Russian Language Journal

Volume 60
2010

Special Issue:

Divergent Thinking: Prospectives on the Language Enterprise in the 21st Century
Presented to Richard D. Brecht by his Students and Colleagues

Special Issue Editor: William P. Rivers

A Journal of the American Council of Teachers of Russian
Russian Language Journal

Editorial Group

Editor:
Maria D. Lekic, University of Maryland and ACTR

Assistant Editor:
Sofia Kasmeridi, ACTR

Associate Editors:
Michael Gorham, University of Florida
William P. Rivers, Integrated Training Solutions/National Language Service Corps

Assistant Editor (English):
Andrea Grenadier
American Councils for International Education

Editorial Consultant:
Dan E. Davidson, ACTR and Bryn Mawr

Editorial Board

Elizabeth Allen, Bryn Mawr, PA
Richard Brecht, CASL/University of Maryland
Robert Channon, Purdue University
Zita Dabars, Friends School (emerita), MD
Martin Doyle, Staten Island Technical H.S.
Ruth Edelman, Tenafly H.S., NJ
Karen Evans-Romaine, University of Wisconsin-Madison, WI
Elena Farkas, Turnagain Elementary School, AK
Thomas J. Garza, University of Texas, Austin
Beverly Ignatovicz, Marlboro H.S., NJ
Olga Kagan, University of California, Los Angeles
Jerzy Kolodziej, Indiana University
V. G. Kostomarov, Pushkin Institute, RF
Natalia Lord, Foreign Service Institute, VA
Byron Lindsey, University of New Mexico

Peter Merrill, Phillips Academy, MA
G. V. Moskvin, Moscow State University, RF
Benjamin Rjikin, Temple University, PA
William P. Rivers, Integrated Training Solutions, VA
K. A. Rogova, St. Petersburg University, RF
Cynthia A. Ruder, University of Kentucky
Elizabeth Sandstrom, Thomas Jefferson H.S. for Science and Technology, VA
Ala Smyslova, Columbia University, NY
Mara Sukholutskaya, East Central University, OK
E. D. Sulejmenova, Kazakh State University
Jane Shuffelton, Brighton H.S., NY
Charles Townsend, Princeton University, NJ
L. A. Verbistskaya, St. Petersburg University, RF
Irwin Weil, Northwestern University, IL

Russian Language Journal (ISSN 0036-0252) is a bilingual scholarly review of research, resources, symposia, and publications pertinent to the study and teaching of Russian language and culture, as well as comparative and interdisciplinary research in Russian language, culture and the acquisition of Russian as a second language. RLJ appears annually beginning with Volume 52 (2005) and is distributed to members of ACTR. It is available by subscription at a rate of $30.00 per year to individuals ($40 outside North America), $50 per year to institutions within North America and $65 outside North America. RLJ is also available to all interested readers via www.russnet.org/rlj in abstract form. Guidelines for contributors are listed on cover three and via the web at www.russnet.org/rlj/guidelines. Payments should be made to Russian Language Journal and forwarded to the subscription office.

Address manuscripts, books for review, and all other content-related correspondence to the appropriate section editor to:

Russian Language Journal
American Councils for International Education: ACTR/ACCELS
1828 L Street, NW, Suite 1200
Washington, D.C. 20036
Tel: (202) 833-7522 Fax: (202) 822-1689
Email: rlj@actr.org
http://modules.russnet.org/rlj/

Address business correspondence, subscription information, back issues, changes of address and related communication to Yulia Ledbetter

Copyright © 2010 by ACTR
Contents

Dedication 1

Introduction 2

Slavic Linguistics

A la Recherche du Temps Perdu Russian -SJA and -EN- in ASBMT 15
Leonard H. Babby

On the Status of Russian Perfective Passives in –sja 39
Jim S. Levine

SLA of Russian

The Overseas Immersion Setting as Contextual Variable in Adult SLA: 55
Learner Behaviors Associated with Language Gain to Level-3 Proficiency in
Russian
Dan E. Davidson, Maria D. Lekić

Adult Learners’ Perspectives on the Acquisition of L2 Russian Pragmatic 79
Competence
Victor Frank

The influence of prior language learning experiences on learning of 97
unrelated or distantly related languages
Ewa Golonka
Blogging, and Tweeting, and Chat, Oh My! Social Networks, Classroom Culture, and Foreign Language Instruction
Thomas Garza

Дисплейный текст как форма сетевого общения
Виталий Костомаров

A Case Study of the Acquisition of Narration in Russian: At the Intersection of Foreign Language Education, Applied Linguistics, and Second Language Acquisition
Benjamin Rifkin

Assessing the Oral Proficiency of Adult Learners, “Heritage” and “Native” Speakers Using the ILR Descriptors and ACTFL Guidelines: Considering the Challenges
Cynthia Martin

Heritage Language Development

Taking the “L” out of LCTLs: The STARTALK Experience
Catherine Ingold, Mary Hart

Heritage is Not Enough: The Changing Demographics of the Chinese Language Field in the United States
Scott McGinnis

Russian Heritage language speakers in the U.S.: A profile
Olga Kagan

Language Policy

Русификация и казахизация как языковая гомогенизация многоязычного Казахстана
Eleonora Suleimenova
Language Matters: Some New Contributions from Sociology (Emanating from Richard Brecht’s Castle)
*John P. Robinson, William Rivers, Cynthia Costell, Jennifer Robinson*

Some Thoughts on the Societal Impact of Linguistics
*Arto Mustajoki*

---

**Standards and Testing**

Assessment Practices in STARTALK Language Programs: A View of Current Language Assessment Literacy
*Margaret E. Malone, Megan J. Montee, Francesca DiSilvio*

A Concurrent Validity Study of Self Assessments and the Federal Interagency Language Roundtable Oral Proficiency Interview
*Charles W. Stansfield, Jing Gao, William P. Rivers*

Examining the Validity of the 2010 Prototype AP Russian Exam through a College Comparability Study
*Camelot Marshall*
Contributors

Leonard H. Babby: Princeton University, Princeton, NJ
Cynthia Costell: Independent Scholar
Dan E. Davidson: American Councils for International Education: ACTR/ACCELS, Washington, DC
Bryn Mawr College, Bryn Mawr, PA
Francesca DiSilvio: Center for Applied Linguistics, Washington, DC
Victor Frank: American Councils for International Education: ACTR/ACCELS, Washington, DC
Thomas Garza: University of Texas, Austin, TX
Ewa Golonka: University of Maryland Center for Advanced Study of Language, College Park, MD
Mary Elizabeth Hart: National Foreign Language Center, College Park, MD
Catherine Ingold: National Foreign Language Center, College Park, MD
Olga Kagan: University of California, Los Angeles, CA
Vitaly Kostomarov: Pushkin State Russian Language Institute, Moscow, Russia
Maria Lekic: University of Maryland, College Park, MD
American Councils for International Education: ACTR/ACCELS, Washington, DC
James S. Levine: George Mason University, Fairfax, VA
Margaret E. Malone: Center for Applied Linguistics, Washington, DC


Cynthia Martin: University of Maryland, College Park, MD


Megan J. Montee: Georgia State University, Atlanta, GA

Arto Mustajoki: University of Helsinki, Finland

Benjamin Rifkin: The College of New Jersey, Ewing, NJ

William P. Rivers: Integrated Training Solutions, Arlington, VA
National Language Service Corps, Arlington, VA

Jennifer Robinson: National Foreign Language Center, College Park, MD

John P. Robinson: University of Maryland, College Park, MD

Charles Stansfield: Second Language Testing, Inc., Rockville, MD

Eleonora Sulejmenova: Al-Farabi Kazakh National University, Almaty, Kazakhstan
There is a tide in the affairs of men,
Which, taken at the flood, leads on to fortune;
Omitted, all the voyage of their life
Is bound in shallows and in miseries.
On such a full sea are we now afloat,
And we must take the current when it serves,
Or lose our ventures.
*Julius Caesar, Act 4, scene 3, 218–224*

**Dedication**

Professor Brecht’s tireless work with students and colleagues includes three central tenets by which he has led his life of leadership and scholarship. Shakespeare encapsulates the first principle above, reminding us that our engagement with the world, and our application of our scholarly work, is in fact best taken as the wave crests. In living this tenet, Richard has led his public life by example, founding organizations, pursuing lines of research and inquiry, and participating in conversations deemed heterodox by the field at the time. It is worth recalling as well that Brutus precipitates his fall in *Julius Caesar* through his inability to understand the circumstances he finds. This leads us to the second tenet Richard has worked to instill in his students and colleagues: the value of острапенние in research and praxis. Richard’s work illuminates a state of mind – the rigorous examination of theoretical, applied, and actionable challenges from many perspectives in order to defamiliarize oneself, to stretch one’s analysis and examine one’s notions of what is and ought to be. The title of the volume alludes to a parable, taken from his fifth grade primer and oft repeated by Richard to his students, in order to illustrate the practical application of острапенние. A boy climbs an apple tree, and puts the biggest apple into his pocket. Once he tries to removes it, he thinks only of tearing the pocket, but his teacher suggests turning and cutting the apple. This volume embodies these two tenets of fearlessness and inquiry in expressions of the third tenet of Richard’s life: the profound engagement by Richard, his students, and his colleagues in his and their vocation: the improvement of the common weal through the expansion of foreign language learning and foreign language use.

Introduction

William P. Rivers

Dr. Richard D. Brecht
This volume is dedicated to an extraordinary leader, scholar, teacher, friend and mentor to many, Richard D. Brecht, in honor of his distinguished career as a researcher, organizer, advocate, and leader of significant and lasting institutions which have served the Less Commonly Taught Languages, the field of Russian, and the language readiness and preparedness of the United States. There is a tendency to count achievements as the measure of a man, more so when the qualities cited above are said to accrue to him. While Richard’s work has always been scientifically rigorous and politically and epistemologically provocative, what is far more telling and more important is the quality and diversity of people he has gathered around him. The works collected here from his students and colleagues reflect the breadth of Richard’s scholarly interests and the impact of his activities over the past forty years insofar as they reflect the countless debates, heated conversations, and long discussions of Slavic linguistics, the acquisition of Russian, language policy in the US and elsewhere, and many other subjects that are intimately and inimitably linked to his vocation. The contributors to this volume represent a distinguished international коллектив of like-minded colleagues, students, and researchers sharing Richard’s vocation: the improvement of the greater good through the expansion of language learning and use. That vocation rests on the integration of theoretical and empirical research, policy development. For the reader who may not be aware of his career and its import, I would like to take note of the significant reasons why we honor Richard Brecht with this Festschrift.

I would like to address Richard’s vocation first. Over a career of more than forty years, he embodies an approach to the academic and scholarly enterprise that we have come to call action linguistics. It is defined by, and proceeds from, taking action to achieve meaningful goals in changing the world around us, with the rigorous development of theory, tested against empirical data, and always improving and then applying that theory in the messy world of praxis. As I noted above, Richard has taken as his goal the improvement of the common weal in the United States and globally through research, praxis, and
advocacy to expand the teaching, knowledge, and use of languages. He has come to that goal through a long period of discernment and purposeful activity in his scholarly life, from which one is called to a vocation. At the same time, Richard has taken the tide, as it were, building networks of colleagues and consequently, organizations, on the principle of common approaches and shared solutions – first and foremost among these being the American Council of Teachers of Russian, followed by a great many other enduring institutions: the American Councils for Collaboration in Education and Language Study, the American Councils for International Education, the National Council of Less Commonly Taught Languages, the regionally-focused Language Resource Centers under Title VI of the Higher Education Act, and the Heritage Language Consortium.

We turn to the evidence that gives us specification of Richard’s influence on language learning and use, which we will come to see as the template for all of Richard’s later networking and building of organizations. While a graduate student in the Slavic Department at Harvard University in the late 1960s, specializing in the application of Jakobson’s (and by extension, Chomsky’s) formal systems analysis to Russian temporal and aspectual phenomena, Richard struck up a friendship with another graduate student whose interests inclined to the connections between 19th century Russian and German literature. The partnership that developed between Richard and Dan Davidson endures some 45 years later, and the fruit of their work, captured by Richard in his encomium for Dan in a festschrift only recently published, endures and continues to grow.1

Among the chief among these must lie the establishment and activities of the American Council of Teachers of Russian, itself the culmination of a bold and perhaps impertinent gamble by two young Ph.Ds, building on two unprecedented conferences on Soviet-American collaboration on the teaching and analysis of Russian.2 ACTR subsequently extended into the other constituent republics of the (former) Soviet Union through the establishment of the American Council for Collaboration in Education and Language Study (ACCELS). As the Cold War wear ended, Richard and Dan led the merger of

---


ACTR and ACCELS into the American Councils for International Education in 1993. The growth of American Councils into the primary US NGO for educational exchanges and development in the new post-Cold War space is ably reported elsewhere (Huber, 2004), as is the development of the "Russian in Stages" series - the first modern instructional complex for Russian to account for contemporary language usage and advances in Second Language Acquisition. However, the greater import in terms of Richard's career lies with fundamental and disruptive work done by Richard, Dan, and their late colleague Ralph Ginsberg on proficiency gains during study abroad. By disruptive we mean only that the combination of rigorous instrumentation of pre- and post-program proficiency levels and multiple demographic, biographic, and cognitive independent variables, with advanced correlational analysis, set a new standard for research in Second Language Acquisition. Specifically, Dick and Dan set out to do two things never before attempted: first, to describe the outcomes of long-term sojourns abroad, and second, to test hypotheses as to how these outcomes derive. The "Predictors" model of research has been replicated and extended to other research questions, from government programs in the US to the analysis of language shift in Belarus, Ukraine and elsewhere.

Rigor and excellence in basic and applied research in the service of action is further represented by the University of Maryland Center for Advanced Study of Language, the national laboratory for language, culture, and cognition, in service of national security. The practice of language learning is represented in Richard’s portfolio by The Language Flagship and LangNet, among others, while his ability to network and support colleagues is amply demonstrated by ASTM

---


Main Committee F43 on Language Services and Products, established by his student and colleague William P. Rivers in 2011. The purpose of this latest enterprise is to organize the $25b language industry and connect it with the government and academic sectors through the development of national and international standards for language testing, teaching, translation, interpreting, and human language technology. The establishment of the National Language Service Corps (NLSC) in 2008 is another example of applying a wide range of work to a practical goal: mobilizing the latent capacity of heritage speakers, naturalized citizens, and academic language learners in a national volunteer organization to serve the nation in times of need. The NLSC too is an idea advocated by Richard and implemented by his colleagues. He and several of the authors in this volume are members of the NLSC.

These organizations rest on and proceed from a solid foundation of empirical knowledge and flexible, evolving frameworks for the description and organization of the language enterprise and in particular, the Less Commonly Taught Language fields, in the US. It is highly instructive to examine these in some detail, as they serve to illumine both the chronological expansion of Richard’s interests as well as the lasting effect he has had on the language enterprise of the United States. We examine several models here – the Field Architecture Model, the Market Forces Framework, and the Core and Multiplier/Pipeline and Reservoir Model. These serve as ready rubrics for organizing the themes presented in this volume by Richard’s students and colleagues.

The Field Architecture model provides a coherent view of the span of institutions, organizations, and programs contributing to the national capacity in foreign languages. As Brecht & Rivers (2012) state:

Because of its importance, it is necessary to appreciate the nature of the academic infrastructure underlying the nation’s language capacity. Essentially, the core of our ability to develop and maintain expertise is the language field, which can be analyzed as comprising, for any given language or language area, foundational elements (expertise base, research, national organization, strategic planning, national resource centers), infrastructure (teacher training programs, in-country immersion programs, publications outlets, assessment instruments, etc.), as well as exemplary national programs (Brecht & Walton, 1994; Brecht & Rivers, 2000)

Indeed, this model was empirically developed through three landmark national surveys of language fields: Japanese, Chinese, and Russian, undertaken by Richard and his colleagues at the National Foreign Language Center in the late
1980s and early 1990s. The culmination of these three, the seminal Russian in the United States: A Case Study of America’s Language Needs and Capacities, presents the first versions of both the Language Field Architecture Model discussed here and the Market Forces Framework for Language.

Figure 1: The Language Field Architecture Model

![Diagram of the Language Field Architecture Model]

The vitality of this model – its repeated implementation to improve the LCTLs in the United States and the ongoing refinement of the model – reveals itself in a careful examination of the multiple programs, organizations, and institutions exemplifying the model and in which Dr. Brecht has played a foundational role. “Foundation” elements include field organizations such as NCOLCTL – co-

---

founded by Richard in 1987, and the American Council of Teachers of Russian, co-founded by him and Dan Davidson, Irwin Weil, and Claire Walker in 1974. Infrastructure elements include study abroad, which we will see below to be one of the enduring contributions Richard has made to the theory and praxis of language learning, as well as Advanced Distance Learning (or blended learning) where Richard is a pioneer, having founded project ICONS at the University of Maryland in 1982, and LangNet at the National Foreign Language Center in 1997. As I noted in the dedication, Richard uses metaphors in abundance to illustrate key points; the fruit on the tree of capacity, as it were, are the programs which teach foreign languages to professional levels of proficiency. Here too, Richard is responsible for sustained excellence as a member of the faculty of the Russian Departments of the University of Maryland and Bryn Mawr College. He was Robert Slater’s partner in the original elaboration of the Language Flagship, which, under the auspices of the National Security Education Program, has funded high level programming in LCTLs for more than ten years.

The second theoretical framework originated by Richard and his college Ron Walton in the mid-90s is the Market Forces Framework. As elaborated by Brecht & Rivers (op cit.), this macroeconomic description of the external forces acting upon the language fields, as exemplified above, is intended to describe these forces, and to provide a theoretical basis for advancing language capacity beyond the traditional measures of student enrollments (and the concomitant programs that meet student demands for instruction). Figure 2, below, presents this model.

Figure 2: The Market Forces Framework for Language
Demand refers to the specific tasks or interactions for which language competence is necessary or desirable; supply refers to the available language competencies (human and technological). Supply and demand are immediate and present in the marketplace; the innovation that Richard introduced with Ronald Walton is that the national interest as it relates to languages requires strategic, long-term considerations. Thus, they introduced the concepts of need and capacity. Need represents the perceived conditions that require, or may require, language competence. Brecht and Rivers (2005) extend this with an analysis that places need squarely in current economic theory as the perceived social marginal value of language – the public benefit accruing from language, or the cost if language needs are not met. The final, and for this discussion, critical element of this framework is capacity: the ability of the nation to produce the supply of linguistic human competence and technology designed to meet demand. We then refer back to Figure 1, which is subsumed by and constitutes the capacity sector in the Market Forces Framework.

First published in the Annals of the American Academy of Political Science in 1994 (Brecht and Walton, 1994) and then in two later monographs (Language and National Security in the 21st Century, and Language and Critical Area Studies after September 11: An Evaluation of the contributions of Title VI/F-H to the National Interest), this model has had tremendous impact on federal support for language programming in the past decade. Simply put, it makes an empirical case for the proposition that language, being vital to the common good (need) requires investment in long-term solutions (capacity), independent of short term fluctuations in enrollments (which have historically driven local decisions in the educational sector on the maintenance, expansion, or contraction of programs), job openings for language professionals, and so forth. Language becomes therefore a common good, on par with clean air, or financial security; the cost of maintaining such a good becomes then a political imperative rather than solely a market driven calculation. The language advocate has a means of advancing

---

beyond that which we hold self-evident – that language is sui generis an inherent good deserving of support – while, in fact, instantiating that argument empirically.

More recently, Richard has developed a roadmap for investment in the language capacity in the US, based on both of the foregoing models, that captures in detail the complexities of the requirements for language in the federal sector, and the multiplicity of sources that provide or support the government’s capacity for language expertise. The essential hypothesis is that no federal, state, or local agency can meet the need for language with core capabilities alone. Rather, a whole range of “multiplier” sources or strategies have to be employed, which include “shared,” “outsourced,” “warehoused,” “localized,” and “reach back.” These sources are summarized in the Pipelines and Reservoirs Model, depicted below in Figure 3 (from Brecht & Rivers, 2012). This model provides, for the first time, a coherent picture that relates all of the many elements of language capacity to specific language needs. Moreover, it details the complex relationships that entwine the language enterprise in the United States. At this writing, Richard is advancing this model as a conceptual framework for improving federal investments in language.

We turn now to the present volume and its relationship to Richard’s career. In roughly chronological order, with some notable omissions, his interests have spread from Slavic Linguistics, Second Language Acquisition and pedagogy, of Russian as well as other Less Commonly Taught Languages, Heritage Language Development; Language Policy and Planning; and Standards and the Language Industry. The contributions in this volume follow this order. The perspicacious reader will no doubt note that several of these contributions span more than one of these strands, reflecting the interrelationships among the scholarly and practical arenas of the language enterprise, as well as Richard’s work in them. The articles in this volume combine theory, empiricism, and praxis in varying degrees; the authors as a collective and in their individual contributions reflect Richard’s emphasis on all three. Richard’s focus on Slavic linguistics and morphosyntax, in particular the deeply entwined problems of grammaticalized meanings like aspect, tense, and mood, began in his graduate work with Horace Lunt at Harvard University.12,13 This strand of work is here represented by

---

12 Figure 3 Abbreviations: CLPs: Command Language Programs; CASL: University of Maryland Center for Advanced Study of Language; DLIFLC: Defense Language Institute Foreign Language Center; DoS: Department of State; FLAP: Foreign Language Assistance Program;
contemporary contributions from Leonard Babby, who extends our understanding of the theoretical basis of voice and reflexivity in the Argument Structure Theory of Morphosyntax, and James Levine, who takes new, corpus-based data to re-examine the theoretical grounding for perfective passives in Russian.

We turn next to Second Language Acquisition – of Russian and other LCTLS – which instantiates the practical application of the work done by Richard and others in the 70s on Russian linguistics, spreading out to entanglements with cognitive psychology, sociolinguists, pragmalinguistics, and many other allied fields. Richard’s work here includes fundamental contributions to the study of immersion and pedagogy. Moreover, Dr. Brecht was a pioneer in the application of the internet to language learning. In the 1990s, Richard began to take an interest in third language acquisition and the cognitive advantages accrued by advanced mastery of a second language, recognizing the commonality between research in third language acquisition and in advanced second language learning. Contributors to this volume include Dan E. Davidson and Maria D. Lekić, working with the vast American Councils of


Teachers of Russian database on immersion outcomes to continue the work begun in the late 1980s on the empirical analysis of language gain during study abroad. Their work in this volume takes a particularly close look at highly advanced learners in the Flagship program and examines their outcomes, finding that high levels of proficiency attainment – above and beyond the ACTFL Superior level – correlate very strongly with extensive time spent in a wide variety of language behaviors in addition to formal instructional settings.

Victor M. Frank provides an in-depth analysis of the acquisition of pragmatics among advanced sojourners in Russia, adding significant texture to the field’s understanding of the ability learners tested as having professional levels of language proficiency to acquire native sociopragmatic skills. Benjamin Rifkin adds an analysis of the acquisition of narration, the critical functional component of basic proficiency in any foreign language. Ewa M. Golonka contributes an article on the cognitive bases of third language acquisition, here focusing on the benefits accruing when a learner moves to an unrelated third language.

In the mid-1980s, Richard turned his attention to Heritage Language Development, as an outgrowth of work on language policy and planning in the US with A. Ronald Walton. From this work grew several key concepts in language capacity as discussed above; among the most important was the concept of Heritage Language Development. This work has continued and grown, with several national conferences, the establishment of a National Language Resource Center under the auspices of Title VI of the Higher Education Act, the Heritage Language Consortium housed at CAL, and the integration of heritage language development into the conceptualization and implementation of language policy in the United States. This volume presents contributions on heritage language development from Scott McGinnis on the state of the Chinese field in the US, extending the work done by Brecht and Walton to provide an evolving description of the field. Olga Kagan adds a careful, empirical description of Russian heritage learners in the US, and the complexities in learner profiles and skills masked by the term ‘heritage.’ Cynthia Martin examines the foundational construct of proficiency in light of heritage language development, provoking a dialogue on the validity of the proficiency construct when applied to heritage learners and native speakers.

---

Language policy and planning became a significant and enduring focus of Richard’s from the mid-1980s on, with and emphasis on the development of fundamental, empirical understanding of how language functions in the United States, and application of the resultant frameworks to the policy challenges of language in the American context and in the interface of work done in the US with work done in other polities. This volume presents a small sampling of the lines of policy work proceeding from the Field Architecture Model, the Market Forces Framework, and the Core & Multiplier Pipeline Model (Brecht & Rivers, 2012). Suleimenova details the complex, rich, and at times tense, balance between Russian and Kazakh, now undergoing significant change as the first fully Post-Soviet generation emerges in the Republic of Kazakhstan. John P. Robinson, William P. Rivers, Cynthia Costell, and Jennifer L. Robinson examine sociological and educational correlates of Americans’ language abilities, continuing a line of large-scale national survey work on language policy begun some dozen years ago with Richard. They draw the conclusion that language abilities in the US have steadily fallen, controlling for immigration, during an era when educational outputs – degree completions in particular – have increased. Arto Mustajoki develops a framework for the societal impact of linguistics – echoing Richard’s call for Action Linguistics.

We turn next to new paradigms for language learning, where Richard has again been a pioneer and inspiration. Vitaly G. Kostomarov provides a theoretical description of a new genre of text, the “display,” which includes the rapidly evolving text typologies found in the new online, global world; Kostomarov points out that communication in the 21st century requires ever more mastery of this genre, and he lays out initial thoughts as to its incorporation in pedagogy. Thomas Garza reports on his pedagogical work in using these new media for teaching Russian. Catherine W. Ingold and Mary Elizabeth Hart describe the history and future of StarTALK, an initiative underway at the National Foreign Language Center to expand LCTL offerings, including Russian, through a combination of teacher training for heritage speakers and other non-teachers, summer programs for junior and senior high school students, and on-line assessments and blended learning for year round support of these teachers and learners.
In our final section, we address an emerging area of work, and another outgrowth of the capacity framework: a focus on developing assessments for language. With Richard’s active involvement, this has grown into a lever to promote the development of international and national industrial standards and the basis for a platform for organizing the $15b language industry in the US. This volume presents new research from Margaret E. Malone, Megan J. Montee, and Francesca DiSilvio, who provide a systematic analysis of the formative and summative assessment practices of the StarTALK program. Charles W. Stansfield, Jing Gao, and William P. Rivers present a concurrent validity study of the “can-do” self-assessments and the Oral Proficiency Interviews used by the National Language Service Corps. They demonstrate, albeit tentatively, that such self-assessments can be used for initial screening of proficiency for high stakes...
purposes. Camelot A. Marshall presents a detailed comparability study of the Russian AP® Test, demonstrating that it is on par with tests developed for much larger student populations.

**Concluding Note: “There is a tide in the affairs of men / Which, taken at the flood, leads on to fortune”: A Prospective on the Language Enterprise**

The title of this volume refers to Richard’s habit of teaching through metaphors, with the subtitle intended to focus our energies on moving forward with our shared vocation of improving the greater good through language learning and use. The past fifteen years have seen sustained growth in the federal interest in language in the United States, as well as geometric growth in the language services industry. Indeed, in many objectively measurable ways, from the public attention paid to foreign language programs and competency in the mass media, to the aforementioned growth in the private sector language market, to the growth in federal funding, (to include programs beyond those in the US Department of Education, to public support for foreign language programs,19 language has never been better positioned in the United States. We say ‘positioned’ because past experience shows that language has been part of the response to crises, and then rapidly forgotten. That this volume is a “prospective,” speaks to the commitment on Richard’s part, and on that of his students and colleagues, to seize the opportunities presented, or to create them if absent; to look always ahead; to continue to learn and build, to strive and achieve, with Richard’s energy, fearlessness, and kindness as a model of how this vocation can be lived.

À la Recherche du Temps Perdu: Russian -SJA and -EN- in ASBMT

Leonard H. Babby

1. Introduction
In 1973-74, Dick Brecht and I wrote an article entitled *The Syntax of Voice in Russian*, which appeared in *Language* in 1975 (hereafter B&B 1975). Since then Dick and I have gone off in different directions, but I have returned periodically to the rich vein of data and problems in our paper and return to them again here, this time within the framework of the argument-structure based theory of morphosyntax (ASBMT) proposed in Babby 2009, 2010a, 2010b, and 2011, which has enabled me, after the passage of 35 years, to pinpoint what was right in the *Syntax of Voice in Russian* and where we went wrong. Our argumentation against the existence of complex, construction-specific transformations like early versions of the passive transformation has stood the test of time. I will concentrate here on the most productive uses of the Russian -sja (sja and s’ after vowels) and -en- (en- ~ n- ~ t-) suffixes, which was the empirical focus of our 1975 paper.

We proposed that the different morphosyntactic properties of -sja and -en-follow from the fact that the former is introduced “transformationally”, i.e., by a syntactic rule, whereas the later is introduced “lexically”, i.e., that -en- is a participle-forming “derivational suffix” which is introduced pre-syntactically in what at the time was called the morphological component (see Halle 1973). I now reject this analysis.

It is demonstrated in Babby 2010a that in ASBMT there are two computational domains rather than the one syntactic computational domain posited in Principles & Parameters theory (P&P) and the Minimalist Program. They are the argument structure (presyntactic) domain (AS), where the verb stem V composes with one or more of the language’s productive bound morphemes, and the syntactic domain. This entails that, contra B&B 1975, the -sja and -en- suffixes in passive and middle (derived intransitive) derivations are both introduced presyntactically in the AS domain and, therefore, that their disparate morphosyntactic properties must be encoded in terms of their AS representations and featural composition; e.g., the -en- suffix has inherent adjectival features, whereas -sja has no categorial features of its own and, therefore, [V-en-] is a deverbal adjective (-en- participle) but -sja in [V-sja] does not determine V’s final category. Syntactic operations (rules) alter the sentence’s syntactic structure; they do not introduce suffixes and cannot affect the internal structure of the word. AS operations
(rules) typically alter V’s initial AS, and the derivation’s final AS representation projects the sentence’s core syntactic structure. Thus, syntactic operations do not have access to AS or the verb’s internal morphological structure, but V’s final, derived AS determines (projects) the sentence’s core syntactic structure, which is why ASBMT can be described as an AS-based theory of morphosyntax.

It was assumed in B&B 1975 and Babby 1975 that all the productive uses of -sja could be explained in terms of a single -sja suffix. I argue below that there is evidence in modern Russian for at least two different, homophonous -sja suffixes; this evidence is based on theta-role assignment, which did not play as significant a role in our 1975 analysis as it does in ASBMT.

This paper is organized as follows: Section 2 is an outline of the relevant tenets of ASBMT that differ from P&P theory and the Minimalist Program. Section 3 argues that -en-, unlike -sja, always derives unaccusative verb stems (V) in standard Russian. Section 4 treats -en- and -sja in passive and middle derivations in the ASBMT framework. In section 5, I present my main hypothesis, namely, there are at least two distinct, productive homophonous -sja suffixes in Russian. Section 6 briefly explores the function of -sja and -en- in impersonal sentences.

2. Argument Structure Based Morphosyntactic Theory

ASBMT is based on isolating (e.g., English), agglutinating (e.g., Turkish), inflectional (e.g., Russian), and polysynthetic (e.g., Mohawk) languages. The following are its most salient properties; empirical evidence for all aspects of ASBMT can be found in Babby 2009, 2010a, 2010b, and 2011.

• It is ASBMT’s two tandem computational domains, with AS-domain operations preceding the syntactic-domain operations, that makes ASBMT a balanced, fully integrated, seamless derivational theory of morphosyntax. P&P theory is at heart a theory of syntax, not morphosyntax.
• AS in ASBMT is the presyntactic, hierarchically structured computational domain in which AS operations (rules) map V’s initial AS into final, derived AS. More specifically, the initial AS of V, which is ‘stored’ in V’s entry in the mental lexicon, and the AS of one or more of the language’s productive affixes compose (amalgamate) to derive V’s final AS. The AS of each affix is responsible for a specific change in the AS it composes with. Languages differ with respect to their inventory of productive affixes (e.g., Russian has neither a causative nor applicative suffix) and to the AS and featural composition of these affixes (e.g., we see below that the AS of -en- in standard Russian is different from the AS of -en- in Ukrainian and northern Russian dialects). Thus each productive suffix has its own AS and the central AS-domain operation is composition (merger) of V’s AS with the AS one or more suffixes. The
symbol “*” denotes AS composition and “>” denotes the result of AS-domain operations; e.g., \( V + Y > [V \ Y] \), i.e., \( V \) composes with \( Y \) to create \([V \ Y]\). See (6) for a specific AS-domain derivation (“\( >> \)” denotes an automatic, configurationally induced AS operation (see (6c) \( >> \) (6d)) and “\( \Rightarrow \)” denotes projection of final AS to initial syntactic structure in non-polysynthetic languages). We see in (1) that “>” has a second function: when not coupled with “*”, it denotes “is realized as”, e.g., \( c > -af \) is read “\( c \) is realized as an affix”. Thus an AS-domain operation is typically an affix-driven ‘rule’ that alters \( V \)’s initial or derived AS representation in highly specific ways (see below), which thus affects the projected sentence’s structure and grammatical relations.

- AS domain operations canonically involve the composition (AS merger) of \( V \)’s AS with the AS of a productive affix. For example, the AS of the causative suffix in Turkish has its own external agentive argument, which projects as subject of the causative sentence, displacing the external argument in \( V \)’s AS when the two ASs merge. \( V \)’s displaced external argument occupies the left-most ‘unused’ position in \( V \)’s initial AS and thus projects to syntax as an object (see below). The passive affix’s AS is responsible for the displacement (dethematization) of \( V \)’s external (agent) theta role.

- AS-domain operations can thus alter the grammatical relations encoded in \( V \)’s initial AS. In contrast, syntactic operations cannot alter the sentence’s projected-from-AS grammatical relations (see below). The primitives of AS-domain operations are thus morphemes, whereas the primitives of syntactic operations are words and the phrases they head. AS-domain derivations culminate in well-formed verbs (words) and their final ASs, whereas syntactic derivations culminate in sentences.

- In non-polysynthetic languages, \( V \)’s final, derived AS projects the sentence’s initial syntactic structure, which is the second computational domain. Syntactic operations (rules) map the initial syntactic structure to the sentence’s final syntactic structure, which is the input into Information Structure, which is responsible for Russian’s discourse-sensitive final “free” word order. Only AS operations can alter grammatical relations; post-AS operations can thus be characterized as grammatical-relation-preserving. Wh-movement is thus a quintessential syntactic rule: the wh-word’s movement does not alter the sentence’s core grammatical relations or case.

- The ordered, tandem relation between the two computational domains derives the Lexicalist Hypothesis: syntactic rules cannot backtrack and introduce, rearrange, or extract the individual components of words, i.e., roots, stems, prefixes, suffixes, etc.

- \( V \)’s final AS in polysynthetic languages like Mohawk does not project core syntactic structure, where \( V \)’s NP\(_q\) arguments occupy dedicated phrase-structure argument positions (\( \theta \) denotes theta role). The arguments in the Mohawk \( V \)’s final AS project
as verb-internal argument affixes (prefixes) i.e., -af- (see Babby 2010a); overt NPs in Mohawk are optional adjuncts, not arguments (see Baker 1996, 2009).

- Verb structure in polysynthetic languages and phrase structure in non-polysynthetic languages are homologous projections of the universal core grammatical relations encoded in AS’s hierarchically arranged positions. There do not appear to be any other kinds of AS projection in human language (see Babby 2010a).

- It is assumed in P&P and the Minimalist Program that AS consists of a hierarchically ordered set of theta roles, which are assigned to their NP arguments in syntax. But the existence of -af- argument affixes entails that V’s theta roles must be linked to their categorical heads in AS, not assigned to NPs in syntax (see Babby 2009 and 2010a for extensive argumentation).

- It follows from this conception of AS that: (i) arguments in AS are bipartite, i.e., a theta role (θ) linked (^) to a categorical head (c), which is represented in ASBMT as {θ^c}; (ii) c in {θ^c} has the two values represented in (1) (“>” here denotes “is realized as”):

\begin{equation}
\begin{aligned}
\text{c} & \text{ > -af- (bound morpheme) and } \{\theta^\text{-af-}\} \text{ arguments project verb-internal affix relations in verb structure. } \{\theta^c\} & \text{ > } \{\theta^\text{-af-}\} \text{ is obligatory in polysynthetic languages and optional in some non-polysynthetic languages (see below).}
\text{or}
\text{c} & \text{ > N (word) and } \{\theta^\text{N}\} \text{ arguments project syntactic relations in phrase (syntactic) structure in non-polysynthetic languages.}
\end{aligned}
\end{equation}

We see in §5 that the distinction in (1) is the basis of my hypothesis that there are two productive -sja suffixes in Russian.

- ASBMT is a shift in focus from concentrating on syntactic computation to determining the proper interrelation and apportionment of computational operations between the AS and syntactic computational domains. This proposal is entirely consonant with the basic goals of the Minimalist Program since it effects an overall simplification of grammar’s organization and computational operations.

