wedding* is a formal event that occurs either in a church or a civil office (but not both) and typically ends with the words, "I now pronounce you husband and wife." The events that follow (reception, honeymoon, etc.) are not considered to be part of the concept WEDDING. Therefore, while the author of (5a) may indeed have had the Lithuanian concept in mind, her use of the plural marker in English is confusing. It is even more confusing because the word *cousins*, which should be a possessive singular noun (*cousin's*), contains no apostrophe, leading to the impression that several cousins were getting married in a large number of weddings more or less at the same time.

The uncountable noun *žinios* can be made countable in Lithuanian simply by dropping the plural marker. When this noun is used in the singular, the emergent meaning of the word is that of a small amount; indeed, the related word *žinutė* "a note; a short text message (e.g., SMS)" is formed from the singular *žinia* and the diminutive marker *-utė*, making the emergent inference explicit. In English, however, the concept NEWS cannot be modified in this way. Either the word must be dropped in favor of an alternative (e.g., *message*, *note*, etc.) or a phrasal construction such as *a piece of news* can be used. What is most interesting is that in the case of (5b), in fact, neither option is necessary; the best alternative would be to simply leave it uncountable: *the*

good news is…. Thus, by relying too heavily on her Lithuanian conceptual base, this student has created an English mismatch.

Generic Space	Lithuanian Input Spaces	English Input Spaces	Blended Space
VESTUVĖS	*vestuvės*	WEDDING	weddings
	PL	"-s"	**Emergent inference:** (mismatch)

Blend 8. "weddings"

Blend 8 illustrates the process for example (5a). Here it can be seen that the generic space contains the Lithuanian-specific concept VESTUVĖS. It is linked to two Lithuanian input spaces in this network, as this seems to best represent how the student has translated both the word and its plural marker separately. This unblending, as in the previous examples throughout this section, is the cause of the mismatch found in the final blended space.

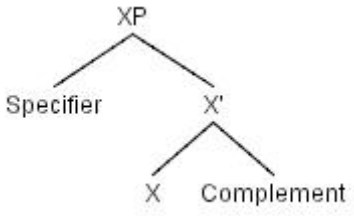
Three students, or 15% of the control group, also translated *vestuvės* as *weddings*. The sentence they were given was *Praėjusią vasarą aš su mama ir sese važiavome į Vokietiją, į pusseserės vestuves*. Note that this sentence also contains the *mes su* NP_{INST} chunk seen in example (2). As might be expected, eight students, or 40% of the group, translated the chunk *aš su mama ir sese* as something like *I with mother and sister*. (The number of students who translated *mes su draugais* as *we with friends* was also eight. Interestingly, it was not the same eight both times.)

Because Lithuanian is a nonconfigurational language, the order of words and phrases within sentences can be varied much more than in English. If students continue to follow the Lithuanian word order when translating their thoughts into English, the results are likely to be mismatches or, in extreme cases, incomprehensible utterances. Example (6) shows how a difference in the word ordering requirements for determiners in Lithuanian and English can lead to syntactic errors in production:

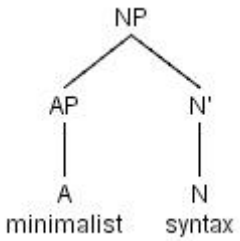
(6) The last our project lingered too long of not enough quantity of money and good ideas how to realise it.

In English, determiners like *our* have two very important properties: they cannot combine, and they almost always come first in the NP[86]. Thus, it is incorrect to say *the our project*, and it is incorrect to say *last our project*; example (6) is therefore wrong on both counts, as it must be simply *our last project*. As a Lithuanian utterance, however, it has no problems: *paskutinis mūsų projektas* is as acceptable as *mūsų paskutinis projektas* (but see the discussion below).

Before proceeding to the analysis, it will be useful to briefly review X' theory, a construct of Minimalist Syntax (Chomsky, 1995; Poole, 2002; Radford, 1997, 2004) that also sheds light on the error exemplified in (6). According to this theory, the underlying structure of any given phrase (NP, VP, AP, etc.) can be reduced to the following simple formula (adapted from Poole, 2002: 48):



where Specifier refers to any sub-level phrase that modifies the head X of the phrase XP, and Complement refers to any sub-level phrase which is joined in a complement relationship to the head X. In English, for example, adjectives commonly act as specifiers for nouns, as in the phrase *minimalist syntax*, which can be represented by the following diagram:



[86] Predeterminers such as *all*, *many*, *quite*, *such*, and *what* can, of course, come before determiners, but the total number of these words is very small.

Here, the adjective *minimalist* belongs to an AP which is the specifier of the head noun *syntax*; because the phrase ends there, no complement branches off from the N'. On the other hand, prepositional phrases take the form PP → P NP, as in *on the shelf* or *on Monday*, where the determiner phrase *the shelf* and the noun phrase *Monday* are complements of the preposition. Heads can naturally have more than one specifier or complement; the phrase *our last project* attaches both a determiner and an adjective as specifiers of the head noun.

The important point for the purpose of analyzing example (6) is that the organizing principle in this crosslinguistic blend is simply the language-specific syntactic rules governing the internal structure of NPs. In the Lithuanian utterance *paskutinis mūsų projektas* the Lithuanian generic space contains the grammatical formula NP → Spec Spec N, as just discussed. The emergent meaning of the blend has once again to do with the fact that Lithuanian is a nonconfigurational language. As mentioned, *paskutinis mūsų projektas* is not erroneous, but neither is *mūsų paskutinis projektas*. The difference between these two phrases in Lithuanian is not grammatical, as it would be in English, but pragmatic. The former expression implies that it is important to the author that this was the most recent (*last*) project in which she took part; the latter could imply either that a) the fact that she was a member of the team which prepared the project is more important than which project it happened to be, or that b) this was the final (*last*) project which she and her team prepared. As can be seen, the word *last* ~ *paskutinis* is polysemous in both languages. Because the author chose the former expression (as evidenced by the word order in the English example), the emergent meaning intended in the utterance is to resolve the ambiguity in favor of recency, rather than finality or group membership.

<table>
<tr><th>Generic Space</th><th>Lithuanian Input Spaces</th><th>English Input Spaces</th><th>Blended Space</th></tr>
<tr><td rowspan="4">NP → Spec Spec N</td><td rowspan="2">paskutinis</td><td>THE</td><td rowspan="4">the last our project
Emergent inference:
(mismatch)</td></tr>
<tr><td>LAST</td></tr>
<tr><td>mūsų</td><td>WE_{POSS}</td></tr>
<tr><td>projektas</td><td>PROJECT</td></tr>
</table>

Blend 9. "the last our project"

The conceptual integration network for the learner construction in (6) can be found in Blend 9. The determiner *the* has been added as a link from *paskutinis* on the assumption that the student is familiar with such English chunks as *the last straw*, *the Last Supper*, etc. This is purely hypothetical. It may simply be that, like many Lithuanian learners of English, articles are a particularly troublesome type of determiner, there being none in Lithuanian. Some students simply avoid using them at all, while others overcompensate and insert them wherever possible. This may therefore be an IDK error, as discussed in Part 1.

In the control group, five students (25%) created the learner construction *last our project*, one of whom also included the determiner *the* as in (6). The sentence they were given was *Paskutinis mūsų projektas tęsėsi per ilgai – pritrūko ir pinigų, ir gerų idėjų*. As can be seen, this sentence also contains another construction which uses *ir* to mean *too*: *ir pinigų, ir gerų idėjų* (e.g., *we ran out of money and good ideas, too*). The student who wrote (6) did not use *and* here, but one of the control group students did: *we shorted and money, and good ideas.*

The next example demonstrates both the transfer of pragmatic word order and the specifics of Lithuanian verb morphology:

(7) Of course [about spiritual or moral things <u>I even will not write</u>] because I do not believe that [things like that human can figure out] by himself.

In the underlined learner construction, the order of morphemes in Lithuanian and English is significantly different, due to the fact that a) English is a

non-null-subject language which requires explicit mention of subject pronouns, and Lithuanian is not; b) English forms the future tense with modal auxiliaries, and Lithuanian forms the future tense with suffixes; c) the negative particle in English is always separated from the main verb, and in Lithuanian it is attached as a prefix; and d) in English, adverbial specifiers such as *even* are usually placed immediately before the main verb, after the negative particle and any auxiliary verbs, while in Lithuanian they come before the verb and cannot be inserted between the various morphemes that it comprises.

In Lithuanian, the underlined phrase could be expressed as *net nerašysiu*. Here we see that the result as produced by this student is neither fully Lithuanian nor fully English, and is therefore an excellent example of the effects of CLI on students' written production. The grammatical blend created by this verb form establishes a counterfactual situation (Fauconnier & Turner, 2002) in which the student does write about spiritual and moral questions, presumably because she has learned all there is to know about them. This is the emergent inference of the blended verb. The grammatical structure in the generic space is simply the negative first person future form of the verb: "negation routinely sets up counterfactual blended spaces, which can be elaborated" (Fauconnier & Turner, 2002: 239).

To properly form a negative future tense verb phrase in Lithuanian, the following morphemes are required: the negative particle *ne*, the verb stem (in this case, *rašy*-), the future tense infix -*s*-, and an appropriate number marker which, as this is a first person singular sentence, in this case should be -*iu*. In addition, the entire VP is being modified by the adverbial specifier *net*, which, as mentioned, comes before the verb in Lithuanian. This order of morphemes is significantly different in English, in which the following order would be acceptable: first, the first person singular pronoun *I*, then the future modal auxiliary *will*, the negative particle *not*, the adverbial specifier *even*, and finally the verb stem *write*. These morpheme orders are presented for comparison in Table 9.

Table 9. Comparison of negative future tense morphemes, LT and EN.