- Given that V can have no more than three arguments, (2a) is the universal representation of a ditransitive V’s AS: i is V’s left-most, ‘external’ theta role, which projects the sentence’s subject and is thus VP-external (i = agent if there is one); j and k are V’s internal, object theta roles (the theme and affectee in Bower’s 2010 terminology); the sub-numbers in (2) have an expository function, enabling us to refer to AS’s four positions; “-“ denotes an unused slot (V itself is never linked to a theta role in initial AS (cf. derived {i^[V-af-]}a in passive AS). Thus AS has two tiers (the theta role selection tier and the categorical (c) selection tier), four positions (three
argument positions and \( V' \)’s position), and, therefore, eight \textit{slots} in all. All ASs, no matter what \( V' \)’s valence may be (0, 1, 2, or 3), have the immutable 2x4 skeletal structure in (2). A crucial tenet of ASBMT is that \( V' \)’s initial and final ASs always have 2x4 structure, i.e., rules cannot increase or decrease the number of \textit{positions} in AS (but AS-level operations can introduce new arguments into \( V' \)’s initial AS (e.g., see causativization below)). (2a-e) represent the cross-linguistically most common initial ASs in non-polysynthetic languages (see Babby 2009 for (2f)).

(2)
a. \{i^N\}_1 \{i^N\}_2 \{k^N\}_3 \{-^V\}_4 \quad \text{(ditransitive} \ V)\\
b. \{i^N\}_1 \{i^N\}_2 \{-^v\}_3 \{-^V\}_4 \quad \text{(monotransitive} \ V)\\
c. \{i^N\}_1 \{-^v\}_2 \{-^v\}_3 \{-^V\}_4 \quad \text{(unergative} \ V)\\
d. \{-^v\}_1 \{-^v\}_2 \{-^v\}_3 \{-^V\}_4 \quad \text{(impersonal/subjectless} \ V)\\
e. \{-^N\}_1 \{j^N\}_2 \{-^v\}_3 \{-^V\}_4 \quad \text{(unaccusative} \ V)\\
f. \{-^v\}_1 \{j^N\}_2 \{-^v\}_3 \{-^V\}_4 \quad \text{(transitive impersonal} \ V)

The AS-to-syntactic-structure derivation of (2a) is schematically represented in (3); the AS-to-verb-structure derivation of a ditransitive in polysynthetic languages is represented in (4).

We see in (3) and (4) that the AS-encoded grammatical relations, the hierarchical positions of the projected arguments, and the theta-tiers are identical in the non-polysynthetic and polysynthetic projections. For example, \([v \ [ -\text{af}i- \ [ \ldots ]]]\) and \([v \ NPi \ Nomi [v' ... ]]])\] are homologous projections of \( V' \)’s initial external argument and thus both -\text{af}i- and NPi are ‘subjects’.

(3)
\[ \{i^c\}_1 \{j^c\}_2 \{k^c\}_3 \{-^V\}_4 \quad > \]
\[ \{i^N\}_1 \{i^N\}_2 \{k^N\}_3 \{-^V\}_4 \quad => \]
\[ \{v' \ NPi \ [V' \ [v \ NPi \ [v' \ tv \ NPi k]]] \} \quad \text{(syntactic structure)} \]

(4)
\[ \{i^c\}_1 \{j^c\}_2 \{k^c\}_3 \{-^V\}_4 \quad > \]
\[ \{i^\text{af}i\}_1 \{j^\text{af}i\}_2 \{k^\text{af}i\}_3 \{-^V\}_4 \quad => \]
\[ \{v' \ [ -\text{af}i- \ [-\text{af}i- \ [-\text{af}i- \ V]]] \} \quad \text{(verb structure)} \]

- The unused slots in \( V' \)’s initial 2x4 structure we see in (2) play a crucial explanatory role in AS-level operations. For example, as we saw above, productive causative affixes have their own external \( \{i^N\}_1 \) argument, which right-displaces \( V' \)’s initial \( \{i^N\}_1 \) external argument when the ASs of the causative affix and \( V \) compose: the
final AS position and, therefore, the syntactic projection, of V’s displaced external argument is exhaustively determined by the left-most unused {-^} position in V’s initial AS (see Babby 2009:45-54). For example, we see in (5) that when the monotransitive AS (2b) is causativized (i.e., the ASs of V and the causative affix compose), V’s left-most unused position is {-^}3, which correctly predicts that V’s displaced external argument {i^[N]}1 is displaced to {-^}3, giving {i^[N]}3, which projects as the sentence’s dative ‘indirect’ agentive object; the causative affix’s {i^[N]}1 argument becomes the external argument of the derived causative AS and projects as the sentence’s nominative subject (see Babby 1983).

(5) Causativized monotransitive Vs:

(a) Polisi Hasan-i bütün paket-ler-in-i aç-tir-di.
   
   Police-NOM Hasan-DAT all package-PL-POSS-ACC open-CAUS-PAST
   
   ‘The police made/had Hasan open all his packages.’ (Turkish)

(b) J’ai fait manger les pommes à Paul.
   
   I have made to-eat the apples-ACC to-Paul
   
   ‘I made Paul eat the apples.’ (cf. *J’ai fait Paul manger les pommes.) (French)

(c) Tanaka-ga John-ni hono-o yomi sase masu.
   
   T-NOM J-DAT book-ACC read CAUS tense/aspect
   
   ‘Tanaka makes John read the-book.’ (Japanese)

• Since verbs are stored in the mental lexicon as stems (V) along with their ASs, when we select a specific V, presumably on the basis of its lexical meaning, we automatically select its initial AS. This means that the projected sentence’s core syntactic structure is imposed on us by the grammatical relations encoded in V’s initial AS (in non-polysynthetic languages). Languages thus typically have a set of AS-bearing affixes that alter V’s initial AS and, therefore, the projected sentence’s grammatical relations and morphosyntactic structure (e.g., applicative, antipassive, and causative sentences (see Xolodovič 1974)), which enables the speaker to adjust the sentence’s syntactic argument focus (see the derivations of passive, middle, and impersonal sentences below). But none of these operations alter AS’s skeletal 2x4 structure (see Babby 2009: chapter 1).

• There are two kinds of productive affixes that compose with V in the AS computational domain: those with their own ASs (e.g., the causative affix) and those like the {θ^[af]-} argument affixes, which do not themselves have ASs.
• There are no operations that convert one theta role into another, e.g., an agent into an experiencer. In sentence pairs like the following, the k-role in the first sentence does not change to an i-role when it becomes the subject in the second: Vodoj.kINST napolnilo N jamuj.ACC ~ Voda.NOM napolinae jamuj.ACC ‘Water filled the-pit’ (see §6 and Babby 1994b and 2010c).

3. The -en- suffix in Russian

3.1 [V-en-] is always unaccusative. The AS of the -en- suffix canonically composes with the AS of a perfective transitive V, deriving an unaccusative stative [V-en-] participle stem, as (6b) > (6c) >> (6d). A primary difference between -en- and -sja is that the former always derives an unaccusative stem in standard Russian, which accounts for the fact that -en- participles never have agentive subjects or direct objects, whereas the AS derived by the latter is not always unaccusative (see §5). Since -en- is the head of the derived participle stem (see Di Sciullo and Williams 1987 for the notion “head of the word”) and since the -en- suffix has inherent adjectival features, i.e., [+V] and [+N], the [V-en-] participle stem inherits -en-’s categorial features and is morphosyntactically a deverbal adjective (cf. §4.3). A stem, which is a morphological primitive, becomes a word, which is a syntactic primitive, only when the stem composes with an inflectional suffix (see (6e) > (6f) below). Short form (SF) -en- participles (see (6f)), which are obligatorily predicates, must compose with a form of the copula, just as underived (lexical) SF adjectives must.

[V-en-] participle stems compose with either the predicate short form (SF) or the attributive long form (LF). Unlike lexical adjectives, -en- participle stems in their predicate function are not normally affixed with the LF (see §4.3).

3.2 -en- passive sentences. What happens to a perfective transitive V’s initial AS when it composes with the AS of -en- in a passive derivation? V’s initial internal {j^N}_2 argument is realized in the projected passive sentence as the subject of the SF -en-participle; V’s initial {i^N}_1 external argument is optionally realized as the by-phrase, which is an instrumental case adjunct NP in Russian, e.g.: Viktor.jN NOM bystro perevelm staunch jU.jF ACC ‘Viktor translated the-article quickly’ ~ Stat.jF.NOM bylae bystro perevedenase.f (ViktoromkINST) ‘The-article was quickly translated (by-Viktor)’. The passive derivation of perfective V is schematically represented in (6); the ‘blank’ slots in { ^ }, which are common in the AS of affixes and auxiliaries, denote inheritance, i.e., a blank slot inherits the contents of the corresponding slot in the immediately preceding AS (e.g., see (6d) + (6e) > (6f)).

(6) Passive derivation of a perfective monotransitive verb in Russian:
a. \{i^N\}_1 \{j^N\}_2 \{-^\sim\}_3 \{-^V\}_4  
\quad (V$'$s initial AS) 
\quad +  
\quad (composes with) 
b. \{-^\sim\}_1 \{^\sim\}_2 \{^\sim\}_3 \{^\sim\text{-en}\}_4  
\quad (-\text{en}$'$s unacc.-forming AS) 
\quad >  
\quad (a+b yields) 
c. \{-^N\}_1 \{j^\sim\}_2 \{-^\sim\}_3 \{i^\sim[V\text{-en}]\}_4  
\quad (unacc. passive participle stem) 
\quad >>  
\quad (automatic operation) 
d. \{j^N\}_1 \{-^\sim\}_2 \{-^\sim\}_3 \{i^\sim[V\text{-en}]\}_4  
\quad (unacc. passive participle stem) 
\quad +  
\quad (composes with in AS) 
e. \{^\sim\}_1 \{^\sim\}_2 \{^\sim\}_3 \{^\sim\text{afst}\}_4  
\quad (AS of the SF inflectional suffix) 
\quad >  
\quad (d+e yields) 
f. \{j^N\}_1 \{-^\sim\}_2 \{-^\sim\}_3 \{i^\sim[V\text{-en}]-\text{afst}\}_4  
\quad (final AS) 
\quad =>  
\quad (projects to syntax) 

Since \{-^N\}_1 and \{j^\sim\}_2 in (6c) do not project well-formed syntactic structures, j in \{j^\sim\}_2 obligatorily relinks (externalizes) to \{-^N\}_1 (see (6c) \textgreater (6d)), which yields the derived unaccusative AS in (6d). \{j^N\}_1 projects the sentence$'$s NP thematic subject, which is the hallmark of unaccusative sentences (for case marking, see Babby 1994a and Woolford 2006). A natural question here is why N in \{-^N\}_1 does not itself relink to \{j^\sim\}_2 instead. The reason is: (i) Unlinked theta roles can relink but there is no empirical evidence that unlinked categorical heads relink; (ii) The relinking of N to \{j^\sim\}_2 in (6c) would reprise the original \{j^N\}_2 argument in (6a), which is an illicit operation; (iii) The relinking of N in \{-^N\}_1 would yield \{-^\sim\}_1 (i.e., \{-^N\}_1 \{j^\sim\}_2 \textgreater \{-^\sim\}_1 \{j^N\}_2): \{-^\sim\}_1 is the external argument of impersonal (subjectless) sentences and Russian does not have \{-^\sim\}_1 \{j^\sim\}_2 \{-^\sim\}_3 \{i^\sim[V\text{-af}]-\text{afst}\}_4 impersonal transitive passive sentences (see §6 for details).

Since the SF of adjectives and participles in contemporary Russian always occurs with a form of the copula, byt$'$ $'$be$'$-introduction is required by the SF and is not a unique property of Russian -en- passive sentences, e.g., Stat$'$ja,F,NOM dolžnas,F bylas byt$'$ perevedenas,F,ADV včera$'$The-article was supposed to have been translated yesterday$'$: byla$'$ $'$was$'$ is the past tense copula of the SF modal adjective dolžna and the infinitive $'$byt$'$ $'$to-be$'$ is the obligatory copula required by the SF participle perevedena (see Babby 2009: chapter 2 and 2010d); perevedenas,F$'$s copula is the infinitive byt$'$ because dolžna selects an infinitive complement (cf. *Stat$'$ja dolžna byla perevedena and the active Viktor dolžen byl perevesti stat$'$ju včera$'$ Vikor was supposed to translate the article yesterday$'$: here perevesti is the infinitive complement of dolžen).

While not all unaccusative verbs in Russian are passive, all passivized verbs are unaccusative, which means that deriving the unaccusative AS in (6d) from (6a) is only half the story. Passivization per se universally involves the right-displacement of V$'$s i theta role from its initial slot in \{i^N\}_1 (which is why initial unaccusative Vs can$'$t pas-
rivier: they do not have an external i). More specifically, -en-’s external argument \(\{^\wedge\}\) in (6b) has the following effect when it composes with \(V\)’s external \(\{^iN\}\) argument: \(V\)’s external \(N\) is inherited, whereas \(V\)’s external i theta role is right-displaced (it deletes in middle derivations (see below)). Thus what is “special” about passive derivations is that \(V\)’s dethematized i is displaced, relinking to the left-most \(\{^\wedge X\}\) theta receptor, where \(X = N\) or \(V\) (see (6b) > (6c)). The only theta receptor in passive ASs is \(\{^\wedge[V\text{-af}]\}\) \(\{\text{-af} = \text{-en}\) or -sja\), yielding \(\{^i[V\text{-en}]\}\) (or \(\{i^[V]\text{-sja}\}\) when \(V\) is imperfective). Speaking in general terms, whenever a theta role (e.g., i in passive derivations) or bipartite argument (e.g., \(\{^iN\}\) in causative derivations) is displaced as the result of composition with the AS of an affix, the displaced element must occupy a vacant slot or position, or the derivation “crashes”, i.e., the resulting AS cannot project to syntax. Displaced theta roles must relink to the first, left-most \(\{^\wedge X\}\) and displaced bipartite arguments canonically occupy the first, left-most \(\{^\wedge\}\) position in \(V\)’s AS. \(\{^\wedge[V\text{-af}]\}\) is the only \(\{^\wedge X\}\) in the passive derivation: in (6c), \(\{j^\wedge\}_2\) and \(\{^\wedge\}_3\) are not theta receptors. i in \(\{i^[V\text{-af}]\}\) in (6) is referred to as an implicit theta role (see Brody and Manzini 1990), which licenses the optional adjunct by-phrase.

This derivation of passive sentences explains why the AS of passivized unergative \(V\)s project impersonal sentences: As we see in (2c), unergative \(V\) does not have an internal j to externalize and thus no external argument to project a subject (e.g., see German Sonntags wird nicht gearbeitet ‘One does not work on Sundays’ (see Roberts 1987:512)). Russian does not have impersonal passives; sentences like the following are not passive (see §6.2): EmuDAT ne rabotaetsja ‘He doesn’t feel like working’ ~ Onnom ne rabotaet ‘He is not working’.

Since the SF inflectional suffix -a in Stat’ja perevedena ‘The-article (has been) translated’ does not have inherent categorical features, -en- is the head of the participle, i.e.: \(\{^i\text{[pereved-en]-}[a]\}\) (see the notion of “relativized head” in Di Sciullo and Williams 1987). \(\{i^[en]\}\) is by definition an argument affix (see (4)). As noted above, all Mohawk’s arguments are \(\{\theta^\wedge\text{-af}\}\) argument affixes and, as demonstrated in Baker 2001, all Mohawk’s overt NPs are optional adjuncts (“argument adjuncts” in Grimshaw’s 1990 terminology), which, I claim, are licensed by the corresponding \(\{\theta^\wedge\text{-af}\}\) argument affixes. Thus the optional by-phrase in non-polysynthetic languages and the optional NP argument adjuncts in Mohawk are essentially the same phenomenon: both are licensed by “implicit” \(\{\theta^\wedge\text{-af}\}\) theta roles. The crucial difference is that all Mohawk’s arguments are obligatorily verb-internal argument affixes, whereas the adjunct by-phrase that passive \(\{i^[en\text{[pereved-en]-}[a]\}\) licenses in non-polysynthetic languages is optional.

Ukrainian impersonal transitive passives like Zapruż-en-IMPERSONAL, ričku,ACC ‘The-river is dammed-up’ illustrate that the Ukrainian -en- suffix also derives a static
participle, but, unlike its Russian counterpart, it is not always unaccusative (see Lavine 2000 and Babby 2002:367 for discussion).

4. Middle and passive sentences
The derivation of -en- passive sentences in (6) simplifies the task of accounting for the derivation and properties of ‘middle’ verbs (i.e., derived intransitives) and of imperfective passive sentences (where -sja is the unaccusative-forming suffix).

4.1 Russian imperfective passive sentences. -sja is used instead in -en- in passive derivations when \( V \) is imperfective (see (7) below): all we have to do in (6) is replace the -en- suffix with -sja and bear the following facts in mind: (i) unlike -en-, -sja has no categorial features and is thus not the head of the verb despite its obligatory verb-final position; (ii) -sja in passive derivations is an unaccusative-forming affix, just as -en- is; (iii) since -sja and -en- are both unaccusative-forming suffixes in standard Russian, they cannot both occur in the same derivation (but see dialect forms like U ego prostuž-en-o-s’ ‘He caught cold’ in B&B 1975:347; here, as in the case of the Ukrainian transitive impersonal passive above, -en- does not derive an unaccusative \( V \); (iv) -sja in modern Russian is a word-internal enclitic suffix, which means that it always occurs at the end of the verb no matter where in the derivation it is introduced. This can be represented as follows: \( V + -sja \rightarrow [...] [[V-] sja]...] \), which indicates that all subsequent suffixes compose with [V-], not -sja, thereby deriving the correct order of suffixes (see (25)). In contrast, \( V + -en- \rightarrow [...][[V-en-]...] \), which encodes the fact that all subsequent suffixes must follow -en- (see Babby 2009 for details).

(7)  
a. Rabočie strojat dom. ‘The-workers are-building a-house.’  
b. Dom stroitsja rabočimi. ‘A-house is-being-built by-the-workers.’

The fact that syntactic rules in ASBMT do not operate on affixes and do not alter the grammatical relations encoded in \( V \)’s AS eliminates from contention any analysis in which -sja is merged in syntax (as in B&B 1975).

4.2 Russian middle sentences. -sja is used in the derivation of middle (derived intransitive) verbs of both aspects. The following are typical imperfective middle sentences:

(8)  
a. Etot otryvok ego knigi legko perevoditsja  
   ‘This excerpt of his book translates easily.’  
b. Metally rasširjajutsja pri nagrevanii.  
   ‘Metals expand when heated (lit. widen upon heating).’
c. Eta tkan’ xorošo stiraetsja.
   ‘This material washes well.’

d. Cyplenok legko perevarivaetsja.
   ‘Chicken is easily digested (lit. digests easily).’

Middle sentences are derived-unaccusative sentences in which V’s initial external theta role i is deleted rather than displaced when V’s external \(i^{\text{^N}}\) argument and the external \(-^\wedge_i\) argument of the unaccusative-forming suffix (-en- or -sja) compose (see (6)). Passive and middle sentences in Russian are thus both derived unaccusative sentences; they differ with respect to the fate of i, which is made implicit in passive sentences and is deleted in middle sentences. This explains why passive verbs license by-phrases and middle verbs do not.

The following question naturally arises: How is this optionality of i encoded in V’s initial AS? It could be represented as \((i)^{\text{^N}}\) for verbs that have both a passive and middle. Something along these lines is needed to encode the fact that V’s like položit’ ‘to put’ have a passive (Kniga byla položena na stol (Ivanom) ‘The book was put on the table (by Ivan)’ but do not have a middle counterpart, i.e., Kniga položilas’ na stol *The book put on the table’, Kniga legko kladetsja na stol *The-book easily puts in the table’ (NB: the form položit’sja exists, but is the perfective of polagat’sja ‘to rely on’). Thus the external argument of položit’ ~ klast’ ‘put’ would be \(i^{\text{^N}}\), indicating that i cannot be deleted. But this is an area where further research is needed: the fact that položit’ does not form the middle may be predictable in terms of its lexical semantics and thus may not have to be encoded in its AS.

Perfective middle sentences canonically denote a completed action without reference to V’s agent (i.e., i in \(i^{\text{^N}}\)). In middle sentences of both aspects, V’s agent is semantically dissociated from the action denoted by V, as in (9b), whereas in passive sentences like (9c), the agent is merely defocused. “Unagentive” (natural-force adversity) impersonal sentences like (9d) are used when the speaker needs to explicitly assert that no agent is involved in the event (see Babby 1994b, 2010c):

(9)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>Myj.pl.NOM perevernuli pl lodkij.F.ACC</td>
<td>‘We turned the-boat over.’ (active)</td>
</tr>
<tr>
<td>b.</td>
<td>Lodkaj.F.NOM perevernulas’f.</td>
<td>‘The-boat turned-over.’ (middle)</td>
</tr>
<tr>
<td>c.</td>
<td>Lodka byla perevernuta (nami).</td>
<td>‘The-boat was turned over (by us).’ (passive)</td>
</tr>
<tr>
<td>d.</td>
<td>Lodkij.F.ACC perevernulon.</td>
<td>‘The-boat got turned over.’ (impersonal)</td>
</tr>
</tbody>
</table>

The [V-en-] participle can be the predicate of stative middle sentence, as in (10) - (12): a passive interpretation is either unlikely, as in (10), or semantically excluded, as in
(10)  a. Rebenok prostudilsja.  ‘The child caught cold.’ (non-stative middle)
    b. Rebenok prostužen.  ‘The child has a cold.’ (stative middle)

(11)  a. U nego ruki smorščilis’. ‘His hands became wrinkled.’
    b. U nego ruki smorščeny. ‘His hands are (have become) wrinkled.’

(12)  a. Ee telo pokrylos’ jazvami  ‘Her body became-covered with (*by) sores.’
    b. Ee telo pokryto jazvami  ‘Her body is covered with (*by) sores.’

The derivation of -en- stative middle sentences is basically the same as the derivation of -en- passive sentences represented in (6) above. The only significant difference is that instead of being made implicit, as in passive derivations, V’s external i is deleted, just as in the derivation of -sja middle sentences in (8) and (10a), (11a), and (12a).

4.3 -en- participles and -enn- adjectives. It is common in Russian for LF -en- participles to be reanalyzed as lexical adjectives, which have their own entries in the mental lexicon and have meanings that differ from V; thus, unlike -en- participles, the -enn- adjective is not a member of V’s paradigm (i.e., stem + productive suffixes), i.e., is not a morphosyntactic form created from V in AS-to-sentence derivation. This reanalysis creates homophonous -en- participle ~ -enn- adjective pairs like the participle rasseqannyj ‘scatter’ from rasseqat’ ‘to scatter’ and the adjective rasseqannaja ‘absentminded, scatterbrained’; see Babby 1993 for details. The following are the distinctive morphosyntactic differences between the -en- participle and the corresponding -enn- adjective: (i) In the SF non-masculine singular forms, the adjectives have -nn- whereas the participle has -n- (participles have -enn- only in their LFs), e.g., the SF feminine singular of the -en- participle of rasseqat’ is rasseqana ‘scattered’, whereas the SF of the -enn- adjective rasseqannaja ‘absentminded, scatterbrained’ is rasseqanna. (ii) -enn- adjectives, unlike the corresponding participles, can form the comparative (rasseqannee, but *rasseqanee). (iii) -enn- adjectives have manner adverbs in -o: rasseqanno, but *rasseqano. (iv) -enn-o is also the predicate of impersonal sentences (and of “personal” sentences with a neuter subject), see (13).

This brings us to the examples in (13), (14), and (15) (= (14c), (15), and (16) in B&B 1975), which we thought were problematic since the -enn-o forms were erroneously identified as participles rather than adjectives (see B&B 1975: footnote 5):

(13)  V nomere bylo teplo i otgoroženno ot vnešnegom mira.
'The room was warm and cut-off (isolated) from the outside world.'

(14) Anna smotrela na nego razočarovanno.

‘Anna looked at him disappointedly.’

(15) * Vanja rastroenno ženozj smotrel na menja.

*‘Vanja agitatedly by-his-wife looked at me.’

These sentences cease to be problematic once it is recognized that the -enn-o forms are derived from -enn- adjective stems, not from -en- participle stems. Adjective stems affixed with -o are common in impersonal sentences: tepló and otgorožennó in (13) are both impersonal SFs of the adjective; razočarovannó in (14) is a manner adverb formed from the adjective razočarovanniy (SF = razočarovanna) not the participle razočarovannyj (SF= razočarovana). (15) is ill-formed because adjectives do not have implicit i and thus do not license by-phrases.

The difference between -en- participles and -enn- adjectives is this: -en-participles are deverbal adjectives, i.e., derived from V’s AS: the -en- suffix’s AS composes with V’s AS in AS-to-sentence derivations and, therefore, the -en- participle is a hybrid verbal category, i.e., it retains V’s AS and is thus ‘internally’ verbal and ‘externally’ adjectival (see Babby 2009: chapter 3 for discussion of hybrid verbal categories in Russian). In contrast, -enn- adjectives are departicipial adjectives, i.e., they are the result of reanalysis, which is a diachronic phenomenon. Unlike -en- participles, they are not hybrid categories (i.e., do not encapsulate V’s AS): they have the same properties as basic, underived adjectives (e.g., form comparatives and manner adverbs). In other words, -enn- adjectives are the result of word formation, whereas -en- participles are formed in the AS domain as a phase of sentence formation (derivation) and, therefore, do not have their own lexical entries in the mental lexicon (and are thus not normally listed separately in dictionaries of Russian). This analysis correctly predicts that not all -en- participles have corresponding -enn- adjectives and the shifts in meaning under participle-to-adjective reanalysis are not predictable.

5. Unaccusative-forming -sja and argument affix -sjaq in Russian
We saw above that when the ASs of V and -en- compose, the AS of the resulting [V-en-] participle stem is always unaccusative in standard Russian; -sja in passive and middle sentences is also an unaccusative-deriving suffix. But, unlike -en-, there are relatively productive uses of -sja that patently do not derive unaccusative AS (see below). This fact was the impetus of my hypothesis that there are two productive -sja suffixes in Russian, whose disparate properties correspond exactly to the two types of productive suffix posited in ASBMT to account for other phenomena in (1); see (16).

(16) The two types of morphosyntactic affixes:
a. Affixes which have their own argument structures and compose with other ASs, e.g., the -en- and -sja unaccusative-forming suffixes discussed above.

b. Argument affixes, which are the selected value of c in {i^n}c (see (3) and (4)); e.g., verb-internal arguments in Mohawk, where {θ^n}c > {θ^-af}-2 => -af-.

More specifically, in addition to the AS-bearing, unaccusative-inducing -sja suffix we encountered above in passive and middle sentences (see (16a)), Russian has a second, homophous -sja suffix whose properties demonstrate that it is an -sja argument affix (i.e., {θ^n}c > {j^-sja}_2 => -sja), which, like all argument affixes, does not itself have an AS (see (16b)). Thus the argument affix -sja is optionally linked to transitive V’s j theta role in AS and does not have its own AS, whereas unaccusative -sja is just the opposite: it is not linked to a theta role and does have its own AS. The following is a useful rule of thumb: -sja affixed to a transitive V is an argument affix (-sja) if the sentence’s subject is NP1 (i.e., the projection-to-syntax of V’s initial external argument). Both unaccusative -sja and argument suffix -sja developed from Old Russian sja, which was a sentence-level accusative enclitic pronoun.

Under certain conditions, which are related to ease of recoverability, -sja can optionally be linked to j in a transitive V’s initial AS instead of to N, thereby deriving an AS whose projection is syntactically intransitive (only V’s external {i^n}N1 argument projects to syntax as an NP1 argument), but is thematically transitive since it still has V’s initial i and j theta roles, which are projected from their initial theta-tier positions, i.e.: {i^n}N1 {j^-sja}_2 ... => [... NP1 [v V-sja]...]. Russian direct object -sja projects verb-internal structure because it is a bound morpheme in modern Russian. However, unlike the obligatory argument affixes in Mohawk, optional {j^-sja}_2 in Russian does not license an NP argument adjunct (see Babby 2010a: §9). This sja1 ~ N1 alternation is schematically represented in (17).

(17)  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>{i^n}c1 {i^-c}_3 {^-V}4</td>
<td>&gt;</td>
</tr>
<tr>
<td></td>
<td>{{i^n}N1 {j^-sja}_2 {^-V}4}</td>
<td>=&gt;</td>
</tr>
<tr>
<td></td>
<td>[v NP1 [v V-sja]...]</td>
<td>(argument affix)</td>
</tr>
<tr>
<td>b</td>
<td>{i^n}c1 {i^-c}_3 {^-V}4</td>
<td>&gt;</td>
</tr>
<tr>
<td></td>
<td>{{i^n}N1 {j^-N}_2 {^-V}4}</td>
<td>=&gt;</td>
</tr>
<tr>
<td></td>
<td>[v NP1 [v V NP1]...]</td>
<td>(argument NP)</td>
</tr>
</tbody>
</table>

Since the -sja1 suffix in (17a) is linked to a theta role in AS and does not have its own AS, it is by definition an argument affix. Given that grammatical relations are encoded in AS and realized as either syntactic relations in phrase structure ((17b)) or as affixal relations in verb structure ((17a)), the optional selection of {j^-sja}_2 rather than {j^-N}_2 does not affect the projected sentence’s grammatical relations: In (17a), NP1 is the syn-
tactic subject and -sja1 is the verb-internal direct object, whereas in (17b), NPj is the syntactic direct object.

For example, -sja1 can optionally be used in place of the accusative direct object reflexive pronoun sebja1 ‘(one’s)self’ to realize a transitive V’s initial {j^c}2 argument when i and j are coreferential. See (18a-b) ((18c) is ill-formed as a translation of He washed himself because it violates the Projection Principle: {j^c}2 is not projected):

(18)  
a. O^n: myl [NP sebja]. ‘He NOM washed himself_{ACC}.’
b. O^n: {v myl-sja}. ‘He washed himself.’
c. *O^n: myl. ‘He washed himself.’
d. Sobaka: ližet sebja1 / ližetsjaj. ‘The dog licks itself.’

The lexical restrictions on -sja1 ~ sebja1 are too complex to go into here. Suffice it to say that the restrictions on -sja1 are far more numerous than those on unaccusative -sja and, therefore, the former is less productive than the latter. Examples like the following suggest that -sja1 ~ reciprocal drug druga; ‘each other’ is the same phenomenon: Oni redko obnimajutsja1 ~ Oni redko obnimajut drug druga; ‘They rarely hug each other’.

-sja1 also occurs in non-reflexive derivations. When V selects a specific noun as its direct object rather than the variable N, this noun is completely recoverable (semantically redundant) and is thus easily replaced by -sja1 without changing the projected sentence’s basic meaning or grammatical relations. For example, the direct object of vysmorkat’ ‘to blow’ can only be nos ‘nose’ and its initial AS is: {θ^c}1 {θ^c}2 {^-^}3 {-^V}4 > {[^iN]}1 {[^^c]}2 {-^}3 {-^[vysmorkaj]-^}4. This AS predicts that j can be optionally linked to -sja1, giving [{[^iN]}1 {[^sja]}2 {-^}3 {-^[vysmorkaj]-^}4]: vysmorkat’ nosi and vysmorkat’sja1 both mean ‘blow (your) nose’. See also skalit’ zuby1 ~ skalit’sja1 ‘to bare (one’s) teeth’; nesti jajco1 ~ nesti1 ‘to lay an egg’; nasupit’ brovi1 ~ nasupit’sja1 ‘to knit one’s brow’; nomorščit’ lob1 ~ nomorščit’sja1 ‘to wrinkle one’s forehead’; tratit’ den’gi1 ~ tratit’sja1 ‘to spend money’; parkovat’sja1 ‘to park the car’.

In the following examples illustrate another, a semi-productive use of -sja1: Sobaka: kusaetsja1 ‘The dog bites (people, not himself)’; Moja koška ne carapaetsja1 ‘My cat doesn’t scratch’; Mal’čik ljagaetsja1 i tolkaetsja1 ‘The-child kicks and pushes (others)’; Krapiva žžetsja1 ‘Nettles sting’.

The derivation of -sja1 is schematically represented in (19); since it is a word-level enclitic suffix, -sja1 always occupies the right edge of the verb, i.e., [v [V-... sja], not *[V-sja-af-]]. Compare (19) with (20), which is the derivation of unaccusative -sja in middle and passive derivations:

(19) The derivation of [V-sja] argument affixes:
The passivity and middle derivation of imperfective V's in Russian:

(a) \{[i^c]_1 [j^c]_2 \{-^\_\}_3 \{-^V\}_4\} > 
(b) \{[i^N]_1 [j^sja]_2 \{-^\_\}_3 \{-^V\}_4\} + 
(c) \{[\ }_1 [\ }_2 [^\}_3 \{^-afINFLECTION \}_4\} > 
(d) \{[i^N]_1 [j^sja]_2 \{-^\_\}_3 \{-^V^-afINFLECTION\}_4\} => 
(e) \[vP NP, [\ V^-afINFLEC}sja]\ VP]

(20) The passive and middle derivation of imperfective V's in Russian:

(a) \{[i^c]_1 [j^c]_2 \{-^\_\}_3 \{-^V\}_4\} > 
(b) \{[i^N]_1 [j^N]_2 \{-^\_\}_3 \{-^V\}_4\} + 
(c) \{[\ }_1 [\ }_2 [^\}_3 \{^-sja\}_4\} > 
(d) \{[i^N]_1 [j^sja]_2 \{-^\_\}_3 \{[i]^[[\ V^-sja]\}_4\} \gg 
(e) \{[i^N]_1 [j^N]_2 \{-^\_\}_3 \{[i]^[[\ V^-sja]\}_4\} + 
(f) \{[\ }_1 [\ }_2 [^\}_3 \{^-afINFLECTION^-\}_4\} > 
(g) \{[i^N]_1 [j^sja]_2 \{-^\_\}_3 \{[i]^[[\ V^-afINFLECTION \} sja]\}_4\} => 
(h) \[vP NP, [\ V^-afINFLEC}sja]\ VP]

6. The function of -sja and -en- in impersonal sentences

This last section is devoted to the most productive uses of -sja and -en- in impersonal (subjectless) sentences and to argumentation that -sja in impersonal sentences is not a third -sja suffix (see Babby 2010c).

6.1 Derived-unaccusative impersonal sentences. -sja in impersonal sentences like (21) - (24) is the unaccusative-stem deriving suffix, not -sja or an impersonalizing suffix; it is the -o suffix that is responsible for the impersonalization (see Babby 1996). The verbs in (21) - (24) are initially monotransitive, i.e., \{[i^N]_1 [j^N]_2 \{-^\_\}_3 \{-^V\}_4\} and both arguments are deleted in the course of the derivation.

(21) My vygnali stado na baxču, gde s dovoennoj pory ne paxalos’ i ne sejalos’.
  ‘We drove the herd onto a field where there had been no plowing or sowing done since before the war.’ (paxat’ ‘to plow,’ sejat’ ‘to sow.’) (A. Leonov)
(22) (about a p’esat’ ‘play’) Sam ne veril, kogda napisalos’. (Trusinovskaja)
  ‘I didn’t believe (it) myself when (it = play) (somehow) got-writtenN.’
  (napisat’ ‘to write’)
(23) Ona pytalas’ ob’edinit’ etix dvux v odnu kompaniju, no u nee ne ob’edinilos’.
  (ob’edinit’ ‘to unite’) (Tokareva).
‘She tried to combine these two (people) into a couple (= get them together), but it didn’t work out (lit. at her (it) didn’t combine/unite).’

(24) V golove zakružilos’. V živote u nego sžalos’.

‘(His) head began to spin (lit. in-head began--spinning). His stomach tightened (lit. in stomach (there) was-contracting)’(zakružit’ ‘to spin, sžat’ ‘to squeeze’).

Impersonalization canonically involves deleting V’s agentive external {i^N}1 argument (e.g., (9d); [NP] OblomkiACC korabljaGEN] vyneslon [PP na bereg] ‘Pieces-of-the-ship (got) washed-up on shore’). But in (21)-(24), V’s initial agent i is deleted during the middle-formation phase and it is the derived unaccusative AS’s externalized {j^N}1 argument that is deleted by impersonalizing -o (see the derivation (25)).

The basic semantic function of middle voice is to dissociate the agent (i) from the action/event denoted by the verb and to focus on the relation between the initially internal {j^N}2 argument and the verb’s lexical meaning (see (8)). When a middle sentence is impersonalized, its derived {j^N}1 external argument is deleted and thus also dissociated from the action denoted by V, the result being maximal focus on the action/state denoted by verb, abstracted away from its arguments, both of which have been deleted (see the notion of “thetic judgement” in Kuroda 1972).

For example, consider the initial monotransitive AS of ob’edinit’ ‘to unite, unify’ in (25a). But ne ob’edinilos’ (i.e., ob’ediniSTEM=1past-o’impersonal-s’ unacc-sja) in (23) has no arguments and means: no unification took place or there was no unification. The [PP u nee] adjunct, which is not projected from AS and not licensed by an implicit theta role, identifies a human affectee and is not confined to impersonal sentences (e.g., U nee často menjaetsja nastroeniεnom ‘(lit.) At her often changes mood’). The derivation of the impersonalized middle sentence in (23) is schematically represented in (25); V = ob’edini-.

(25) a. {i^N}1 {j^N}2 {-^}-3 {-^V}4 + (monotransitive V)
b. {-^}1 {^-}2 {^}3 {-^sja}4 > (unaccsative -sja)
c. {j^N}1 {j^-}2 {-^}-3 {-^[V-]sja}4 >> (a & b compose)
d. {j^N}1 {^-}2 {-^}-3 {-^[V-]sja}4 + (middle AS)
e. {^}1 {^}2 {^}3 {^}1past-}4 > (inflectional suffix)
f. {j^N}1 {-^}-2 {-^}-3 {-^[V-]sja}4 + (d & e compose)
g. {-^}1 {^}2 {^}3 {^}0’impersonal}4 > (impersonalizing -o)
h. {-^}1 {-^}2 {-^}3 {-^[V-0-o-]sja}4 => ob’edini-l-o-s’

6.2 Emu ne rabotaetsja. The following sentences, in which V is intransitive (unaccusative and unergative), illustrate another type of productive impersonal sentence involving
unaccusative -sja: On NOM.M ne rabotalaM ‘He wasn’t working’ ~ Emu DAT.M ne rabotalaS.N ‘He didn’t feel like working. I argue that Emu ne rabotaloS’ is not derived from On ne rabotal. The following are additional examples: Emu ne igraloS’ ‘He didn’t feel like playing’; Emu ne ležaloS’ v posteli ‘He didn’t feel like lying in bed’; see also: guljaloS’ ‘stroll’, xodiloS’ ‘walk’, sideloS’ (doma) ‘sit (home)’, spaloS’ ‘sleep’, tancaloS’ ‘dance’, xotelos’ + infinitive ‘feel like doing something’, Raz nenaviditsja - nenavidj ‘If you feel like hating, go ahead and hate’. The AS-level derivation of this type of impersonal sentence is schematically represented in (26).

(26)  
a. [i^N]1 {^-}2 {^-}3 {^-V}4 +  
b. [^-]1 {^-}2 {^-}3 {^sja}4  >  
c. [^-N]1 {^-}2 {^-}3 {^-[V-]sja}4 +  
d. {^}1 {^}2 {^}3 {^-lpast-}4 >  
e. [^-N]1 {^-}2 {^-}3 {^-[V-l-]sja}4 +  
f. [^-]1 {^}2 {^}3 {^-Oimpersonal}4 >  
g. [^-]1 {^-}2 {^-}3 {^-[V-l-o-]sja}4 =>  
h. (Emu ne) rabota-l-o-s’ ‘He did feel like working, etc.’

A basic tenet of ASBMT is that there are no operations of any kind in any domain that change the value of the theta roles specified in V’s initial AS. This entails that the agent on in On NOM.M ne rabotalaM ‘He wasn’t working’ is not converted into the dative experiencer emu in Emu DAT ne rabotalaS ‘He didn’t feel like working’. My analysis is that an unergative V’s initial external [i^N]1 and an unaccusative V’s externalized [j^N]1 are deleted by the sequence of middle-formation and impersonalization, just as in the derivation of the impersonal sentences in (21) - (24). The difference between the derivations in (25) and (26) is this: unaccusative -sja’s AS in (26b) deletes only V’s external i because there is no internal N linked to j here for it to delete and, therefore, the [j^N]2 > [^-]; component of -sja’s unaccusativizing AS applies vacuously, i.e. [^-]2 in (26a) + [^-]2 in (26b) > [^-]2 in (26c).

The optional dative experiencer in sentences like Emu ne rabotaloS’ has essentially the same status as [PP u + NPGEN] in (23), i.e., it is an optional adjunct, which is merged in syntax, not projected from AS, and is found in many other kinds of impersonal sentences, e.g.: UčenikUM.DAT (emuUM.DAT) otrezaloN.impersonal palecM.ACC ‘The student (he) got (his) finger cut-off’; Emu tam skučn-o ‘He is bored there (experiences boredom)’ (cf. On skučen ‘He is boring (others are bored by him)’); see Babby 1994b. See also Ja nastupil emu na nogu (*...na emu nogu) ‘I stepped on his foot (lit. *I stepped him on the foot)’.

À la Recherche du Temps Perdu: Russian -SJA and -EN- in ASBMT
Leonard H. Babby
My guess is that the ungrammatical nature of sentences like *Emu ne rabotalo, etc. has a semantic explanation: impersonal sentences like Emu otrezalo palec are anti-agentive, i.e., they assert that no human agent is involved. This construction is naturally limited to verbs denoting actions that can occur with or without the participation of an agent (see (9)), whereas the verbs in the Emu (ne) rabotalos’ construction cannot make such an assertion since they all involve actions whose lexical-semantic meaning is inherently human (animate) and thus describe an event that lexically precludes deagentification; cf. Emu otrezelo palec ‘His finger got severed’ vs. *Emu amputirovalo palec ‘His finger got amputated’.

6.3 [V-en-] in impersonal sentences. The analysis of the impersonal sentences in (21) - (24) correctly predicts that the same type of impersonal sentence should be possible with [V-en-] stems. We see in (27) that this prediction is correct, although this construction is lexically limited (see Babby 2010c: 33). The derivation of sentences like (27a-d) is schematically represented in (28). Compare (25) and (28): the only significant difference is the introduction of the tense suffix in (25): SF lexical adjectives and -en- participles express tense by means of the copula, which I argue in Babby 2011 is merged in AS not syntax.

(27) a. Ob etom vo vsex učebnikax napisano. (< napisa-(e)n-o) ‘This has been written about in all the textbooks (lit. About this (has been) written in all textbooks)’ cf. (22)
b. On dejstvoval, kak bylo obgovorenno. (< obgovori(e)n-o) ‘He acted as was agreed-upon’
c. K razbiratel’stvu bylo pristupleno nemedlenno. ‘The-investigation was undertaken immediately (lit. To the-investigation was undertaken immediately)’
d. Ob etom už e rasskazano. ‘This has already been related (lit. About this already told)’

(28) a. \{i^N\}_1 \{j^N\}_2 \{^-\}_3 \{^-V\}_4 +
b. \{^-\}_1 \{^-\}_2 \{ ^-\}_3 \{ ^-en\}_4 >
c. \{^-N\}_1 \{j^-\}_2 \{^-\}_3 \{^-[V-en]\}_4 >>
d. \{j^-N\}_1 \{^-\}_2 \{^-\}_3 \{^-[V-en]\}_4 +
e. \{^-\}_1 \{ ^-\}_2 \{ ^-\}_3 \{ ^-Oimpartial\}_4 >
f. \{^-\}_1 \{^-\}_2 \{^-\}_3 \{^-[V-en-o]\}_4 \Rightarrow napisano, etc.
The fact that the -en- suffix always derives an unaccusative AS automatically explains why impersonal transitive sentences like (29) are not possible in standard Russian:

(29) *U Salindžera_{GEN} (*Salindžerom_{INST}) bylon_{N} napisan_{N} vsego odor roman_{ACC}.
‘Only one novel was written by Salinger.’

Since -en- always derives an unaccusative [V-en-] stem in standard Russian, V’s initial j in \(j^N\) always externalizes, giving \(j^N\), which may subsequently be deleted if impersonal -o’s AS is merged (see (28a-f)). Thus -en- participles in standard Russian can never project a direct object or an agentive subject in any type of sentence, including sentences like (30a):

(30) a. V pogrebe ne bylo obnaružen_{N} [NP] nikakix dokumentov_{GEN.PL}.
‘There were no documents found in the basement’

b. *V pogrebe bylo obnaružen_{N} [NP] dokumenty_{ACC.PL}.
‘Documents were found in the basement.’

c. V pogrebe byli_{PL} obnaruženy_{PL} [NP] dokumenty_{NOM.PL}.
‘Documents were found in the basement’.

Evidence is presented in Babby 2006 and 1980 that nikakix dokumentov in (30a) is the genitive subject, not the direct object (cf. (30b-c)). It follows from this that in nonstandard (dialect) Russian, which allows transitive -en- participles, -en- has a different AS and derivation, and u + NP_{GEN} has a different function (see Timberlake 1976).

References


On the Status of Russian Perfective Passives in –sja

James S. Levine

Introduction

In an early and important study of voice in Russian, Babby and Brecht (1975) introduced significant theoretical revisions to prevailing transformational analyses of passives in Russian. In their article, B&B demonstrated the inadequacies of previous analyses, which posited a single “passive transformation” for Russian. Instead, they proposed an analysis of voice that achieved maximum generality by accounting in a principled way for the syntactic relations between active, passive, and middle sentences. In particular, they argued that passives formed with the suffix –en- must be derived differently from passives in –sja, the former resulting from a lexical process that derived –en-adjectives from perfective transitive verbs, while the latter were derived from imperfective verbs by syntactic rules.

Further, and most importantly, they showed that while both –sja and –en- could derive passives, they are also used in the derivation of middles (and other non-passive constructions), and consequently, neither could be considered a “passive morpheme.” Rather, they characterized –sja as the “voice morpheme” in Russian, and they defined voice as “the relationship between a verb’s subcategorization feature and the realization of this feature in the surface structure of the sentence” (364). In their analysis, -sja functions as a morphological indicator of a marked voice relation, for example, it signals that a verb’s subcategorized direct internal argument (direct object) appears in the surface structure as the external argument (subject).\(^1\) The theory of –sja elaborated by B&B accounts for the formation of active, middle, and passive sentences of both imperfective and perfective verbs,\(^2\) the “complete paradigms”

---

\(^1\) B&B point out that as the indicator of a marked voice relation, -sja may also signal the non-occurrence of the subject in the surface, or its appearance in a case other than the Nominative (see also Brecht and Levine 1984:124).

\(^2\) The significant generalization achieved by this analysis of –sja, along with its intuitive appeal, led to the adaptation of the theory for teaching the various –sja constructions in Russian. For example, see the treatment of –sja in Russian Stage Two (1985), a textbook authored by R. Brecht (with N. Baranova, D. Davidson, and N. Kostromina).
of which they illustrated with the examples in (1a-d) and (2a-d):³

(1) IMPERFECTIVE PARADIGM
   a. Oleg otkryval kalitku.  Active
      ‘Oleg was opening the gate.’
   b. Kalitka otkryvalas’.    Middle/Passive
      ‘The gate was opening/was being opened.’
   c. Kalitka otkryvalas’ Olegom.  Passive
      ‘The gate was being opened by Oleg.’
   d. *Kalitka byla otkryvana Olegom.  Passive

(2) PERFECTIVE PARADIGM
   a. Oleg otkryl kalitku.  Active
      ‘Oleg opened the gate.’
   b. Kalitka otkrylas’.    Middle
      ‘The gate opened.’
   c. *Kalitka otkrylas’ Olegom.  Passive
      ‘The gate was opened by Oleg.’
   d. Kalitka byla otkryta Olegom.  Passive
      ‘The gate was opened by Oleg.’

Note that in the imperfective paradigm, Kalitka otkryvalas’ (1b) shows a middle/passive ambiguity. This ambiguity, characteristic of imperfective verbs in –sja, is resolved by an agent adjunct in the Instrumental case (1c), which can occur in passives, but not in middles. In the perfective paradigm, Kalitka otkrylas’ (2b) is identified as a middle, while the passive is expressed by the short passive particle (2d) formed with the suffix –en (which may be realized as –n-, -en- or –t-). Examples (1d) and (2c) are both marked unacceptable as passives. B&B state (342): “In Contemporary Standard Russian

³ The analysis proposed in Babby and Brecht 1975 is an elaboration of Babby 1975, which treated –sja as the marker of “syntactically derived intransitivity,” as it signals that a verb’s subcategorized direct object has been removed from its initial position by a rule of preposing or deletion. During this period of generative theory, it was widely assumed that movement rules and other transformations were the only way to capture generalizations over constructions like actives, passives, and middles. In Babby’s more recent work on voice (e.g., Babby 1993), he has argued that morpholexical operations on argument structure may account for much of what in earlier generative theory was assumed to be determined syntactically. In other studies (Babby 1994, 1998), he has proposed that –sja is affixed to a verb as part of a set of morpholexical rules that alter a verb’s initial argument structure, or diathesis (See section 3). On this view, what are traditionally called active, passive, and middle “voices” are assumed to be particular diathetic alternations, which taken together, produce a given verb’s “diathetic paradigm.”
(CSR), perfective verbs do not form their passives with –sja, and imperfective verbs do not form their passives with –en,...”. The purpose of this paper is to re-examine the first part of this statement regarding the non-occurrence of perfective passives in –sja. In Section 1, I will give a brief overview of the long-standing controversy regarding the existence of passives formed from perfective verbs in –sja. In Section 2, I will review recent studies that document the use of such passives on Russian Internet sites, as well as in edited literary, scientific, and journalistic texts. I will also present examples elicited from native informants. Section 3 provides a brief summary and conclusion, which speculates on how an emergent perfective passive may be altering the aspectually determined means of expressing passive voice in Russian, and raises some implications for theoretical treatments of voice in Russian.

1. Historical Overview
The existence of perfective passives in –sja in Contemporary Russian has been a topic of long-standing debate. Some scholars take the strict view that perfective verbs cannot form passives with the affix –sja (Isačenko 1960: 362), and even exclude the possibility of a passive interpretation for perfectives in –sja (Padučeva 2001:53). Others acknowledge the existence of perfective passives in -sja but place some qualification on their occurrence. For example, Vinogradov (1972: 497) claims that such passives are possible, but only in the future tense. According to Harrison (1967: 17), the perfective form of “reflexive verbs” of certain semantic classes can express passive meaning, but unlike passive participles, these perfective passives are “not used with an animate agent” in the instrumental case. In contrast, Janko-Trinickaja (1962: 148) asserts that perfective passives in –sja may occur in both the past and future tense, and with an animate agent, but views them as limited in their usage. Similarly, Švedova (1970: 352) states that perfective verbs may form a passive in –sja, but characterizes them as “unproductive,” while Apresjan (2002: 19) asserts that “episodically” cited examples of perfective passives in –sja, such as Vse èto potom zabetoniruetsja ‘All this will later be paved with concrete’, do not extend beyond the “occasional” occurrence.

Disagreement over the existence and range of usage of perfective passives in –sja in Contemporary Russian is perhaps at least in part due to the history of this voice construction. From the late eighteenth to the first half of the nineteenth century, passives in –sja formed from perfective verbs were widely used and within the norms of

---

4 It should be noted that many of the semantic groups of verbs that Harrison characterizes as expressing perfective passives, such as verbs denoting “covering” (Ploščad’ pokrylas’ snegom ‘The square became covered in snow’), “filling” (Komnata napolnilas’ dymom ‘The room filled up with smoke’), and “illuminating” (Ulica osvetilas’ fonarjami ‘The street was illuminated by/with street lights’) are now widely assumed to derive middles, not passives. For a different opinion, see Wade 2011:332.
standard Russian.\textsuperscript{5} As Gradinarova (2009) demonstrates, V. Dal’s four-volume *Tolkovyj slovar’ živogo velikogo russkogo jazyka* (Dal’ 1956), originally published 1880-1882, provides abundant illustrative examples of “reflexive perfective passives,” many with an Instrumental agentive adjunct.\textsuperscript{6} Among the numerous examples she cites from Dal’ are: *Obvinilsja ljud’mi, da opravdalsja ppered Bogom* ‘Accused by people, but vindicated before God’; *Delo uviditsja nami, kogda uže budet pozdno* ‘The matter will be seen by us, when it will already be too late.’

Throughout this period, however, perfective passives in –*sja* were already in the process of being reinterpreted as middles (decausatives).\textsuperscript{7} As Gradinarova points out, passive meaning in this construction was weakening, especially in the past tense, and in sentences involving non-agentive causers, such as in *Tvoim učen’em nam svet otkrylsja* ‘The world (was) opened up for us through your teaching.’ Also during this period of reinterpretation, the Instrumental (causer) adjunct NP characteristic of the passive construction was being replaced by the PP *ot + NP Gen*, an adjunct found in middles: *Gajka otvinčivaetsja trjasko* (Instr)/*ot trjaski* (Gen) ‘The nut gets loose (lit. ‘unscrews’) from shaking.’ Illustrative examples in Dal’ cited by Gradinarova suggest that at this time perfective passives in –*sja* still occurred with an Instrumental NP denoting natural force, e.g. *Borota vetrom rasxlestnulis’* ‘The gates were whipped opened by the wind,’ but these sentences too were being supplanted by alternative constructions, such as the “demiactive” *Vorota raspaxnul veter* ‘The wind flung open the gates’), or “adversity impersonal” *Vorota raspaxnulo vetrom* ‘The gates got opened by the wind’). In addition, Gradinarova suggests that perfective passives in –*sja* were being forced out of use by the passive participle in –*n/-t* (*Vorota paspaxnuty vetrom* ‘The gates have been opened by the wind’) because the participles were “semantically transparent” and unambiguous.

### 2. The “Revival” of a Voice Construction

With their widespread reinterpretation as middles, and replacement by competing constructions, the perfective passives in –*sja*, once a prevalent and fully acceptable voice category, receded to the periphery. However, as noted in section 1, perfective passives in –*sja* have survived to the present day and there are indications that this voice category is in the process of a revival. In their 1975 article, Babby and Brecht generally ruled out the formation of passives from perfective verbs in –*sja*, but they did

---

\textsuperscript{5} On the passive interpretation of perfective verbs in –*sja*, and examples of their use in the first half of the 19th century, see Bulaxovskij (1954:119-20; 315).

\textsuperscript{6} The adjective “reflexive” is used here to refer to any verb formed with the morpheme –*sja*. Thus the term “reflexive perfective passive” (vozvratnyj perfektivnyj passiv) is used by Gradinarova and many others to refer to the perfective construction that is the subject of this paper.

\textsuperscript{7} A term for middle often used in the Russian grammatical tradition is “sredne-vozvratnoe značenie.” Regarding the reinterpretation of passives as middles, Gradinarova references Švedova 1964.
acknowledge that a small number of perfective verbs that take experimenter subjects (e.g. *zabyt‘to forget*, *počuvstvovat‘to feel*, *uslyšat‘to hear*) can form passives in –*sja*. The sentences in (3), cited in Jones and Levine 2010: 295 (fn. 1), are two such examples:

(3)  
(a) Èffekt ètogo lekarstva počuvstvuetsja vami srazu.  
‘The effect of this medicine will be felt by you immediately.’  
(b) Èto sobytie ne zabudetsja nami.  
‘This event will not be forgotten by us.’

In fact, perfective passives in –*sja* are not limited to the class of experimenter verbs, but occur across a range of verb classes, and show up regularly in forums on Russian Internet sites, as well as in edited scientific, literary, and journalistic texts. In sections 2.1-2.3, I will introduce a sampling of examples taken from these sources, including those adduced by scholars in recent studies, and by native informants.

2.1. Distinguishing passives and middles
As is well known, passives and middles formed with the morpheme –*sja* are intransitive constructions that are normally characterized by their relationship to a corresponding transitive with respect to both the syntactic and semantic roles of the NPs associated with the verb. Specifically, when compared with their transitive counterparts, both passives and middles share the characteristic that the direct object (patient) argument of the transitive appears as the grammatical subject of the passive or middle. The two constructions differ in the fate of the subject (typically the agent) argument of the initial transitive: in passives, the subject is either implied, or appears as an argument adjunct in the Instrumental case, while in middles the initial subject is eliminated.

(4)  
(a) Menedžer otkryvaet dveri.  
‘The manager is opening the doors.’

(b) Dveri otkryvajutsja.  
‘The doors are opening.’  
[middle reading]  
‘The doors are being opened.’  
[passive reading]

(c) Dveri otkryvajutsja menedžerom.  
‘The doors are being opened by the manager.’

As the examples in (4) indicate, when formed from an imperfective verb, the derived intransitive is ambiguous (4b), susceptible to either a middle or passive interpretation.
This ambiguity is resolved in favor of the passive by the presence of an Instrumental adjunct (4c).

When a derived intransitive in –sja is formed from a perfective verb, the normal interpretation is that of a middle, as in (5a). However, as indicated by the example in (5b) taken from a Russian Internet site, a perfective derived intransitive like otkryt’sja can occur in Contemporary Russian with passive meaning.

(5)  
a. Dver' otkrylas.’  
‘The door opened.’

b. Dver’ otkrylas’ kem-to v parallelnom mire.  
‘The door was opened by someone in a parallel world.’

The function of the middle voice, according to Babby (1993:12), is to “dissociate the agent from its initial AS [Argument Structure]” and to present the action from the point of view of the direct internal argument (direct object). Further, middles “do not assert the absence of an agent”; rather their function is to present an action without reference to the participation of an agent. Babby gives the example in (6) to illustrate his point.

(6)  
Ja staralas’ est’ konfetu pomeldennee, no ona kak-to nezametno s’elas’, očen’ bystro.  
‘I tried to eat the candy slowly, but it somehow got eaten [lit. ‘ate-up’] without my noticing, very quickly.’

Babby comments (12) that “the middle voice s’elas’ is used here rather than the active or passive to underscore the dissociation of the agent from the action of eating the candy.” While this may be an apt characterization of the middle, it appears that a middle reading of a perfective in –sja may be cancelled out in the same way as it is in imperfective verbs, by an agentive adjunct in the Instrumental case. Consider the examples in (7) - (9).

(7)  
a. Vse šokoladki bystro s’elas.’

‘All the chocolates quickly got eaten [lit. ‘ate-up’].’

b. Vse šokoladki uže s’elis’. Interesno kem.  
‘All the chocolates already got/were eaten. I wonder by whom.’

(8)  
Včera večerom byli bliny. Sdelalis’ i tut že vsej sem’ej s’elis’.  

44
‘Last night there were pancakes. They got made and were immediately eaten by the whole family.’

(9) ...ne xotelos’, čtoby vsja butylka vypilas’ by v pervyj že den’ odnim čelovekom.

‘... didn’t want a whole bottle to be drunk on the very first day by one person.’

The sentence in (7a) contains the middle s“elis’. However, in (7b), suggested by a native informant, the expected middle reading of s“elis’ in the first sentence is supplanted by a passive interpretation, due to the agentive instrumental NP in the following sentence, Interesno kem. In (8) and (9), examples taken from the Internet, the presence of the Instrumental adjunct NPs vsej sem”ej ‘by the whole family’ and odnim čelovekom ‘by one person,’ respectively, make the verbs s“elis’ and vypilas’ in these sentences unambiguous perfective passives.

2.2. Recent Studies
In the recent literature, there has been a lively discussion among Russian linguists surrounding the intransitive use (with –sja) of perfective verbs such as napisat’ ‘to write’, pročitat’/pročest’ ‘to read’, perevesti ‘to translate’, sostavit’ ‘to calculate, put together’, kupit’ ‘to buy’, počinit’ to repair’, etc., with opinions varying as to whether or not these forms have passive meaning (see Nikitina 2006). The examples cited in these studies do not occur with an Instrumental agent NP, which leaves their interpretation open to question. For example, Gavrilova (2001: 163) discusses the examples in (10)-(12), which she characterizes as constructions with “quasi-passive” meaning:

(10) Tekst uže otredaktirovalsja.
‘The text already got edited.’

(11) Stat’ja uže perevelas’.
‘The article already got translated.’

(12) Smeta skoro sostavitsja.
‘An estimate will soon calculate.’

As my glosses of these examples indicate, sentences of this type, outside of context, are likely to be interpreted as middles (Gavrilova’s quasi-passives). Gavrilova argues that sentences like (10)-(12) are widely used in spoken Russian, and she attributes the increasing use of such forms to the application of computer technology to
processes such as accounting, the editing and translating of texts, i.e., processes that had previously been performed by an agent without the assistance of technology. On the other hand, verbs such as \textit{napisat’sja, pročest’sja, počinit’sja}, etc. as used in examples (13)-(15) are characterized by Padučeva (2001: 73) as having “passive-potential” meaning, a term she attributes to V.A. Plungjan:

(13) ‘Stat’ja napisalas.’  
    ‘The article got/was written.’

(14) ‘Roman pročelsja.’  
    ‘The novel got/was read.’

(15) ‘Časy počinilis.’  
    ‘The watch got/was repaired.’

According to Padučeva, the actions denoted by verbs like those in (13)-(15) require the exertion of effort to achieve the desired result and therefore presuppose an agent. In fact, the verbs in (13)-(15), as well those in (10)-(12) characterized by Gavrilova as “quasi-passive” in meaning, can occur with an overt agent in the Instrumental case, that is, in the form of the prototypical passive. Consider the following examples, all taken from the Internet:

(16) Esli v stat’e budet zamečeno takoe, ona otredaktiruetsja perevodčikom…  
    ‘If such a thing is noticed in the article, it will be edited by the translator…’

(17) Nu, pesn’ja “My vmeste” napisalas’ mnoj v Moskve…  
    ‘Well, the song “We are together” was written by me in Moscow…’

(18) V takom vide kniga pročtetsja vzroslymi, kotorye sami vospolnjat nedostojuščee.  
    ‘In this form the book will be read by adults who will themselves fill in what is missing.’

---

8 The examples in (10)-(12), as well as Gavrilova’s (2001) characterization and explanation of the use of these forms, are found in Percov 2003:46-47; 69.

9 To illustrate her point, Padučeva (p. 73) offers the following contextualized example: \textit{Iskal vo vsex uglax – vse naprasno: ona ne otyskalas’} ‘I searched in every corner, all in vain: it didn’t turn up.’ She distinguishes agentive perfective V-\textit{sja} of this type from decausatives (e.g. \textit{Okno razbilos’} ‘The window broke’), since the latter “exclude an agent.”
Kстати, билеты уже покупаются, точнее, 2 уже точно купил мною.
‘By the way, the tickets are already being bought, in fact 2 were just bought by me.’

‘…отдай мне тачку утром в ремонт, спокойно делаем свои дела, знаем что всё подгоняется и всё исправляется хорошо механиком…
‘…I take my wheels in the morning for repairs, calmly go about my affairs, knowing that everything will be repaired and fixed by a good mechanic.’

The growing use of perfective verbs in -сja with passive meaning, like those in (7b; 8-9) and (16)-(20), have led Percov (2003: 47) to remark that “the reflexive perfective passive – moreover sometimes with an accompanying agent in the instrumental case – in recent decades has been experiencing a kind of renaissance in the Russian language.” Percov sees this construction taking hold primarily in informal, conversational style, both oral and written, as well as serving a “poetic function” in the language. He acknowledges (46) that no firm norm has been established for usage of these forms, and therefore, informants’ judgments of the acceptability of individual examples of perfective passives in -сja can diverge significantly. The examples in (21)-(24) (Percov’s examples 4-6) are among several Percov discusses, and that he personally considers fully acceptable in the contemporary spoken language.

Этот фильм раскупится фанами, да и не только, минимум в 10 миллионов копий.
‘This film will be bought up by fans, and not only [by fans], 10 million copies at a minimum.’

Хорошо ли работает ваша программа бухгалтерского учета?
– Ну как вам сказать? Обычно оценка расходов составляет ей минуты за две.
‘Does your accounting software work well?
– Well how can I put it? The usual estimate of expenses will be calculated by it in about two minutes.’

Такой ордой все это молоко выпьется за два дня.
‘By such a horde all this milk will be drunk in two days.’

Based on his observations, and polls he has conducted with several philologist-colleagues, Percov argues that reflexive perfective passives, though sometimes judged
acceptable to varying degrees by his informants, nevertheless, in his view, are an “indisputable fact” in Contemporary Russian.

Percov suggests several factors contributing to the use of the reflexive perfective passive in contemporary speech. First, he points out that there is historical precedence for this voice category, given its acceptability in the nineteenth century, which Percov considers to be well established, and amply illustrated by Janko-Trinickaja (1962). Second, he suggests a stylistic factor, namely, that in relaxed-colloquial style (“v neprinuždenno-razgovornom stile”) a more compact form is generally preferred to a more diffuse form, and this is exemplified by the synthetic vs. analytic passives, respectively. Specifically, he notes that in all forms except the masculine past tense there are 1-2 fewer syllables to pronounce in the passive derived from the finite verb in –sja, than in the periphrastic passive participle in –en, e.g., four syllables in the future tense so-stav-sja vs. six syllables in bu-det so-stav-le-na. Finally, he perceives a semantic distinction between the synthetic and analytical passives, which may be a factor contributing to the use of the former. Thus, according to Percov, in the sentences Èkzempljary raskupjatsja and Èkzempljary budut raskupleny ‘The copies will be bought up’, the situation described may be the same, but attention is focused on different aspects of the event: in the first sentence there is a focus on the “actional” aspect, on the future “buying up” of the books, while in the second sentence the focus is on the result of the action, on the state arising after the books have been bought up. Thus a speaker’s intent to depict an event from an actional (rather than resultative) point of view may be a factor motivating the use of the perfective passive in –sja.

An important study substantiating the use of the reflexive perfective passive in literary, scholarly, and journalistic texts, is the corpus investigation of Kolomatskij (2007, 2009), whose primary goal was to identify the parameters for the distribution of “passive markers” in Russian. Kolomatskij examined Russian passive forms and lexemes in nearly 6,000 sentences compiled from the databases of two corpora: the Uppsala Russian Corpus of Uppsala University in Tuebingen and the National Corpus of the Russian Language. One surprising finding of his investigation is what he characterizes (2009: 10) as the “large percentage” (velika dolja) of perfective verbs that form a passive with the reflexive affix –sja, as well as the passive participle in –en, such as the verb vyrabotat’, as in (24 a and b):

(24) a. V rezul’tate rjada èksperimentov nami vyrabotalas’ predlagaemaja
sxema dyxaniya biatlonista.
‘As a result of a series of experiments the proposed scheme of a
biathlete’s breathing was developed by us.’
b. V itoge nami byla vyrabotana «gibridnaja» metodologija ... 'In the end, a «hybrid» methodology was developed by us ...

According to Kolomatskij (2007: 19), 30% of the passive constructions in his database are formed from perfective verbs in –sja, though if one considers only those examples occurring with an agentive adjunct in the Instrumental case, the percentage falls to about 15%. These findings suggest to him that the reflexive perfective passive is undergoing a process of «rehabilitation». He points out (p. 18) (as does Percov 2003) that unlike the «statal» meaning characteristic of the participial passive, the finite passive form in –sja has an «actional» (processual) nuance, and therefore, “successfully occupies its own niche“ among the possible means of expressing passive voice in Russian.

3. Summary and Conclusions
Based on the findings of current research on the passive voice discussed in this article, it appears that the morpheme –sja may indeed be affixed to perfective verbs in Contemporary Russian to express passive meaning. Evidence of this usage derives from several sources: from forums on the Internet, from the informal, conversational speech observed and used by educated native speakers, as well as from modern Russian literary and informative texts. There are several factors that may be contributing to this development. First, as discussed in section 1, in the nineteenth century the morpheme –sja was used to form passives from both perfective and imperfective verbs. Thus Babby and Brecht (1975: 363) aptly note that “there does not appear to be any semantic motivation for the non-occurrence of perfective passives in –sja.” They point out that the cognate morphemes of Russian –sja and –en in certain other Slavic languages are used to form passives in both aspects. For example, in Bulgarian the cognates of –sja (se) and –en are used to form passives from both perfective an imperfective verbs, so that each aspect may have two passives, one in se and one in –en. By contrast, Polish –się is not used in the formation of passives in either aspect, while –en is used in both. Second, as discussed in section 2, there is the possibility, suggested by Percov, that when compared with the –en passive participle, the shorter syllable structure of perfectives in

10 Not all the examples that Kolomackij characterizes as passive may qualify as such. For instance, in his example (6), given below, the italicized phrase in the second sentence is arguably a middle, not a passive. Vpervye perešagnut vuzovskij porog 654 tysjači junоšej i devušek – stol’ko že i v prošlom gody. Kem popolnilas’ studentёskaя сем’ja? ‘For the first time 654 thousand boys and girls will step over the threshold into a higher educational institution – the same number as last year. With whom did the student family fill up (replenish)?”
11 For numerous textual examples culled from the National Corpus of the Russian Language, see Gradinarova 2008.
–sja, favors their use in colloquial speech. Finally, the perceived “actional” nuance of the perfective –sja passives, in contrast to the “resultative” meaning of –en passives, may provide a semantic rationale for their occurrence.

The emergent perfective passive in –sja poses challenges to theoretical accounts of voice phenomena in Russian, specifically, the formation of passives and middles. Based on the data presented in section 2, it is clear that the morpheme –sja may be affixed to a perfective verb in the derivation of both middle and passive sentences, resulting in a middle/passive ambiguity analogous to that manifested in imperfective verbs in –sja (as in 4b above). Consider, for example, sentences (25-27) cited by Percov (2003:48).

(25) Publika dovolo’no bystro raskupila knigu. Active
‘The public rather quickly bought up the book.’

(26) Kniga dovolo’no bystro raskupilas’. [Middle reading]
‘The book rather quickly bought up.’
‘The book was rather quickly bought up.’ [Passive reading]

(27) Kniga dovolo’no bystro raskupilas’ publikoj. Passive
‘The book was rather quickly bought up by the public.’

At the same time, perfective verbs regularly passivize via affixation of the morpheme –en. Thus, the same verb may derive two passive constructions, the passive participle in –en (28), as well as a perfective passive in –sja.

(28) Kniga byla raskuplena. Passive
‘The book was bought up.’

It is not clear how these facts are to be accounted for in a theoretical account of voice in Russian. In the theta-theoretic analysis of voice developed by Babby (e.g., 1993, 1994, 1998), middle and passive sentences are viewed as diathetic alternations of the base verb’s initial argument structure, or diathesis.¹² The key element in the formation and interpretation of middles and passives is the fate of the initial verb’s characteristic external argument. In the derivation of middles, the external argument is “dissociated”

¹² Babby and Brecht (1975:364, fn 23) note that their definition of voice, “the relationship of a verb’s subcategorization feature and the realization of this feature in the surface structure of the sentence,” is related to the concept of diathesis proposed in Xolodovič 1970 and Mel’čuk and Xolodovič 1970. This point is also made in Brecht and Levine 1986: 135 (fn. 2).
from the action by eliminating it. In contrast, in the derivation of passives, the external argument is “defocused” or deemphasized, by making it an implicit argument, which licenses an optional adjunct NP in the Instrumental case. In both middle and passive formation, the direct internal argument is externalized to subject position, and either -sja or -en is affixed to the verb stem.

In the analysis of passives proposed by B&B 1975, and more recently by Babby (1998), the aspect of the verb is assumed to determine which morpheme is affixed to the verb: imperfective verbs form a passive with –sja, while perfective verbs with –en. However, as the data discussed above indicate, this distribution does not hold for all verbs. For some perfective verbs, and apparently a growing number, it would seem necessary to stipulate that in deriving a passive, either –en or –sja, may be affixed to the verb. Such an approach assumes that –sja, and –en, already used in the derivation of middles from perfective verbs, do the same work in deriving perfective passives. Any perceived differences in interpretation, such as the “resultant state” reading of the –en passive participle, versus the “actional” reading of the –sja passive, would need to be accounted for elsewhere.

An alternative theoretical treatment would be to suppose that the two morphemes operate in different ways, giving rise to different structures, which happen to share the resultant property that the external argument is (more or less) out of the picture. This is the line of development pursued in Jones and Levine 2010. This kind of treatment holds out the possibility (certainly not exhaustively worked out in J&L) that the differing interpretive preferences, as well as the peculiarities of the differences between decausatives and middle constructions (with their apparently obligatory attendant adverbial modification) might be tied to differences in structure. Assuming these interpretive differences are real, at this point it seems safe to say that it is not clear where an account of them should be located.

---

13 B&B (1975 :356) argue that not only –sja, but also –en can be affixed to the stem of a perfective verb to form a middle sentence. They show that certain verbs, presumably a small number, can form two middle constructions, as in (i) and (ii) below (B&B’s example (29 a and b):

(i) Ruki (starika) smorščilis’ ‘The hands (of the old man) became wrinkled.’
(ii) Ruki (starika) byli smorščeny’ ‘The hands (of the old man) were wrinkled.’

As the glosses of these sentences indicate, they differ in interpretation: the middle in –sja focuses on the action, while the –en middle has a stative meaning. Note that this semantic contrast, and the affix distribution of perfective middles is parallel to that of the perfective passives, such as Kniga raskupilas’/byla raskuplena publikoj ‘The book was bought up by the public’. Thus, perfective verbs in Russian have rich derivational possibilities, being capable of deriving, via the morphemes –sja and –en, a pair of passives and a pair of middles.
References


nauk. Moscow: Moskovskij gosudarstvennyj universitet imeni M.V. Lomonosova.