Lithuanian					English				
net	*ne-*	*rašy*	*-s-*	*-iu*	I	will	not	even	write
even	not	write	will	I	*-iu*	*-s-*	*ne-*	*net*	*rašy*

The author of (7) has become confused about the exact order of these morphemes. Blend 10 shows a conceptual blending network in which, as discussed above, the generic space is meant to provide the VP schematic for the grammatical blend. By separating the five morphemes and translating them individually, however, the student has created an erroneous blend. Clearly the English word order arrived at by following the Lithuanian morpheme order, as demonstrated in Table 9, is incorrect; but the correct order is also difficult to determine when faced with five competing morphemes. To resolve this competition, the student seems to have chosen a rather expedient solution: knowing that in English subject pronouns almost always come first, she has taken the standard order S–Aux–Neg–V and inserted *even* into the first possible spot after S. The emergent meaning of this blend is a mismatch, but one which can be resolved with relatively little cognitive effort.

Generic Space	Lithuanian Input Spaces	English Input Spaces		Blended Space
VP$_{\text{NEG FUT 1SG}}$	*net*	EVEN	S - Aux - Neg - V	I even will not write **Emergent inference:** (mismatch)
	ne-	NOT		
	rašy	WRITE		
	-s-	WILL		
	-iu	I		

Blend 10. "I even will not write"

In the control group, only one student (5%) created this learner construction. This rather small number, however, may simply be due to the fact that 13 of them (a full 65%) did not translate the word *even*. One possible explanation for this is that they were unsure where to include it and simply avoided the problem altogether. Had they included it, the number of students producing *I even will not write* may well have increased.

The internal structure of the verb phrase is not the only problem with (7), however. Two structural errors are marked in brackets in the example. Both of these reflect the strong influence exerted on the developing FL language system by the language-specific norms of the student's L1. Only the first example will be analyzed here. Two factors seem to be at work: the pragmatic word order of Lithuanian, and the English-specific grammatical blending operations encoded by phrasal verbs.

Structurally, the generic space encodes a structure for the utterance based on Lithuanian pragmatic word order rules: the THEME is brought to the front and foregrounded. The ACTION follows, in its counterfactual form as seen in the discussion above. The AGENT is probably only encoded as the first person singular morpheme on the verb. To keep this structure in English one could use a passive form of the verb, as in *Spiritual and moral issues will not be written about*. This leads to the second factor affecting this learner construction. The English verb *write about* is a phrasal verb which points to lemmas such as [DESCRIBE] and [DISCUSS]. At the same time, the non-phrasal verb *write* can take PP complements (e.g., *write to me*); it would therefore be possible to find sentences where *write about* is non-phrasal, as in *I wrote about an hour and then went to sleep.* In Lithuanian, phrasal verbs are formed using prefixes, thus distinguishing them clearly from their non-phrasal counterparts: thus *rašiau apie medžius* and *aprašiau medžius* have two distinct senses, even though both could be translated as *I wrote about trees*. In the former, the [DISCUSS] sense is highlighted, while the latter highlights the [DESCRIBE] sense.

The example in (7) separates *about* from *write* by a distance of nine words. This is likely to interfere with the interpretation of this verb as phrasal. In the conceptual integration network shown in Blend 11, there are two generic spaces corresponding to the Lithuanian pragmatic word order structure discussed above. The ACTION/AGENT blend was discussed above as Blend 10. For the THEME argument, the Lithuanian input spaces are broken into syntactic phrase units to demonstrate that the phrase *about spiritual or moral things* is

being treated as a single unit, a prepositional phrase, which is composed of sub phrases; these, represented as individual Lithuanian words, are finally matched to English lexical items and blended together with the blend seen in example (6) to produce the final learner construction.

The sentence the control group was given to translate was *Aišku, apie dvasinius ar moralinius dalykus net nerašysiu, nes nemanau, kad žmogus gali visą tai išsiaiškinti pats*. Seven students (35%) followed the Lithuanian pragmatic structure and broke up the English verb *write about*, creating learner constructions in which the PP beginning with *about* was placed before the subject and verb. Interestingly, recalling the discussion of CLI in Part 3[87], five students (25%) also translated *žmogus* as *human*.

<table>
<tr><th>Generic Spaces</th><th colspan="3">Lithuanian Input Spaces</th><th>English Input Spaces</th><th>Blended Space</th></tr>
<tr><td rowspan="4">THEME</td><td rowspan="4">PP</td><td>P</td><td>apie</td><td>ABOUT</td><td rowspan="5">about spiritual or moral things
VP_{NEG FUT 1SG}

Emergent inference:
(mismatch)</td></tr>
<tr><td rowspan="3">ConjP</td><td>dvasinius</td><td>SPIRITUAL</td></tr>
<tr><td>ar</td><td>OR</td></tr>
<tr><td>moralinius</td><td>MORAL</td></tr>
<tr><td>NP_{ACC}</td><td>dalykus</td><td>THING_{PL}</td></tr>
<tr><td>ACTION_{COUNTERFACTUAL}
AGENT</td><td colspan="2">VP_{NEG FUT 1SG}</td><td>*</td><td>*</td></tr>
</table>

Blend 11. "about spiritual or moral things I even will not write"
*see Blend 10

The last example also combines several errors, though the error in syntax is the most serious; indeed, this sentence, like (4) above, probably results in incomprehension: not simply a mismatch, but a true error.

(8) <u>To prove my as dancers possibilities</u> and to wide knowledge which I have now about technique of dances

This sentence comes from a student's curriculum vita, in which she describes her experience as a dancer and expresses her desire to continue working in this field. The underlined section closely follows an acceptable Lithuanian

[87] See example (29).

syntactical pattern and can be translated as *Įrodyti savo, kaip šokėjos, galimybes*. When this pattern is directly translated into English, however, the result is difficult, if not impossible, to understand.

There are three types of learner constructions in (8), and while the focus of the discussion will be on the syntactical blend, brief mention will first be made of the others. One is orthographic: *dancers* should contain an apostrophe to help the reader understand that the author of the sentence is also a dancer, to whom the *possibilities* belong. (Actually, as shall be seen below, the word shouldn't be possessive at all, and thus the word should really be rendered *dancer*.) Secondly, the word *possibilities* is itself inappropriate in this context: a search of Google[88] revealed no instances whatsoever of the phrase *prove my possibilities*, and only 730 instances where the words *prove* and *possibilities* were joined by a pronoun of any sort (usually *its*). A search of the phrase *prove* ____ *potential*, however, turned up nearly 14,000 results, suggesting that, in this example, *potential* would be a more typical lexical choice.

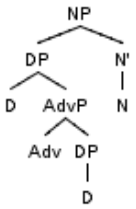
Leaving these troubles aside, the main problem with (8) is the insertion of the phrase *as dancers* between *my* and *possibilities*. In the Lithuanian version, this is done as a kind of parenthetical remark which establishes *šokėjos* as a synonym of *savo*. The Lithuanian pronoun *savo* is "light" in the sense that its meaning depends on the context – it can replace any other possessive pronoun. For this reason it is often modified with a complement (in the form of an AdvP beginning with *kaip*) to make its meaning more specific[89]. In English, on the other hand, possessive pronouns are person-specific and do not take complements themselves; they are determiners and, as such, are put in the specifier position of NPs.

Savo, kaip NP_{GEN} is a well-established Lithuanian grammatical blend, and a search of Google reveals more than 50,000 instances of constructions formed

[88] On the use of Google as a searchable reference corpus, see, for example, Mittelberg, Farmer & Waugh (2007), Taylor (2006), and Barnden (2006). The Google searches described in this section were performed in January 2009.

[89] In the lexical concepts and cognitive models theory proposed by Evans (2006), this process is called **adjustment**. It is the emergent meaning of the blend.

with *savo* followed by a parenthetical phrase containing the preposition *kaip* and a noun or noun phrase in the genitive case. It is therefore a chunk with one slot, and the only grammatical blending that needs to be done online to produce this phrase in Lithuanian is to attach the genitive suffix to the chosen NP.

Generic Space	Input Space 1	Input Space 2	Blended Space
V_{INF} O_{DEF} AdvP		[ĮRODYTI] [SAVO] [GALIMYBĖ] [KAIP] [ŠOKĖJA]	*Įrodyti savo, kaip šokėjos, galimybes* **Emergent inference:** ADJUSTMENT

Blend 12. "*Įrodyti savo, kaip šokėjos, galimybes*"

Blend 12 illustrates the Lithuanian-specific grammatical blend. The generic space codes for an infinitive verb, a definite object, and an adverbial phrase. The first input space contains the phrase-structure rules allowing the adverbial phrase to specify the noun by attaching to the determiner *savo* as a complement. The second input space contains the lemmas required for the blend[90]. The emergent meaning of adjustment arises in the final, blended space.

It should be noted that another possibility exists in Lithuanian for the expression of this blend, namely, to leave the adverbial after the object, rather than inserting it into the middle. Then the utterance would be *įrodyti savo galimybes kaip šokėja*. This once again demonstrates the pragmatic word order of Lithuanian, in which foregrounding (in this case, adjustment) is performed by moving phrases closer to the front of the sentence.

Now consider again the English learner construction in (8). English word order rules do not allow determiners to be separated from their nouns by adverbials; phrases such as **the last night party* or **put your quickly shoes on* are inappropriate at best. A more acceptable way to express (8) would be *To prove my potential as a dancer*, where the adverbial *as a dancer* is a complement of the object, *potential*. This eliminates the need for the

[90] As this is a Lithuanian-specific grammatical blend, and not a crosslinguistic blend, it works with lemmas during grammatical encoding.

possessive form of *dancer* seen in (8). The question remains of whether the emergent meaning attached to the Lithuanian original has transferred to this new English-specific blend. In the original, the phrase was intended to highlight the fact that the author is a dancer by making *šokėjos* a synonym of *savo*. In this English alternative, the phrase describes what kind of *potential* is being proven, and does not necessarily require that the author be a dancer already. Indeed, it could be understood that she is not a dancer now, but has the potential to become one in the future, which fact she wants to prove. Thus the English blend ends up encoding a slightly different meaning than the Lithuanian original. It is, however, still better to arrive at a slightly different meaning than at no meaning at all, especially when dealing with language-specific grammatical blends. Unfortunately, because of the large number of problems in (8), no meaning is what any English NS is likely to take from the utterance, unless he or she has a very good working knowledge of Lithuanian.