The Overseas Immersion Setting as Contextual Variable in Adult SLA: Learner Behaviors Associated with Language Gain to Level-3 Proficiency in Russian

Dan E. Davidson and Maria D. Lekic

Overseas immersion study for second language (L-2) acquisition is widely regarded not only as valuable educationally, but also essential for the attainment of upper-level proficiencies. Unfortunately, the mechanisms though which overseas language and cultural immersion training may (or may not) foster successful adult acquisition are still not well understood (e.g., Rivers, 1998; Ginsburg & Miller, 2000, Kinginger, 2011). As a result, researchers across a range of disciplinary backgrounds are examining cognitive, psycholinguistic, and cultural factors affecting linguistic outcomes within the study abroad environment, including increased efforts to account for the notable variation that has been observed in actual levels of L-2 attainment by students returning from study abroad, as previously reported in Freed (2004), Davidson (2010) and Kinginger (2009).

In one of the first ethnographic studies of foreign language gain in the overseas context, Richard Brecht and Jennifer Robinson (1995) investigated the relationship between formal study, target-language utilization, and current proficiency levels of young adult learners of Russian engaged in study abroad. Relying on a large-scale macro-investigation of second language gain (Brecht, Davidson, Ginsburg, 1994), Brecht et al (1995) presented empirical data on L-2 utilization in the semester-long advanced programs in Russia for students with measured oral proficiencies at the 1+ and 2 levels. Using calendar diaries and

1 Among his wide-ranging services to the study and teaching of languages in the U.S., Dr. Richard Brecht contributed substantially over the past ten years to the conceptualization and implementation of the Language Flagship Programs, sponsored by the National Security Education Program (NSEP). He was also one of the first scholars to draw attention to the relationship between language gains and the measurement of time-on-task in the overseas context, e.g., Brecht, 1996.
time-place maps, the authors calculated actual L-2 use by adult subjects, while Freed (2004) and Magnon (2007) have reported corresponding baselines for intermediate-level learners of French, as measured for in-country summer and semester programs. Similarly, a recent study by Martinsen et al (2011) examined target language utilization patterns within a domestic immersion context, the academic-year language-specific residential houses at a major American university, in order to relate language gains, as measured by pre-/post-OPI testing, to language utilization levels in the immersion dormitory context.

Previous quantitative studies (Brecht, Davidson, Ginsburg (1995), Davidson (2007, 2010) have provided large-scale, multi-institutional analyses of language gain by American students (N = 1,881) engaged in the formal study of language and regional studies at U.S. partner universities in Russia at the intermediate and advanced levels. Subjects represented statistically robust numbers at different initial levels of proficiency engaged in typical study durations of 8-weeks (summer), 15-weeks (semester), or 30 weeks (academic year).

The present study is focused on the advanced to superior levels (and above), providing updated outcomes and correlation data for this level (the Russian Overseas Flagship Program), noted for a smaller sample in Davidson (2010), and new data on the effects of different levels and varieties of on-program language utilization on ultimate proficiency outcomes.

Given the considerable variation in post-program L-2 outcomes from immersion students across languages and regions, even for students within the same programs, with similar backgrounds, and relatively equal initial levels of language proficiency, predicted proficiency results from study abroad are normally stated in terms of ranges, defined by the duration of the immersion program, pre-program proficiency, and a relatively small number of individual linguistic and other variables (Davidson, 2007; Wilkinson, 1998). On-program language utilization (formal and informal, interpersonal, presentational, and interpretive, within in and out the classroom) is one such variable which is broadly assumed to affect learning outcomes in the study abroad context. Levels and parameters of actual L-2 usage have relatively rarely been documented and linked to proficiency attainment, particularly at the "Superior" level; the present study is intended to address this limitation.

**RESEARCH QUESTIONS**
Given the increased importance of study abroad in U.S. education and the emphasis in the work place on professional-level ("Superior"/ILR-3) competence
in world languages, the current study will examine the relationship between the amount of target-language spoken in the overseas context and the participant's ultimate success in reaching the desired goal. The following research questions have guided the research.

- What is the typical level of L-2 utilization of an adult learner engaged in an intensive/immersion year-long overseas study program at the professional level?
- Is there a relationship between levels of language utilization and ultimate proficiency outcome for students at or near Level 3?
- Are there particular contexts of formal and/or informal language utilization associated with ultimate outcome, beyond sheer time-on-task measures?

The study is limited to analysis of U.S. overseas programs focused on training to ILR Level 3 ("Superior"), the Language Flagship Programs funded by the National Security Education Program (NSEP), specifically, the Russian Overseas Flagship (ROF) at St. Petersburg State University.

**METHOD**

The present study makes use of global measurements of outcomes to compare adult learners in terms of ultimate levels of proficiency, attained in the course of a year-long intensive/immersion academic program in Russia, with systematically reported levels and varieties of informal and professional target-language utilization over the course of the same academic year. The approach, therefore, is both product- and process-focused, taking the measured ultimate proficiency outcome as the dependent variable and the levels and varieties of documented language utilization within the overseas context as independent variables. Its goal is to provide baseline information on L-2 acquisition related to successful language-acquisition behaviors, and, more specifically, to account for, if possible, some of the variation that has been observed by researchers in the measured outcomes of adult learners, who had otherwise undertaken training from comparable starting points.

The term, “outcome,” reflects standard practice in the larger SLA community to denote progress along a sequence of stages that characterize the acquisition of a second language (Ellis, 2006). Within the present study, outcome is understood to refer to measured oral, reading and listening proficiency of
participant-subjects at the beginning and end of a nine-month overseas immersion program.

Participants
Undergraduate participants in the present data set are drawn from five designated undergraduate Flagship programs (Bryn Mawr / Haverford College, Portland State, University of Maryland, University of Wisconsin, and UCLA). In addition, at-large students who hold a B.A. from any U.S. institution and who meet the uniform admissions criteria may also apply for the Flagship program. Admission to the Russian Overseas Flagship (ROF) is competitive and based on the prospective participant’s ability to demonstrate ILR-Level 2 ("Advanced") proficiency or higher in speaking and at least one other skill at the time of the application in January of each year. Selection committees consider academic background, faculty recommendations, seriousness of purpose, motivation for the course of study, and successful prior participation in an academic program in Russia in accepting both domestic and at-large students into the Russian Overseas Flagship (ROF). A participant’s ability to pay for the program is not a criterion for participation. Therefore, the findings reported in the current study may be generalized for other overseas programs at this level (Flagship and non-Flagship) of similar design to the extent that they place comparable constraints on the applicant pool.

Table 1.

<table>
<thead>
<tr>
<th>Demographic Data on ROF Subject Population</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
</tr>
<tr>
<td>• 66 Russian language learners</td>
</tr>
<tr>
<td>• Representatives of 43 U.S. universities and colleges</td>
</tr>
<tr>
<td><strong>Age</strong></td>
</tr>
<tr>
<td>• 65% of students were 22-24 years old</td>
</tr>
<tr>
<td>• Range: 21-36</td>
</tr>
<tr>
<td>• Average: 24</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td>• 34 (52%) male students</td>
</tr>
<tr>
<td>• 32 (48%) female students</td>
</tr>
</tbody>
</table>
Table 2.

<table>
<thead>
<tr>
<th>Language Demographics of Student Population</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of students that studied Russian in High School</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Number of students that studied Russian in College</td>
<td>64</td>
<td>97</td>
</tr>
<tr>
<td>Number of Students with knowledge of other Slavic languages</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>Number of students with knowledge of 1 or more other Slavic languages</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Number of students with knowledge of 1 or more other non-Slavic languages</td>
<td>52</td>
<td>79</td>
</tr>
</tbody>
</table>

Main Components of the Russian Overseas Flagship Program

The subject group comprises 7 consecutive classes of ROF students (N = 66), whose program of study at St. Petersburg State University (SPBSU) has been developed in cooperation with American Councils/ACTR and utilizes a dedicated curriculum focused on the integration of speaking, listening, reading, writing, and cultural modalities at the professional level, ACTFL “Superior,” ILR - 3. The structure of the Flagship Program can be seen in the outline below, which includes formal and informal language training and utilization activities.

- Direct enrollment subject course for credit at SPBSU (selected by the student)
- Core language/culture study cluster
- Language course work in small groups
- Independent research projects
- Individual language tutors
- Internships placements requiring presentational work and weekly reports
- Discussion groups
- Homestays

---

• Integrated cultural program (bi-weekly, tied to thematic units of the Flagship language courses)
• On-going evaluation (testing, site visits, teacher/tutor reports, portfolio development, self-evaluation)
• Bi-Weekly Language Utilization Reports (LUR) (time-place mapping and self-management template)

Outcomes Measures
Multiple outcomes measures have been used, given the need to establish generally recognized baselines for the Flagship program, both in the US and in Russia. The multiple measures include pre- and post-program Oral Proficiency Interviews (OPI), pre- and post-program American Councils Standardized Reading and Listening Proficiency Tests, based on the ILR proficiency scale, as well as the Test of Russian as a Foreign Language (TORFL), based on the Common European Framework and serving as the official state certified examination of Russian for foreigners in the Russian Federation. Pre- and post-program test results using the TORFL Certification Level -3 (corresponding to CEF Threshold Level C-1) and Certification Level – 4 (corresponding to CEF Threshold Level C-2) are juxtaposed below with ACTFL and ILR “Superior (ILR-3) and “Distinguished” (ILR-4) for the same population of Russian Flagship participants3.

The Language Utilization Reporting (LUR) System
As part of a strong emphasis on self-managed learning and development of executive function in the advanced learner, each participant in the Flagship programs administered by American Councils in Russia, Central Asia, the Middle East, China, and Africa, is required to complete a bi-weekly language utilization report (LUR), which provides both an accounting of how the students spend time using the language in and outside of class, as well as an opportunity to analyze and reflect on successful and unsuccessful speech situations in which they have found themselves during the reporting period. The LUR, therefore, serves both as a diary for documenting language use, as well as a vehicle for personal reflection and communication with an academic advisor (in-country or

---

3 The “TORFL” (Test of Russian as a Foreign Language), a standardized, competency-based battery, based primarily on the Common European Framework (CEF) for Foreign Languages, is available at the “elementary,” “basic” (bazovyj), and certifications levels 1 through 4. For detailed description and sample tests for Certification Level 3, see Ivanova (2009), Andrjushina (2010).
at-home) for the purpose of goal setting for the next reporting period. (See Appendix 1 for the complete LUR template.)

The LUR is a monitored self-reporting online survey instrument, for which students are provided explicit deadlines and instructions to ensure comparability of responses. LUR’s are not submitted by students during breaks or for a week that is interrupted by travel outside the host language area. The same reporting schedule has been observed on the LURs over the seven consecutive years represented in the present report, beginning from September 2004 through May of 2011.

Target language activities are listed on the LUR using 13 rubrics, one of which is “other.” While all students within the Flagship program devote similar amounts of time to formal classroom activity, peer tutors, and internships, these activities are nonetheless listed on the bi-weekly report, as they constitute significant target-language performance areas which become part of the overall picture of each student’s Russian language use while on-program. The full set of rubrics is as follows:

1. Homework 8. Public Transportation
2. Internships 9. Other activities
3. Tutoring 10. Reading for pleasure
4. Academic Reading 11. Reading the press
5. Cultural Events 12. Public transportation and shopping
6. With friends 13. Local TV or radio
7. In class

Classroom attendance is required of all ROF students and monitored carefully by on-site resident directors. Quantitative data, based on the American Councils LUR provide the principal focus of the analysis that follows; sample narrative material from open-ended responses have also been included here to assist the reader in better understanding the language utilization behaviors reported here.

**ANALYSIS**

Analysis has proceeded along two lines: a) review of L-2 proficiency outcomes of the Russian Overseas Flagship for the current subject group in order to confirm post-program outcomes; b) presentation of data relating outcomes to overall levels and types of L-2 utilization activities of participants over the same period.
Learner outcome variables
The proficiency-based outcomes report of the ROF to date can be viewed in the following tables, which compare pre- and post-Flagship program test performances across modalities. Tables 3 and Figures 1-7 below update the Davidson 2010 report on ROF proficiency-based outcomes through the inclusion of two additional years of Flagship Program data.

Table 3.

Comparison of Pre- and Post-Program OPI Scores: 2004-2010 ROF participants

<table>
<thead>
<tr>
<th>Prof. Level</th>
<th>Pre-Program</th>
<th>Post-Program</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of Students</td>
<td>% of Students</td>
</tr>
<tr>
<td>1+</td>
<td>5</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>53</td>
</tr>
<tr>
<td>2+</td>
<td>24</td>
<td>36.4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>66</td>
<td>100</td>
</tr>
</tbody>
</table>

Note: ALL OPI testing was performed by ACTFL-certified or FSI certified testers.

Over the past seven years, approximately 90.01% of participants in the ROF have achieved the stated program goal of ILR Level – 3 (ACTFL “Superior”) proficiency (or higher) and a TORFL – “Level 3” (or higher). While the overall sample size of the ROF program is still small, the consistency of the results from two entirely independent but parallel systems of performance-based testing is notable.

Predicting L-2 gain in the Flagship environment
Previous studies have identified certain linguistic correlates of L-2 language gain for US students of Russian which have remained reasonably consistent predictors over the past 15 years for American students of Russian participating in any of the American Councils/ACTR programs (Brecht, Davidson, Ginsburg, 1994, 1995; Davidson, 2007, 2010). For example, pre-program reading skills and
pre-program control of language structure have served as good predictors of ultimate oral proficiency gain for adult-learner study abroad at various levels⁴.

FIGURE 1

**Comparison of Pre- and Post-Program OPI Scores: Russian Overseas Flagship Program, 2004-2010 (N = 66)**

The gender effect was not significant in the 2010 study as well.

⁴ The gender effect was not significant in the 2010 study as well.
FIGURE 3

Russian Overseas Flagship Pre- and Post-program Listening Scores, 2009-10/2010-11 (N=19)

FIGURE 4

Pre- and Post-Program TORFL Scores: Speaking (N = 66)
FIGURE 5

Pre- and Post-Program TORFL Scores:
Reading (N = 66)

FIGURE 6

Pre- and Post-Program TORFL Scores:
Writing (N = 66)
### Table 4.
Summary of Multiple Regression Analyses for Variables Predicting Gain in Oral Proficiency: ROF (N = 66)

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>2.30</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>Pre-Program Speaking Proficiency Level</td>
<td>-0.76</td>
<td>0.11</td>
<td>-0.76*</td>
</tr>
<tr>
<td>Pre-Program Reading Proficiency Level</td>
<td>-0.11</td>
<td>0.11</td>
<td>-0.09</td>
</tr>
<tr>
<td>Pre-Program Listening Proficiency Level</td>
<td>0.54</td>
<td>0.14</td>
<td>0.48*</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.53</td>
<td>0.34</td>
<td>-0.13</td>
</tr>
<tr>
<td>High School Russian</td>
<td>-0.38</td>
<td>0.44</td>
<td>-0.08</td>
</tr>
<tr>
<td>Knowledge of Other Slavic Language</td>
<td>0.52</td>
<td>0.49</td>
<td>0.09</td>
</tr>
<tr>
<td>Knowledge of Non-Slavic Language</td>
<td>-0.37</td>
<td>0.42</td>
<td>-0.07</td>
</tr>
<tr>
<td>Grammar Score</td>
<td>0.08</td>
<td>0.02</td>
<td>0.53*</td>
</tr>
</tbody>
</table>

**Notes:** \( R^2 = .50 \), adjusted \( R^2 = .45 \), \(^* p < .001\)
While the inter-correlation of reading and grammar skills is well documented in the literature on language gain, the emergence of listening comprehension as a predictor of oral proficiency gain at the professional level was first documented in Davidson (2010). The next section examines turns to certain non-linguistic variables that, like the above-identified variables, also appear connected to positive L-2 development in the study abroad context.

**Target language activities**
The full cohort of 66 Flagship participants submitted on average 16 bi-weekly reports per year describing a total of 64,355.56 hours of target language use, distributed across the above ten activity rubrics. (See Appendix 2: LUR Descriptive Statistics)

Overall, L-2 language utilization increased for the group from the early weeks of the fall semester into the spring semester. No significant differences in utilization patterns by gender were observed, nor were there significant recorded differences in language utilization patterns between fall and spring semesters, other than that accountable in terms of the substitution of a direct-enrollment content course in the fall with a day-long internship in the spring. Finally, the year of program participation had no significant or observable effect on the language utilization pattern of the participants.

Noteworthy differences in language utilization were seen to distinguish three categories of learners within the overall group: those who completed the program with oral proficiency scores of 2 or 2+ (low gainers), those who completed the program at OPI Level 3 (gainers), and those who completed the program at OPI Level 3+ or higher (high gainers). Over the course of the academic year, the mean total hours per week of using Russian for the low-gain group was 50.45 (S.D. 14.31), the mean for the Level 3 group was 65.10 (S.D. 20.89), and the mean for the high-gain group was 68.57 (S. D. 18.88). In short, the amount of weekly target-language utilization differentiated the low-gain group from the gain/high gain cohort to a significant degree, and, to a lesser extent, the gainers from the high gainers, as well.

Turning now to the report of specific activities, summarized below, it is clear that participants completing the program at Levels 3 and Level 3+ invested their time somewhat differently than those who did not, keeping in mind that all students entered the program at 2 or 2+ in September.

Fixed components of the program, such as tutoring, internships, class-room related academic reading, and cultural excursions generally show similar time commitments by Flagship students over the seven years of reporting, regardless
of their ultimate attainment in the program. Discretionary activities, by contrast, varied considerably, such as the amount of time spent by students on homework, in discussions with the homestay family, in reading for pleasure, following the press, including local TV or radio, and passing time with friends, varied among the groups, some cases significantly, as the next charts makes clear.

Table 5.
Mean Weekly Time-on-Tasks (in Hours) by Activity Type

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework</td>
<td>0.9</td>
<td>4.9</td>
<td>7.9</td>
</tr>
<tr>
<td>Internship</td>
<td>6.6</td>
<td>6.2</td>
<td>4.8</td>
</tr>
<tr>
<td>Tutoring</td>
<td>3.1</td>
<td>3.1</td>
<td>3.0</td>
</tr>
<tr>
<td>Academic Reading</td>
<td>4.7</td>
<td>2.0</td>
<td>2.4</td>
</tr>
<tr>
<td>Cultural Events</td>
<td>1.2</td>
<td>2.3</td>
<td>2.7</td>
</tr>
<tr>
<td>Host Family</td>
<td>5.3</td>
<td>6.8</td>
<td>8.8</td>
</tr>
<tr>
<td>Reading for Pleasure</td>
<td>0.7</td>
<td>3.7</td>
<td>4.2</td>
</tr>
<tr>
<td>Following the Press</td>
<td>2.1</td>
<td>2.6</td>
<td>3.0</td>
</tr>
<tr>
<td>Local TV/Radio</td>
<td>3.3</td>
<td>5.3</td>
<td>5.0</td>
</tr>
<tr>
<td>Time Spent with Friends</td>
<td>6.9</td>
<td>11.2</td>
<td>10.9</td>
</tr>
</tbody>
</table>

Table 6.
Percentage of Time Spent on Various Activities

<table>
<thead>
<tr>
<th>Activity</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friends</td>
<td>15%</td>
</tr>
<tr>
<td>Host Family</td>
<td>13% (3+), 9.5(3), 8.6(2+)</td>
</tr>
<tr>
<td>Academic</td>
<td>40%</td>
</tr>
<tr>
<td>Reading</td>
<td>14% (3+), 11%(2+/3)</td>
</tr>
<tr>
<td>Cultural events</td>
<td>4%(3+), 3.1% (3), 1.7% (2+)</td>
</tr>
<tr>
<td>Internships</td>
<td>10%</td>
</tr>
<tr>
<td>Other</td>
<td>4%</td>
</tr>
</tbody>
</table>
Host family interactions appear to generate a regular pattern of language utilization situations reflecting the basic functional demands and pressures of daily family life, where the expression of greetings, gratitude, agreement/disagreement, paying and responding to compliments, requests, apologies, consolation, congratulations, and leave taking, become not only routinized, but also necessarily adapted and expanded to accommodate a range of socio-linguistic situations arising in the course of a year in the life of a well educated, professional family in St. Petersburg. One student described her language situation early on in the Flagship program in terms of the kinds of tasks and social situations that arise routinely in the home setting:

“When we started making the pie, I realized how little kitchen vocabulary I know. It was more like a game of charades at first, but K. [the host sister] kept giving me the necessary vocabulary, and listened as I repeated instructions multiple times. I don’t see this as a particular deficiency in my language, but the best way to minimize the charades would be to cook as often as possible with Russians.”

LUR Russian Flagship Report. (October)

Student reflection and self-management strategies are in frequent evidence in the LUR’s, as vocabulary diversity and premium on extended discourse are noted by the student, as she compares speech performance in the spring with a similar situation in the fall.

“Today I was talking to my host mom about her Easter cake and I felt like my command of the language was reflected in my ability to produce a diverse range of compliments about her cooking. I got a dialog going with her, all based on compliments on the cake, whereas in September I would have just said the case was “tasty”, and left it at that.”

LUR Russian Flagship Report (April)

The student’s ability to self-assess with respect to her overall ability to express a range of opinions and views resulting from her interactions with the host family is further evidence of her management of her L-2 learning, in this case in the homestay situation.

A male Flagship student reports on the importance of staying attuned to the life of his host family, for improving his language and cultural understanding, not relying entirely on media and outside sources for this important form of linguistic and cultural input.

“It seems to me that the only approach here is to be as involved as possible in the everyday life of my host family, which I have already been trying to do. The main
thing to remember is taking an interest in the life in front of my eyes here at home."


Reading proficiency also affects, positively or negatively, the pace of student progress at Level 3, as Table 5 indicates. Lower proficiency readers require more time to complete academic reading tasks, and, therefore, appear to have less time for reading for pleasure.

**Language Utilization and Program-Final Proficiency Attainment**

While the above utilization reports are of inherent interest for program designers and Flagship advisors, not all rise to the level of statistical significance. For that reason, a number of standard statistical procedures were performed on the LUR data. Regarding the overall level of L-2 usage, the greater the mean number of hours per week spent using the target language (regardless of the type of activity), the more likely the learner was to attain Level 3 proficiency or higher, \( r (63) = .25, p < .05 \).

While the overall number of classroom hours and tutorials was roughly the same for all participants, investment of time in academic activities, i.e. attending language classes and tutorial sessions and completing homework assignments, was positively related to ultimate gain in the program, \( r (63) = .31, p < .05 \). It should be noted that homework assignments within the Russian educational system are modest by US Standards, 1-2 hours per day on average. The data shows, however, that an investment of two hours rather than one hour per day on homework paid considerable dividends for the learners, \( r (63) = .36, p < .01 \). Likewise, spending more time with host family was positively associated with higher levels of language skills, \( r (62) = .26, p < .05 \).

**CONCLUSIONS**

The findings presented in the study underscore the importance of overall levels of L-2 utilization in the context of a nine-month structured Immersion training and professional internship program. Both overall utilization and, in particular, certain “high-yield” target-language (academic and informal) activities were associated with post-program proficiency attainment in Russian at ILR Levels 3, 3+, and 4 across modalities. By comparison, lower overall levels and smaller investments by students in key activities were associated across the seven generations of students covered by this study with lower post-program proficiency outcomes.
Regarding the first two research questions, the typical level of L-2 utilization for an adult learner engaged in an intensive/immersion study program at the professional level, such as the ROF, is documented over several years of Flagship students (N=66) at 50.45 hours per week for those who do not reach 3-level proficiency, 65.10 hours per week for those who do reach the professional level, and slightly higher still (68.57 hours per week) (for those exceed 3-level proficiency). The relationship between the level of language utilization and attainment of level-3 proficiency in the course of the 9-month program is noteworthy and statistically significant.

By comparison, typical utilization patterns for students engaged in lower levels of study over the same reporting period have been observed elsewhere at 30-40 hours per week for students training to Level 2, and 15-20 hours per week for students training at Level 1 (Davidson, forthcoming).

With respect to the third research question, L-2 gain from L-2 utilization activities emerges in the current body of LUR data as statistically significant. In particular, the study indicates that while participants’ discretionary L-2 usage over the course of the program varied considerably, those activities involving a more sustained engagement in homestay interactions, academic reading at Level 3, preparing homework assignments, professional internships, and following the press were associated with success in attaining the highest levels of proficiency.

Further research is needed using more ethnographically-oriented tools to explore the quality and quantity of individual L-2 language production in exchanges with native interlocutors that comprise the observed “high-yield” activity types noted above. For example, attention has been focused recently on the nature and effect of direct and indirect error correction feedback by native speakers and teachers (See SSLA, 2010), frequency of self-correction (Golonka, 2000), and the impact of study abroad on the modeling of written and oral L-2 socio-pragmatic skills (Shardakova, 2006; Moskala-Gallaher, 2011). L-2 listening comprehension (noted above) should also be considered in analyzing the role of corrective feedback, modeling, and the selection of communicative strategies at this level. Insufficient L-2 listening comprehension can seriously hamper learner uptake of culturally situated speech models and many forms of corrective feedback essential for student progress at this level.

The homestay environment, when functioning well, is a “high-yield” SLA environment for supporting, in a highly contextualized way, the development of a broad range of functional and socio-pragmatic skills in response to the normal round of family life. Some communication is routine and rapidly becomes automatic for the student, but unexpected communicative and cultural
challenges can also arise, as, for example, in the following student report concerning the unexpected death of a host family relative and the student’s efforts to console them:

“What do I do, keep silent? No. Pile on the optimism that made me famous? Not so much.... I understand that these types of comments are not what is expected by a Russian. So I listened. I found that they responded very well to my encouraging them that they had done the right thing, everything possible... It is difficult to console anyone in any language, to be sure, but when so many cultural factors affect the way an individual responds to death, an ordinary hill becomes the Cliffs of Insanity—without the rope.”

LUR Russian Flagship Report (January)

The student’s reflection on strategy selection as well as on the underlying cultural differences that inform the proper expression of consolation and condolences are notable. Because this particular set of interactions played out in the homestay situation over several days, the learner was able to reflect on the situation and on her own position within the family, and evaluate the effectiveness of her speech efforts.

The homestay environment has the capacity to support the emergence of a broad range of functional and socio-pragmatic skills in the learner in response to the pressure of daily communicative needs that arise there. As Kinginger (2011) has noted, however, the homestay environment is successful only to the extent that it fosters observation, participation, and reflection opportunities for the learner. The student comments included in the present report provide examples of all three, in which the visiting student appears to have found effective levels of inclusion with the host family to enhance and support the improvement of the learner’s socio-pragmatic skills as well as broader linguistic and cultural growth.

By contrast, students who spend too much time hanging out with friends, a behavior that undoubtedly benefitted their language development from Level 1 to Level 2, will discover that informal conversation cannot replace the professional language one requires in academic work, internships, or closely tracking professional media sources in one’s field.

The overseas immersion context, represented in this case by the Overseas Russian Flagship program, is best regarded as first and foremost an academic context for learning and knowledge generation, in which good study habits, sound self-management skills, pursuit of related intellectual and professional interests, and the willingness of both visitors and hosts to take advantage of a range of opportunities for socialization within the target culture (host family,
professional circles, cultural groups, as well as informal friendships) are most likely to help the learner attain the desired program results. Students acculturation to the academic environment, home stay environment and internship environment, to observe it, participate in it, and reflect on the language learning process are critical for the successful immersion learning at this level.

As Pellagrino-Avani (2004) has shown, adult learners undertake study abroad with expectations, images of self, and models of the host culture which invariably change over time. Managing the image of self and the need for external validation of one’s position and worth is a challenge for most young adults, as they carry on with their study and lives overseas:

“And then we talked about her boyfriend and what’s wrong with him. Maybe this is just a normal conversation for anyone to have, I only had to ask one word, and I was pretty mistake-free, not a big deal, except that I was trying to be convincing and be really supportive. She [the host family sister] thanked me for all the great advice and support and [said] she felt much better having talked to me about it. Among my friends at home I often play this role, and to have been unable to play this part until now is to have effectively become another person or, at the very least, to have not been fully me. It’s like the American-me and the Russian-me are beginning to mesh better. I like that. I like that a lot. Being able to be (almost) completely yourself in a foreign language in a foreign land makes one feel a whole lot less like a monkey on a tricycle.”

LUR Russian Flagship Report (February)

The preceding LUR narrative depicts not only careful self-monitoring and use of direct feedback, but a remarkable appreciation of what Pellagrino-Avani has termed “identity competence.”

The present study has examined the manifestations of some of these factors in the form of documented time-on-task, participant investment in specific categories of activities, and narrative comments supplied by participants at the time of their filing of the bi-weekly LUR’s. If certain structural and cognitive factors have been shown over the past 15 years to correlate with learning success in the overseas environment, the present study provides initial evidence of the role of on-program behaviors and activity choices for ultimate proficiency attainment in this critical form of language acquisition.
Appendix 1  Language Utilization Report (Specimen)

Student Name
Started: Not yet started — Completed: Not yet submitted.
Due: Sep 16

I. Language Use
How many clock hours during the past 7 days did you spend using Russian in the following activities (remember that your individual language classes last for 90 minutes each). Please use decimal-point notation, i.e. 1 hr. — ”1”; 45 min. — ”0.75”; 30 min. — ”0.5”; 20 min. — ”0.30”; 15 min. = ”0.25”; 10 min. — ”0.15”. Please only record the activities for ONE week. Do not Add the time for the two week period of the report.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thur</th>
<th>Fri</th>
<th>Sat</th>
<th>Sun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formal language learning classes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language tutorial sessions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internship or specialization coursework</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In public transportation or while shopping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With friends</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cultural events</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Russian radio or television</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading the press</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional/or academic reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading for pleasure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In homework and other preparation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>for formal language learning classes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Please list other activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
II. Communication Challenges
A. What challenges did you encounter during the past week in operating in Russian?
(Consider interpersonal, interpretive, and presentational modes in responding, e.g., scanning detailed text for relevant information, understanding detailed written instructions or spoken advice, note taking, letter writing, holding the floor in a discussion, reacting appropriately to a culturally "sensitive" situation, etc.) Describe the situation briefly.

B. In view of the challenges described above, what language/cultural resources will you need to master or have greater control over to deal more effectively, when a similar situation arises in the future?

III. Communication Successes
Describe a linguistic interaction in Russian during the past week which might be seen to characterize your command of Russian at its best.
(This may be an interaction of any kind which you consider you handled well.)

IV. Goals for the Month(s) Ahead
Identify 2-3 thematic areas, language functions, or speech genres which you plan in the coming weeks to make stronger in your own Russian.
(E.g., getting better at making requests or offering apologies, understanding certain recurrent colloquialisms in the speech of friends, catching references to contemporary Russian culture and/or politics, mastering abstract expressions, learning how to keep up a conversation, improving phonetics/intonation, becoming more proficient at offering toasts, responding to compliments, invitations, leading a discussion, chairing a meeting, etc.)

V. Additional Comments
Please note any other observations that come to mind about your use of Russian during the past week?

Appendix 2  Language Utilization Reporting 2004-2011: Descriptive Statistics

<table>
<thead>
<tr>
<th>Activity</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Sum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of forms completed</td>
<td>65</td>
<td>7.00</td>
<td>32.00</td>
<td>1043.00</td>
<td>16.05</td>
<td>6.77</td>
</tr>
<tr>
<td>Formal language learning classes</td>
<td>65</td>
<td>105.00</td>
<td>265.50</td>
<td>11686.65</td>
<td>179.79</td>
<td>35.85</td>
</tr>
<tr>
<td>In homework and other preparation for formal</td>
<td>65</td>
<td>0.00</td>
<td>410.50</td>
<td>6277.70</td>
<td>96.58</td>
<td>94.20</td>
</tr>
<tr>
<td>Activity</td>
<td>Hours</td>
<td>Rate</td>
<td>Total</td>
<td>Time</td>
<td>Average</td>
<td>Percentage</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td>----------</td>
<td>------------</td>
</tr>
<tr>
<td>Internship or specialization coursework</td>
<td>65</td>
<td>9.00</td>
<td>225.00</td>
<td>5218.42</td>
<td>80.28</td>
<td>40.83</td>
</tr>
<tr>
<td>Language tutorial sessions</td>
<td>65</td>
<td>7.00</td>
<td>113.90</td>
<td>2994.50</td>
<td>46.07</td>
<td>14.97</td>
</tr>
<tr>
<td>Professional or academic reading</td>
<td>65</td>
<td>0.00</td>
<td>161.00</td>
<td>2212.60</td>
<td>34.04</td>
<td>37.26</td>
</tr>
<tr>
<td>Cultural events</td>
<td>65</td>
<td>0.00</td>
<td>104.50</td>
<td>2304.40</td>
<td>35.45</td>
<td>24.32</td>
</tr>
<tr>
<td>Host family</td>
<td>65</td>
<td>0.00</td>
<td>345.25</td>
<td>7514.68</td>
<td>115.61</td>
<td>83.59</td>
</tr>
<tr>
<td>In public transportation or while shopping</td>
<td>65</td>
<td>0.00</td>
<td>175.90</td>
<td>2980.57</td>
<td>45.85</td>
<td>44.53</td>
</tr>
<tr>
<td>Other</td>
<td>65</td>
<td>0.00</td>
<td>116.50</td>
<td>1427.10</td>
<td>21.96</td>
<td>28.59</td>
</tr>
<tr>
<td>Reading for pleasure</td>
<td>65</td>
<td>0.00</td>
<td>161.10</td>
<td>3623.45</td>
<td>55.75</td>
<td>38.66</td>
</tr>
<tr>
<td>Reading the press</td>
<td>65</td>
<td>0.00</td>
<td>114.00</td>
<td>2535.69</td>
<td>39.01</td>
<td>26.67</td>
</tr>
<tr>
<td>Local radio or television</td>
<td>65</td>
<td>5.30</td>
<td>220.00</td>
<td>4798.75</td>
<td>73.83</td>
<td>43.65</td>
</tr>
<tr>
<td>Time spent with friends</td>
<td>65</td>
<td>3.00</td>
<td>628.50</td>
<td>10761.05</td>
<td>165.55</td>
<td>107.52</td>
</tr>
<tr>
<td>Total hours in all activities</td>
<td>65</td>
<td>368.80</td>
<td>2140.60</td>
<td>64335.56</td>
<td>989.78</td>
<td>349.55</td>
</tr>
<tr>
<td>Total hours spent on academic activities</td>
<td>65</td>
<td>147.30</td>
<td>720.50</td>
<td>23171.45</td>
<td>356.48</td>
<td>143.31</td>
</tr>
<tr>
<td>(language class, homework, tutorial, and</td>
<td>65</td>
<td>130.00</td>
<td>696.00</td>
<td>20958.85</td>
<td>322.44</td>
<td>129.25</td>
</tr>
<tr>
<td>academic read)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total hours spent on three academic activities</td>
<td>65</td>
<td>130.00</td>
<td>696.00</td>
<td>20958.85</td>
<td>322.44</td>
<td>129.25</td>
</tr>
<tr>
<td>(language class, homework and tutorial</td>
<td>55</td>
<td>130.00</td>
<td>696.00</td>
<td>20958.85</td>
<td>322.44</td>
<td>129.25</td>
</tr>
<tr>
<td>sessions)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total hours spent on reading activities</td>
<td>65</td>
<td>7.00</td>
<td>301.00</td>
<td>8371.74</td>
<td>128.80</td>
<td>71.50</td>
</tr>
<tr>
<td>(professional/academic reading, reading for</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pleasure, and reading the press)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**References**


Golonka, Eva M. (2000). Identification of salient linguistic and metalinguistic variables in the prediction of oral proficiency gain at the advanced–level


Adult Learners’ Perspectives on the Acquisition of L2 Russian Pragmatic Competence

Victor Frank

Introduction
Decades of research have highlighted the central role that instruction, interaction, and feedback play in the acquisition of a foreign or second language, whether in the classroom or in naturalistic contexts, particularly in the domain of interlanguage grammar. The role that these factors play in the acquisition of interlanguage pragmatics has come under rigorous investigation only in the last decade (Barron 2002). In this article, I will discuss the degree to which adult learners of Russian acquire native-like pragmatic competence, and present their own unique perspectives on its acquisition, both in the domestic classroom and during study abroad.

Background
The present article stems from the American Council of Teachers of Russian/National Foreign Language Center studies of gains in L2 Russian proficiency during study abroad in Russia and the former Soviet Union (Brecht, Davidson, and Ginsberg 1993; Brecht and Robinson 1993). These studies, using L2 proficiency scores, language learning histories and demographic data, time/place maps, and interview data from over 600 study-abroad sojourners, established the key predictors of in-country proficiency gain and the optimal conditions under which such gain occurs. For example, multiple regression analyses highlighted the significant role that the learner’s L2 grammar knowledge plays in listening, reading, and speaking proficiency gains (Brecht, Davidson, and Ginsberg 1993). Additionally, Brecht and Robinson (1993) determined through analyses of the time/place data concerning language use that the learners who made the greatest proficiency gains were those who spent the most out-of-class time speaking Russian with close and sympathetic native-speaker interlocutors; for example, caretakers, such as boyfriends, girlfriends, or host family members. This last finding in particular highlighted the need for
future studies of how contextual factors of linguistic interaction (for example, social distance) influence L2 acquisition in naturalistic contexts.