Generic Space	Lithuanian Input Spaces	English Input Spaces	Blended Space
ADJUSTMENT	*įrodyti*	PROVE$_{INF}$	To prove my as dancers possibilities **Emergent inference:** (none)
	savo	I$_{POSS}$	
	kaip	AS	
	šokėjos	DANCER$_{POSS}$	
	galimybes	POSSIBILITY$_{PL}$	

Blend 13. "To prove my as dancers possibilities"

The crosslinguistic blend for the learner construction in (8) is provided in Blend 13. It begins in the generic space with the emergent inference of the Lithuanian grammatical blend, adjustment. The already inflected, morphologically rich Lithuanian words are each taken as separate inputs and linked to English lexical items.

The control group was given this sentence to translate: *Noriu įrodyti savo, kaip šokėjos, galimybes ir pagilinti turimas žinias apie šokių techniką*. Six of them (30%) inserted either *as a dancer* or *like a dancer* between *my* and the head noun. This noun, incidentally, was translated as *possibilities* by six

students, as *potential* by four, as *opportunities* and *abilities* by three each, and as *skills* by two, indicating that students were trying to lexically accommodate the change in meaning between the Lithuanian and English versions, as discussed above. Two students chose not to translate this sentence at all. Perhaps they realized the inappropriateness of a construction like (8) in English, but were unable to think of a suitable alternative.

4.3 Summary and Discussion

Part 4 has analyzed eight learner constructions written by AnRK students. It was shown that the learner constructions found in these samples of students' written English production can be seen as evidence of crosslinguistic blending, a conceptual operation akin to grammatical blending but which typically involves unblending language-specific constructions and reblending them using target language lexical items, often resulting in structures that may be difficult to comprehend for anyone unfamiliar with the source language.

It has been stressed several times above that written production is inherently different from spoken. Because speaking is an online process that is highly resource-costly and affected by variables such as the speed of production, anxiety, audience comments, etc., the many different types of CLI seen in Part 3 are able to creep in, as it were undetected. However, because writing is offline and can be undertaken at leisure, it may be assumed that where CLI can be found in the written corpus, the students cannot correct it on their own; it stems not from online production constraints but from the structure of the developing English language systems.

This indicates that for these students English is an FL, and English concepts, chunks, syntactic rules, and pragmatic conventions may all still be unfamiliar or even unknown to them. As such, they must rely on their knowledge of their primary language of communication, Lithuanian, as a scaffolding upon which to hang the English words and phrases they are familiar with. Recall Kecskes & Papp (2000: 108, 117): "L2 language production is *heavily influenced by the L1-dominated conceptual base*…. Until

multicompetence occurs, a typical language learner will *think in the L1*" (emphasis added). It is precisely the fact that they think in Lithuanian that causes these students to create learner constructions.

In these eight examples, the language-specific nature of the CUCB plays a role in the creation of errors. One of the more pervasive errors is caused by the pragmatic word order of Lithuanian, which allows phrases to be moved to the front of sentences for the purpose of foregrounding or focusing the topic. As seen above, this language-specific feature causes problems time and again. Similarly, many concepts are plural in Lithuanian but singular in English, a type of language-specific encoding that also leads to crosslinguistic blending. According to Kecskes & Papp (2000), language-specific concepts such as these must be neutralized in order to become available through both language channels in the CUCB. Neutralization is a process which requires some cognitive effort to perform, but saves time and prevents error in the long run. For example, the student who wrote *weddings* could be encouraged to neutralize the concept by, e.g., spending time thinking about the differences underlying these two concepts in the two languages (see the discussion above), and practicing saying and writing *wedding* in the singular (e.g., through imagined dialogues or other repetitive spoken or written tasks), until such time as the word *weddings* in the plural begins to sound incorrect to her. This effort would be rewarded in the future by more appropriate usage of this word. The LME process just described differs in no essential way from the resonance-enhancing activities described in MacWhinney (2005).

The learner constructions analyzed in this section show that these students, rather than neutralizing concepts, are approaching the task of producing written English as a kind of linguistic problem to be solved through decomposition. In effect, they are overanalyzing their own language production processes by unblending constructions into individual words and then linking these one by one to words in the FL. It would be more useful and appropriate for them simply to learn English chunks as equivalents of

Lithuanian chunks. The crosslinguistic blending seen in these examples actually requires more cognitive effort than simply recalling stored chunks. As an example, consider again the learner construction seen in (1):

(1) I gripped her hand so tightly, that's why she screamed not into her voice.

This construction was created by a student's attempt to translate a Lithuanian chunk, *nesavu balsu*, one morpheme at a time. This task is particularly difficult for two reasons: first, the word *savo* has no direct English equivalent, and must be translated with an appropriate pronoun; second, the fact that both *nesavu* and *balsu* are in the instrumental case, which doesn't exist in English, means that the student will have to essentially guess at an appropriate preposition to use instead. Meanwhile, several English idioms, e.g., *like crazy*, *her head off*, *loud enough to wake the dead*, etc., are both conceptually and semantically appropriate to use as an equivalent. In other words, *ne savo balsu* can be neutralized simply by learning to express it as any one of the above options. Neither chunk needs to be unblended for this to happen, and no translation of individual words or grammatical cases is necessary.

The examples analyzed in this section clearly show the effects of CLI in students' written English production. Indeed, a conceptual blending analysis such as this shows that both languages are in direct contact at the conceptual level – through the CUCB – no matter what the final language of production may be. A student writing in English nonetheless continues to conceptualize in and translate from Lithuanian. As a result, English is adapted to suit the requirements of Lithuanian: English words are used in place of Lithuanian ones, in structures that do not exist in English, in ways that mean the opposite of what is intended, etc. The rules of English are bent by the translation process, which is a process of conceptual blending in which English words are combined with Lithuanian concepts to produce utterances that range from unusual to incomprehensible.

CONCLUSIONS

This dissertation addressed the following hypotheses:

1. **The linguistic attitudes of the multilingual subjects affect their communicative competence in FL.**

This hypothesis was confirmed. Attitudes were tapped through a detailed sociolinguistic survey which confirmed that AnRK students have strong allegiances to Lithuanian and Russian cultures, yet remain estranged from English-speaking cultures. Subjects' linguistic identities correlate with their ethnic identities. This cultural self-identification affects both linguistic competence and linguopragmatic behavior. Communication in FL generally follows L1-appropriate patterns.

2. **The L1 linguistic and sociolinguistic competence of the multilingual subjects affects their FL competence.**

This hypothesis was confirmed. The analysis of the 25,000-word corpus of spoken and written discourse collected for this dissertation confirmed that L1 lexical, semantic, grammatical and even orthographic competence are important factors in the formation of communicative competence in an FL. The social and linguistic usage strategies formed through years of L1 experience are unavoidably brought to bear on FL competence. Language production in the FL thus continues to evidence L1 sociocultural traits.

3. **Crosslinguistic influence takes the form of transfer of L1 skills into the FL.**

This hypothesis was partially confirmed. While transfer is indeed a common form of CLI in both spoken and written production, it is only one of the seven types found in the corpus. Hesitation, code-switching, neologisms and

foreignizing appear to be caused less by transfer than by the presence of (personal and/or cultural) lexical gaps.

4. The mechanism underlying the transfer of L1 skills into the FL is conceptual blending.

This hypothesis was confirmed. Conceptual blending is a productive skill used within languages to create grammatical blends that encode particular implicit constructions used for high-speed communication. When these constructions are unblended and translated word-for-word, however, the resulting crosslinguistic blends can be difficult to parse by anyone unfamiliar with L1 linguistic norms.

REFERENCES

1. Acton, W. and Walker de Felix, J. (1986) Acculturation and Mind. In J. Valdes (ed) *Culture Bound*, pp. 20–32. Cambridge: Cambridge University Press.
2. Adamson, H. D. (1988) *Variation Theory and Second Language Acquisition*. Washington, D.C.: Georgetown University Press.
3. Agarwal, A. and Bhattacharyya, P. (2006) Augmenting WordNet with Polarity Information on Adjectives. In *Third International WordNet Conference (GWC 06)*, Korea.
4. Alptekin, C. (2002) Towards Intercultural Communicative Competence in ELT. *ELT Journal* 56 (1): 57–64.
5. Austin, J. L. (1962) *How to Do Things with Words*. Oxford: Clarendon Press.
6. Barnden, J. A. (2006) Artificial Intelligence, Figurative Language and Cognitive Linguistics. In G. Kristiansen, M. Achard, R. Dirven, and F. J. R. de Mendoza Ibáñez (eds) *Cognitive Linguistics: Current Applications and Future Perspectives*, pp. 431–459. New York: Mouton de Gruyter.
7. Bates, E. and MacWhinney, B. (1982) Functionalist Approaches to Grammar. In E. Wanner and L. Gleitman (eds) *Language Acquisition: The State of the Art*, pp. 173–218. New York: Cambridge University Press.
8. Beattie, G. (1983) *Talk: An Analysis of Speech and Non-Verbal Behavior in Conversation*. Milton Keyes: Open University Press.
9. Bialystok, E. (2005) Consequences of Bilingualism for Cognitive Development. In J. F. Kroll and A. M. B. de Groot (eds) *Handbook of Bilingualism: Psycholinguistic Approaches*, pp. 417–432. Oxford: Oxford University Press.
10. Blum-Kulka, S., House, J. and Kasper, G. (1989) *Cross-Cultural Pragmatics: Requests and Apologies*. Norwood, NJ: Ablex.
11. Bock, K. and Levelt, W. J. M. (1994) Language Production: Grammatical Encoding. In M. A. Gernsbacher (ed) *Handbook of Psycholinguistics*, pp. 945–984. San Diego, CA: Academic Press.
12. Boers, F. and Lindstromberg, S. (2006) Cognitive Linguistic Applications in Second or Foreign Language Instruction: Rationale, Proposals, and Evaluation. In G. Kristiensen, M. Achard, R. Dirven, and F. J. R. de Mendoza Ibáñez (eds) *Cognitive Linguistics: Current Applications and Future Perspectives*, pp. 305–355. New York: Mouton de Gruyter.
13. Braun, A. (2008) It Sounds Clumsy in English: Conceptual Blending and Lithuanian-English Translation. *Respectus Philologicus* 13(18)A: 186–200.
14. Brown, P. and Levinson, S. D. (1978) *Politeness*. Cambridge: Cambridge University Press.

15. Brown, R. and McNeill, D. (1966) The "Tip of the Tongue" Phenomenon. *Journal of Verbal Learning and Verbal Behavior* 5: 325–337.
16. Bull, M. and Aylett, M. (1998) An Analysis of the Timing of Turn-taking in a Corpus of Goal-orientated Dialogue. *Proceedings of ICSLP-98, Sydney, Australia* 4: 1175–1178.
17. Cenoz, J. (2001) The Effect of Linguistic Distance, L2 Status and Age on Cross-linguistic Influence in Third Language Acquisition. In J. Cenoz, B. Hufeisen and U. Jessner (eds) *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*, pp. 8–20. Clevedon: Multilingual Matters, Ltd.
18. Cenoz, J. (2003) The Intercultural Style Hypothesis: L1 and L2 Interaction in Requesting Behaviour. In V. Cook (ed) *Effects of the Second Language on the First*. Buffalo, NY: Multilingual Matters, Ltd.
19. Cenoz, J., Hufeisen, B. and Jessner, U. (eds) (2001) *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*. Clevedon: Multilingual Matters, Ltd.
20. Chafe, W. (1998) Language and the Flow of Thought. In M. Tomasello (ed) *The New Psychology of Language*, pp. 93–111. Mahwah, NJ: Lawrence Erlbaum Associates.
21. Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
22. Chomsky, N. (1995) *The Minimalist Program*. Cambridge, MA: MIT Press.
23. Clark, H. and Schunk, D. (1980) Polite Responses to Polite Requests. *Cognition* 8: 111–143.
24. Clark, R. (1974) Performing without Competence. *Journal of Child Language* 1: 1–10.
25. Cook, V. (1991) The Poverty of Stimulus Argument and Multi-Competence. *Second Language Research* 7: 103–117.
26. Cook, V. (ed) (2003) *Effects of the Second Language on the First*. Buffalo, NY: Multilingual Matters, Ltd.
27. Costa, A. (2005) Lexical Access in Bilingual Production. In J. F. Kroll and A. M. B. de Groot (eds) *Handbook of Bilingualism: Psycholinguistic Approaches*, pp. 308–325. Oxford: Oxford University Press.
28. Coulson, S. (1996) Menendez Brothers Virus: Blended Spaces and Internet Humor. In A. E. Goldberg (ed) *Conceptual Structure, Discourse and Language*, pp. 67–81. Stanford: CSLI Publications.
29. Coulson, S. (1999) *Conceptual Blending and Discourse Irony*. Paper presented at the annual meeting of the Western Humanities Association, San Diego, CA.
30. Coulson, S. (2001) *Semantic Leaps: Frame-shifting and Conceptual Blending in Meaning Construction.* Cambridge: Cambridge University Press.

31. Coulson, S. (2006) Conceptual Blending in Thought, Rhetoric, and Ideology. In G. Kristiensen, M. Achard, R. Dirven, and F. J. R. de Mendoza Ibáñez (eds) *Cognitive Linguistics: Current Applications and Future Perspectives*, pp. 187–208. New York: Mouton de Gruyter.
32. Coulson, S. and Fauconnier, G. (1999) Fake Guns and Stone Lions: Conceptual Blending and Privative Adjectives. In B. A. Fox, D. Jurafsky, and L. A. Michaelis (eds) *Cognition and Function in Language*, pp. 143–158. Stanford: CSLI Publications.
33. Čubajevaitė, L. (2009) Daugiakalbystė Lietuvos miestuose: Kauno atvejis. Metodologiniai aspektai. *Žmogus kalbos erdvėje* 5: 53–61.
34. Dąbrowska, E. (2004) *Language, Mind and Brain: Some Psychological and Neurological Constraints on Theories of Grammar*. Washington, D.C.: Georgetown University Press.
35. De Angelis, G. and Selinker, L. (2001) Interlanguage Transfer and Competing Linguistic Systems in the Multilingual Mind. In J. Cenoz, B. Hufeisen and U. Jessner (eds) *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*, pp. 42–58. Clevedon: Multilingual Matters, Ltd.
36. de Bot, K. (1992) A Bilingual Production Model: Levelt's "Speaking" Model Adapted. In Li Wei (ed), *The Bilingualism Reader*, pp. 420–442. New York: Routledge.
37. de Bot, K., Lowie, W., and Verspoor, M. (2007) A Dynamic Systems Theory Approach to Second Language Acquisition. *Bilingualism: Language and Cognition* 10(1): 7–21.
38. de Groot, A. M. B. (1992) Bilingual Lexical Representation: A Closer Look at Conceptual Representations. In R. Frost and L Katz (eds) *Orthography, Phonology, Morphology, and Meaning*, pp. 389–412. Amsterdam: Elsevier Science.
39. Dechert, H. W. (1987) Analysing Language Processing through Verbal Protocols. In C. Faerch and G. Kasper (eds) *Introspection in Second Language Research*, pp. 96–112. Clevedon: Multilingual Matters, Ltd.
40. Dell, G. S. (1986) A Spreading Activation Theory of Retrieval in Sentence Production. *Psychological Review* 93: 283–321.
41. Dewaele, J-M. (1996) Les phénomènes d'hésitation dans l'interlangue française: analyse de la variation interstylistique et interindividuelle. *Rassegna Italiana da Linguistica Applicata* 28: 87–103.
42. Dewaele, J-M. (2001) Activation or Inhibition? The Interaction of L1, L2 and L3 on the Language Mode Continuum. In J. Cenoz, B. Hufeisen and U. Jessner (eds) *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*, pp. 69–89. Clevedon: Multilingual Matters, Ltd.
43. DuBois, J. W. (1974) Syntax in Mid-Sentence. *Berkeley Studies in Syntax and Semantics, Volume 1*. Berkeley, CA: University of California.

44. Ehrich, J. F. (2006) Vygotskian Inner Speech and the Reading Process. *Australian Journal of Educational & Developmental Psychology* 6: 12–25.
45. Ellis, N. C. (2002) Frequency Effects in Language Processing. *Studies in Second Language Acquisition* 24 (2): 143–188.
46. Ellis, N. C. (2003) Constructions, Chunking, and Connectionism: The Emergence of Second Language Structure. In C. J. Doughty and M. H. Long (eds) *The Handbook of Second Language Acquisition*, pp. 63–103. Oxford: Blackwell.
47. Ellis, N. C. (2007) Dynamic Systems and SLA: The Wood and the Trees. *Bilingualism: Language and Cognition* 10 (1): 23–25.
48. Erman, B. and Warren, B. (2000) The Idiom Principle and the Open Choice Principle. *Text* 20: 29–62.
49. Evans, V. (2006) Lexical Concepts, Cognitive Models and Meaning-Construction. *Cognitive Linguistics* 17 (4): 491–534.
50. Evans, V. and Green, M. (2006) *Cognitive Linguistics: An Introduction*. Mahwah, NJ: Lawrence Erlbaum Associates.
51. Fauconnier, G. (1994) Mental Spaces: Aspects of Meaning Construction in Natural Language. Cambridge: Cambridge University Press.
52. Fauconnier, G. (1997) *Mappings in Thought and Language*. Cambridge: Cambridge University Press.
53. Fauconnier, G., and Turner, M. (1994) Conceptual Projection and Middle Spaces. UCSD Cognitive Science Technical Report 9401.
54. Fauconnier, G., and Turner, M. (1996) Blending as a Central Process of Grammar. In A. E. Goldberg (ed) *Conceptual Structure, Discourse and Language*, pp. 113–130. Stanford: CSLI Publications.
55. Fauconnier, G., and Turner, M. (2002) The Way We Think: Conceptual Blending and the Mind's Hidden Complexities. New York: Basic Books.
56. Ferguson, C. (1959) Diglossia. In C. B. Paulston and G. R. Tucker (eds), *Sociolinguistics: The Essential Readings*, pp. 345–358. Oxford: Blackwell.
57. Fillmore, C. J., Kay, P. and O'Connor, M. C. (2003) Regularity and Idiomaticity in Grammatical Constructions: The Case of *Let Alone*. In M. Tomasello (ed) *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure, Volume 2*, pp. 243–270. Mahwah, NJ: Lawrence Erlbaum Associates.
58. Fishman, J. A. (1967) Bilingualism With and Without Diglossia; Diglossia With and Without Bilingualism. In C. B. Paulston and G. R. Tucker (eds), *Sociolinguistics: The Essential Readings*, pp. 359–366. Oxford: Blackwell.
59. Fodor, J. A. (1998) *Concepts: Where Cognitive Science Went Wrong*. Oxford: Clarendon Press.
60. Ford, C. E., Fox, B. A., and Thompson, S. A. (2003) Social Interaction and Grammar. In M. Tomasello (ed) *The New Psychology of Language, Volume 2*, pp. 119–143. Mahwah, NJ: Lawrence Erlbaum Associates.