Following the initial studies of listening, reading, and speaking proficiency gain, Frank (2002, 2009), Owen (2002), and Shardakova (2007) focused more narrowly on the acquisition of target language pragmatic competence by L2 Russian learners by investigating their production of target language speech acts. Frank (2002, 2009) analyzed L2 and native Russian requests, using a triangulation of data collection methods: (a) the collection and analysis of speakers’ performance of the speech act of request in role-play situations, varying by contextual factors (social distance, relative power, and degree of imposition); (b) the examination of non-native speakers’ individual characteristics potentially related to differences in speech act performance, including personal background and target language proficiency data; and (c) the collection and analysis of retrospective interviews with non-native speaker participants concerning their planning, performing, and assessment of the role-plays. The analyses drew a picture of the non-native speakers’ state of second-language pragmatic competence, and shed light on the factors associated with the development of this competence.

Frank (2002, 2009) found a few weak between-group differences by contextual factors on the levels of text size and request move choice. However, large between-group differences were found on the level of sentence-level internal modifications, especially concerning the use of grammatical voice and 2nd person (interlocutor) reference. Significant between-group differences were also found in minimal social distance situations where speakers made requests of friends or peers; in these situations, L2 Russian speakers inserted several justifying statements before the request compared to native speakers, who were much more direct. Additionally, a strong within-group correlation was found between the learners’ grammatical competence and their use of pragmatically significant sentence-level morphosyntactic features. In other words, a non-native

---

1 Other, related studies of L2 Russian learning during study abroad investigated acquisition of verbal aspect (Boots-Ebenfield 1995) and metalinguistic processes that predict proficiency gain (Golonka 2000, 2006).
2 For definitions of the contextual factors, derived from Brown and Levinson (1987) and Hudson, Detmer, et al. (1995), see Appendix A.
3 While the triangulation of data cannot pretend to present an objectively true picture of the situation addressed by the research (due at least in part to the multiple artifacts of the data), the use of multiple sources of data may strengthen the validity of certain interpretations by including the participants’ perspectives (Silverman 1993).
speaker’s grammar score well-predicted the proportion of that individual’s moves marked for interlocutor reference; the higher the grammar score, the greater proportion of moves including an interlocutor reference, \( r_s = .52 \) (\( n = 23 \), \( p < .05 \)).

The present article discusses the findings from the retrospective interviews conducted with in-country learners after performing the request role-plays. The retrospective interview protocol is included in Appendix B. The interviews were partially transcribed and analyzed qualitatively to reveal dominant themes.

**Findings**

The themes from the interviews identified in the qualitative analyses relate to some of the most basic developmental issues in the interlanguage pragmatics research agenda (Kasper and Schmidt 1996): planning processes and performance; the inter-relationship of interlanguage grammatical and pragmatic competences; the role of personality on developing pragmatic competence; and the influence of motivation and attitudes.

**Planning**

In the retrospective interviews, many learners were acutely aware that their current inventory of both grammatical and lexical L2 forms was inadequate to encode their intended pragmatic meanings. This awareness had a certain impact on planning for the role-play requests, whether planning in English or Russian. As Participant 016 put it: “I know what words I know, I go with the words I know, but I figure it out, for sure, I figure out what I know in English.” Participant 011 faced similar limitations but planned in Russian: “…it’s harder to try to plan something in English first, because there are particular words that you don’t know in Russian, so it’s easier to start in [Russian].” Participant 024 explained further:

I have such a limited base that I can use anyway, that I know what they are. A lot of times, if I’m in a stressed situation, I’ll forget words beyond my base vocabulary, and so I pretty much just jump in [without planning]. That way I don’t feel bad about having improper grammar or mistakes, or having to use a lot of small words to explain the situation if I just jump in and do it. And so, when I approach someone for directions in the metro or whatever I just do it as quickly as possible, just to get it over with, to save
as much face as possible. Hopefully they’ll understand it on the first try, and I’ll get a positive response…

The awareness of interlanguage limitations led to simplifications in planning for many learners:

I just tried to think of the simplest possible way [to make the request] and not to add anything that would trip me up… I tried to build on the “у вас есть,” “можно,” “нужно” kind of, you know, constructions and just add whatever possible words I knew. (Participant 052)

Other learners felt the need for grammatical modifications to their requests, but lacked the linguistic means. Participant 004, for instance, was looking for a way to incorporate a marker of conditionality:

…like “would you mind?” “do you think you could take me home?” I really wanted to put that conditional in there but I just didn’t know how, like “[ты] мог бы”… I do know “я хотела бы”… I don’t know how to use it, we didn’t study it very much… I can’t feel it [on the level of syntax] even down to just word order, I can’t feel where it needs to go in the sentence. I can feel that I need it but I just have no idea where to use it.

In other instances, limited knowledge of another pragmatically significant form, the imperative, affected pragmatic performance, as Participant 056 explains:

At this point I do not use any imperatives, because I don’t really know strongly how to do them. I know “читай,” “представь”, I don’t use them ever, ever. I don’t know how to form them well enough to use them. That’s a really good reason not to use them. I always say “можно мне что-то” because it didn’t even occur to me to say “дай мне”. Maybe that would be like too direct or something, I don’t know. But, you will find, maybe it’ll be different when you see me in Petersburg at the end of the semester, but you will find no imperatives in anything I just said.

**Interlanguage and grammatical person**

A frequent topic of the interviews was the learners’ use of grammatical person in Russian, especially second-person reference. Problems arose on several levels: fossilized forms incorporating second-person reference; inadequate control over pronominal forms; and inadequate syntactic control. Participant 001 was
particularly stymied by syntax in a role-play situation in which he asked a fellow student for a ride home:

I couldn’t think of how to ask for a ride, I was talking about myself, saying “я”, but I didn’t know how to switch it around, how to get her as the subject, because she was always the object...

Two conventionalized politeness forms, “простите” and “извините,” caused many problems for learners trying to act appropriately in minimum social distance situations. Participant 018 explained:

Usually I keep it straight. If I’m speaking with someone I’m on “ты”, I “ты” them all the way through, and if it’s someone I’m on “вы” with I “вы” them all the way through. But for expressions like “простите” or “извините” I’m used to hearing them in “вы” form so I accidentally do that... If it’s something I’ve heard a lot in the “вы” form, I do that. If I have more time to think, I’ll correct myself...

Several learners mentioned that their incomplete control over pronominal declension led them to revise their choice of second-person reference, switching between “ты” and “вы” forms. For example, Participant 011 noted about a situation where she switched from “вы” to “ты”:

I think I used “тобой” because I couldn’t off the top of my head think of “вами”. You’re going for better grammar a lot of times [more] than even politeness, just to sound more clean.

Participant 052 also switched between “ты” and “вы” when unable to recall the possessive form “ваша”:

I was thinking “вой” or “воя” or something. I remembered “твоя” easily, so I just switched back to that... I knew “твоя” was right, whereas I didn’t know whether “ваша” was right, so I just switched to [“твоя”].

Participant 053 used “ты” in a role play situation asking a university administrator for expedition of a bureaucratic document, explaining that his utterance was motivated by both phonology and grammatical control:

I get confused, ‘cos I use the “ты” more often than the “вы”. Also I’m scared of making like “вы” ‘cos I’m so conscious of the “ты” [sound]. I know it’s the same sound but I’m not as nervous about it... when I say the word “вы” I feel this great obligation to get the “ты” [sound] right and I don’t when I say the word “ты”... maybe because [“вы” is] more formal.
Like when I’m talking to someone that I’m using the word “вы” on then I should get it right, whereas when I’m using “ты” it’s not such a big deal. I don’t like using [the “вы”] conjugation unless I’m going to drop [the pronoun] and just use the conjugation, just be like “можете” ...

**Transfer and pragmatic performance**

Many learners indicated that transfer of their native language and culture (L1/C1) pragmatic representations (or conversely, the lack of analogous pragmatic representations to transfer) played a large role in how they carried out the request role-plays. Participant 019 explained the transfer process clearly, saying:

Maybe it’s more a matter of personality than anything. I’m just transferring my personality at home, what I would say in English and how I would go about it into the situation in Russia. I’m not really taking a lot of cultural things [differences] in mind...

To a great extent, the lack of a difference in English for second-person reference similar to that in Russian caused a great deal of confusion for the learners. Participant 008 noted that “I don’t realize how odd it sounds [using “вы” in an ostensibly “ты” situation]... having to distinguish between formal and informal forms of address is not something that comes naturally.” Participant 009 said that she and her 17-year-old host sister still used “вы” with each other, but added

Sometimes we say “ты”. I don’t know why, I guess we don’t see each other that often... it’s hard really to comprehend why it’s “ты” and “вы” here... I don’t think I’d feel uncomfortable but I don’t want to insult her by using “ты” with her.

Many learners like Participant 023 confirmed this observation, and frequently mentioned that they generally use “вы” with almost all interlocutors in order to avoid unintentionally offending them:

...I figure until I’m used to [“ты” and “вы”] and can make a distinction real easily it doesn’t hurt be too polite rather than being insulting. Like saying something to an older person or someone that I don’t know very well and them take offense. So, to be on the safe side.
Other learners talked about instances in which they had transferred their L1/C1 concepts about politeness, with varying results. This transfer especially concerned American English conceptions about requests and thanking. Participant 033 cited a situation that arose soon after he arrived in Moscow:

When I first got here, when people would say “пожалуйста” after giving me something, I thought I was being rude. In America, sometimes if you forget to say “thank you,” they’ll say “you’re welcome.” I thought it was the same situation [in Russia], that they were saying “you’re welcome” before I could say “thank you”… and then for a couple weeks when people were walking up to me I would say “thank you” [“спасибо”] before they could say “you’re welcome”, so I could beat them to saying “пожалуйста” because I thought I was saying it at the wrong time. And then someone explained to me that it meant nothing, just like “here, take it”… I thought they were just saying “you’re welcome” because they were upset that I didn’t say “thank you.”

Participant 019 cited a somewhat different example, also concerning requesting and pragmatic force, which revolved around a situation with his host family:

I used “будьте добры” once, …they said “whaaat???” And I heard somebody else say sometime that was excessively polite, it was like something you don’t normally say unless you want a big favor. I think I read somewhere that Americans tend to use “пожалуйста” and all the “please” type words way more than Russians… Because we say them so much in our standard vocabulary and they’re kind of meaningless sometimes, but I mean they’re just kind of standard. We over-translate, I mean not over-translate but… [Transfer?] Yeah transfer. I mean because I am transferring my English words into Russian.

This example clearly illustrates that encoded forms, especially those resulting from L1/C1 transfer, may convey undesirable (or at least unintended) pragmatic meanings in terms of the target culture (in other words, pragmatic failure).4

---

4 Two main sources of transfer have been identified: sociopragmatic transfer, or the transfer of the speaker’s native language and culture (L1/C1) sociological values; and pragmalinguistic transfer, or the transfer of forms related to pragmatic force and politeness values from the speaker’s native language (Thomas 1983). Kasper (1992) further notes that pragmatic transfer may be positive or negative. Positive pragmatic transfer, or the performance of native-like pragmatic strategies, typically facilitates communication, and causes miscommunication or pragmatic failure only when such behavior is considered to be inappropriate for non-native speakers, due to their
Two learners, Participant 034 and Participant 056, voiced particularly interesting perspectives on transfer of their L1/C1 competence. Participant 034, a bilingual Polish-English speaker, had a very detailed view of her pragmatic competence and its application to Russian:

I don’t think in words, I tend to think in pictures and images. So for instance if I learn a word, there’s an image associated with it. And in situations, the situations you’re giving me, basically I wanted to say either I wanted something from them, or I want to apologize, and it’s all the same thing. I just basically have to say I’m sorry and I didn’t mean it… Well I mean it comes from [working in] two languages at the same time. You cannot translate situations or meanings from one language to another, because the nuances don’t match. But if you have [a picture], the pictures can overlap with different words. Like if I have a picture of a concept, and then I’m looking for it in English, and I know exactly what I want to expect and it’s not in English, a word in Polish will pop out and I know that’s exactly what I want to say. And then I have to go around and say well how can I say this in English again. You know, look in my old dictionaries and find some ancient word that means the same thing. But it’s a description that has to fit the picture rather than the picture fitting the description. [Italics added]

Participant 056, bilingual in English and Spanish, displayed a similarly high level of metalinguistic awareness concerning pragmatic competence, but emphasized L1/C1 transfer more strongly than did Participant 034:

Basically overall my strategy was “what’s the emotion, and how can I communicate that emotion.” Because even if I don’t know the exact words if I can get the emotion across, they know, you’re trying to hit the right register with them… I know how to be appropriate in English… like they say about language learning, you take the utterance, you take what’s underneath the utterance, you put it in the other language, same register, same energy. That’s what I did. I just tried to make it be my emotion, and my expression match the energy that underlies the same interaction that I
would have in English. So that’s basically my strategy, I took the core and tried to transpose it into Russian. …On the one hand, I could say the wrong things, like I could say something that’s slightly wrong, or say the wrong words, but if I’m being in a certain way, they’re gonna get what I’m trying to say, for the most part. There are certain ways to be that are inappropriate, and you have to find those out, but a lot of it is just the being behind the message that you’re trying to get across.

Several participants mentioned the role of classroom training on how they planned and carried out the role-play requests, mentioning both instruction on pragmatically significant forms, such as imperatives, and second-person reference. As Participant 011 noted, the lack of formal instruction concerning imperatives affected her pragmatic performance:

I don’t use the imperative a lot at all. I don’t know why… We were always taught to use [the imperative] for commands, so I can’t think of that in situations where I would [use it]. The way they taught it to us, we didn’t use it often in class. You know and even in conversation class here [in Russia], we use, very rarely, use an imperative form. So I don’t use it much.

The instruction of second-person reference, especially in the American classroom, seemed to generate confusion about appropriate use of the second-person pronoun in target language contexts. Participant 053 noted that he used the “вы” form with everyone in his US classroom, including the native speaker instructor, adding:

I was using “вы” and I kept getting tripped up and I was getting confused and [the instructor] was like “Алёша [his classroom Russian name], в Русс я «вы», здесь я «ты».” Like, “don’t worry about it, we’re in America, it’s less formal.” Here [in America] the way professors are treated it’s ok to use “ты”.

The inherent difficulties in the instruction of native-like pragmatic competence in the domestic classroom setting was not lost on students, who viewed the study abroad experience as an opportunity to acquire this competence in a “rich and real” environment:

I would guess that [native speakers] do [requests] much differently, but that’s why I’m going to Russia… I’m really unsure… We’ve done dialogs

---

See Brecht, Walton et al. (1997).
in class similar to all these sorts of things, but applying them to real life situations is much different... It’s different because you don’t have a teacher that you know is going to correct you... it’s less real [in class] than actually having to ask someone or actually having to speak to someone who if you get in a bind, you might not be able to speak English, or have something explained to you... (Participant 054)

**Pragmatic competence and impression management**

Many learners displayed a high level of metalinguistic knowledge concerning their relative pragmatic performance in English and Russian. For instance, Participant 036 recognized his lack of knowledge about colloquial and formal speech in Russian, but focused his efforts on communication without reference to the social context:

Like in America, obviously I have very different ways of speaking to adults or people I don’t know very well than I do to my friends, but here I can’t really do that yet. I can get the idea out, that’s still the most important thing. [Italics added] There’s no difference yet.

For many learners, impression management meant displaying to native speakers their control of Russian vocabulary and grammar, and not making mistakes; sounding “clean,” in Participant 011’s words: “I’ve found that like with Russians, they appreciate clean speaking more than they do that you use the exact right verb, they like it better that you conjugate it whatever verb [sic] that you did use right.” From this perspective, grammatical and lexical accuracy carry more importance, relatively, than appropriateness in interactions with native speakers: “I haven’t noticed anybody get really offended by me messing that up [“вы” where “ты” is more appropriate]. I think people just figure that’s the smallest mistake I’m going to make, so that’s ok” (Participant 029). Oliver, however, attempted to balance both demands (accuracy and appropriateness) when carrying out the requests: “I wanted to be polite about it, but I was also

---

6 Impression management: “the way learners make use of their L2 resources in interaction to create social meanings favourable to themselves” (Ellis 1994, p.161).
7 In this respect, learners’ beliefs closely correspond to native speakers’ perspectives on error gravity: “Grammatical accuracy, as opposed to other issues of language production, was considered one of the most important issues in learners’ spoken Russian by all the respondents (and all subclasses) in this study” (Rifkin 1995).
trying, which I’m always trying, to be grammatically accurate... so I was trying to balance a few, wear a few hats I guess.”

For other learners, the chief goal of impression management entailed displaying their core L1/C1 values, even when they realized that their performance diverged from NS norms. Participant 037 host brother, for example, corrected her several times for saying “спасибо” too often. Nonetheless, she was not inclined to change her speech behavior:

I don’t think I’ve slowed down in my usage of the word because it comes naturally, like habit, to say “thank you” for everything. And I don’t want people to think that I don’t appreciate everything... here I feel like I have to show that I really appreciate, and so I think I use it appropriately, in appropriate contexts... maybe I do say it too much, but it’s habitual. I don’t think I’ll slow down.

Learners explained that their speech behavior in Russian arose from a deeply ingrained sense of rights and obligations, a core part of their personality. Participant 001, emphasizing that he was raised in New England, whose residents he characterized as very independent and self-reliant, was adamant on this point:

I think a lot of these answers (have to do with) the way you were brought up [taking responsibility for own actions, relying on oneself]... I feel that I’m putting somebody out, so I have to make it worth their while... by asking for a favor, I immediately wanted to offer a favor in return.

The learner’s pragmatic competence, when viewed as an intrinsic part of personality, may not be available to change or remediation, as Participant 014 explained: “When in Rome do as the Romans do, but I can’t... [it’s my] personality, maybe.” When questioned whether she might have altered her requests to the native norms if she were more fluent, she answered, “probably not... it’s me!” (Italics added)

Perhaps also in reaction to their reduced sense of rights and obligations on account of their status as foreigners and limited proficiency Russian speakers, some learners said that they preferred to diverge from native speaker pragmatic norms, even when they knew the norms. This perspective especially concerned situations where speakers felt that they were imposing on someone:

Because to me it’s so obvious that I’m not from [here]... I just feel awkward that I’m... I can’t do it the way [native speakers] do it because
I’m not from here and there’s something about that that I wouldn’t want to do it. (Participant 013)

The same student additionally noted the effectiveness of identifying oneself as a foreigner:

Because people feel sorry for you then... Because people expect foreigners to not know what they’re doing... they don’t expect that you’ll know what you’re doing, they expect that you’ll be confused a lot. They can tell from my accent, anyway.

Other learners also noted goal-oriented reasons for divergence from native speaker norms: “I would be like “может быть, I don’t want to be any trouble”... [that approach] tends to get results more instead of angering the other person.” (Participant 008)

Impression management, gatekeepers, and caretakers
For language learners in country, interacting every day with a wide range of native speakers in a multitude of public and private contexts, much of the language-learning task (at least outside of the classroom) boils down to developing appropriate target language verbal behavior for two main purposes. First, the learners must develop effective verbal behavior to obtain necessary or desired resources, such as food, accommodations, and travel arrangements, in interactions with not infrequently unsympathetic or impatient native speakers (gatekeepers). Second, the learners must develop appropriate target language verbal behavior to develop a social network of sympathetic native speakers on whom they can rely on for support, guidance, and friendship (caretakers). The caretakers further assist the learners in improving their target language proficiency.

It is essential to note that the language-learning task described above is vastly different from the task inherent in the classroom of the learner’s home institution, where the focus is generally on academic accomplishment; specifically, attaining a good grade. As discussed earlier, the domestic classroom learning experience does not prepare learners to interact with a wide range of native speaker interlocutors. As one learner noted early on during her study abroad sojourn, “There seems to be this dichotomy between public and private life here [in Russia], the way people act, and I can't quite figure it out.” (Participant 042)
For learners (especially those with limited proficiency), the difficulties of interacting with gatekeeper native speakers should not be underestimated. In these situations the challenge for the learners is to develop strategies to achieve their goals when communication problems arise, as Participant 019 described partway through his semester-long study abroad experience:

…it’s when they lose their patience, and there’s that “что? что? что?” and it starts to fluster you, and it’s hard to really come across, and already you’re self-conscious enough talking, so it’s difficult when someone just gets impatient with that process.

The same student noted a typical interaction with a gatekeeper native speaker in a ticket booth, when attempting to buy theater tickets:

…we got all the way to the end [of the transaction] and then she said something at me kind of quickly and I didn’t understand it and I asked her to repeat it, and she was impatient and said it again like 3 or 4 times. She was just asking me which tickets I wanted… and it really wasn’t very complicated what she said but after 4 repeats and it getting louder and louder and faster instead of slower and more helpful, it had the absolute opposite effect. She could have repeated it a hundred times, and it would’ve just [not helped], cause my mind starting shutting down a little bit. There’s more anxiety where there’s less contact [physical separation by glass of the касса, in addition to social distance]. There’s kind of this wall, even if it’s transparent, there’s kind of this wall between the two people.

In these situations, the learner must control the impression of himself as an effective and full-fledged agent that he wants to make on the interlocutor, by remaining calm and mustering the linguistic resources (both forms and pragmatic formulae) at his disposal, however limited.

Using the target language in country to build social networks also differs vastly from the language-learning task of the domestic classroom, as noted earlier. The distinction between the two tasks is far from lost on study abroad sojourners. They realize that building social networks and forming friendships with native speakers is perhaps the sine qua non of the study abroad experience. Half-way through her semester in Moscow, one student put it more succinctly than any other:
[Making Russian friends] brings your hopes up and your confidence up and it makes you want to learn... [with friends] you want to be able to communicate and get your point across. In class you want to do the same thing... [but with friends] you want to be able to show your personality, whereas in class you want to be able to show your intelligence and stuff. Here [in Russia] it’s like, I want to make some friends, I want them to know who I am. (Participant 037; italics added)

Discussion
As noted in Brecht and Robinson (1993), those study-abroad sojourners who made the greatest proficiency gains spent a large proportion of their free time with a close native speaker friend. However, building such relationships requires that learners possess sufficient target language competence to show their friends “who they are,” in the words of the learner above. Yet the quantitative analyses of their speech act performances (Frank 2002, 2009) showed that they lacked sufficient L2 grammatical competence to interact effectively with a broad range of native speaker interlocutors, who may be unused to interacting with non-native speakers.

The retrospective interviews provide evidence that L2 speakers focused much of their attention on L2 grammatical forms, where their performance was far from native-like. Conversely, they devoted much less attention on pragmatic appropriateness and choice of request or speech act move type (for example, stating the justification for the request, versus a more conventionally direct query of the speaker’s ability or possibility of performing the request).8,9 This relative difference in focus during request speech act planning and performance lends support to the hypothesis that L2 processing related to pragmatic appropriateness was automatic (and largely accurate, perhaps due to serendipitous L1/C1 transfer). L2 morphosyntactic processing, in contrast, was effortful and deliberate. This hypothesis is further supported by the inter-relationship of L2 grammatical and pragmatic competences reflected on level of pragmatically important sentence-level modifications, such as grammatical person, especially 2nd person reference.

8 For a full taxonomy of request speech act move types, see Frank (2002, 2009).
9 As noted earlier, the L2 speakers’ selection of speech act move type did not significantly differ from the native speaker baseline. On the level of speech act move, L2 performance diverged from native speakers only by delaying the request in minimum social distance situations.
These findings lead to further questions regarding classroom-based instruction in L2 pragmatics, and learners’ threshold level of L2 competences and the timing of study abroad. Instruction in L2 pragmatics has been found to be effective for features typically missing from classroom discourse, although “consciousness-raising” activities may be more effective than explicit instruction (Kasper 1997; Kasper and Rose 1999). As the learners indicated in the interviews that they had received little classroom L2 pragmatic instruction, L2 Russian pedagogues and curriculum designers have an opportunity to prepare their students for in-country language use by incorporating pragmatically significant features into existing activities. Specifically, the key finding from Brecht, Davidson, et al. (1993) that grammatical competence predicts in-country proficiency gains, and the finding from Frank (2002, 2009) regarding the inter-relationship of grammatical and pragmatic competences, point to the importance of integrating pragmatic features into grammatically-oriented activities. However, despite this interrelationship, it remains unclear whether accuracy increases and processing speed decreases with similarly developing trajectories in the two domains or not, and whether this occurs similarly in both the domestic classroom and overseas L2 immersion environments.

Additionally, given the evidence for a threshold level of grammatical competence for optimal in-country proficiency gains, an analogous threshold may exist for pragmatic competence. Identifying specific criteria that overlap both competences, such as near automatic control of grammatical person (especially 2nd person reference), will help both learners and pedagogues to plan for effective overseas language learning experiences.

**Appendix A: Definitions of Contextual Factors**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social distance</td>
<td>The distance between the speaker and the hearer. In effect, the degree of familiarity and solidarity they share as represented through in-group or out-group membership. (Hudson, Detmer, and Brown 1995, p. 5)</td>
</tr>
<tr>
<td>Relative power</td>
<td>The power of the speaker with respect to the hearer. In effect, this is the degree to which the speaker can impose his or her will on the hearer due to a higher rank within an organization, professional status, or the hearer’s need to have a particular duty or job performed. (p. 4)</td>
</tr>
<tr>
<td>Degree of imposition</td>
<td>The imposition in the culture, in terms of the expenditure of good and/or services by the hearer, or the obligation of the speaker to perform the act. This will vary depending upon whether the speech act is a request, a refusal, or an apology. (p. 4)</td>
</tr>
</tbody>
</table>
Appendix B: Retrospective interview protocol

We will play back the role play situations and then I will ask you a few questions about how you performed each situation. When we listen to each situation, try to remember how you planned to perform it, and why you chose to perform it the way you did.

What language did you plan to perform the situation in?

Did you consider any other ways of performing this situation? What were your other options? Why did you discard those options? E.g., didn’t know word, didn’t know how to say it in Russian, didn’t know if it was appropriate to say in the given situation?

Was this a realistic situation? Have you ever been in a similar situation in Russia? What happened? Were you able to appropriately negotiate the situation?

Would you have performed the situation any differently if the other person were not (your friend, sister, boss, stranger)? How?

Do you think a native speaker of Russian would have performed differently? How?

Is this how you would have done it in real life?

References


The Influence of Prior Language Learning Experiences on Learning of Unrelated or Distantly Related Languages

Ewa Golonka

Existing research provides evidence that learning one foreign language, and learning it well, makes learning another much easier. Learning another foreign language as an adult is easier both for those who acquired a second language (L2) in childhood and for those who acquired a second language as adolescents or adults in school. The learning of multiple languages within school settings in European countries, such as Benelux, Germany, Switzerland, and Austria, is a common experience. In the United States, deliberate adult third language (L3) instruction—which accounts for learner’s knowledge of other foreign languages—has been primarily in U.S. government training institutes. In these training institutes, two types of third language instruction are practiced: conversion and cross-training. Conversion refers to the retraining of individuals with demonstrated proficiency in one foreign language into a closely related language; for instance, L2 speakers of Russian learning Serbian/Croatian. Cross-training refers to L3 instruction in which learners acquire an unrelated language, e.g., L2 speakers of German learning Arabic.

Where does the evidence come from?
The bulk of empirical evidence supporting the claim that learning L3 is easier than L2 comes from studies that demonstrate the positive and facilitative effects of bilingualism on the acquisition of subsequent languages. For instance, several studies found that bilingual subjects showed a higher level of L3 proficiency than monolingual ones while learning a third language for the same period of time (Cenoz & Valencia, 1994; Lasagabaster, 2000; Muñoz, 2000; Sanz, 2000). Other evidence of advantages of bilinguals over monolinguals while learning L3 comes

---

1 I wish to thank Fred Eckman, Cathy Doughty, Bill Rivers, and Kathy Rhoad for their contributions to earlier versions of this article.
2 In higher education, L3 learning is usually not deliberate; students enrolled in the same courses often have background in other languages, but the instruction does not account for that.
from a broad scope of investigations that included metalinguistic awareness (Jessner, 1999), writing (Sagasta Errasti, 2003), pragmatic competence (Safont Jorda, 2003), vocabulary acquisition (Keshavarz & Astaneh, 2004), and language learning aptitude (Eisenstein, 1980).

There are some documented cases (and many more anecdotal, self-reported instances) of adult L3 courses that were shorter than regular L2 courses but produced similar results. For example, Corin (1994) reported on a Serbian/Croatian conversion course at the Defense Language Institute that retrained 40 Czech linguists in Serbian/Croatian in a three-month period, rather than the typical nine-month period that Category II languages usually require. Based on the outcomes of the course reported using the Interagency Language Roundtable (ILR) scale, the median oral proficiency score in this course was level 2 and the mode was level 1+, which was comparable with the results of typical nine-month courses. Rivers (1996) reported that cross-training courses in Georgian and Kazakh (with Russian as L2) were one third the length of basic courses in these Category III languages.

**What do we know about L3 acquisition?**

The evidence currently available suggests that previous linguistic knowledge and previous language learning experience affect the individual’s learning of consecutive languages. Learners tend to establish links between the target language (TL) and previously learned languages and to look for similarities among known languages (Rivers & Golonka, 2009 a, b). This ability to establish links between languages is possible due to cross-linguistic similarities among languages, both actual and perceived. Perceiving similarities, and in fact using them, are necessary for linguistic transfer to occur (Jarvis, 2000; Kellerman, 1995; Odlin, 2003; Ringbom, 2001). There is a general agreement among specialists that the more closely related the languages, the more transfer is possible. Specialists also agree that transfer from L2 to L3 is most salient in the lexical domain because learners can rely on a large number of cognates in related languages.

Most recent studies (e.g., Ringbom, 2007; Ringbom & Jarvis, 2009) distinguish different levels of transfer—item transfer (transfer of individual forms, such as sounds, morphemes, or words); procedural transfer (transfer of functions of the words or grammar rules); and overall transfer (transfer of both the form and the function). If the languages are closely related, quick and effective learning of items can occur, even at early stages of learning the L3 and at lower L2 proficiency. Procedural transfer, however, requires higher L2 proficiency levels in a closely related language. Additionally, the perceived
similarities can facilitate L3 comprehension regardless of the learner’s level of L2 proficiency. However, to facilitate L3 production, a high level of proficiency in a closely related language is needed. If that is not the case, the perceived similarities can cause errors in production. All in all, learners with higher L2 proficiency benefit more from conversion training.

Relevant literature has very little, if anything, to say about whether any positive linguistic transfer is possible in a situation when there is a zero or near-zero relation of one language to another (e.g., English and Chinese). There is, however, evidence that even if languages are unrelated, the experience gained in L2 learning can still make learning a third language easier, because experienced language learners have a much better idea of how to learn a new language. In fact, metacognitive awareness is often cited as a critical factor acquired during L2 learning that can be beneficial in learning a new language; however, little empirical evidence exists on the topic (Thomas, 1988). In addition, L3 learners show a high tendency towards learner autonomy—that is, they tend to take the entire learning process into their hands (Rivers, 2001). Students who have prior experience with language learning understand the language learning process, build awareness of their learning styles and preferences, and develop appropriate learning strategies (Ramsay, 1980; Wenden, 1999).

What do we know about L3 instruction?
In spite of the growing body of literature on L3 acquisition, the field still lacks a systematic investigation of L3 instruction. Publications concerning L3 instruction often have not had a strong research base; they have been based on teacher observations or expert opinions rather than on empirical data. In general, the authors of these studies tend to suggest that L3 instruction should maximize the students’ prior language knowledge and learning experience.

For closely related languages, there are documented efforts of developing instructional materials that take students’ prior languages into account based on the principle that there is a high degree of mutual intelligibility between the languages (Gribble, 1987; Jensen, 1989; Jordan, 1991; Townsend 1995). However, the efficacy of these courses, which are highly contrastive in nature, has not been measured. In fact, contrastive analysis is often cited as a type of method that has potential to enhance L3 acquisition. However, contrastive analysis has been shown to be ineffective in L2 acquisition and its efficacy in L3 instruction is unproven.

Additionally, student and teacher selection for L3 courses is crucial to the success of those courses. Brecht et al. (1998) argues that the guiding principle for
student selection should be to place students in courses where the TL is as closely related to student L2 as possible and to select students with the highest possible L2 proficiency. Other self-reported data suggests that it is beneficial for the teachers to have a background in the students’ L2s or at least be familiar with what language backgrounds their students come from.

Because no specific methodologies or materials exist for teaching L3 as an unrelated language, cross-training courses are often taught as regular L2 courses, and they only take advantage of the students’ prior knowledge and experience through rigorous student selection into the program, e.g., only experienced learners selected. Anecdotal, self-reported data suggests that even in those cases students perceived their knowledge of another language as beneficial in learning L3.

Method
Using qualitative methodology, this study examines the effects of prior language knowledge and prior language learning experience on the acquisition of unrelated or distantly related languages.

Participants
Participants were 25 students enrolled in U.S. government cross-training courses and 10 instructors either teaching such courses or having backgrounds in teaching experienced language learners. All students were adult native speakers of English who ranged from 20 to 60 years of age, with the majority of students ranging in age from 26 to 30. Twenty female and five male students participated. All students demonstrated at least an ILR level 2 proficiency in a language that was unrelated to the language of their study (also called here a qualifying language). Table 1 below represents the language experiences of each student including the student’s target language, qualifying language, and, if applicable, background in any other languages. Out of the 10 instructors, six were teaching in the cross-training program (henceforth called current instructors), and four were instructors with multiple years of experience teaching experienced language learners at various U.S. government training programs (henceforth called expert instructors). All of the current instructors were foreign nationals, with an average of 7 years experience as language instructors. All of the expert instructors were native speakers of English. Seven male and three female instructors participated. All instructors spoke English, although it was the native language for only four of them. Arabic and French were spoken by at least three of the instructors, and German, Italian, Iraqi Arabic, and Levantine Arabic were
spoken by two or more of the instructors. All of the current instructors taught either their native language or a language that is similar to their native language. On the other hand, all of the expert instructors taught a non-native language.

Table 1: Student Language Background

<table>
<thead>
<tr>
<th>Student</th>
<th>Target Language</th>
<th>Qualifying Language</th>
<th>Background in languages other than TL and qualifying language</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Arabic</td>
<td>French</td>
<td>Italian, Slovak, Spanish</td>
</tr>
<tr>
<td>7</td>
<td>Arabic</td>
<td>Spanish</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Arabic</td>
<td>American Sign Language</td>
<td>Japanese, Latin</td>
</tr>
<tr>
<td>9</td>
<td>Arabic</td>
<td>French</td>
<td>Spanish, Latin, Dutch</td>
</tr>
<tr>
<td>10</td>
<td>Arabic</td>
<td>French</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Korean</td>
<td>French</td>
<td>American Sign Language</td>
</tr>
<tr>
<td>14</td>
<td>Korean</td>
<td>Bulgarian</td>
<td>French, Serbo-Croatian, Macedonian, Korean</td>
</tr>
<tr>
<td>15</td>
<td>Korean</td>
<td>Spanish</td>
<td>Korean</td>
</tr>
<tr>
<td>16</td>
<td>Korean</td>
<td>Spanish</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Korean</td>
<td>Spanish</td>
<td>Italian, Japanese, Yoruba</td>
</tr>
<tr>
<td>1</td>
<td>Chinese</td>
<td>German</td>
<td>Spanish, Latin</td>
</tr>
<tr>
<td>2</td>
<td>Chinese</td>
<td>French</td>
<td>Latin, Spanish, Arabic</td>
</tr>
<tr>
<td>11</td>
<td>Chinese</td>
<td>French</td>
<td>Spanish, Hebrew</td>
</tr>
<tr>
<td>12</td>
<td>Chinese</td>
<td>French</td>
<td>Italian, Japanese</td>
</tr>
<tr>
<td>19</td>
<td>Chinese</td>
<td>French</td>
<td>Japanese, Spanish</td>
</tr>
<tr>
<td>17</td>
<td>Farsi</td>
<td>Russian</td>
<td>Spanish</td>
</tr>
<tr>
<td>18</td>
<td>Farsi</td>
<td>Spanish</td>
<td>German, Hebrew</td>
</tr>
<tr>
<td>21</td>
<td>Farsi</td>
<td>French</td>
<td>Italian, Russian, German, Dari</td>
</tr>
<tr>
<td>22</td>
<td>Farsi</td>
<td>Russian</td>
<td>German, Ukrainian, Bulgarian, Japanese, French</td>
</tr>
<tr>
<td>24</td>
<td>Farsi</td>
<td>German</td>
<td>Arabic</td>
</tr>
<tr>
<td>3</td>
<td>Pashto</td>
<td>Spanish</td>
<td>Basque, Portuguese</td>
</tr>
<tr>
<td>4</td>
<td>Pashto</td>
<td>Serbian</td>
<td>Spanish, Greek, Bulgarian, Macedonian, German</td>
</tr>
</tbody>
</table>
Data collection instruments and procedures

Protocols for the Learner Individual Interview and the Instructor Individual Interview (Appendix A) were developed based on insights from relevant literature. The format of the interviews was semi-structured in nature, meaning that the protocols served as a guide for the interviewers, permitting a focused conversation with the flexibility to probe and explore certain areas when relevant. The Instructor Interview Protocol focused on the instructor’s background, the current or former cross-training courses, and the cross-training students. The Learner Interview Protocol included questions about the participant’s language background, comparisons between the current program and any previous language learning experience, and student self-assessment about individual learning styles, strategies and overall characteristics as language learners.