61. Fouser, R. J. (2001) Too Close for Comfort? Sociolinguistic Transfer from Japanese into Korean as an L≥3. In J. Cenoz, B. Hufeisen and U Jessner (eds) *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*, pp. 149–169. Clevedon: Multilingual Matters, Ltd.
62. Fromkin, V. A. (ed) (1973) *Speech Errors as Linguistic Evidence*. The Hague: Mouton.
63. Gerken, L. (1994) Child Phonology: Past Research, Present Questions, Future Directions. In M. A. Gernsbacher (ed) *Handbook of Psycholinguistics*, pp. 781–820. San Diego, CA: Academic Press.
64. Gleick, J. (1987) *Chaos: Making a New Science*. New York: Viking.
65. Goddard, C. (2004) "Cultural Scripts": A New Medium for Ethnopragmatic Instruction. In M. Achard and S. Niemeier (eds) *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching*, pp. 143–163. Berlin: Mouton de Gruyter.
66. Goddard, C. (2006) Ethnopragmatics: A New Paradigm. In C. Goddard (ed) *Ethnopragmatics: Understanding Discourse in Cultural Context*, pp. 1–30. New York: Mouton de Gruyter.
67. Goddard, C. and Wierzbicka, A. (2004) Cultural Scripts: What Are They and What Are They for? *Intercultural Pragmatics* 1 (2): 153–166.
68. Goldberg, A. E. (1995) *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
69. Goldberg, A. E. (1998) Patterns of Experience in Patterns of Language. In M. Tomasello (ed) *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, pp. 203–219. Mahwah, NJ: Lawrence Erlbaum Associates.
70. Goldman-Eisler, F. (1961) A Comparative Study of Two Hesitation Phenomena. *Language and Speech* 4: 18–26.
71. Goldman-Eisler, F. (1964) Hesitation, Information, and Levels of Speech Production. In A. V. S. de Reuck and M. O'Connor (eds) *Ciba Foundation Symposium on Disorders of Language*, pp. 96–111. Boston: Little, Brown & Co.
72. Green, D. W. ([1986] 2000) Control, Activation, and Resource: A Framework and a Model for the Control of Speech in Bilinguals. In Li Wei (ed) *The Bilingualism Reader*, pp. 407–419.
73. Green, D. W. (1993) Towards a Model of L2 Comprehension and Production. In R. Schreuder and B. Weltens (eds) *The Bilingual Lexicon*, pp. 243–279. Philadelphia: John Benjamins.
74. Green, D. W. (1998) Mental Control of the Bilingual Lexico-Semantic System. *Bilingualism: Language and Cognition* 1: 67–81.
75. Grosjean, F. (1988) Exploring the Recognition of Guest Words in Bilingual Speech. *Language and Cognitive Processes* 3: 233–274.
76. Grosjean, F. (1997) Processing Mixed Language: Issues, Findings, and Models. In Li Wei, (ed), *The Bilingualism Reader*, pp. 443–469. New York: Routledge.

77. Grosjean, F. (2001) The Bilingual's Language Modes. In J. L. Nicol (ed) *One Mind, Two Languages: Bilingual Language Processing*, pp. 1–22. Oxford: Blackwell.
78. Grumadienė, L. (2005) Language Policy and the Sociolinguistic Situation in Lithuania. *Mercator – Working Papers* 19. Barcelona: CIEMEN.
79. Gumperz, J. J. (1982) Contextualization Conventions. In C. B. Paulston and G. R. Tucker (eds), *Sociolinguistics: The Essential Readings*, pp. 139–155. Oxford: Blackwell.
80. Gumperz, J. J. and Levinson, S. C. (eds) (1996) *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press.
81. Hammarberg, B. (2001) Roles of L1 and L2 in L3 Production and Acquisition. In J. Cenoz, B. Hufeisen and U. Jessner (eds) *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*, pp. 21–41. Clevedon: Multilingual Matters, Ltd.
82. Hatch, E. and Lazaraton, A. (1991) *The Research Manual: Design and Statistics for Applied Linguistics*. Boston, MA: Heinle & Heinle Publishers.
83. Hatzivassiloglou, V. and McKeown, K. (1997) Predicting the Semantic Orientation of Adjectives. In *ACL-1997*: 174–181.
84. Herderson, A., Goldman-Eisler, F. and Skarbek, A. (1966) Sequential Temporal Patterns in Spontaneous Speech. *Language and Speech* 9: 207–216.
85. Herdina, P. and Jessner, U. (2002) A Dynamic Model of Multilingualism: Perspectives of Change in Psycholinguistics. Clevedon: Multilingual Matters, Ltd.
86. Heredia, R. R. (1996) Bilingual Memory: a Re-revised Version of the Hierarchical Model of Bilingual Memory. *The Newsletter of the Center for Research in Language* 10: 3–6. University of California, San Diego.
87. Heredia, R. R. and Brown, J. M. (2004) Bilingual Memory. In T. K. Bhatia and W. C. Ritchie (eds) *The Handbook of Bilingualism*, pp. 225–249. Oxford: Blackwell.
88. Herwig, A. (2001) Plurilingual Lexical Organization: Evidence from Lexical Processing in L1-L2-L3-L4 Translation. In J. Cenoz, B. Hufeisen, and U. Jessner (eds) *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*, pp. 115–137. Clevedon: Multilingual Matters, Ltd.
89. Hockett, C. F. (1967) Where the Tongue Slips There Slip I. *To honor Roman Jakobson, Volume 2*. The Hague: Mouton.
90. Hogan-Brun, G. and Ramonienė, M. (2004) Changing Levels of Bilingualism across the Baltic. *Bilingual Education and Bilingualism* 7(1): 62–77.
91. Hogan-Brun, G. and Ramonienė, M. (2005) Perspectives on Language Attitudes and Use in Lithuania's Multilingual Setting. *Journal of Multilingual and Multicultural Development* 26(5): 425–441.

92. Hudson, A. (1991) Toward the Systematic Study of Diglossia. In C. B. Paulston and G. R. Tucker (eds), *Sociolinguistics: The Essential Readings*, pp. 367–376. Oxford: Blackwell.
93. Hymes, D. ([1986] 2003) Models of the Interaction of Language and Social Life. In C. B. Paulston and G. R. Tucker (eds) *Sociolinguistics: The Essential Readings*, pp. 30–47. Oxford: Blackwell.
94. James, D. (1972) Some Aspects of the Syntax and Semantics of Interjections. *Papers from the Ninth Regional Meeting*. Chicago Linguistic Society.
95. Jessner, U. (2003) A Dynamic Approach to Language Attrition in Multilingual Systems. In V. Cook (ed) *Effects of the Second Language on the First*, pp. 234–246.
96. Jones, H. G. V. and Langford, S. (1987) Phonological Blocking in the Tip of the Tongue State. *Cognition* 26: 115–122.
97. Jurafsky, D. (2003) Probabilistic Modeling in Psycholinguistics: Linguistic Comprehension and Production. In R. Bod, J. Hay and S. Jannedy (eds) *Probabilistic Linguistics*. Cambridge, MA: MIT Press.
98. Kachru, B. (1985) Standards, Codification, and Sociolinguistic Realism: The English Language in the Outer Circle. In R. Quirk and H. Widdowson (eds) *English in the World: Teaching and Learning the Language and Literatures*, pp. 11–30. Cambridge: Cambridge University Press.
99. Kachru, B. (1991) Liberation Linguistics and the Quirk Concern. *English Today* 25: 3–13.
100. Karaliūnas, S. (1997) *Kalba ir visuomenė: Psichosociologiniai ir komunikaciniai kalbos vartojimo bruožai*. Vilnius: Lietuvių kalbos institutas.
101. Kasper, G. and Blum-Kulka, S. (1993) Interlanguage Pragmatics: An Introduction. In G. Kasper and S. Blum-Kulka (eds) *Interlanguage Pragmatics*. Oxford: Oxford University Press.
102. Kay, P. and Fillmore, C. J. (1999) Grammatical Constructions and Linguistic Generalizations: The *What's X doing Y* Construction. *Language* 75: 1–34.
103. Kecskes, I. and Papp, T. (2000) *Foreign Language and Mother Tongue*. Mahwah, NJ: Lawrence Erlbaum Associates.
104. Kellerman, E. (1983) Now you see it, now you don't. In S. Gass and L. Selinker (eds) *Language Transfer in Language Learning*, pp. 112–133. Rowley, MA: Newbury House.
105. Kellerman, E. and Sharwood Smith, M. (eds) (1986) *Crosslinguistic Influence in Second Language Acquisition*. New York: Pergamon Press.
106. Kempen, G. and Hoenkamp, E. (1982) Incremental Sentence Generation: Implications for the Structure of a Syntactic Processor. In J. Horecky (ed) *Proceedings of the Ninth International Conference on Computational Linguistics*. Amsterdam: North-Holland.