Each semi-structured interview lasted between 45 and 75 minutes. They were conducted either by one or two interviewers. In the cases where participants consented to audio recording (30 participants), the interviews were digitally recorded and subsequently transcribed. In the cases where the participants did not consent to audio recording (five participants), the researchers took field notes which were later typed up and prepared for analysis.

Data analytic procedures

A qualitative analysis software program, ATLAS.ti, was employed to code and analyze data. The starting point for this was the development of concept maps that organize the coding process (Figures 1 & 2, see below).

By reviewing the background literature on L3 acquisition, the interview protocols, and by examining the completed interview transcriptions, the analysts designed two concept maps, one to be used for instructor interviews and the other for student interviews. There is some degree of overlap between these two figures as they attempt to organize data that come from different sources, while focusing on the same phenomenon. This heuristic served as the basis for conceptualizing relationships among known and potentially new factors that are significant in adult language acquisition, such as students’ and teachers’ language background, education, and L3 learner characteristics. These factors were then further organized into two categories, factual and inferential, from which factual and inferential codes were developed. Factual codes focus primarily on low-inference, factual information of each student or instructor, e.g., level of education, number of languages studied, and length of study. Inferential codes require a high-inference level of judgment from the coder to identify the
phenomena, e.g., to represent potential variables influencing language learning, such as learning strategies or self-related beliefs.

**Figure 1:** Instructor Interviews: Possible Factors Influencing L3 Acquisition in a Cross-Training Setting.

The codebook for this study contains 318 factual codes (189 for learner interviews and 129 for instructor interviews) and 175 inferential codes (101 for learner and 74 for instructor interviews). The factual coding was completed by two analysts and the inferential coding was completed by four analysts (five at the beginning of the process). In general, relatively high reliability is expected in factual coding, although lower levels of reliability are accepted in inferential coding. After the coding socialization period (when all analysts coded portions of interviews together, then separately, then refined code definitions, and added new codes), the analysts began coding the remaining interviews independently. Each time the coding of an interview was completed, a second analyst reviewed the codes and discussed disagreements with the first analyst, who then finalized the file based on the outcome of the discussion.
**Figure 2:** Student Interviews: Possible Factors Influencing L3 Acquisition in a Cross-Training Setting.

**Inter-coder reliability**

Inter-coder reliability was calculated mainly to develop the most adequate code book by refining the codes and code definitions while gaining experience with the interviews. Inter-coder reliability for factual coding was calculated using joint probability of agreement of the factual coding of both student and instructor interviews. There were four categories used in assessing the reliability: total: total number of codes; same: codes that were coded identically; different: codes that were coded differently; missing: codes that were present in one analyst’s portion, but were not present in the other’s portion (Appendix B). This process of testing reliability was repeated in six out of the 25 student interviews and five out of the ten instructor interviews. The achieved reliability for the former was 77% and for the latter, 83%. This was due to one analyst who added and utilized five new codes during coding procedures, while the other analyst did not, thereby causing the largest level of disparity in reliability.
Inter-rater reliability for inferential coding was calculated using the internal reliability for each rater. The internal reliability of each rater was calculated according to the equation: \( R(x) = N(x)/N_{(tot)} + E(x) \), where

- \( R(x) \) = a single coder’s reliability score for one interview portion
- \( N(x) \) = total correct codes by rater \( x \)
- \( N_{(tot)} \) = total number of correct codes determined by consensus
- \( E(x) \) = extra codes by rater \( x \)

The results (Appendix B) allowed progress in reliability to be charted across time in two ways. First, each individual coder’s performance from week to week could be compared. Second, the product of all the internal reliabilities, i.e., the inter-rater reliability score, could be compared from subject to subject, based on the equation:

\[
R_{(tot)} = \sum \frac{N_{(xi)}}{N_{(xtot)}} + E_{(x)}
\]

Two aspects of reliability should be noted. First, because each full interview was coded by one person, checked exhaustively by another person, and then corrected by the original coder, the potential for missed codes or otherwise anomalous coding was low. For that reason, the reliability scores do not necessarily reflect the quality of the actual coding that led to analysis. Second, the primary benefit derived from collective coding and discussion of the portions was refinement of the code list and code definitions according to the coders’ experiences with the interviews. Consequently, the code list became well-tailored to the interview protocol and the types of answers that subjects gave, while remaining grounded in existing language-acquisition theory and empirical findings.

**Results**
The interviews revealed that learners from this study attempted to establish links between L3 and previously learned languages specifically related to their linguistic and metalinguistic L2 knowledge, as well as their experience in language learning.

**Knowledge of an L2**
Although the students’ L2 and L3 may have been genetically related only distantly, or even unrelated, and further, even though all students perceived their L3 as unrelated to their L1 and L2, twelve (out of 25) students indicated that
knowledge of an L2 facilitated their acquisition of the L3. While each interview addressed transfer of linguistic knowledge from L2 to L3, only students who were learning Indo-European languages—Pashto and Persian Farsi—were able to identify specific instances of non-native transfer [i.e., transfer from a language other than their native language (L1)]. These students pointed to a number of distant, genetically related languages, such as Spanish or German, as being helpful in learning their target language—especially in learning its lexicon and to some extent its morphology and syntax. The following quotation from a Pashto student illustrates his willingness to look for similarities between Pashto and a number of distant, genetically-related languages in the area of lexicon:

“For example if you count to five in Pashto: ‘Yow, Dwa, Dre, Salor, Pinsa.’ Yow is not too far different from ‘uno’, or ‘one’, or ‘eins’, or whatever you want too... all the other Indo-European languages. Not all of them, but many of them, at least the Romance languages and some Germanic languages. Dwa, two, dos. Dre, tres. And you go to different words, ‘mother’ and ‘father.’ ‘Mother’ is ‘mor’, which is not too different from ‘madre’, from ‘mother’, from ‘mater’ or whatever. Father is ‘plar’, ‘pater’, ‘father’, ‘padre’. Water is ‘uba’, ‘agua’, ‘eau’ in French. When you know what to look for, you don’t always know what to look for, when you sort of have some clues of maybe what to look for, then sometimes you can find some similarities.” (Student 3, Pashto)

Some students also said they learn more quickly when they can identify similarities between languages because then they do not need explanations of certain grammatical concepts. As an example of such a “shortcut,” one student discussed word order learned in German, which was helpful in learning similar concepts in Pashto:

“German is one of the languages where it’s subject, object, verb. So, when it came to Pashto and I saw subject, object, verb, it wasn’t like a great enormous shock for me, because it’s like ‘OK, I’ve seen this before so we move on.’” (Student 4, Pashto)

Conversely, another group of eleven students learning non-Indo-European languages, specifically, Arabic, Chinese, and Korean, did not identify instances of linguistic transfer. They stated that because of huge differences between the TL and any other languages they previously learned, linguistic transfer is impossible. All of the students who believed that their previous knowledge did not affect their learning of a new language cited the distance
between the languages as the main reason for the fact that the previously learned languages had no effect on their L3. The following example shows that a student of Chinese believed that a related language was helpful in learning another foreign language in the past, but the unrelated one did not:

“I don’t know that the language itself did (help), in that Chinese has nothing to do with it. You take everything you think you know about language and throw it out the window and start fresh. French had a big impact on my ability to learn Italian, but I didn’t feel that sense of that degree of transfer in French to Chinese.” (Student 12, Chinese)

While the majority of students claimed that there were advantages of previous language knowledge, seven students indicated that their prior language knowledge interfered with their learning or use of the TL. It should be noted that all of the interference was described by the students as slight or minor. This is how one Chinese student discussed an example where interference occurred, because she already understood a particular grammar concept in Spanish:

“It was one of those things where it took me a while to learn the difference. [...] Because I had this preconceived notion in my head, I was expecting it to be in certain places where it wasn’t, and I was putting it in certain places where it wasn’t, or I was not expecting it to be in certain places where it was.” (Student 11, Chinese)

**Metalinguistic knowledge**

All expert instructors and the majority of the students pointed to metalinguistic knowledge as the most facilitative factor influencing the learning of a new, unrelated language.

**Global awareness about language**

The interviews with students and expert instructors revealed that L3 students possess the knowledge of what languages are and how they are constructed. Without being explicitly asked, all of the students demonstrated some aspects of metalinguistic knowledge by reflecting on the nature or functions of the language. While the interview protocol did not explicitly seek information about specific languages, in all of the interviews, students discussed some grammatical peculiarities of the TL. In answering a question whether previous knowledge of or about language helps or hinders the students in their learning, 16 students
expressed the view that the metalinguistic awareness has helped them significantly in L3 learning. As expressed by one Farsi student:

“I think knowing a specific language hasn’t helped as much as kind of (having) a global awareness of how languages work. I think that’s really it.” (Student 18, Farsi)

Several students stated that the previous study of other foreign languages gave them the grammatical basis they needed to study a new language. In fact, all of the students stated that they only started studying grammar while studying their first foreign language. One Chinese student said:

“So when I learned French grammar, that’s my grammar basis. When something comes up in Chinese, that’s when I need to apply it to French because that’s my only place I understand grammar. So, I guess having a base in the grammar helps.” (Student 2, Chinese)

**Importance of English grammar**

In addition to metalinguistic knowledge of other foreign languages, and of grammar in general, several students listed knowing English grammar as an important factor, claiming that English grammar learned through studying an L2 was beneficial in learning their L3. They indicated that English grammar gave them a needed foundation and stressed its importance and facilitative role in learning a new language. The following citations supporting this finding are from interviews with students of Chinese and Farsi:

“I actually learned the most grammar, the most English grammar in my foreign language classes. So, because of that, you know, when I’m learning Chinese I already have that foundation.” (Student 19, Chinese)

“I think this is one of the things maybe my Russian has helped in my Farsi is that there have been times that I never really understood English grammar. And I think that sometimes, and I have only referenced Russian, not because I understood Russian grammar better, but I knew it better because I was actually taught the grammar formally. So I think that that has helped me. […] I think it has helped me to learn the grammar concepts, because even now I understand English grammar better.” (Student 17, Farsi)
Refined expectations about language
Students mentioned a number of other advantages of previously acquired language knowledge. For instance, some students listed an advantage of “having an open mind” while learning a new, unrelated language, which can be related to the discussion in Zobl (1992) and Klein (1995) on less conservative learning procedures by experienced learners. One Farsi student expressed:

“I just think that having learned a language made it easier for me to understand how to learn another language. You know, because other people are like ‘well it doesn’t make sense what they are saying in Farsi.’ It doesn’t make sense to us, but it makes sense; otherwise they would not say it. I think I learned that from learning German. [...] And I think from learning other languages I had an open mind about that and some people don’t have an open mind about that.” (Student 24, Farsi)

Another advantage that came out in the interviews was the fact that the L3 learners come to the classroom with set expectations about foreign languages that they had formed based on their knowledge of other languages. One Arabic student explained:

“But the structure is very different, and I think if I had only been exposed to English, I would have to get into my head the idea that the structure of another language is very different before I could actually learn the structure. Here it’s like I already got that the structure is going to be different than English...” (Student 5, Arabic)

Experience with language learning
While the linguistic knowledge that the L3 learners bring to the learning task is language-specific, the experience of learning another language is not. Rather, the experience gives the learners an understanding of the language learning process and builds their awareness of their own learning styles and preferences. Through the experience of learning other languages, the students are exposed to different learning situations and are able to develop strategies that work best for them.

Awareness of learning styles and preferences
All students interviewed in this study expressed the opinion that their prior language learning experience provided them with some learning strategies and study skills that they now use to support their new learning tasks. All of the interviewed students showed an awareness of their learning process, as well as
an understanding of what learning a language means and how it can be achieved. All of the students said that they know what works for them in terms of language learning. In particular, they all believed that they can accurately assess how they learn best. In most cases, they demonstrated knowledge of learning styles in general, and were familiar with relevant terminology.

**The use of metacognitive and cognitive strategies**
The majority of students showed that they understand their language learning needs and that they actively took steps toward fulfilling them. The interview protocol probed whether the students requested or proposed any changes to the course content or structure. Twenty-one of the students indicated that they had wanted, or asked for, changes to the current program. Below are two excerpts from student interviews discussing changes or improvements to the course content and structure:

“Initially, the way the program was presented was that we were supposed to focus on all four skills; it wasn’t supposed to just be reading and listening because it is understood that you just can’t take two skills of a language and just learn it. You learn a language best by doing everything. We started out that way, but the speaking and the writing became less and less emphasized and there were some of us who continually asked for more speaking practice.” (Student 12, Chinese)

“We did ask that things slow down a little bit. We were rushing through the chapters very quickly and we just couldn’t remember it. You learn a new set of vocab every two days, you use it a few times, and you move on. We wanted some review, maybe some practice once and a while to try the grammar patterns out and to make sure we got them.” (Student 6, Korean)

All of these students who had not made requests in previous courses asked for changes in the current cross-training program. Four students indicated that they did not ask for changes in the past because they lacked the confidence to do so. Below is a conversation between a student and an interviewer after the student discussed changes she proposed to the current course:

“Interviewer: Have you suggested or requested any changes in your French class or Italian class?
Student: No.
Interviewer: Why not? Why now?”
Student: Well, I think because I’m older and maybe I’m more confident. In this situation, I felt like I was a professional, and somewhat peers with my teachers.” (Student 12, Chinese)

Another metacognitive strategy discussed in interviews was selective attention, sometimes called directed attention, which relates to focusing on specific parts of the language input that could help learning the important material in a learning task and ignoring irrelevant material (O’Malley, J. & Chamot, A. (1990). Those students who used selective attention as a strategy for learning did so mainly in reading and translation. Their strategy for reading something for the first time was not to look up every single word they did not know, but rather to look at the words they did know and try to get the meaning of what they were reading from context. One Arabic student felt that she had an advantage over L2 learners because she already knew that omitting some irrelevant information might help in performing some language tasks. She attributed this knowledge to her previous experience learning foreign languages:

“I’m not trying to translate every word. A lot of times when I see people who have not learned another language before, or who are for the first time trying to learn another language, they are trying to go ‘word, word, word’. I think I already know that it’s not going to work out that way. It’s just ‘how do you express the thought’, and that saves some time.” (Student 5, Arabic)

The interviewed students discussed the use of at least two cognitive strategies in their language learning: analytical thinking and analyzing contrastively. Some students in the cross-training program noted that they see language as a type of puzzle: “I always look at language as kind of a puzzle, so I actually enjoy figuring things out.” (Student 9, Arabic). They usually look at the patterns and systems when learning the TL as a result of their metalinguistic knowledge. Speaking about defining patterns, one Farsi student compared learning a language to learning math:

“It is almost like math: A plus B equals C. You can’t ever have A plus B and equal F. You know, like it is very systematic... so I think anyone who can define patterns and identify them and utilize them is the best. Because it is like that in every language.” (Student 17, Farsi)

The interview protocol probed the students to find out whether they compare the languages they know without prompting from their teacher. Those who said that they try to compare their L2 and L3 stated that they do it to understand concepts, to have a point of reference, or to help to remember. Of
those who indicated that they compare the L2 and L3, several said that they find those comparisons useful.

**Cross-training instruction**
The results of this study indicated that cross-training classes, which foster the learning of a new TL after learning one or more unrelated languages, were taught as regular L2 classes. In particular, the instruction did not have specific cross-training goals that would take advantage of prior language training and experience that students bring to the learning situation. The course goals were usually proficiency oriented or specific-skill oriented. In all cases, teaching methods and materials were not adjusted to address the L3 learner. Very often, the instructors did not know the language backgrounds of their students. With the exception of sporadic L2-L3 comparisons, current instructors did not make any changes specific to L3 learning and teaching. The program mainly took advantage of the students’ prior knowledge and experience by very rigorous selection to the program.

**Discussion and Conclusions**
The results of this study show that previous exposure to learning one foreign language can positively influence the learning of a new, unrelated or distantly related language. Based on student and expert instructor interviews, it can be concluded that metalinguistic knowledge about languages, including English, facilitates language learning of a distant language to a larger extent than the knowledge of a particular language. Experience of learning another language also plays an important role in learning a distant language, because the skills acquired during the years of studying foreign languages are not language-specific and can be transferred from one learning experience to another.

In the cognitive theory of learning, prior knowledge is listed as one of the key principles of learning (Ausubel, 1968; Hartley, 1998). In the field of foreign language learning, while the same idea is of interest, the extent to which it may influence learning a new language mainly depends on language proximity. If the languages of interest are unrelated or distantly related, not much prior knowledge is relevant (Ringbom, 2007). Because the participants in this study were students enrolled in cross-training courses, the challenge of this study was

---

3 Ringbom (2007) distinguishes between language proximity, i.e., similarities between the languages and language distance, i.e., differences between the languages.
to establish whether there is any prior knowledge to which learners can refer to in a situation where the languages of interest are unrelated or distantly related.

Not surprisingly, the topic of linguistic transfer was discussed by all of the participants of this study. The interviews with students revealed that their views regarding non-native transfer were consistent with the findings from previous research. These studies demonstrate common agreement among researchers regarding the cross-linguistic influence in closely related languages (Cenoz, 2001; Hammarberg, 2001; Ringbom, 2007). But even when languages are unrelated, or only distantly related, they can still share some characteristics that might be helpful to the students. The only students who discussed linguistic influence from languages like Spanish and German were students of two Indo-European languages: Pashto and Farsi. Conversely, all of the interviewed students stated that for languages that have little or no relationship to the languages the students already knew, it was not the students’ knowledge of another language that helped them the most, but rather the students’ knowledge about language in general, as well as the students’ knowledge of grammar, including English grammar.

Ideally, in the L3 classroom, these learner-internal factors such as proficiency in one foreign language and metalinguistic knowledge can be put to use so that the students’ prior experience is not only disregarded, but is in fact maximized. This can be achieved in part by the learners themselves and in part enhanced by instruction. The results of this study revealed that the learners believe that they are taking advantage of their prior learning experience, regardless of whether or not the instruction helps them in this process. Learners also revealed characteristics of autonomy, demonstrated in the use of a variety of metacognitive and cognitive strategies. Interviewed students showed the ability to take charge of their own learning by doing independent work, and they felt confident to propose changes to the course content or structure.

Results also demonstrated, however, that in the cross-training courses in which the interviewed students were enrolled, very few techniques specific to cross-training could be discerned. In fact, relatively few comparisons between the TL and another language (other than English) were mentioned. In all of those cases, the students expressed the view that those language comparisons facilitated their L3 learning. No other methods related to cross-training were used; the courses did not have specific cross-training goals, and in general, these courses were not taught differently than regular L2 courses. Thus, the interviews revealed that the program largely took advantage of the students’ prior knowledge and experience through rigorous student selection into the program.
This way, all of the students in a classroom were experienced learners who had previously demonstrated that they were able to learn a foreign language. In addition, these students also possessed metalinguistic knowledge and knowledge of learning styles, strategies, and study skills beneficial to their learning of a new language.

As the review of relevant literature demonstrates, the field of L3 acquisition has not yet undertaken the systematic investigation of L3 instruction, especially in those cases where the relationship among languages known by an individual are typologically and genetically unrelated. Because no specific methodologies exist for teaching L3 as an unrelated language, it remains a challenge to determine how to best design the instruction in cross-training courses so that it can take full advantage of students’ prior language knowledge and experience.

Appendix A: Interview Protocols

Learner Individual Interview Protocol
(Semi-Structured Interview)
Interview Questions: (1 hour)
Prioritized Questions:
A. We would like to get a sense of your background and experiences learning and using foreign languages. Could you talk a bit about:
   a. The foreign languages you’ve learned? Which is strongest?
   b. Any proficiency tests/OPIs you’ve had in your FLs?
   c. How long have you studied your strongest FL?
   d. Where did you study your strongest FL?
   e. In what context did you study it: classroom, immersion, study abroad?
   f. Have you lived abroad?
   g. Are you a native speaker of English? Did you learn other languages at home as a child?
   h. Have you ever used your foreign language professionally? That is, has anyone paid you to use it?
   i. Have you ever taught language: TL, English, other?
   j. What language do you study here? When did you start this program?
   k. Was the TL your choice?
B. Now, we’d like you to think about your experiences learning other languages and compare them to your current experience. In particular:
   a. Is your learning experience in this program different compared to foreign language courses/programs you took in the past? If yes, in what ways?
   b. Does the fact that you’ve studied other languages in the past and that you know them now influence you in any way in this program?
      i. Is it advantageous?
      ii. Is it a hindrance?
   c. Is your learning experience different now compared with when you first started this program?

C. Next, we’d like to talk about you as a language learner:
   a. I know this program is very intensive. What motivates you to keep going?
   b. You are spending several hours in class every day, and then you need to do your homework, study for tests, etc. How do you manage your learning on a day-to-day basis? How do you plan your day?
   c. Do you set (immediate and long-term) goals for yourself?
   d. Can you accurately assess how, where, and when you learn best? In other words, do you know what works for you?
   e. Do you know more about your learning styles/preferences and needs now than when you were learning L2?
   f. Have you requested changes to this course content or structure? Either individually or collectively?
      i. What changes? Can you give examples?
      ii. Why? What goals did you have in mind?
   g. Have you requested/suggested changes to any other course you took prior to this program?
   h. Given a reading task, how do you approach new words? Why?
      How do you know to apply such a strategy?
   i. How do you prefer to learn new vocabulary?
   j. Do you find it easier or harder to learn TL vocabulary now than when you first learned L2?
   k. Do you find it easier or harder to learn TL grammar now than when you first learned L2?
1. Do you sometimes look through material ahead of time to see how it is organized in order to prepare for an upcoming task? If yes, do you study new material ahead of time?

m. Are you concerned about making mistakes while speaking in TL? In another language?
   i. Do you find it useful to be corrected by your teachers? Colleagues?
   ii. How do you prefer to be corrected?
   iii. Do teachers tend to correct all/most of the errors?
   iv. Does this sound like a good practice?
   v. What do you then do following their corrections?

n. Do you monitor your own speech: look for mistakes and correct them after they occur, or even before they occur?

D. Now we’d like to talk about this program.
   a. Does this program take advantage of your background knowledge of another foreign language?
   b. Do teachers compare your L2 and TL in class? English and TL?
      i. What areas: (syntax, morphology, phonology, lexis, culture, etc.)?
      ii. Do you find such comparisons useful?
   c. Do you look for differences/similarities between L2 and TL with no prompt from your teacher?
   d. Are there more/fewer grammar explanations in this course than in your previous language courses?

E. Now we’d like to talk a bit about this program in particular, without referring to your past language learning experiences.
   a. Which skills are you learning faster/slower: reading, listening, speaking, or writing?
   b. How do you think you are progressing compared with other learners in class? Why do you think so?
   c. If you are doing better/worse than others, what accounts for it?
   d. As you know, you had to demonstrate proficiency in a second language in order to enroll in this program. Is this a valid criterion for screening applicants? Does this make sense to you?
   e. If it were up to you, what might be the background of an ideal student for this program?
   f. If it were up to you, what might be the background of an ideal teacher for a course in this particular program?
g. If it were up to you, how would you structure the program? What might be done a little bit differently? Any ideas for improvement?

F. Wrap Up:
   a. Has anything occurred to you that we haven’t asked about that you feel is relevant?
   b. Do you have any questions for us?

Instructor Individual Interview Protocol
(Semi-Structured Interview)
Prioritized Questions: (1 hour)

1. Can you tell me a little bit about how you learned the language that you’re teaching now, and any other languages that you know? (Follow up and see how much of it is formal coursework and at what level, how much is native exposure, and how much is immersion experience.) Also, what kind of teacher training have you had that you can call upon to help you with this current position?

2. Have you taught any cross-training courses in the past, or is this your first time? If you have taught courses like this before, what were they and how many times have you taught them?

3. Now I’d like to talk about the different ways that courses like the one you’re currently teaching are developed and implemented. I have four different aspects of development and teaching that I want to talk about one by one: the goals of the course, the methods use, the role of linguistic similarities and differences, and the materials that you use.
   a. Course goals
      i. What are the goals of the class you’re teaching now? Are any of them specifically cross-training in nature? (In other words, are there goals that are taking into account the fact that the students know at least one other language?)
      ii. What course goals are explicitly communicated to the students?
      iii. What are the proficiency goals of the course? In what skills?
      iv. Which skills are emphasized in this course?
   b. Methods
i. It’s probably a little different to teach students who already have studied a foreign language than it is to teach students who are learning a foreign language for the first time. Have you changed the way you teach at all for students in this kind of cross-training class? If so, how?

ii. What are the kinds of instructional methods that you’re using in this class?

iii. Can you talk about the proportion of language use in the class? How much of the time are you and the students using English, and how much of the time are you using the target language (the language that you’re teaching them in the class).

c. There are always similarities and differences between languages, and the fact that students are familiar with other languages besides English may both help them and cause problems with their learning. Do linguistic similarities and differences between languages play a role in this class? Do you know what other languages the students you’re teaching already know? And do any of the following play a role in your teaching?

   i. Linguistic transfer (syntax, morphology, phonology)
   ii. Lexical transfer
   iii. Cultural transfer

iv. Are these kinds of transfer helpful for teaching or learning in this class? Or do they cause problems? Can you give some examples?

v. Do you think students are transferring more from English (or their native language if it’s not English)? Or do you think they’re transferring more from their second language? Can you give some examples?

d. Materials

   i. Do you have materials that have been adapted to take into account that your students aren’t learning a foreign language for the first time? If so, what are they? And if it comes through in your teaching but not written materials, in what ways?

   ii. Do you use any materials that were written for native speakers by native speakers? Like newspaper articles or
texts from the Internet? If so, what do you use, and how do students seem to respond to these materials?

4. What does the learning curve look like for students who are learning a third language as compared to students who are learning a second language?
   a. High, fast, initial gain and then leveling off?
   b. What does the learning curve look like for different skills?
   c. Is there a relationship between how well students speak, write and listen in their second language and how well they speak, write, and listen in the target language that you’re teaching in this class? Do you think that if they are more proficient at their second language they’ll acquire the target language faster?

5. Could you compare the way students behave in this class as compared to students in classes where they’re learning a second language for the first time? [Of course, this should be skipped if the interviewee has no basis for comparison.] Can you please talk about these areas?
   a. Rate of acquisition
   b. Autonomous learning
   c. Motivation and attitude toward the learning process
   d. Linguistic risk-taking
   e. Error self-corrections
   f. Metacognitive learning strategies (e.g., planning, setting goals, self-evaluation, monitoring)

6. If you have enough experience to make an assessment, can you talk about how the outcomes of classes like these compare with the outcomes of basic L2 courses for this language?

7. Whenever I teach, I’m always thinking about ways that the course could be done a bit better the next time. Do you have any ideas of what you might do a little bit differently the next time you teach this class or a class like it? Any ideas for improvement?

Appendix B: Reliability Measures

Table 2: Factual Reliability for Coded Student Interviews

<table>
<thead>
<tr>
<th>Interview Number</th>
<th>Number of Lines</th>
<th>Total</th>
<th>Same</th>
<th>Different</th>
<th>Missing Codes</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>30</td>
<td>22</td>
<td>20</td>
<td>2</td>
<td>0</td>
<td>91%</td>
</tr>
<tr>
<td>23</td>
<td>30</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>100%</td>
</tr>
</tbody>
</table>
Table 3: Factual Reliability for Coded Instructor Interviews

<table>
<thead>
<tr>
<th>Interview Number</th>
<th>Number of Lines</th>
<th>Total</th>
<th>Same</th>
<th>Different</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>30</td>
<td>22</td>
<td>16</td>
<td>3</td>
<td>73%</td>
</tr>
<tr>
<td>7</td>
<td>30</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>8</td>
<td>30</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>75%</td>
</tr>
<tr>
<td>9</td>
<td>30</td>
<td>26</td>
<td>21</td>
<td>5</td>
<td>81%</td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>17</td>
<td>17</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>Mean Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>83%</td>
</tr>
</tbody>
</table>

Table 4: Inferential Reliability for Coded Student and Instructor Interviews

<table>
<thead>
<tr>
<th>Coder</th>
<th>STU 18</th>
<th>STU 23</th>
<th>INS 2</th>
<th>INS 5</th>
<th>INS 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>69%</td>
<td>57%</td>
<td>56%</td>
<td>67%</td>
<td>74%</td>
</tr>
<tr>
<td>2</td>
<td>62%</td>
<td>56%</td>
<td>70%</td>
<td>72%</td>
<td>52%</td>
</tr>
<tr>
<td>3</td>
<td>54%</td>
<td>51%</td>
<td>54%</td>
<td>67%</td>
<td>44%</td>
</tr>
<tr>
<td>4</td>
<td>62%</td>
<td>50%</td>
<td>71%</td>
<td>72%</td>
<td>59%</td>
</tr>
<tr>
<td>5</td>
<td>62%</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Reliability</td>
<td>62%</td>
<td>54%</td>
<td>63%</td>
<td>69%</td>
<td>57%</td>
</tr>
</tbody>
</table>

Works Cited


language and multilingualism (pp. 73-96). New York: New York Academy of Sciences.


Blogging and Tweeting and Chat, Oh My!
Social Networks, Classroom Culture, and Foreign Language Instruction

Thomas Garza

Introduction: The Age of the Cyber-Student
The first decade of the 21st century has witnessed an unprecedented global explosion in the number of online social networking sites (SNSs) available to anyone with a computer and an Internet connection. SNSs, such as Facebook, LiveJournal, MySpace, Twitter, YouTube, LinkedIn, Bebo, and SixDegrees, among scores of others1, have transformed the mode and medium of person-to-person communication, by making instant, direct – and free – contact with tens of thousands of interlocutors at one time a reality. A 2007 study of the National School Boards Association showed that American school-age children were spending nine to twelve hours per week not simply on the computer, but specifically on social networking sites – about the same amount of time as they spent watching television (Karlin, 1). Significantly, according to more recent research, that number increases to fifteen to twenty hours per week for university students, according to the Chronicle of Higher Education, nearly twice the average time per week that a college student spends studying and preparing for classes (Deresiewicz, 2).

In the past year, the authors of Online Social Networking on Campus Understanding What Matters in Student Culture, Ana Martinez Alemán and Katherine Wartman, claimed that “Higher Education professionals should recognize that online social networking sites like Facebook are part of a larger generational development in computer mediated communication that epitomizes most students on our campuses today.” (67). As Alemán and Wartman further responded in an interview on the subject of students and computer-mediated communication:

“College students use these sites to engage socially in a manner that is not conceptualized as ‘virtual’ and thus not ‘real’, but rather as a digital exchange of cultural norms and their

---

transgressions broadly conceived. Communication that is exchanged on Facebook, for example, is understood by students as ‘real’ with a complex web of rules that guide playful misrepresentation, for example” (Jaschik 2).

Such a claim is crucial to understanding both the extent to which students are willing to reveal certain things – sometimes very personal details – about themselves and about their lives to an open, mostly un-moderated, cyberspace that almost anyone in the world with a computer can freely access, and the amount of time they are willing to devote to these activities as part of their daily life experience. Given that popular SNSs have quickly become multimedia venues in which users can add text (in a wide and growing variety of international alphabets), photos, sound, video, or even stream live video, it was only a matter of time before the use of these sites would become the leading means of student interaction on and off campus. Indeed, initial forays into the use of social media, notably Twitter and Facebook, in foreign language and culture classes at the University of Texas are already being implemented and are discussed in this work.

The extensive use of these SNSs globally cannot be overemphasized in the larger picture of world culture and current events. The 2008 election of Barack Obama as President of the United States, as well of many other political figures in the U.S. Congress and state and local governments, was facilitated and reported to a large extent via social network sites such as Facebook and Twitter, which informed, persuaded, and mobilized large numbers of American voters, especially younger voters, to go to the polls in November 2008, and once again in 2010. In an age of political apathy, the inclusion of such popular modes of communication and networking proved to be a much-needed impetus to move record numbers of young Americans to political action – on their terms.

The use of SNSs for political purposes is not limited to the U.S. Who can forget the striking images in the global media of massive protests in Iran during the 2009 presidential elections, or the even more recent demonstrations on Tahrir Square in Cairo, Egypt, which ultimately led to the resignation of former president Hosni Mubarak and the establishment of new government? Both of these events were largely organized and conducted remotely on SNSs; in particular, via Twitter. Such use facilitated both far-reaching communication and calls for action and organization, while maintaining the anonymity of its users. It is doubtful that the mass gatherings of protestors and international press at these remotely organized protest meetings all over Tehran and the surrounding areas
would have been possible at all without the facilitation of social networks as a means of instantaneous, widespread, and anonymous communication. Such specialized uses – and their subsequent outcomes – are hallmarks of a technology and a medium that can change the face of our societies, our cultures, and our world. Given such potential impact, imagine the potential effect such media might have on the process and products of education around the world.

The Darker Side of the Cyber World
In the extended global community of the new millennium, the need to be and stay connected to one another seems to have risen to the forefront of world consciousness. In the U.S., there was without a doubt a palpable rise in this sentiment following the attacks of 9/11, in which millions of lives from around the world, literally, were at stake, and the need to be in immediate communication with others became of paramount importance. However, between the ubiquitous use of cellular telephones – especially the growing number of “smart” phones – and the Internet, it seems that SNS use in the last five years has taken on a somewhat different importance. No longer merely a means of being or staying in touch, online SNSs now replace much of social interaction that was previously the domain of face-to-face contact in our cultures. The act of “hooking up,” or making a connection, online has become shorthand in our students’ parlance for the larger world of “cyber-dating” in our university culture. Entire relationships can be initiated, conducted, and ended online, eliminating the need for actual face-to-face human contact. William Deresiewicz, for example, comments: “With the social networking sites of the new century, the friendship circle has expanded to engulf the whole of the social world, and in so doing, destroyed both its own nature and that of the individual friendship itself” (8).

Further, some students have taken online dating to the next extreme level of soft-core pornography through posting inappropriate photographs or other materials, and through “sexting,” the sending of instant messages with embedded provocative-to-pornographic photographs into cyberspace. These images sometimes find themselves onto dubious websites, well beyond the scope of social networks. Indeed, the laws in many states are finding it difficult to keep up with the changing face of online interaction and communication, while not violating the First Amendment right to freedom of speech.

The appearance of such inappropriate materials on SNSs has become an unfortunate feature of university cyber culture in the 2000s, to such an extent that some universities, such as American University in Washington, D.C., have added
dedicated institutional websites to combat the inappropriate use or exploitation of SNSs for illegal or improper purposes. Further, since many legitimate future employers and educational providers also now turn to SNSs to get a more complete profile of a potential employee or student, American University gives its students advice on how to “de-tag,” or remove questionable material from one’s personal pages on SNSs, including an admonition on the hubristic posting of dubious photos on oneself on one’s personal pages: “Consider the image your postings project. What would a prospective employer think of you? A scholarship awards committee? De-tag or remove inappropriate pictures of yourself.” Another warns:

Watch what you say and post about others. False statements can subject you to charges of defamation or libel; photos posted without someone’s knowledge and consent can be construed as an ‘invasion of privacy.’ Generally, follow the ‘Golden Rule’ – Treat others as you want to be treated (See “Saving Face: Face Saving Tips,” 1).

Ironically, the very cyber-medium that first purported to liberate the student from the physical constraints of interacting with his or her peers was now at the center of controversy in facilitating illegal use of the Internet, to the detriment of some of its users.

**Questioning the Effect of SNSs on University Academic Performance**

Further, a growing number of U.S. educators are now concerned that the skyrocketing rise in use of the SNSs of its university students will certainly lead to a corresponding proportionate drop in academic performance. At the very least, they worry that students will simply find less time to concentrate on their studies, due to the attractive alternative of being online with friends. Indeed, Aryn Karpinsky of Ohio State University conducted a study that suggests those students who used Facebook on a regular basis had overall lower grade point averages than those who did not. Further, the study concluded that overall academic performance was lower for over half of the students surveyed who use the SNS. Some students in this limited survey actually used the term “addiction” to describe their relationship to Facebook in their daily lives, being unable to go through an entire day without checking in online with their friends on the site (Karpinsky 2 - 3).

Very recent research, however, adds a different, seemingly contradictory, dimension to the growing indications that activity on SNSs such as Facebook has
no adverse impact on academic performance, and might indeed have certain developmental advantages in terms of a student’s ability to interact with his peers. One study conducted at Northwestern University on one thousand students at the University of Illinois at Chicago found no connection between the amount of time spent online on SNSs, such as Facebook and MySpace, and academic performance (Truong). Northwestern University researcher, Eszter Hargittai, states that the benefits of using social-networking sites might cancel out the distractions they pose:

You could go on there and waste your time. On the other hand, you can connect with your classmates, get information about homework assignments, get to know people better, and feel more comfortable engaging with them on academic matters (Truong, 1).

Another current study suggests that SNS use might actually be expanding the opportunities and the means for college students to interact in an increasingly large sphere of personal communication. In his 2010 survey of over nine hundred current college students and recent college graduates across the U.S., professor of media studies at the University of Texas, S. Craig Watkins, reports that their findings indicate that Facebook is actually not replacing face-to-face interactions that students have with friends, family, and colleagues. Watkins states, “In fact, we believe there is sufficient evidence that social media afford opportunities for new expressions of friendship, intimacy and community” (1). He contends:

Whereas engagement with Facebook four years ago was principally about connecting to a small sphere of friends, the use of the platform today includes a broader range of activities, such as communicating with friends and family, collaborating on school work, browsing photos and videos, playing games and quizzes, consuming news, and participating in civic life (1).

Watkins also offers some evidence to support the use of the Facebook platform for intercultural and international communication. As he reports on the actions of the students in his research:

When asked about the type of communication they engage in on Facebook, forty-seven percent of survey participants cited their communication with friends who live in a different state or country as ‘very important,’ while 28 percent cited
communicating with friends who live in the same city as ‘very important.’ Thirty-five percent of survey participants cited communicating with family members, such as parents, aunts and uncles, as very important’ (2).

This “softening” image of Facebook among other SNSs in regard to its impact on the academic and social lives of students raises new questions about the efficacy of using these social media in the classroom. Five years ago, such a discussion of increasing SNS use would have seemed to be anathema to academic success in the university setting. However, with growing evidence that there is no causal link between SNS activity and poor academic performance, the potential for incorporating social networking sites into course curricula is becoming a reality, especially in view of their remarkable presence in other spheres of current events and public discourse.

**SNSs in the American Classroom**

Interestingly, a growing number of experts in the area of SNSs use and education, such as Alemán and Wartman, indicate precisely such a trend in education:

SNS will become an instructional tool soon. Facebook has already partnered with a course management system; some faculty have begun to use Facebook groups to foster peer learning, conduct group projects, etc. Computer mediated technologies have already made it necessary for academic faculty to modify or simply transfer traditional modes and norms of real-life academic and pedagogical communication online.” (Jaschik 5).

Certainly, this view is supported by the increasing trend in the U.S. educational culture to incorporate into their courses online management systems, such as Blackboard, WebCT, Virtual-U and Learning Space, to deliver course content materials ranging from course syllabi to lecture notes, to keeping attendance records, and handling student communications. In fact, SSNs have already begun to be incorporated into our university courses at all levels, bringing the technology and the culture of the sites to our day-to-day classroom operations.

My own recent classroom experience in attempting to turn this seemingly unstoppable tide of our students’ incessant use of social networking technology –
even during class – began after a conversation with another foreign language colleague of mine, who was commiserating with me over the growing lack of students’ attention in our larger classes due to the use of the Internet during class. In particular, we were deriding the relative newcomer to the social networking family: an instant messaging and posting site called Twitter. While we both joked that any format for personal communication that reduced normal conversation to telegraphic utterances of only 140 characters, called “tweets,” signaled the certain demise of Western civilization, my colleague half seriously suggested that the only way we might ever change the culture of the new American classroom was to begin to use this kind of technology as part of our instruction. As we both laughed at this suggestion, I suddenly wondered aloud, “Why not?” My colleague countered with, “You could use it during your larger lectures to see what students are thinking!” Suddenly, the seemingly ludicrous idea became a genuine pedagogical challenge: Could it actually work? Could a social networking device become an instrument for pedagogical innovation in a standard large lecture classroom?

That same semester, in spring 2009, I was halfway through the term of teaching a large lecture course on science fiction in Soviet and Russian literature and film to a group of nearly one hundred undergraduates at the University of Texas. At the beginning of class, I broached the idea by first asking how many of them used Twitter on a regular basis. About half of the class answered affirmatively. When I then asked how many used their computers to send emails or instant messages during my class, nearly all raised their hands, without embarrassment or hesitation. I then asked if they would like to participate in an experiment in using the social networking site Twitter during class as a way of commenting on the content of the material being presented and asking questions; nearly every student’s hand went up. Before the next class meeting, I set up a Twitter account for the class, and began my lecture with a brief explanation of the process. The basic rules for using Twitter in my class were simple: No profanity or obscenities could be used, and no personal attacks on other students or their comments could be made. Students could comment on any part of the lecture: the lecture itself, the slides, the film clips, or the music that I presented during the class or as part of any homework assignment. The auditorium in which I taught was equipped with two projection screens, so I dedicated one of them to the Twitter page so that students could read each other’s comments in real time throughout the entirety of the class. The lights went down, and then we began.
The lecture that first day was on Andrei Tarkovsky’s film *Stalker* and the Strugatsky Brothers’ original story “Roadside Picnic,” which inspired the cinematic adaptation. I preceded the beginning of the lecture with a short clip from the film in which the main characters enter “The Zone” for the first time. As the film was playing, immediately comments began to appear on the Twitter site, which I could read in the dark. Some of the original student comments included: “Creepy. Reminds me of Mad Max,” wrote one student. “Doesn’t the Zone mean something in Russian?” another added. “Why doesn’t anyone shave in these movies?” Responding to one of the earlier tweets, yet another student responded, “It’s more depressing than Mad Max, more like Solzhenitsyn.” And as the clip came to an end, one student opined, “What’s up with the blue filter the director’s using??” And so they went; the flow of tweets on the screen was about five to ten every minute, more than five times the numbers of questions that I might get after an entire lecture, much less after a single film segment.

By the end of the four-minute clip, I was ready to start my planned lecture, but I had in front of me a strange new kind of informed entry to the material I had planned to talk about. Any fear I had that I would not be able to cover the material that I had planned disappeared instantly. Not only was I able to say everything that was already in my lecture notes, but I could adjust my comments to respond directly to the students’ postings in real time. As expected, a few of the postings were meant only to be humorous, others missing the point of the material entirely; but the majority of comments were germane and interesting commentaries or questions on the film, many of which would never have been expressed in the typical way of responding to the usual request for “Questions?” at the end of every class.

The comments the students made carried a great deal of substance and served as content organizers for my lectures. I honestly can say that I never really changed the basic content or message of my lectures, but found that I was able to craft them in to tailor-made disquisitions that fit the general associations that the class was bringing to the topic. This revelation was possibly the most striking for me in my experiment; I had been concerned that my lectures might become diffuse and disorganized as I tried to respond in real time to the plethora of questions and commentaries for the students. Instead, I found that the comment only helped me to focus my delivery on the very ideas and arguments that the students themselves had in mind. The addition of Twitter had actually *facilitated* my own presentation of the course content in a way that I believe improved the students’ comprehension of the material.
In fall 2009, I used the same technique of incorporating Twitter throughout the entirety of my larger 150-person course on the vampire myth in Slavic cultures, which is actually more of an introduction to Slavic civilization, religion, history, literature, and film than an exegesis on vampires and vampirism. In this class, in spite of the significantly larger size, students used the dedicated Twitter site with great frequency and enthusiasm. (See Figure 1 for additional comments from the actual class.)

While such attempts to use SNSs in a truly academic setting may appear to be purely acquiescent – giving in to popular student habits of their social culture – I contend that these efforts to use social networking media may be the best way to win back the integrity of our lecture halls, and to combat a growing tide to cancel smaller language programs due to falling enrollments and lagging funding. As Alemán and Wortman state, “Though not social online media, online course technologies add another dimension to students’ online habits and customs, and explicitly inject faculty into online interactions with students” (6). This assertion is quite correct, and one might further add that without such engagement with our students’ culture, we may lose the ability to teach them as effectively and efficiently. Indeed, in the current economic climate, without the benefits that technology might offer us, we may not have foreign language classes to teach in the near future.

These first experiments have involved courses with substantial cultural content in and about the Slavic and Russian-speaking worlds; the next logical step would be to integrate Twitter and other SNSs in Russian language and culture courses. However, due to the initial Latinate orientation of the Twitter site, the classes had all been conducted in English, and the messaging language used to communicate on Twitter has been English as well. Now that most social media, including Facebook and Twitter, are fully compatible with non-Latinate alphabets, the next step is to see if the use of an interactive SNS would have a similar positive effect in the teaching of a language-based course. Before 2009, such experiments using Twitter in foreign language classes would have been limited to languages with Latinate alphabets, such as English, Spanish, German, and French. But Twitter and most other SNSs have incorporated the use of other character sets, such as Cyrillic, making the use of the medium in a Russian language classroom fully possible. With this added functionality, Twitter now can also be used in language courses and in content courses in the language to facilitate student input, while allowing the instructor to monitor not only the questions and comments raised by the students, but also the quality of the language used. Twitter can now be used both for students-instructor interaction,
as well as student-student interaction – in and after class. Thus, for example, in an advanced conversation course on Russian film conducted in Russian, Twitter could be used not only to survey and monitor student input during a given class session, but also to facilitate group work in class – all in Russian, or to conduct short-answer surveys during class. The limit to such in-language activities are limited only by the creativity of the instructor – and the 140-character limit of the Twitter medium!

“For All Intensive Purposes...” (sic)
If necessity truly is the mother of invention, then the global economic crisis is a godsend for classroom policies and practices. As budgets at the University of Texas tightened, and the maximum number of students in each class increased, language programs found themselves on the horns of a dilemma: find ways to cut costs in the delivery of language instruction – without sacrificing proficiency goals – or risk losing the language entirely from our course offerings. Such dire choices did, however, engender inspired ideas. Hearkening back to the ideas of optimizing the teaching and learning of foreign language skills from the 1980s and ‘90s, Soviet foreign language pedagogues such as Galina Kitajgorodskaja, Efim Passov, and Aleksej Leonitev are once again relevant to the contemporary foreign language classroom.

The old Soviet notion of intensifying classroom instruction with a focus on the exposure of students to authentic multimedia materials from day one of instruction, and orienting instruction on active reception and production of language on communication are at the center of the new iteration of the intensive classrooms being inaugurated in current language classes attempting – at the University of Texas – to fulfill a two-year foreign language requirement in one, while increasing the number of contact hours from the traditional semester classes, and not sacrificing the proficiency gains of a conventional four-semester program of instruction.

The 21st century model of the Intensive Course has changed significantly from its Soviet predecessors, although the fundamental premises of instruction remain the same: make the student the focus of classroom activity, reserve the bulk of grammatical and lexical study for home, and make the classroom environment fully interactive, using realia, authentic texts, and extensive media, including online resources, to facilitate numerous and varied encounters with the language and culture in context. By giving students multiple opportunities and environments to engage with the language, the chance of getting the student to “buy in” to instruction and the intensive methodology is greatly increased. In the
case of Russian, the first such intensive course was offered in fall 2010. Twenty-five students signed up for the course; at the end of the drop-add period in the third week, there were twenty-six students in the course.

Crucial to the success of this Intensive model is a collegial and supportive *esprit de corps* among the students, especially if the number of participants in the course is large. While in-class activities should facilitate the day-to-day interaction of students in the course, the out-of-class use of course-related SNSs, such as a dedicated blog site, is essential. Besides interacting with authentic website material for homework each day, students also have access to a dedicated course blog site, which is set up for peer interaction. Here, users can ask each other questions about class, homework, or any other assignments or materials used in the course. They can use either English or Russian, as they see appropriate, although use of Russian-only significantly increased by the third month of class. For the University of Texas Intensive Russian course, such interaction is provided via a dedicated Facebook page, which allows both the instructor and individual students to post issues, questions, and comments on a wide variety of course-related topics and then to share their responses with all of the other students taking the course. These exchanges can range from simple questions about homework assignments, to much more complex comments on strategies to read a Russian text. [See Figure 2.]

In this first year of what will undoubtedly be an ongoing curricular sea change of adding more and more intensive-format language courses to our offerings, the Russian experiment is going well. Although assessment from the first semester of the intensive program is still underway, some initial indicators are creating a positive portrait of the students’ language gains. The first round of pro-achievement assessments and Oral Proficiency Interviews put the twenty-six students in the intensive course significantly ahead of their peers in the “regular” two-semester sequence of first-year Russian. While most students in the intensive class, like other students completing the initial year of instruction, fall into the “Intermediate” (ILR S1) category in their OPI assessment, several other indicators of progress in the language outpace those of the “regular” course students. For example, listening comprehension appears to be better in the intensive group, as does lexical breadth. But most significantly is the affective factor of willingness to speak on the part of the intensive group, even on unfamiliar topics and/or in unfamiliar circumstances. Such communicative ability might serve these students very well, especially in an immersion or study-abroad context. As this initial group enters into its second semester of intensive
instruction, much more data will be collected and presented to document their progress.

**Parting Tweets**

As foreign language and culture instructors and faculty begin to increasingly understand the potential educational applications of social networking sites and other online resources to enhance their teaching and research, there is every indication that we can actually begin to change the current, predominant ingrained student culture of using these sites for purely social purposes. By appropriating even a portion of the twenty hours a week that students spend online using SMS instead for academic applications, we may actually have the ability to have an impact, if not change completely, the culture of students using Internet services. Given that it is highly unlikely for our students to tire of or discontinue the extensive use of SNSs during their college days, it is now up to us to enjoin the battle on the same cyber field. We must look at these computer sites not as competitors for our students’ time and attention, but rather as unlikely allies in our efforts to educate them in increasingly difficult circumstances.

If we return to those prophetic words of Marshall McLuhan and his notion of mass media creating a “global village” in our near future, the explosion of social networking sites in the cyber world of our students has certainly contributed to a new global culture of student life in the U.S., a culture that we as educators must learn to use to our best advantage. Given that more than eighty percent of current Internet users reside outside of the U.S. (*Computer Industry Almanac*), this issue is not only for American foreign language educators, but for everyone. The future of online social networking is being determined as you read this sentence; what that future will look like and how it might change foreign language education is entirely in our hands. *Continued on the next page.*
Figure 1:

Names and photographs concealed for privacy considerations.

---

2 Names and photographs concealed for privacy considerations.
Figure 2:

[Twitter conversation discussing Russian language proficiency and reading challenges]
References


Figure 1. Actual Page from Class Twitter Site, “The Vampire in Slavic Civilizations,” September 29, 2009.
Дисплейный текст как форма сетевого общения

В.Г. Костомаров

Нынешняя техногенная эпоха подарила людям чудесную возможность хранить знания не только в письме и чтении, но и более естественно, наглядно и красочно. Аудио-видеозаписи, кино и телевидение, мобильная телефония, компьютер, и особенно интернет коренным образом меняют исторически сложившееся языковое существование людей. Библиотеки стали хранилищами отнюдь не только библиос (книг), но и карт, рисунков, фильмов, кассет, дисков, флешек, различных оцифровок аудио- и видеоизданий.

Весь мир теперь фиксирует, хранит, воспроизводит информацию в виде живых актов общения, создавая даже иллюзию соучастия. Телевидение вслед за кино открыло законы органичного сплава языка (звучка и письма, интонации, характера произнесения и пр.) с «иконкой» – изображением, цветом, движением и прочими носителями смысла. Многие видят в этом угрозу существованию книги (дети не хотят читать!), медиологи полагают, что в истории человечества господство звучащего слова и затем рукописного и печатного текста сменяется звукоизобразительностью. Однако не следует забывать, что письмо и рожденная им богатейшая изощренная книжность отнюдь не погубили первородное звучащее слово. Может быть, лучше думать, что, наряду с устной и письменной коммуникацией воцаряется новая ее форма – сетевая. В настоящее время очень важно добиться органического соотношения этих трех сфер коммуникации.

Наблюдая за телевидением, как наиболее сложившейся аудио-графо-видео сферой, можно признать наиболее очевидным признаком сетевого общения тексты принципиально нового строения. Их отличают: вовлечение языковых (звуковых и письменных) и вневыраженных носителей смысла, упорядоченное чередование единиц разного потенциала (информем и экспрессем), блочный синтаксис и расчленение на удобовосприимаемые клипы (по образцу монтажа кадров в кинофильмах), жесткая конструктивность, своеобразная анаомия (ограничение значения слова, поддерживаемое сопроводительной картинкой), другие признаки.
Стремясь быть интересными и доступными, массовокоммуникативные тексты имитируют межперсональный акт общения элементами разговорного диалога, приметами интимности, непринужденности, доверительности. В то же время, не ограничиваясь повседневно-житейской и художественно-литературной тематикой, они вынуждены обращаться к книжному языку высоких материй – научных, официальных, государственно-политических.

Положение этих текстов среди таких объектов стилистики, как письменная и устная формы текста, книжная и разговорная разновидности самого языка весьма своеобразно. Авторы, обратившие на них внимание, именуют их синтезирующими, синтезированными и даже синтетическими (что разумно подчеркивает две их стороны – «полученные путем синтеза» и «искусственные»), диффузными, креолизованными, экранными. Последнее не кажется удачным: хотя они предъявляются на экране, он чаще служит носителем, наряду с бумагой, привычных буквенно-алфавитных текстов, как например, в «электронных книгах». Нам кажется наиболее удачным назвать их дисплейными.


В принципе любой текст может быть осуществлен в любой форме. Так, разговор (нормально почти исключительно устный, но сегодня нередко ведущийся в эсэмыках и интернет-чатах; забудем, что многие видят в этом «издевательство» над привычной орфографией) или книжный текст (чаще
всего письменный, но сегодня не реже, пожалуй, реализуемый в звуке) могут порождаться и существовать в дисплейной форме (будь то радио, экран, компьютер, цифровая запись). Хотя эта форма до сих пор преимущественно ассоциируется с массовой коммуникацией.

Неприятие нового, особенно старшими поколениями, – отнюдь не миф. Показательно недавнее сообщение о судьбе творческого наследия И.Бродского: «С текстами все в порядке... Продолжается работа с аудиозаписями и другими нетекстовыми материалами». Школьная традиция признает текстом исключительно книжные тексты, а не вообще любые фиксации опредмечивания коммуникативного действия и полученной в нем содержательной и чувственной информации. Абсурдно не считать текстом стихи, которые поэт декламировал перед микрофоном или видеокамерой, но не удосужился записать на бумаге, да еще подвергнуть их редакции. Даже юристы признают доказательством слухозрительные записи устных разговоров.

Дисплейные тексты – очевидное порождение массовой коммуникации, однако невозможно прямо и однозначно увязать их с нею. Сегодня то же техническое опосредование служит чатам, скипу интернета, звуковым разговорам мобильных телефонов и эс-эм-эс, в которых письмо все меньше соблюдает существующую орфографию. Сетевая коммуникация властно ворвалась в повседневное групповое и межличностное общение, что сейчас особенно очевидно в блогосфере. Разумеется, это сопрягается со сдвигами в функционировании и даже составе языка, которые пока еще не только не исследуются, но просто не замечаются или с негодованием отвергаются. Из плана сопоставления дисплейных текстов и книжных, оформленных букво-алфавитным письмом и даже озвученных, читаемых вслух, жизнь требует перейти в план соотношения персональной, которая перестала быть преимущественно устной и контактной, и массовой коммуникации, которая уже очевидно не только письменная.

Иными словами, сегодняшнее «языковое существование» общества не ограничивается личностной (отчасти групповой) и массовой (во многом тоже групповой) коммуникацией. В нем бурно развивается именно сетевое общение, как массовое, так и индивидуальное. Соответственно следует ожидать нового соотношения разговорной и книжной разновидностей литературного языка, и их взаимодействие и взаимопроникновение (не только в сетевой коммуникации!) уже с очевидностью наблюдается.
Ощущая все же некую неловкость от своеобразия языка в дисплейных текстах массовой коммуникации, мы склонны обвинять журналистов в отсутствии стилистического слуха, в порче литературного языка. Наивно полагая, будто всегда и всюду действуют одни и те же законы применения языка, мы недоумеваем, сталкиваясь с нелепыми переносами специфических законов его функционирования в mass-media, столь мощно и бесконечно вторгающихся каждый день в нашу жизнь. Мы злорадно выискиваем в газетах, радио- и телепередачах ошибки, несуразности, действительные или мнимые неудачи, но нам не приходит в голову собирать, цитировать их удачные фразы, находки, тексты, которых, безусловно, не меньше. По своей общественной природе именно mass-media всегда, а в нашу карнавальную эпоху особенно – это опытная (чаще всего смеховая) лаборатория опробования потенций языка. В ходе экспериментирования неизбежны удача, но встречаются и чудеса, которые – не менее, чем художественная литература, – обогащают и общий литературный язык.

В то же время нельзя одобрить модное сегодня бездумное перенесение языковых особенностей дисплейных текстов на тексты иного оформления. Законность своеобразного функционирования языка в дисплейных текстах и незаконность бездумного, неосторожного распространения характерных особенностей такого рода текстов на тексты в других формах кажутся очевидными, потому что они находятся сегодня в разных весовых категориях. Оказавшись в зоне абсолютного влияния mass-media, мы все – вольно или невольно ориентируемся на их язык как на образец для всеобщего подражания. Определенная деформация языка в текстах дисплейной формы, так же как, скажем, обособление языка в письменных текстах науки, производства, бизнеса, делопроизводства, и устных текстах непринужденно-естественнного быта или в художественных, поэтических текстах, в конечном счете представляется неизбежной.

Напомним, что принято даже говорить о разных «языках» (о «книжном языке» с его прагматическими стилями, о «разговорной речи», о «языке художественной литературы») или, что достойнее, о разновидностях применения единого русского языка. Не породят ли дисплейные тексты (подобно тому как письменные тексты породили книжность) еще одну разновидность применения единого русского языка – «массовокоммуникативную»? Главной ее особенностью, как нетрудно предугадать, будет принципиальная техническая опосредованность и принципиальная способность к синтезу (во всех значениях этого термина,
давшего определения синтезированный, синтезирующий и синтетический) текстовых форм, разновидностей применения языка, неязыковых носителей и передатчиков смысла, прежде всего звука и изображения.

Вряд ли при этом утеряется книга и книжный язык потому что дисплейные и книжные тексты различаются по характеру восприятия. Стоит вспомнить, что звуковое общение не исчезло с появлением письменной книжности.

Книжные тексты передают информацию косвенно, возбуждая в читателе разум и воображение, знания и впечатления собственного жизненного опыта, домысливание, догадку, критическую оценку. Чтение книг, то есть их дешифровка, – это напряженная работа, которой надо специально обучать, но зато получаемые сведения становятся частью сознания и мировоззрения личности. Лучше усваивается полученное трудом.

Разумеется, характер чтения зависит и от того, что читается. Ведь чтение – это и восприятие массокоммуникативных текстов, прежде всего газеты, являющейся родоначальницей массовой коммуникации в нынешнем понимании слова. Еще И. Кант придавал чтению газет (в отличие от чтения серьезных книг) отвлекающее значение, для отдохновения. Надо думать, что «чтение» таких массокоммуникативных источников, как газета, тем более реклама, отличается от настоящего чтения. В то же время несомненно, что отнюдь не все тексты, получаемые с экрана (который служит и просто носит элемент письменных текстов, целых книг для истинного чтения), – это тексты в дисплейной форме.

Учить восприятию дисплейных текстов, имитирующих контактное общение, не надо, зато получаемая информация часто скользит мимо сознания, не вызывая скепсиса, слабо задерживается в памяти так же кратковременной и редко переходит в долговременную, связанную с составлением собственного мнения. Затемняется сиюминутность, преходящность, часто мимолетность и поверхностность сведений, не прошедших поверхку разумом; они явно не становятся органичным элементом сознания. Этим дисплейные тексты отличаются от естественного прямого устного общения, где есть непосредственное участие, и от книжных текстов, в которых для постижения смысла требуется разум, воображение, усидчивость. Привычка получать информацию в легкой для восприятия форме невольно отучает от умственного напряжения, необходимого при чтении. Потребитель дисплейных текстов («user») забывает, что они крайне опосредованы, что на самом деле он видит лишь то, что ему показывает
Дисплейный текст как форма сетевого общения
В.Г. Костомаров

оператор, режиссер, редактор, монтажер, в конце концов, владелец канала. Даже если они и далеки от сознательной манипуляции его сознанием, он лишен существующей при чтении возможности критически обдумать, трезво взвесить ценность и истинность ярких клипов, рекламно-красочных сведений, эмоционально поддержанных музыкой, цветом, движением, изображением. Он не может даже просто переспросить что-либо как при непосредственном живом общении, которое имитируют телеведущие. Трудно сказать, меняют ли ситуацию разные способы компенсации отсутствующей «обратной связи» — звонки в студию, ток-шоу с приглашением телезрителей, «поля чудес» и тому подобные приемы.

Таким образом, информация из дисплейных текстов получается непосредственно, без специальной подготовки. Благодаря легкой доступности, краткой дробности, яркости дисплейные тексты создают своеобразную мыслительную сферу (некое «клиповое мышление»). Они отличны от естественных разговорных и от книжных, но в то же время зависимости от тех, и от других. Форма их осуществления своеобразна, подчеркнуто связана с техническим опосредованием. Однако, при всей своей доступности, легкой воспринимаемости и других достоинствах дисплейные тексты не заменяют книгу в воспитании личности, они противопоказаны для философии и поэзии, для души и религии.

Совершенно очевидно, что тексты в разном оформлении не взаимозаменяемы. У них есть свои достоинства и недостатки (вспомним еще «тексты» театрального, балетного, мимического действия, живописи, музыки), и они должны мирно сосуществовать и дополнять друг друга для воспитания полноценных людей. В этом свете ясно, что волнующая людей диллемма «книга или дисплей» некорректна. Книга и Дисплей (как и разговорное звучащее слово) обречены жить и развиваться в содружестве. Задача ввести эти соображения в образование становится принципиальной.

Сейчас модно ругать детей, которые мало читают, но они не стали хуже, они просто предпочитают меньше напрягаться и веселее жить. В данном случае необходимо возрастное опережение: сначала научить читать, а потом допустить до компьютера, причем отдавая себе отчет, что дети в обоих случаях будут играть, а не вдумчиво постигать серьезные материи. Важно создать общее мнение, что грамотно писать столь же престижно, что и владеть компьютером, что быстро искать нужное в Google так же достойно, как и в словаре или энциклопедии. Задача научить совмещать эти навыки имеет общецелевой характер и активно осмысляется. Обход фазы умственного напряжения и критичности восприятия открывает
в потенции возможность отвлечения от подлинного смысла, а то и случайного или сознательного искажения его, незатрудненного навязывания чужих идей, обольщения неопытного зрителя. Но и без злого умысла легкость и мнимая убедительность экранной информации ведет к болезненному дефициту внимания – ADD (Attention Deficit Disorder). В США активно ищут способы борьбы с этим ранее редким заболеванием, сейчас охватывающим массы детей и подростков.

Конечно, и у книги есть схожие недостатки (вроде «чтения взапой»), но они не столь острый. Не подменяя собою чтение, телевидение и Интернет обязаны поднимать его престиж, обучать ему. Ведь и сами электронные средства организационно и содержательно опираются на письменную книжность.

В заключение можно сказать, что утрата письменностью роли единственного и потому обожествляемого способа фиксации и хранения добытого в общении людей знания не может игнорироваться. В то же время вербально-картинно-образная передача информации во многих случаях оперативнее и желательнее при нынешних быстротекущих событиях и автоматизации. Разве плохо быстро получать нужные сведения и справки в краткой и яркой форме?

Дисплейные тексты, не покушаясь на книжные, завоевали свое место в нашей жизни, а их отрицательные (впрочем, так уж обязательно отрицательные, просто непривычные?) свойства сожительствуют и с несомненно положительными. Добиться мирного сосуществования и сотрудничества всех каналов умственного обогащения – важнейшая задача семьи и школы, всего общества, обязанного во имя здоровья будущих поколений строго структурировать современное информационное пространство, создавая единую систему хранения и передачи знания.

Литература
Кара-Мурза Е.С. Русский язык и современная потребительская реклама // Журналистика и культура на переломе тысячелетия. – М., 2002.
A Case Study of the Acquisition of Narration in Russian:
At the Intersection of Foreign Language Education,
Applied Linguistics, and Second Language Acquisition\(^1\)

Benjamin Rifkin

Introduction

Studies of students’ foreign language proficiency—including Carroll (1967), Magnan (1986), and Thompson (1996)—have shown that students in their fourth year of language study typically demonstrate oral proficiency in the intermediate range. Thompson’s study found the median score of students’ oral proficiency in the fourth year to be at the intermediate high/advanced threshold, but her subjects at this level of instruction were students at Middlebury’s summer language program and were tested in the last week of the program. In my own teaching practice, I have found that most of the students majoring in Russian at the University of Wisconsin-Madison graduate with an oral proficiency rating at the intermediate-mid level, sometimes at the intermediate-high level, and very rarely at the advanced level. This usually occurs only with very gifted students who have spent an entire year in Russia. The majority of students in our graduate program who enroll directly upon completing their BAs demonstrate the same range of levels of oral proficiency as those completing their BAs at our university, regardless of their undergraduate institution (excluding from consideration graduate students who are native speakers of Russian or another Slavic language).

Research conducted on the basis of data collected from participants in study abroad programs sponsored by the American Council of Teachers of Russian (ACTR) suggests that most of the students entering into the study abroad program have only intermediate level oral proficiency, and approximately two thirds of them fail to achieve advanced-level oral proficiency

---

\(^1\) I thank the anonymous referee who helped me with this paper’s organization, Dianna Murphy of the University of Wisconsin-Madison and Olga Yokoyama and Olga Kagan of UCLA for their suggestions. Of course, I take responsibility for any errors of fact herein. Lastly, I am grateful to the University of Wisconsin-Madison Teaching Academy for its generous support of the transcription of interviews analyzed in this project.
by the end of their study abroad program. (Brecht, Davidson and Ginsberg: Table 5, p. 10).

This research project began with the question: Given that students can only achieve advanced-level oral proficiency during a study abroad program or in a summer intensive immersion program such as Middlebury’s, and given that not all students can achieve advanced-level oral proficiency even in those contexts, what can we do in the non-immersion Russian-language classroom to help prepare students to make that leap? This is a question in the discipline of foreign language education, but its answer must be informed by research not only in foreign language education, but also by the fields of applied linguistics and second language acquisition.

**Preliminary Investigation in Applied Linguistics**

The ACTFL Oral Proficiency Guidelines (OPG) foreground two language functions, narration and description, as critical or core functions for the advanced level. I selected narration as the function of interest for this research project because it seemed to me that narration, especially past-tense and present-tense narration, emerge in the spontaneous speech of learners of Russian before description. Having selected narration as the function of interest, I then set out to explore the nature of narration in the speech of native speakers of Russian, especially with regard to the frequency of complex sentences in their discourse, an examination within the domain of the discipline of applied linguistics (the study of language in the context of its use). In order to elicit narration, I asked several native speakers to retell the plot of a film or television program they had seen recently or one that had made a great impression on them. I recorded their speech and analyzed it for the balance of complex vs. simple or compound sentences, considering as a complex sentence any sentence with any kind of dependent clause. The results of my analysis are summarized in Table 1 (in which all names are pseudonyms):

**Table 1. Native Speaker Narrations**

<table>
<thead>
<tr>
<th>Name</th>
<th>Total Sentences</th>
<th>Complex Sentences</th>
<th>Relative Clauses</th>
<th>Frequency of Relativization</th>
<th>Frequency of Complex Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhora</td>
<td>19.0</td>
<td>10.0</td>
<td>4.0</td>
<td>.210</td>
<td>.526</td>
</tr>
<tr>
<td>Tania</td>
<td>17.0</td>
<td>6.0</td>
<td>0.0</td>
<td>.000</td>
<td>.353</td>
</tr>
<tr>
<td>Lena</td>
<td>12.0</td>
<td>7.0</td>
<td>9.0</td>
<td>.751</td>
<td>.583</td>
</tr>
<tr>
<td>Kira</td>
<td>14.0</td>
<td>10.0</td>
<td>10.0</td>
<td>.714</td>
<td>.714</td>
</tr>
<tr>
<td>Aliosha</td>
<td>30.0</td>
<td>16.0</td>
<td>14.0</td>
<td>.467</td>
<td>.533</td>
</tr>
<tr>
<td>Mean</td>
<td><strong>18.4</strong></td>
<td><strong>9.8</strong></td>
<td><strong>7.4</strong></td>
<td><strong>.403</strong></td>
<td><strong>.533</strong></td>
</tr>
</tbody>
</table>
As depicted in Table 1, native speakers of Russian provided an average narration of 18.4 sentences in length, of which 9.8 were complex sentences; in other words, roughly slightly more than half of their utterances were complex sentences. Of the 9.8 complex sentences, an average of 7.4 sentences featured relativization; in other words, slightly more than three quarters of the complex sentences or slightly more than 40% of all the utterances in the average narration featured relativization. By relativization, for the purposes of this project, I am referring to a complex sentence featuring a relative clause introduced by the relative pronoun который; I recognize that Russian allows for relative clauses introduced by other pronouns, such as что or кто (Timberlake, 1993), but I restricted the focus of this study just to the use of relative clauses with который, because they were by far the most frequent relative clauses in the speech of the native speakers, and in the speech of the students in the control and experimental groups described below. The frequencies of complex sentences and relative clauses in the speech of the native speakers in this experiment’s findings are close to those observed by Zemskiaia and Kapanadze (1978). Most of the texts recorded and transcribed by Zemskiaia and Kapanadze are dialogic in nature, but in those that were monologic and of paragraph-length, 90-93% of the utterances featured complex syntax, but only 20-39% of the utterances featured relative clauses (alternatively, 22-42% of the complex sentences were complex due to the fact that they featured a relative clause).

I then compared the native speakers’ narrations with narrations produced by two of the best (i.e., most communicative) students in the fourth-year Russian class of the University of Wisconsin-Madison, as depicted below in Table 2.

<table>
<thead>
<tr>
<th>Name</th>
<th>Total Sentences</th>
<th>Complex Sentences</th>
<th>Relative Clauses</th>
<th>Frequency of Relativization</th>
<th>Frequency of Complex Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dina</td>
<td>15.0</td>
<td>2.0</td>
<td>1.0</td>
<td>0.067</td>
<td>0.133</td>
</tr>
<tr>
<td>Rita</td>
<td>9.0</td>
<td>1.0</td>
<td>0.0</td>
<td>0.000</td>
<td>0.125</td>
</tr>
<tr>
<td>Mean</td>
<td>12.00</td>
<td>1.5</td>
<td>0.5</td>
<td>0.034</td>
<td>0.129</td>
</tr>
</tbody>
</table>

These two students had intermediate-mid level oral proficiency, so their narrations, of course, could not be expected to compare functionally to the narrations of the native speakers. Nonetheless, I was surprised to find how syntactically weak were their narrations. One of the students had a narration of only nine sentences in length, only one of which was complex; the other student
offered a 15-sentence narration including two complex sentences (one of which had a relative clause.) Clearly, and not at all surprisingly, there was a significant disjuncture between the speech of the native speakers and our students in terms of their control of Russian syntax.

Given the stark contrast between the performance of students in an upper-level Russian course and the performance of native speakers, it was clear that the investigation of a different approach to teaching speaking skills at this level was warranted. The experiment I would design and conduct would need to measure the acquisition of one particular structure, the subset of relative clauses introduced by the relative pronoun который, in students’ narrations, because the use of который-clauses is closely linked to advanced-level oral proficiency in the provisional proficiency guidelines for Russian, and in the interpretation of many ACTFL-certified OPI testers of the generic oral proficiency guidelines for advanced level function. Accordingly, improvement in this aspect of spontaneous speech, while not necessarily in and of itself sufficient for a determination of advanced level proficiency, would be indicative of students’ progress towards that goal.

**Theoretical Background**

In reflecting on the state of the students’ speaking skills, I considered recent research in the area of second language acquisition. Second language acquisition research has long focused on the notion of input, in part due to Krashen’s work in this area (1982). However, even as early as Wagner-Gough and Hatch’s seminal paper in 1975 on the importance of input data, the field began to move toward a careful examination of output as an important variable in second language acquisition. Swain’s 1985 study, one that involved a research design with narration tasks based on the retelling of a film, was critical for the understanding of output. In the conclusion of her analysis of a study of French immersion students in Canada, Swain wrote:

> Comprehensible output...is a necessary mechanism of acquisition independent of the role of comprehensible input. Its role is, at minimum, to provide opportunities for contextualized, meaningful use, to test out hypotheses about the target language, and to move the learner from a purely semantic analysis of the language to a syntactic analysis of it. Comprehensible output is, unfortunately, generally missing in typical classroom settings...(Swain, 1985: 252).
Since 1985, some second language acquisition researchers have focused on output (that is, learners’ oral discourse) as a means of targeting syntax in the context of classroom instruction. The importance of Swain’s argument about comprehensible output is underscored by the fact that it has been examined in numerous studies: Swain (1993); Izumi, Bigelow, Fujiwara and Fearnwo (1999); Pica and Doughty (1985); Pica and Doughty (1988); Pica, Holliday, Lewis and Morgenthaler (1989); Ellis (1992) who devotes an entire chapter to this topic; Nobuyoshi and Ellis (1993); Swain and Lapkin (1995); and Nagata (1998). Nobuyoshi and Ellis’s study describes the value of “focused communication tasks” and is particularly interesting in that the authors conclude that “pushing learners to improve the accuracy of their production results not only in immediate improved performance, but also in gains in accuracy over time” (Nobuyoshi and Ellis: 208). Gass (1997: 139) notes that the idea that output or language use could be part of the learning mechanism was not seriously contemplated prior to Swain’s (1985) important paper, in which she (Swain) introduced the notion of comprehensible or ‘pushed’ output: “...using the language...may force the learner to move from semantic processing to syntactic processing.... [P]roducing the target language may be the trigger that forces the learner to pay attention to the means of expression needed in order to successfully convey his or her own intended meaning” (Swain, 1985: 249). According to what Ellis calls the output hypothesis, “production will aid acquisition only when the learner is pushed. Opportunities to speak may not in themselves be sufficient” (Ellis, 1990: 118). This suggests that the output-focused classroom cannot merely present opportunities for speech, but must create situations in which learners are compelled to speak, to negotiate meaning, and to achieve communicative goals. The intriguing notion of output focused instruction—the notion of pushing and stretching the students in their production of the target language, a conceptual parallel to Krashen’s i + 1 (providing students input that is generally comprehensible plus some input for which they must reach and work to understand) that by analogy might be called o + 1 – motivated the structure of this experiment.