107. Kempen, G. and Hoenkamp, E. (1987) An Incremental Procedural Grammar for Sentence Formulation. *Cognitive Science* 11: 201–258.
108. Kilborn, K. (1994) Learning a Language Late: Second Language Acquisition in Adults. In M. A. Gernsbacher (ed) *Handbook of Psycholinguistics*, pp. 917–944. San Diego, CA: Academic Press.
109. Kroll, J. F. and Dussias, P. E. (2004) The Comprehension of Words and Sentences in Two Languages. In T. K. Bhatia and W. C. Ritchie (eds) *The Handbook of Bilingualism*, pp. 169–200. Oxford: Blackwell.
110. Kroll, J. F. and Stewart, E. (1994) Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections between Bilingual Memory Representations. *Journal of Memory and Language* 33: 149–174.
111. Kroll, J. F. and Tokowicz, N. (2001) The Development of Conceptual Representation for Words in a Second Language. In J. L. Nicol (ed) *One Mind, Two Languages: Bilingual Language Processing*, pp. 49–71. Oxford: Blackwell.
112. Kroll, J. F. and Tokowicz, N. (2005) Models of Bilingual Representation and Processing: Looking Back and to the Future. In J. F. Kroll and A. M. B. de Groot (eds) *Handbook of Bilingualism: Psycholinguistic Approaches*, pp. 531–553. Oxford: Oxford University Press.
113. Lakoff, G. (1987) *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
114. Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By*. Chicago: University of Chicago Press.
115. Langacker, R. (1987) *Foundations of Cognitive Grammar, Volume I*. Stanford, CA: Stanford University Press.
116. Larsen-Freeman, D. (1997) Chaos/Complexity Science and Second Language Acquisition. *Applied Linguistics* 18 (2): 141–165.
117. Larsen-Freeman, D. (2007) On the Complementarity of Chaos/Complexity Theory and Dynamic Systems Theory in Understanding the Second Language Acquisition Process. *Bilingualism: Language and Cognition* 10 (1): 35–37.
118. Lee, M. and Williams, J. M. (2001) Lexical Access in Spoken Word Production by Bilinguals: Evidence from the Semantic Competitor Priming Paradigm. *Bilingualism: Language and Cognition* 4: 233–248.
119. Leech, G. (1983) *Principles of Pragmatics*. London: Longman.
120. Levelt, W. J. M. (1989) *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
121. Levelt, W. J. M. (1983) Monitoring and Self-Repair in Speech. *Cognition* 14: 41–104.
122. Li Wei (ed) (2000) *The Bilingualism Reader*. New York: Routledge.
123. Lindblom, B. (1983) Economy of Speech Gestures. In P. F. MacNeilage (ed) *The Production of Speech*. New York: Springer.
124. Lithuanian Constitution (1992) <http:// www.litlex.lt/Litlex/Eng/Frames/ Laws/Documents/CONSTITU.HTM> Access date: 19 Mar. 2008.

125. Lithuanian Radio (2008) Lietuvos nacionalinis radijas ir televizija, <http://www.lrt.lt/ schedule.php?c=1353> Access date: 7 Aug. 2008.
126. Maclay, H. and Osgood, C. E. (1959) Hesitation Phenomena in Spontaneous English Speech. *Word* 15: 19–44.
127. MacWhinney, B. (2004) A Multiple Process Solution to the Logical Problem of Language Acquisition. *Journal of Child Language*, *31*: 883–914.
128. MacWhinney, B. (2005) A Unified Model of Language Acquisition. In J. F. Kroll and A. M. B. de Groot (eds) *Handbook of Bilingualism: Psycholinguistic Approaches*, pp. 49–67. Oxford: Oxford University Press.
129. Mahl, G. F. (1956) Disturbances and Silences in the Patient's Speech in Psychotherapy. *Journal of Abnormal and Social Psychology* 53: 1–15.
130. Mandelblit, N. (1997) Grammatical Blending: Creative and Schematic Aspects in Sentence Processing and Translation. Ph.D. dissertation. University of California, San Diego.
131. Meringer, R. (1908) *Aus dem Leben der Sprache.* Berlin: Behr.
132. Michael, E. B. and Gollan, T. H. (2005) Being and Becoming Bilingual: Individual Differences and Consequences for Language Production. In J. F. Kroll and A. M. B. de Groot (eds) *Handbook of Bilingualism: Psycholinguistic Approaches*, pp. 389–407. Oxford: Oxford University Press.
133. Miller, G. A. (1956) The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review* 63: 81–97.
134. Miller, G. A. (1978) Semantic Relations among Words. In M. Halle, J. Bresnan and G. A. Miller (eds) *Linguistic Theory and Psychological Reality*. Cambridge, MA: MIT Press.
135. Milroy, L. and Gordon, M. (2003) *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell.
136. Mittelberg, I., Farmer, T. A. and Waugh, L. R. (2007) They Actually Said That? An Introduction to Working with Usage Data through Discourse and Corpus Analysis. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, and M. J. Spivey (eds) *Methods in Cognitive Linguistics*, pp. 19–52. Philadelphia: John Benjamins.
137. Morton, J. (1979) Word Recognition. In J. Morton and J. Marshall (eds) *Psycholinguistics: Series 2. Structures and Processes*. London: Elek.
138. Myers-Scotton, C. (1993) *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Oxford University Press.
139. Myers-Scotton, C. (2005) Supporting a Differential Access Hypothesis: Code Switching and Other Contact Data. In J. F. Kroll and A. M. B. de Groot (eds) *Handbook of Bilingualism: Psycholinguistic Approaches*, pp. 326–348. Oxford: Oxford University Press.
140. Myers-Scotton, C. and Jake, J. L. (1995) Matching Lemmas in a Bilingual Competence and Production Model: Evidence from

Intrasentential Code Switching. In Li Wei (ed) *The Bilingualism Reader*, pp. 281–320. New York: Routledge.
141. Myers-Scotton, C. and Jake, J. L. (2001) Explaining Aspects of Code-Switching and Their Implications. In J. L. Nicol (ed) *One Mind, Two Languages: Bilingual Language Processing*, pp. 84–116. Oxford: Blackwell.
142. Nooteboom, S. (1980) Speaking and Unspeaking: Detection and Correction of Phonological and Lexical Errors in Spontaneous Speech. In V. A. Fromkin (ed) *Errors in Linguistic Performance*. New York: Academic Press.
143. Oller, J. W., Jr. (1997) Monoglottosis: What's Wrong with the Idea of the IQ Meritocracy and Its Racy Cousins? *Applied Linguistics* 18: 467–507.
144. Paradis, M. (ed) (1987) *The Assessment of Bilingual Aphasia*. Hillsdale, NJ: Lawrence Erlbaum.
145. Paulauskienė, A. (1994) *Lietuvių kalbos morfologija: Paskaitos lituanistams.* Vilnius: Mokslo ir enciklopedijų leidykla.
146. Pavlenko, A. (1996) Bilingualism and Cognition: Concepts in the Mental Lexicon. In A. Pavlenko and R. Salaberry (eds) *Papers in Second Language Acquisition and Bilingualism*, pp. 49–85. Ithaca, NY: Cornell University Press.
147. Perret-Clermont, A.-N., Carugati, F. and Oates, J. (2004) A Socio-Cognitive Perspective on Learning and Cognitive Development. In J. Oates and A. Grayson (eds) *Cognitive and Language Development in Children*, pp. 305–332. Oxford: Blackwell.
148. Pinker, S. (1994) The Language Instinct: How the Mind Creates Language. New York: HarperCollins.
149. Poole, G. (2002) *Syntactic Theory*. New York: Palgrave.
150. Poplack, S. (2000) Sometimes I'll Start a Sentence in Spanish *y Termino en Español*: Toward a Typology of Code-Switching. In Li Wei (ed) *The Bilingualism Reader*, pp. 221–256. New York: Routledge.
151. Poulisse, N. (1990) The Use of Compensatory Strategies by Dutch Learners of English. Dordrecht: Foris.
152. Prodromou, L. (2007) A Sort of Puzzle for English as a Lingua Franca. In B. Tomlinson (ed) *Language Acquisition and Development: Studies of Learners of First and Other Languages*, pp. 225–246. London: Continuum.
153. Quinting, G. (1971) *Hesitation Phenomena in Adult Aphasic and Normal Speech*. The Hague: Mouton.
154. Radford, A. (1997) *Syntax: A Minimalist Introduction*. Cambridge: Cambridge University Press.
155. Radford, A. (2004) *English Syntax: An Introduction*. Cambridge: Cambridge University Press.
156. Roelofs, A. (1998) Lemma Selection without Inhibition of Language in Bilingual Speakers. *Bilingualism: Language and Cognition* 1: 94–95.

157. Rosch, E. ([1978] 1999) Principles of Categorisation. In E. Margolis and S. Laurence (eds) *Concepts: Core Readings*, pp. 189–206. Cambridge, MA: MIT Press.
158. Rosch, E., Mervis, C., Gray, W., Johnson, D. and Boyes-Braem, P. (1976) Basic Objects in Natural Categories. *Cognitive Psychology* 8: 382–439.
159. Sapir, E. (1921) *Language*. New York: Harcourt, Brace and World.
160. Scovel, T. (1998) *Psycholinguistics*. Oxford: Oxford University Press.
161. Scovel, T. (2001) Learning New Languages: A Guide to Second Language Acquisition. Boston: Heinle & Heinle.
162. Searle, J. (1975) Indirect Speech Acts. In P. Cole and J. L. Morgan (eds) *Syntax and Semantics, Volume 3: Speech Acts*, pp. 59–82. New York: Academic Press.
163. Sebastián-Gallés, N. and Bosch, L. (2005) Phonology and Bilingualism. In J. F. Kroll and A. M. B. de Groot (eds) *Handbook of Bilingualism: Psycholinguistic Approaches*, pp. 68–87. Oxford: Oxford University Press.
164. Selinker, L. (1972) Interlanguage. *International Review of Applied Linguistics* 10(3): 209–231.
165. Sharifi, M. and Cohen, W. (2008) Finding Domain Specific Polar Words for Sentiment Classification. In *Language Technologies Institute Student Research Symposium (SRS 2008)*, Pittsburgh, PA, USA.
166. Skehan, P. (1998) *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
167. Slobin, D. I. (1996) From "Thought and Language" to "Thinking for Speaking." In J. J. Gumperz and S. C. Levinson (eds) *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press.
168. Statistics Lithuania (2008) Official Website of the Department of Statistics to the Government of the Republic of Lithuania, <http://www.stat.gov.lt/en/> Access dates: 25 Feb. 2008, 19 Mar. 2008.
169. Swan, M. & Smith, B. (eds) (2nd edn 2001) *Learner English: A Teacher's Guide to Interference and Other Problems*. Cambridge: Cambridge University Press.
170. Taylor, J. (1998) Syntactic Constructions and Prototype Categories. In M. Tomasello (ed) *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, pp. 177–202. Mahwah, NJ: Lawrence Erlbaum Associates.
171. Taylor, J. R. (2006) Polysemy and the Lexicon. In G. Kristiansen, M. Achard, R. Dirven, and F. J. R. de Mendoza Ibáñez (eds) *Cognitive Linguistics: Current Applications and Future Perspectives*, pp. 51–80. New York: Mouton de Gruyter.
172. Thornburg, L. and Panther, K.-U. (1997) Speech Act Metonymies. In W.-A. Liebert, G. Redeker, and L. Waugh (eds) *Discourse and Perspectives in Cognitive Linguistics*, pp. 205–219. Amsterdam: John Benjamins.

173. Thorne, G. (2009) Graphomotor Skills: Why Some Kids Hate to Write. Center for Development and Learning website: <http://www.cdl.org/resource-library/articles/graphomotor.php> Access date: Aug. 2009.
174. Tomasello, M. (2000) The Item-Based Nature of Children's Early Syntactic Development. In M. Tomasello and E. Bates (eds) *Language Development: The Essential Readings*, pp. 169–186. Oxford: Blackwell.
175. Tomasello, M. (2006) Acquiring Linguistic Constructions. In W. Damon, R. M. Lerner, D. Kuhn, and R. S. Siegler (eds) *Handbook of Child Psychology, Vol. 2: Cognition, Perception, and Language*, pp. 255–298. New York: Wiley.
176. Tomlinson, B. (2007) Introduction: Some Similarities and Differences Between L1 and L2 Acquisition and Development. In B. Tomlinson (ed) *Language Acquisition and Development: Studies of Learners of First and Other Languages*, pp. 1–12. London: Continuum.
177. Tomlinson, B. and Avila, J. (2007) Seeing and Saying for Yourself: the Roles of Audio-Visual Mental Aids in Language Learning and Use. In B. Tomlinson (ed) *Language Acquisition and Development: Studies of Learners of First and Other Languages*, pp. 61–81. London: Continuum.
178. Truscott, J. and Sharwood Smith, M. (2004) Acquisition by Processing: A Modular Perspective on Language Development. *Bilingualism: Language and Cognition* 7 (1): 1–20.
179. Turner, M. (2006) Compression and Representation. *Language and Literature* 15(1): 17–27.
180. Välimaa-Blum, R. (2005) *Cognitive Phonology in Construction Grammar: Analytic Tools for Students of English*. Berlin: Mouton de Gruyter.
181. van Geert, P. (1994) Dynamic Systems of Development: Change between Complexity and Chaos. New York: Harvester Wheatsheaf.
182. van Geert, P. (2007) Dynamic Systems in Second Language Learning: Some General Methodological Reflections. *Bilingualism: Language and Cognition* 10 (1): 47–49.
183. von Knorring, C. (2007) Perceptions of Culture by British Students Learning French. In B. Tomlinson (ed) *Language Acquisition and Development: Studies of Learners of First and Other Languages*, pp. 247–266. London: Continuum.
184. Vygotsky, L. S. (1978) *Mind in Society: The Development of Higher Psychological Processes*, ed. M. Cole, V. John-Steiner, S. Scribner and E. Souberman. Cambridge, MA: Harvard University Press.
185. Vygotsky, L. S. (1986) *Thought and Language*, ed. A. Kozulin. Cambridge, MA: MIT Press.
186. Waara, R. (2004) Construal, Convention, and Constructions in L2 Speech. In M. Achard and S. Niemeier (eds) *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching*, pp. 51–76.

187. Waugh, L. R., Fonseca-Greber, B., Vickers, C., and Eröz, B. (2007) Multiple Empirical Approaches to a Complex Analysis of Discourse. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, and M. J. Spivey (eds) *Methods in Cognitive Linguistics*, pp. 120–148. Philadelphia: John Benjamins.
188. Whorf, B. L. (1956) *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*, ed. J. B. Carroll. Cambridge, MA: MIT Press.
189. Wiebe, J. (2007) Subjectivity and Sentiment Analysis. In Proceedings of the 19th Conference on Computational Linguistics and Speech Processing (ROCLING '07), Taipei, Taiwan.
190. Wierzbicka, A. (1985) Different Cultures, Different Languages, Different Speech Acts: Polish vs. English. *Journal of Pragmatics* 9 (2/3): 145–178.
191. Wierzbicka, A. (1998) The Semantics of English Causative Constructions in a Universal-Typological Perspective. In M. Tomasello (ed) *The New Psychology of Language*, pp. 113–153. Mahwah, NJ: Lawrence Erlbaum Associates.
192. Wierzbicka, A. (2006) Anglo Scripts against "Putting Pressure" on Other People and Their Linguistic Manifestations. In C. Goddard (ed) *Ethnopragmatics: Understanding Discourse in Cultural Context*, pp. 31–63. New York: Mouton de Gruyter.
193. Wilson, T., Wiebe, J. and Hoffmann, P. (2005) Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. In Proceedings of Human Language Technologies Conference/Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005), Vancouver, Canada.
194. Wong Fillmore, L. (1991) Second-language Learning in Children: a Model of Language Learning in Social Context. In E. Bialystok (ed) *Language Processing in Bilingual Children*, pp. 49–69. Cambridge: Cambridge University Press.

Dictionaries:

1. General Inquirer (2000) <http://www.wjh.harvard.edu/~inquirer/spreadsheet_guide.htm>
2. Piesarskas, B. (1998) Alkonas: Anglų–lietuvių kalbų kompiuterinis žodynas. Version 1.0.2.190.
3. Tildė (2006) Tildės kompiuterinis žodynas.

Appendix 1

Sociocultural Language Use Survey

1. What is your native language?
1.1 In your opinion, can people have two (or more) native languages?
1.2 What is your mother's native language?
1.3 What is your father's native language?

2. List all the languages you can speak or write in:

3. Perhaps you know some words or phrases of other languages, but don't really speak them. List those languages, too:

4. Which of the languages you know is your favorite language, and why?

5. In your opinion, which two languages are the most similar:
Lithuanian & Russian Russian & English Lithuanian & English

6. Which language do you use most often at home (where you live now)?
6.1 Which language do you use most often with your relatives?
6.2 Which language do you use most often with your friends?
6.3 If you have a job, which language do you use most often at work?
6.4 Which language do you use most often on the Internet?
6.5 Which language did you use most often during the summer holidays?
6.5.1 Did you use any other languages during that time? Which?
6.6 Which language do you use most often when thinking?
6.7 Which language do you use when you dream?
6.7.1 Have you ever dreamed in any other languages?
6.7.2 If so, were they good or bad dreams?

7. Which language do you prefer to read in?
7.1 Which language do you prefer to listen to music in?
7.2 Which language do you prefer to watch films or TV in?

8. The funniest jokes are (choose one): Lithuanian Russian English

9. Do you have an accent when you speak…
Lithuanian? Russian? English?

10. What is your nationality?
11. How old are you?
12. Which country were you born in?
13. How long have you lived in Lithuania (in years)?

14. In which school year did you begin to study Lithuanian?
14.1 In which school year did you begin to study Russian?
14.2 In which school year did you begin to study English?

15. Which was your "best" language in school?
16. Which is your "best" language now?

17. Which translation direction(s) is/are easiest for you (choose up to three):
LT→RU LT→EN RU→LT RU→EN EN→LT EN→RU

18. Choose the adjective that best describes you:
monolingual bilingual trilingual multilingual

19. Write 2 or 3 adjectives that describe the stereotypical Lithuanian style of interaction:

19.1 Write 2 or 3 adjectives that describe the stereotypical Russian style of interaction:

19.2 Write 2 or 3 adjectives that describe the stereotypical British/American style of interaction:

19.3 Bearing this in mind or in spite of this, which style feels closest to you?

Cultural scales: 1 = strongly disagree, 2 = partially disagree, 3 = maybe, 4 = partially agree, 5 = strongly agree.

20. Lithuanian culture is a part of my life.	1……2……3……4……5
Russian culture is a part of my life.	1……2……3……4……5
British/American culture is a part of my life.	1……2……3……4……5
21. I feel like a Lithuanian.	1……2……3……4……5
I feel like a Russian.	1……2……3……4……5
I feel British/American.	1……2……3……4……5
22. Lithuanians would think I'm Lithuanian.	1……2……3……4……5
Russians would think I'm Russian.	1……2……3……4……5
Brits/Americans would think I'm British/American.	1……2……3……4……5
23. Most of my friends are Lithuanians.	1……2……3……4……5
Most of my friends are Russians.	1……2……3……4……5
Most of my friends are Brits/Americans.	1……2……3……4……5

24. I have trouble expressing myself in Lithuanian. 1……2……3……4……5
I have trouble expressing myself in Russian. 1……2……3……4……5
I have trouble expressing myself in English. 1……2……3……4……5

25. I want to speak Lithuanian fluently. 1……2……3……4……5
I want to speak Russian fluently. 1……2……3……4……5
I want to speak English fluently. 1……2……3……4……5

26. I'm comfortable speaking to native Lithuanians. 1……2……3……4……5
I'm comfortable speaking to native Russians. 1……2……3……4……5
I'm comfortable speaking to native English speakers. 1……2……3……4……5

27. I am good at learning languages. 1……2……3……4……5

Appendix 2

Adjective polarity

In analyzing the results of the sociolinguistic survey described in Part 2, the question arose of how to organize the answers to the questions (19–19.2) about cultural stereotypes. A total of 291 answers yielded 115 different adjectives or adjective-like phrases (e.g., *I like them*, a phrase comparable to the adjective *likeable*). It was decided to arrange the adjectives according to their **polarity**, a technique common in **subjectivity and sentiment analysis** (SSA; see, for example, Wiebe, 2007; Agarwal & Bhattacharyya, 2006; Wilson, Wiebe, & Hoffmann, 2005).

There are a large number of words ascribed with **prior polarity**, meaning that they carry positive or negative connotations without any reference to their use in discourse (Wilson et al., 2005; Sharifi & Cohen, 2008). Many of these words have been collected in electronic databases such as General Inquirer (2000), compiled at Harvard University and containing nearly 12,000 words tagged for membership in 182 categories, including the polarity categories "Positive" and "Negative." However, as discussed in Sharifi & Cohen (2008), Wilson et al. (2005), and Hatzivassiloglou & McKeown (1997), polarity can change depending on the context of usage. Thus, "the *contextual polarity* of the phrase in which a word appears may be different from the word's prior polarity" (Wilson et al., 2005: 1).

As General Inquirer (2000) was used as a prior polarity lexicon in both Wilson et al. (2005) and Sharifi and Cohen (2008), the first step was to compare the 115 terms provided by students with the Positive and Negative category lists contained therein. This resulted in 53 positive, 29 negative, and 33 unmarked terms, of which four (*bilingual*, *broad-minded*, *"chameleonic"*, and *snobbish*) were not included in the Harvard lexicon at all. The unmarked terms were not all considered to be neutral descriptors, however. Although the form of the adjectives as provided was apparently context-free (most answers consisting of a list of one, two, or three adjectives separated only by commas),

the context was in fact that of the question itself: students had been asked to evaluate the stereotypical behavior of Lithuanian, Russian, and English-speaking people. As discussed in Hatzivassiloglou & McKeown (1997: 175), adjectives can be assigned polarity depending on "whether the use of this adjective ascribes in general a positive or negative quality to the modified item, making it better or worse than a similar unmodified item." In other words, the contextual polarity of these adjectives could change because of their specific use as modifiers of the concepts *Lithuanian people*, *Russian people*, and *English-speaking people*. Thus, unmarked adjectives such as *active*, *hardworking*, and *strong* were labeled as positive, while *jealous*, *silent*, and *slow* were labeled as negative. The remaining terms were labeled as neutral.

The following charts contain all of the adjectives given by students in answer to Question 19–19.2 of the sociolinguistic survey described in Part 2. The adjectives are divided into positive, neutral, and negative for each culture described. Furthermore, the charts are subdivided by year of study: AnRK7 were first-year students, while AnRK4 were fourth-year students.

Chart 1: Adjectives describing Lithuanians

	AnRK7	AnRK6	AnRK5	AnRK4		
Positive			aware		1	**Total positive adjectives: 32**
	beautiful girls				1	
	broad-minded				1	
	calm (2)				2	
	cheerful				1	
		communicative			1	
		easy			1	
	friendly (2)	friendly (2)	friendly (3)		7	
		funny (2)	funny		3	
	hardworking				1	
	humorous				1	
			kind (2)		2	
			open-minded		1	
	polite				1	
		relaxed			1	
	self-confident				1	
		sociable			1	
	speak when necessary				1	
	straightforward (2)				2	
		strong			1	
		understandable			1	

Neutral			bilingual		1	**Total neutral adjectives: 10**
	emotional				1	
		normal			1	
	quiet				1	
	reserved (2)			reserved	3	
	serious	serious			2	
		simple			1	

Negative	aggressive (2)				2	**Total negative adjectives: 63**
		angry	angry		2	
	boring (3)	boring	boring		5	
	closed	closed	closed (2)		4	
	cold (4)		cold (3)		7	
			conservative re: language		1	
	difficult words				1	
	egotistical				1	
	envious				1	
	greedy (2)				2	
	impolite				1	
	intolerant				1	
		jealous	jealous (2)		3	
	passive				1	
	pessimistic	pessimistic	pessimistic		3	
	poor				1	
	rough				1	
	rude (5)				5	
			selfish		1	
	shy (3)	shy (4)			7	
	silent				1	
	silly				1	
	snobbish				1	
	stubborn	stubborn			2	
	stupid				1	
			uncommunicative		1	
	unfriendly (2)	unfriendly (2)	unfriendly		5	
	unhappy				1	

Chart 2: Adjectives describing Russians

	AnRK7	AnRK6	AnRK5	AnRK4		
Positive	active				1	**Total positive adjectives: 66**
		charming			1	
	communicative (2)	communicative (3)	communicative		6	
		curious			1	
	friendly (9)	friendly (5)	friendly (5)		19	
	funny (2)		funny (3)		5	
	generous (2)	generous			3	
	good				1	
	happy	happy			2	
	helpful				1	
	honest	honest			2	
	intelligent				1	
	interesting				1	
	kind		kind (2)		3	
	modern				1	
	nice (2)				2	
			open (2)		2	
	optimistic				1	
		outgoing			1	
	passionate				1	
		relaxed			1	
	reliable				1	
	sincere (2)				2	
	smart (2)				2	
	smiling				1	
		sociable			1	
		speak when necessary			1	
			warm		1	
			welcoming		1	
Neutral	emotional (2)				2	**Total neutral adjectives: 8**
		quiet			1	
	simple				1	
	speak quickly				1	
	talkative	talkative			2	
				vivid	1	
Negative	alcoholic (3)				3	**Total negative adjectives: 18**
	clingy				1	
		difficult			1	
	hot-blooded				1	
		rude	rude		2	
	speak loudly				1	
	stereotypical				1	
			strange (2)		2	
		strict			1	
			stubborn		1	
		uncomfortable			1	
	unfriendly		unfriendly		2	
	unintelligent				1	

Chart 3: Adjectives describing English-speakers

	AnRK7	AnRK6	AnRK5	AnRK4		
Positive			active		1	**Total positive adjectives: 76**
			attentive		1	
	attractive				1	
	calm				1	
		careful listeners			1	
	communicative	communicative (2)	communicative		4	
	cultured				1	
		excited			1	
		flexible			1	
		focused			1	
	friendly (5)	friendly (4)	friendly (4)		13	
	funny (4)		funny (2)		6	
	happy (3)	happy			4	
	helpful		helpful (2)		3	
	honest	honest			2	
	I like them				1	
				informative	1	
	intelligent	intelligent			2	
	kind				1	
	modest				1	
		not shy			1	
			open (2)		2	
			open-minded		1	
		optimistic	optimistic		2	
		outgoing (2)			2	
	polite (4)				4	
	proud				1	
	quick				1	
	refined				1	
	relaxed (2)				2	
	satisfied				1	
	self-confident	self-confident			2	
	smiling	smiling	smiling		3	
		sociable			1	
			warm		1	
	welcoming				1	
	well-mannered	well-mannered (2)			3	
Neutral	black humor				1	**Total neutral adjectives: 7**
	emotional (2)		emotional		3	
	formal				1	
	neutral				1	
	speak quickly				1	
Negative		"chameleonic"			1	**Total negative adjectives: 11**
	cold				1	
	difficult pronunciation				1	
	hard to understand				1	
	"never mind"				1	
	rushed				1	
	slow				1	
	speak a lot				1	
	speak loudly				1	
			strange		1	
			unfeeling		1	

Appendix 3

Translation task for control group

Translate these sentences into English. Use any materials you have – dictionaries, grammar books, the Internet, etc.

1. Taip smarkiai suspaudžiau jos ranką, kad ji suklykė nesavu balsu.
2. Mes su draugais nusprendėme pabandyti sušokti hiphopą.
3. Tarptautinės santuokos turi ir blogąją pusę – vienam iš poros teks gyventi toli nuo gimtosios šalies.
4. Ji nugriuvo ant laiptų ir vos nenusilaužė kojos.
5. Praėjusią vasarą aš su mama ir sese važiavome į Vokietiją, į pusseserės vestuves.
6. Paskutinis mūsų projektas tęsėsi per ilgai – pritrūko ir pinigų, ir gerų idėjų.
7. Aišku, apie dvasinius ar moralinius dalykus net nerašysiu, nes nemanau, kad žmogus gali visą tai išsiaiškinti pats.
8. Noriu įrodyti savo, kaip šokėjos, galimybes ir pagilinti turimas žinias apie šokių techniką.

Appendix 4

Published Scientific Articles

Braun, A. (2005) The Effects of Bilingualism on Child Language Acquisition. *Žmogus kalbos erdvėje* 4: 286–294. ISBN 978-9986-19-798-8.

Braun, A. (2008) It Sounds Clumsy in English: Conceptual Blending and Lithuanian-English Translation. *Respectus Philologicus* 13 (18) A: 186–200. ISSN 1392-8295. (MLA, EBSCO Humanities International, CEEOL databases)

Braun, A. (2009) Sociolinguistic Survey of Lithuanian-Russian-English Trilinguals. *Žmogus kalbos erdvėje* 5 (1): 30–38. ISBN 978-9955-33-373-9.

Braun, A. (2009) Crosslinguistic Influence as a Problem for the Language Learning Classroom. *Kalbų studijos* 14: 44–48. ISSN 1648-2824. (MLA, CEEOL databases)

Braun, A. (2009) Lithuanian Grammar, English Words: Cross-Linguistic Influence and Students' Written Errors. *Respectus Philologicus* 16 (21): 183–192. ISSN 1392-8295. (MLA, EBSCO Humanities International, CEEOL databases)

Algis Braun (b. 1972) completed a Bachelor of Arts degree in 1994 at the University of Puget Sound, Tacoma, WA, with a major in English and a minor in Psychology. In 1998 he entered the Master's degree program in English Philology at Vytautas Magnus University, graduating in 2000 with an emphasis in English literature. In 2004 he became a lecturer at Vilnius University, Kaunas Faculty of Humanities, Department of Foreign Languages. Since 2005 he has been a doctoral student in Philology (Humanities) with an emphasis in English linguistics.