Design of the Experimental Class: A Project in Foreign Language Education
Elsewhere (Rifkin, 1998; Rifkin, 2000) I have described an advanced-level conversation course for Russian in which I cast task-based learning (based on Murphy, 1993) in a framework with consciousness-raising activities. I wrote that the consciousness-raising activities (student reflections on their own speech) “put
the burden of learning squarely on the shoulders of the learners” (Rifkin, 2000: p. 67) and had a positive correlation with student engagement in the learning process. The experimental class I designed was one based on film, rather than written texts, as the primary source of “input” for the learners. Film and video have long been important in the foreign language classroom. Many scholars have written about the contribution of film and video to foreign language instruction in the context of either listening or speaking or both: Baltova, 1994; Danahy, 1985; Donchin, 1985; Edasawa, Takeuchi and Nishizaki, 1990; Garza, 1991; Liskin-Gasparro and Véguez, 1990; MacWilliam, 1986; Parker, 1976; Rueberg-McCoy, 1990, and Terrell, 1993. Moreover, some scholars have written about criteria for selecting films for use in the foreign language classroom: Joiner, 1990; Mount, Mount and Toplin, 1988; Secules, Herron and Tomasello, 1992; and Voller and Widdows, 1993.

I decided to use the films selected for the course as the basis for pushing learners to communicate in the conversation class. In reviewing the literature on output-focused instruction, I speculated that if I could teach this output-focused conversation class for Russian-language students, a class in which students would be compelled to speak, then the students could practice narration tasks while reflecting on the syntax they were using. I hoped that the students in this course might come closer to the native speaker norms for narration and might approach, if not actually cross the intermediate/advanced threshold according to the OPG. If I were to compare narrations in the speech of students in one or more sections of the experimental class (“experimental group”) and students in one or more sections of a conventional conversation class (“control group”), I thought I would be able to see differences in the syntax they used. Accordingly, the research hypothesis for this project can be stated as follows:

**Research Hypothesis:** Students in an experimental group with output-focused instruction and consciousness-raising activities will, after instruction, use complex syntax more frequently and more accurately in their Russian-language speech and, specifically, in their narration tasks, than peers in a control group with more conventional instruction.
The next steps of the experiment were to design the conversation class, and collect and analyze data from the narrations of students in the experimental and control groups.

The design of a conversation class may seem deceptively easy: it might seem that nothing could be simpler than bringing students together to talk with an instructor. However, as noted by Santoni (1975), the conversation class can be very challenging to design: teachers and students without appropriate preparation can lose track of the purpose that brings them together. In my description described as an advanced-level conversation class (Rifkin, 2000), I argued that film can be used very productively as the basis for learning tasks. To summarize my description of this class (Rifkin, 2000: p. 65-66), students in the experimental group watched a series of films over the course of the semester. Each week they saw a single film, but viewed it twice (both times for homework), with different assignments. After each viewing, they come to class ready to describe the major characters, the settings in which these characters live and work, to retell the plot of the film from the perspective of various characters and various points of time, and to argue and hypothesize about the film from various assigned perspectives.

All of these conversation tasks are derived from the OPG themselves. While completing these conversation tasks in class, the students worked in pairs and triads, talking with one another, trying to make each description and narration longer than the previous one. The students tape-recorded their descriptions and narrations; each student has his or her own audiocassette to tape his or her own monologues. After they had gone through a single set of descriptions and narrations with one partner, they changed partners and taped their next set of descriptions and narrations with a second partner. The students were assigned to help one another make their descriptions and narrations as long and as full as possible, using complex syntax (and the devices of coherence and cohesion, see Appendix A) to do so. The students reflected on their tapes periodically throughout the semester, considering most carefully their use of complex syntax and the devices of coherence and cohesion, and writing a brief self-analysis (in English) about their spoken Russian as evidenced on the tapes they have made of themselves. The students’ own reflection on their learning, their analysis of their own tapes, was of critical importance to the success of this course.
Experimental and Control Group Treatments
Students in the experimental and control groups had pre-test and post-test oral proficiency interviews in which they were asked to retell the narrative of a film or television program they had recently seen. These interviews were transcribed and analyzed with regard to the frequency with which the students used relative clauses in narration tasks. Of course, the two groups differed in their instructional treatment. Students in the experimental group were ‘pushed or stretched’ in their speech performance, because they were constantly asked to narrate and describe, argue and hypothesize, using the list of devices of coherence and cohesion (Appendix A), especially the relative pronoun, in pairs and triads, all semester long. Their in-class oral performance was recorded every day and they were required to reflect on the recordings of their speech and write analyses from time to time.

In the control group, students used the textbook Let’s Talk about Life (Tall and Vlasikova), read the texts and completed the exercises in several chapters, including those on university life, sexuality, and substance abuse. Students in that class were assigned to see two of the same films viewed by the students in the experimental class, Malen’kaia Vera and Chastnaia zhizn’. Students in the control group class spent almost all of their class completing the exercises in the textbook and practicing retelling the plot of the two films time in pairs, triads, and small groups. In the pre- and post-test oral proficiency interviews, students were asked to retell the stories of one of the two films.

In the experimental group students were given a list of conjunctions and other ‘connecting words and phrases’ (Appendix A) and asked to use those words and phrases in their speech as they completed narration, description, argument and hypothesis tasks. This class had 45 contact hours during the semester; 30 of these hours were dedicated to narration and description tasks the students were required to complete in pairs or triads for the duration of most of the class session. A total of 15 of the class sessions were devoted to extended argument and hypothesis, superior level functions according to the oral proficiency guidelines. All student speech was tape-recorded for self-analysis and for analysis for this investigation. The students wrote critiques in English of their speech performance in Russian on a regular basis, thus focusing their attention on the length and complexity of their discourse and the frequency with which they used relativization in their speech. These critiques constitute the ‘consciousness raising’ component of this experiment. Ellis (1992: 238) argues that “consciousness-raising facilitates the acquisition of grammatical knowledge
needed for communication by means of three processes: noticing, comparing, and integrating.” In listening to their own tapes and engaging in self-reflection, students were expected to notice and compare, according to Ellis’s terms, and it was hoped that they would then proceed to integrate.

The class from which the control group was drawn had 30 contact hours over the course of the 15-week semester; during those 30 hours, students were asked to complete listening, speaking, reading and writing tasks, but did not spend nearly as much time in extended discourse in the classroom because they were not asked to ‘push and stretch’ one another in the same way as their peers in the experimental class; they were not asked to focus their attention on complex syntax and were not provided a list of devices to facilitate relativization in their speech. Students in the control group studied the structure of relative clauses and wrote grammar exercises with the relative pronoun, in order to demonstrate conceptual mastery of this syntactical structure, while the students in the experimental group were not required to do any mechanical writing or grammar exercises, although they did write two compositions. Lastly, the students in the control group did not record their speech for self-reflection, and therefore did not reflect on their in-class speech performance to the same degree that their peers in the experimental group did.

An Investigation in Second Language Acquisition

Subjects

The experimental and control groups were composed of students in three distinct heterogeneous classes of students at various stages of the completion of their bachelor’s, master’s or doctoral degrees. Students in both the experimental and control groups demonstrated intermediate-mid or intermediate-high oral proficiency in oral proficiency interviews conducted before the experiment. The students in the experimental group were enrolled in two different sessions of the course I described earlier. The first of the experimental classes, conducted in the spring of 1997, consisted of seven American-born women, five of whom agreed to participate in the study (identified as Ira, Sveta, Raisa, Sonia, Nina—all pseudonyms—in Table 3). Of these five, one was an undergraduate who had studied in Russia for an entire year while in high school, while the other four were in various stages of graduate studies in Russian literature and other fields at the University of Wisconsin-Madison, and all of them had studied in Russia for at least a semester. The second of the experimental classes, conducted in the
spring of 1999, consisted of eight students, half of whom were native speakers of Russian.

Table 3. Student Narrations: Experimental Group

<table>
<thead>
<tr>
<th>Name</th>
<th>Pre-Total</th>
<th>Pre-Comp</th>
<th>Pre-Rel</th>
<th>Pre-Comp-Freq</th>
<th>Pre-Rel-Freq</th>
<th>Post-Total</th>
<th>Post-Comp</th>
<th>Post-Rel</th>
<th>Post-Comp-Freq</th>
<th>Post-Rel-Freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ira</td>
<td>6.0</td>
<td>4.0</td>
<td>2.4</td>
<td>.667</td>
<td>.667</td>
<td>18.0</td>
<td>16.0</td>
<td>10.0</td>
<td>.556</td>
<td>.889</td>
</tr>
<tr>
<td>Sveta</td>
<td>6.0</td>
<td>4.0</td>
<td>0.0</td>
<td>.667</td>
<td>0.0</td>
<td>10.0</td>
<td>8.0</td>
<td>5.0</td>
<td>.500</td>
<td>.800</td>
</tr>
<tr>
<td>Raisa</td>
<td>7.0</td>
<td>7.0</td>
<td>5.0</td>
<td>1.00</td>
<td>.714</td>
<td>8.0</td>
<td>6.0</td>
<td>4.0</td>
<td>.500</td>
<td>.750</td>
</tr>
<tr>
<td>Sonia</td>
<td>26.0</td>
<td>4.0</td>
<td>2.0</td>
<td>.154</td>
<td>.077</td>
<td>9.0</td>
<td>1.0</td>
<td>0.0</td>
<td>.000</td>
<td>.111</td>
</tr>
<tr>
<td>Nina</td>
<td>6.0</td>
<td>3.0</td>
<td>0.0</td>
<td>.500</td>
<td>0.0</td>
<td>10.0</td>
<td>8.0</td>
<td>3.9</td>
<td>.910</td>
<td>.800</td>
</tr>
<tr>
<td>Rodia</td>
<td>17.0</td>
<td>4.0</td>
<td>1.0</td>
<td>.235</td>
<td>.059</td>
<td>30.0</td>
<td>15.0</td>
<td>2.0</td>
<td>.066</td>
<td>.500</td>
</tr>
<tr>
<td>Ania</td>
<td>16.0</td>
<td>9.0</td>
<td>3.0</td>
<td>.563</td>
<td>.188</td>
<td>22.0</td>
<td>14.0</td>
<td>7.0</td>
<td>.318</td>
<td>.636</td>
</tr>
<tr>
<td>Mila</td>
<td>14.0</td>
<td>4.0</td>
<td>0.0</td>
<td>.286</td>
<td>0.0</td>
<td>18.0</td>
<td>6.0</td>
<td>1.0</td>
<td>.055</td>
<td>.333</td>
</tr>
<tr>
<td>Kostia</td>
<td>6.0</td>
<td>1.0</td>
<td>0.0</td>
<td>.167</td>
<td>0.0</td>
<td>37.0</td>
<td>21.0</td>
<td>5.0</td>
<td>.135</td>
<td>.568</td>
</tr>
<tr>
<td>Mean</td>
<td>11.6</td>
<td>4.4</td>
<td>1.7</td>
<td>.379</td>
<td>.147</td>
<td>18.0</td>
<td>10.6</td>
<td>4.8</td>
<td>.337</td>
<td>.677</td>
</tr>
</tbody>
</table>

Table 4. Student Narrations: Control Group

<table>
<thead>
<tr>
<th>Name</th>
<th>Pre-Total</th>
<th>Pre-Comp</th>
<th>Pre-Rel</th>
<th>Pre-Comp-Freq</th>
<th>Pre-Rel-Freq</th>
<th>Post-Total</th>
<th>Post-Comp</th>
<th>Post-Rel</th>
<th>Post-Comp-Freq</th>
<th>Post-Rel-Freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vadim</td>
<td>3.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>10.0</td>
<td>3.0</td>
<td>3.0</td>
<td>.300</td>
<td>.300</td>
</tr>
<tr>
<td>Dima</td>
<td>8.0</td>
<td>2.0</td>
<td>0.0</td>
<td>.250</td>
<td>0.0</td>
<td>25.0</td>
<td>6.0</td>
<td>6.1</td>
<td>.240</td>
<td>.040</td>
</tr>
<tr>
<td>Mean</td>
<td>5.5</td>
<td>1.0</td>
<td>0.0</td>
<td>.182</td>
<td>0.0</td>
<td>17.5</td>
<td>4.5</td>
<td>2.0</td>
<td>.257</td>
<td>.114</td>
</tr>
</tbody>
</table>

Of those students who were not native speakers of Russian, three were native speakers of English and one was a native speaker of another European language. All four of these students agreed to participate and they are identified as Rodia, Ania, Mila and Kostia in Table 3. Nine students were enrolled in the control

---

2 Includes two errors in the declension of который in an attempt to create relative clauses in this narration.
3 Includes one error in the declension of который in an attempt to create a relative clause in this narration.
4 One of the attempts to create relative clauses in this narration resulted in an error in the declension of который.
5 Includes two errors.
6 In addition to the one correct relative clause, this speaker failed to use the relative pronoun for relative clauses in which it was required, e.g., Лена встретила, встретила с молодой человек, называется, не называется, зовут Сергей, и она, они влюбились (sic).
class, another conversation class available for undergraduate or graduate credit conducted in the spring of 1997; of these nine, two students agreed to participate in the study. Both were men, and both were graduate students, one in Russian literature, the other in another field. Both had studied in Russia before, one for a summer, the other for a semester, and both were American-born native speakers of English.

Data Analysis and Discussion
While all of the students in both the control and experimental groups demonstrated intermediate-mid or intermediate-high oral proficiency in the pre-test OPI, two of the students in the experimental group demonstrated advanced-level oral proficiency in the post-test OPI. (One of these had a pre-test rating of intermediate-high, but the other had a pre-test rating of intermediate-mid.) These ratings are only advisory, because they were not second-rated by another ACTFL-certified OPI tester. However, the ratings of the students’ oral proficiency are, in fact, not the focus of the present analysis. As we look at the data in Tables 3 and 4, we can see that the speech of the students in the experimental group features consistently more frequent and more accurate use of relativization than the speech of the students in the control group. The learners in the experimental group demonstrated frequencies of relativization approaching the range of native speakers in most cases, while the learners in the control group did not. (I will not argue that the sample size in this research project is large enough to warrant the conducting of t-tests to compare the means, but the numbers are interesting in and of themselves as a case study.) While the learners in the experimental group could not, in the course of a mere 45 contact hours, master both speech functions critical to advanced-level oral proficiency (narration and description in paragraph-length discourse in all time frames), the improvement in their control of relativization suggests that they are better positioned to achieve advanced-level oral proficiency in the context of a

7 I have chosen not to present data on T-units, the minimal terminal unit (as introduced by Hunt in 1965) because, as Bardovi-Harlig argues (1992), sentence analysis more accurately reflects learner performance. In examining T-unit calculations for this set of data, as well as coordination indices (a concept introduced by Bardovi-Harlig), neither measurement adequately conveys the information I wish to present here. T-unit analysis reduces measures of complexity, especially with regard to subordinate clauses, while Bardovi-Harlig’s coordination index fails to capture the use of subordination (attending primarily to coordination, which can be seen to decline as learners use more subordination in their speech and writing).
semester- or year-long study abroad program than learners with less control of relativization. Arguably, the students in the experimental group would show an increased likelihood of achieving advanced-level oral proficiency after a study-abroad experience due to the nature of the output focused instruction precisely because the output focused instruction improved their control of relativization, a defining characteristic of advanced level speech. The students were very keen observers of their own speech, as evidenced by what they wrote in their critiques (see Appendix B).

**Caveats, Constraints and Questions for Further Research**

The sample size in this study is, of course, too small to allow us to generalize to the broader population of students of Russian. This is a case study; it is presented here primarily to spark a discussion of approaches to teaching at the advanced level. With only nine students in the experimental group, two in the control group, and five native speakers with whose speech the student discourse is compared, it would be impossible to argue that the findings reported here are anything more than interesting. Moreover, I am mindful of the fact that the students in the control group were all men. Gender may well be a factor in the students’ performance, and only a larger study, with mixed-gender groups, would clarify the degree to which gender is related to oral performance in output-focused instruction.

Furthermore, the narrations studied only focused on пересказ or retelling tasks based on films the students (and native speakers) had seen in the recent past. It is possible that this task is inherently more complex (or less complex) than other narration tasks for some speakers. It is important to note that students in the experimental class signed up for a more advanced class than those in the control class, as well as for a class whose very name included the word ‘film,’ while students in the control group signed up for a class required for the major in Russian which did not feature the word film in its title. It is possible that students in the experimental group were simply more interested in film as an art and, therefore, performed better on narration tasks related to film texts. Furthermore, students in the experimental class were required to watch Russian films every week throughout the semester, while students in the control group were required to watch only two films in the semester, but were required to listen to audiotapes every week. Students in the experimental group were not asked to listen to audiotapes, and were not provided audiotapes to which they could listen. However, students were not asked to report whether they had seen
any Russian films in the course of the semester other than those assigned in each class (likely), they were not asked to report whether they had listened to any audiotapes other than those assigned in each class (highly unlikely), and they were not asked to report whether they had engaged in Russian-language conversation in any other contexts, such as with friends, in other classes, at ‘Russian Table,’ or in other places (quite possible). It is important, therefore, to revisit this project not only with a larger group of participants, but also with a broader range of tasks representative of both advanced-level language functions, narration and description, including also the function of comparison, which can be used in the context of an oral proficiency interview to elicit either narration or description (and is, indeed, sometimes considered the third function at the advanced level, although it is not a ‘core’ function at this level.)

Despite all these confounding variables, this research project suggests important directions for future study to help move the foreign language profession toward greater success, especially at the intermediate/advanced threshold, a threshold that seems so impenetrable for many of our students. Indeed, not only does our work in this regard affect our students’ foreign language education, but their liberal education more generally, as noted by one of the students in the experimental group: “Believe it or not, I actually think that I am speaking better in English due to all this practice in Russian. Listening back through my tapes I have noticed that I am talking more, longer, and in a more ordered manner. Honestly, I think just getting used to hearing myself speak (in any language, let alone Russian) for more than five minutes straight is a feat.” In order to verify findings here, it would be useful to consider studies with larger subject pools not only in Russian, but in other Slavic languages (and, indeed, in non-Slavic languages), as well as in other instructional contexts.

**Appendix A** Devices of Coherence and Cohesion (Rifkin, 2000)

<table>
<thead>
<tr>
<th>Безусловно</th>
<th>В виду того, что</th>
<th>Действительно</th>
</tr>
</thead>
<tbody>
<tr>
<td>Благодаря тому, что</td>
<td>В конце концов</td>
<td>Дело в том, что</td>
</tr>
<tr>
<td>Ведь</td>
<td>В отличие от чего</td>
<td>До сих пор</td>
</tr>
<tr>
<td>Вместе того, что</td>
<td>В первую очередь</td>
<td>До тех пор, пока не</td>
</tr>
<tr>
<td>Вообще говоря</td>
<td>В результате чего</td>
<td>До того, как</td>
</tr>
<tr>
<td>Во-первых, во-</td>
<td>В связи с тем, что</td>
<td>Если</td>
</tr>
<tr>
<td>вторых</td>
<td>В соответствии с чем</td>
<td>За – до того, как</td>
</tr>
<tr>
<td>Всё-таки</td>
<td>В таком случае</td>
<td>Значит</td>
</tr>
</tbody>
</table>
Из-за чего       Однако       Судя по тому, что
Кажется, что     Одним словом     С одной стороны.... с
Как будто       Перед тем, как     другой стороны
Как говорится    Подобен (подобна,     С тех пор, как
                 подобно, подобны)   С точки зрения
Как ни           чему            (кого?)
Как оказалось    После того, как     Таким образом
Когда            Потому что        Так сказать
Который          По крайней мере     Так, как (так, что)
Кроме того, что  По мере того, как     Тем более
К сожалению, К    По следующим      Тем не менее
счастью          причинам           То, как (то, что)
Мало того, что    По сравнению с чем     Хотя
 Между тем        Прежде всего       Через – после того,
 Например         Прежде чем (+     как
 На всякий случай  инфинитив)        Честно говоря
 На самом деле    Сначала            
 Несмотря на то, что    Согласно чему, 
 Несомненно       Согласно с чем

Appendix B: Excerpts from Student Self-Critiques

Nina: “I think I’ve improved a lot since the beginning of the course, and I actually feel like my speaking ability is back at the same level as it was when I left Russia last summer, and that in itself is quite an accomplishment.... Sometimes I even feel like I can speak a little better than I could at the end of my year in Russia, because I’m getting used to speaking in paragraphs, ordering my thoughts, and using connecting devices, whereas in Russia I just had conversations and was never required to say more than a few sentences at a time. I did start preparing better after my last tape analysis, and I definitely improved my fluency and use of devices of coherence and cohesion on those days when I was well-prepared....My use of connecting devices has increased since the beginning of the course, and maybe even since my last analysis...I’m using more sophisticated devices like с моей точки зрения, несомненно, с одной стороны/с другой стороны, безусловно, по сравнению с, none of which were in my vocabulary before taking this class. I have started using them even more in my writing...”
Sonia: “I feel like I am doing better with descriptions and narrations when I had prepared a certain amount of time beforehand, however, if I prepared too much, I ended up trying to stick too closely to a prepared script. I found myself wasting time trying to remember words I’d looked up, rather than circumlocuting so I could get on with the topic. I heard myself on tape making a fair amount of mistakes in aspect, but then catching and correcting them. I still seem to be doing a lot of pausing and stammering. I repeated как сказать rather often, but did not use который often enough, and when I did use it, it was sometimes unnecessary, for example, “У нее такое лицо, которое выражает...” instead of simply, “Ее лицо выражает...” In general, I don’t think I made as much progress as I wanted to.”

Sveta (whose post-test [unofficial] OPI was rated at the advanced level): “I truly feel that there has been noticeable improvement in my language over the course of the semester. My use of devices of coherence and cohesion has...significantly increased. Furthermore, I feel that I am speaking more comfortably and using more complex sentences. I have grown much more comfortable arguing and debating in Russian, too.”

Raisa: “I found I used the word который [relative pronoun required, in Russian to introduce a relative clause] a lot in my discussion today. I was really pleased I could carry out this sort of abstract discussion in Russian. I remember trying to have these sorts of discussions in Russia (debating political questions, etc.) and failing miserably. I spent so much time formulating every sentence that I would lose my train of thought altogether before I could utter two sentences. Here, although in places I am not particularly graceful, I do manage to keep some sort of argument together. Of course, this is due mostly to our favorite words [devices of coherence and cohesion]; while I am trying to work in the words, I necessarily have to structure my speech logically. Believe it or not, I actually think that I am speaking better in English due to all this practice in Russian. Listening back through my tapes I have noticed that I am taking more, longer, and in a more ordered manner. Honestly, I think just getting used to hearing myself speak (in any language, let alone Russian) for more than five minutes straight is a feat.”

References
A Case Study of the Acquisition of Narration in Russian
Benjamin Rijkin


Assessing the Oral Proficiency of Adult Learners, “Heritage” and “Native” Speakers Using the ILR Descriptions and ACTFL Proficiency Guidelines: Considering the Challenges

Cynthia L. Martin

The ILR Descriptions – Speaking and the ACTFL Proficiency Guidelines – Speaking trace their genesis to an oral proficiency assessment system first developed at the Foreign Service Institute in the 1950s to assess “foreign language users’ abilities according to a consistent scale.” The original purpose was clearly aimed at assessing the abilities of non-native, non-heritage adult foreign language learners, and this genesis is still reflected in the current ILR Descriptions and the ACTFL Guidelines. However, tools based on these descriptions/guidelines are increasingly being used to assess the abilities of language users who are not what we would consider to be “learners” of the “foreign” language being tested, but rather, who have commonly been described as “native” or “heritage.”

1 I offer comments in this essay as an individual, drawing upon my experience as an ACTFL and ACTFL/ILR OPI tester over the past 25 years, primarily in Russian (and in English for nearly a decade), as well as in training and mentoring testers in a wide variety of languages (80+) to use both ACTFL and ILR assessments. The opinions expressed here are entirely my own, and I do not intend to suggest endorsement by any organization, agency or company. The ILR Descriptors may be found at http://www.govtilr.org; the ACTFL Guidelines at http://www.actfl.org.


3 The last 12 years or so have seen a great deal of work in the US on the concept of “heritage” speakers, primarily with a focus toward learning, but less so toward systematic evaluation of a speaker’s functional ability, especially for the professional workplace. For overviews of the notion of “heritage” language, primarily in the US, the reader is directed to the following two volumes:


Very little research has been done to date on the use of ACTFL/ILR assessment tools to assess the non-foreign-language learner. As stated in the essay, Olga Kagan at UCLA with the help of ACTFL is currently conducting such research on assessing heritage profiles in Russian and Spanish using the ACTFL OPI.
Assessing the Oral Proficiency of Adult Learners, “Heritage” and “Native” Speakers  
Cynthia L. Martin

both sets of guidelines, ILR and ACTFL, focus on the speaker’s functional ability to perform certain tasks, there is no reason why assessment tools based on them should not be used to assess the proficiency of any language user, regardless of his or her profile and of how or where the language being assessed was learned or acquired. This focus on function, I would argue, is responsible in large part for the resiliency and flexibility of these guidelines, allowing them to endure for decades.

It is not the purpose of this essay to offer new theoretical conceptualizations or definitions of these terms, but rather to discuss some typical profiles of ACTFL and ILR OPI candidates who are clearly not traditional foreign language learners of the languages being assessed; to summarize current issues being faced in the ILR/ACTFL OPI testing community; and to encourage fruitful discussions about future directions. I shall discuss the following questions currently facing the ACTFL/ILR testing community:

1. What kinds of speakers who are NOT foreign language learners of the language being assessed are currently encountered in the context of ACTFL/ILR OPI testing?
2. What do the guidelines say about “native” speakers?
3. How do the profiles of speakers achieving the same overall ratings differ in terms of the rating criteria depending on whether they are learners, or what we might refer to as “native,” or “heritage speakers”?
4. Should there be a different set of assessments or criteria for “native” speakers, especially at the higher levels?
5. Might a given speaker’s profile matter to the end-user of the ratings? If so, what mechanisms are currently used and/or may be developed to provide this information?

1. What kinds of speakers who are NOT foreign language learners of the language being assessed are currently encountered in the context of ACTFL/ILR OPI testing?

The definition or categorization of different types of speakers has become increasingly complex in recent years, and as stated above, I do not presume here to offer new definitions. Sweeping generalizations, of course, can be quite dangerous and misleading, yet currently, there seem to be three distinct profiles of adult speakers/test-takers who are being assessed using ACFTL/ILR oral proficiency scales 1) adult foreign language learners; 2) what I will refer to as “heritage” speakers and 3) what I will refer to as “native” speakers. Since the profiles of speakers who fall into the two latter categories differ significantly
from the profile of a “typical” American adult foreign language learner for whom the scales were originally developed, the application of either the ILR or ACTFL scales in practice has raised a number of interesting issues for the ACTFL/ILR testing community. In the subsequent discussion of questions no. 2-5 below, I shall use the following “working definitions” of these three diverse profiles.

a. A major segment of the test-taking population continues to be composed of traditional adult learners of the foreign language being tested (i.e., through explicit study outside the target culture, such as at a university or in a special language school, such as the Foreign Service Institute or Defense Language Institute here in the U.S.). The majority, but not all, of these candidates have English as their “native” language. It was for this kind of speaker that both scales were originally designed, so applying the assessment criteria to these profiles is generally straightforward.

b. A second, growing group of test-takers is composed of those I will call “native” speakers who acquired the language as a first language while living inside the target culture, and who completed most or even all of their formal education (usually at least through adolescence, or 14-15 years old) in the target culture in institutions where the primary language of all instruction was the target language being tested. For many languages (but not all), this means the speaker is fully literate in the language and has reasonably high proficiency in all the skills: speaking, reading, listening, and writing. Many of these speakers grew up in bilingual or multi-lingual environments, but usually consider the testing language to be, if not literally their first or “mother” tongue, then at least their strongest and most sophisticated language.

The term “native speaker” has become increasingly problematic as the diversity of linguistic profiles encountered in the ACTFL/ILR testing community grows. I offer here two very important cautions: 1) we must not conflate “native” speaker with ethnicity; 2) we cannot hold the “native” speakers to one uniform standard of pronunciation. Take, for example, the case of English. Speakers growing up in America, Great Britain, Australia, New Zealand, and even in countries such as India, may have vastly different profiles and certainly different pronunciations, though all may consider themselves to be “native” speakers of English. Consider also the case of speakers who grew up in the former Soviet republics, such as residents of Central Asia or the Caucasus region. Regardless of the actual ethnic origins of speakers from these regions, they essentially grew up in a bilingual environment. For many, regional languages may have been spoken in the home, but Russian (not Kazakh, Uzbek, Tadzhik, Georgian, etc.) was the
primary language of education, as well as of professional and public life. These speakers may have regional pronunciations that differ from that of speakers from Moscow, but many may consider their “native” language to be Russian (and their regional languages to be “heritage”) and would certainly score higher on our proficiency scales in Russian than in their “home” languages. Despite the presence of regional accents, would not these speakers be considered “native” speakers of Russian? These are just a few examples, but we could cite many others of the difficulty of linking a “normative pronunciation” to the definition of “native,” as in the case, for example, of the vast diversity of “native” speakers of Spanish, Arabic, Portuguese, and the list goes on. The sheer number of variants among those speakers whom we might consider “native” means that pronunciation alone cannot be the ultimate criterion by which someone’s linguistic identity, especially in terms of “native,” is determined.

c. “Heritage” speakers are generally considered to be those who grew up outside the target culture (or left at a very young age, usually before becoming fully literate in the language) in an environment where the language was spoken in the home, by friends and relatives, but not in the broader surrounding community. Generally, speakers with this profile who are being encountered in the ACTFL/ILR testing context differ from the first group, in that they did not receive their formal education inside the target culture or in the target language, and themselves often consider their “heritage” language to be a second language, or at least not as strong as their “main” language, or language of the society and educational system where they primarily grew up and received formal education.

I wish to reiterate here that the “native” and “heritage” profiles as described above are narrow, and I am using them here exclusively in the context of common profiles currently encountered within the ACTFL/ILR testing context.

2. What do the guidelines say about “native” speakers?

ACTFL Proficiency Guidelines – Speaking

Throughout the ACTFL Guidelines - Speaking, references to “native” speakers are limited to comments about the “native” speaker as the receiver of the message. The ACTFL Guidelines divide “native” speakers into those “accustomed to dealing with non-natives” and those “unaccustomed to dealing with non-natives.” As criteria-referenced guidelines, the ACTFL descriptions are very careful not to have as a criterion any comparison whatsoever of a speaker’s performance or features required to perform certain functions with an idealized or normative performance of an imagined “native” speaker. Nowhere in the ACTFL
guidelines are there suggestions that a speaker must perform a given function “the way a native speaker would” in order to be communicatively successful; quite the contrary, in fact, and the ACTFL guidelines clearly avoid drawing such parallels, focusing instead on how well a given speaker is able to accomplish a linguistic task using the language inside the target culture, communicating with native speakers. Of course, as we move up the ACTFL scale, the kinds of functions required necessarily mean that the speaker is likely increasing his/her ability to function in a target-language environment in more linguistically and culturally appropriate ways, employing the linguistic and cultural norms of that language and society.

**ILR Skill Level Descriptions Speaking**

In ILR Speaking Proficiency Descriptions for levels 0+ through 3, references to native speakers refer either to how the native speaker would understand the person being assessed, or how the native speaker might have to adjust his or her own speech to be understood by the speaker. It is only in the descriptions for 3+ (General Professional Proficiency Plus) and above that we find references comparing a speaker’s performance or linguistic features to that of a “native” speaker:

ILR 3+ says that the speaker “has discourse competence in a wide range of contexts and tasks, often matching a native speaker’s strategic and organizational abilities and expectations.”

ILR 4: “Organizes discourse well, employing functional rhetorical speech devices, native cultural references, and understanding. Language ability only rarely hinders him/her in performing any task requiring language; yet, the individual would seldom be perceived as a native.”

ILR 4+: “Speaking proficiency is regularly superior in all respects, usually equivalent to that of a well-educated, highly articulate native speaker... However, the individual would not necessarily be perceived as culturally native.”

ILR 5: “Speaking proficiency is functionally equivalent to a highly articulate, well-educated native speaker and reflects the cultural standards where the language is natively spoken. The individual uses the language with complete flexibility and intuition, so that the speech on all levels is fully accepted by well-educated native speakers in all of its features, including breath of vocabulary and idiom, colloquialisms, and pertinent cultural references. Pronunciation is typically consistent with that of well-educated native speakers of a non-stigmatized dialect.”
Hence, at the high end of the ILR scale, the descriptions do contain comparisons between the speech of a test-taker and a “native” speaker, raising the kinds of questions already mentioned about how we determine the standards for “native” performance. In question no. 3 below, I shall discuss the concern raised by how these descriptions are perceived by test-takers who self-identify as “native” speakers.

3. How do the profiles of speakers achieving the same overall ratings differ in terms of the rating criteria depending on whether they are learners, or what we might refer to as “native,” or “heritage speakers”?

Since both sets of oral proficiency guidelines are primarily focused on function, it is obviously possible for learners, native or heritage speakers to receive the same global rating while commanding very different skills and having very diverse profiles. Empirical linguistic analysis is required to confirm these observations and others, and there are some investigations into various aspects of different speech profiles currently underway at the Center for the Advanced Study of Language at the University of Maryland, for example, or a study on heritage speakers of Spanish and Russian currently being conducted by UCLA Heritage Center under the direction of Olga Kagan with the participation of ACTFL that may shed light on these differences (see footnote 3). My intention here is to highlight some of the most common differences that have been observed in some testing languages among adult foreign language learners, “heritage,” and “native” speakers in terms of the four general broad categories that form the core of the assessment criteria for the oral interview protocols currently in widespread use: Global Tasks and Functions, Content/Context, Comprehensibility (including Accuracy), Text Type.

**Global Tasks and Functions**

The key component in this assessment category is the ability to accomplish a given task linguistically. All speakers, regardless of profile, who receive the same rating have one thing in common: they can all perform the required tasks for the level assigned and cannot sustain performance of the required tasks at the next higher levels. The other rating criteria serve to support the functions of a given level. Hence, for this particular criterion, the same standard must be met by all speakers achieving a given rating, i.e. they must be able to sustain performance of the functions associated with that level. How they do so may differ dramatically from speaker to speaker. These qualitative differences are manifested in the other rating criteria categories.
All speakers achieving an ACTFL Novice or ILR 0+ must demonstrate the ability to produce memorized, rehearsed language in highly contextualized situations, and can often produce lists of related items. ACTFL Intermediate and ILR 1/1+ ratings are awarded to speakers who can demonstrate the ability to create with the language in order to engage in simple conversation, handle a simple social transaction, and ask questions.

Speakers at ACTFL Advanced and ILR 2/2+ may be learners, heritage or native speakers, but the required tasks, such as narrations, descriptions, handling situations in everyday social transactions, must be met by all. For learners and "heritage" speakers alike, there may be emerging awareness of the culturally appropriate ways to perform various functions, but the required tasks can clearly be accomplished successfully in ways that are not entirely "native" or even "native-like." Native speakers achieving these ratings are most likely able to perform the Advanced/L2 functions required in natural, native or native-like ways across all the tasks, though as we have established, this is not a criterion specified either in the ACTFL or ILR scales at these levels.

The highest rating on the ACTFL scale is Superior, but as in the case of Advanced-level ratings, this does not mean that there is a convergence of linguistic profiles at this level; rather a non-native learner, a heritage speaker, and a native speaker who are all rated Superior may have very different profiles, indeed, while fully meeting the criteria to be awarded a Superior rating. The same may be said of the ILR ratings of 3/3+. The functions that a speaker must demonstrate to achieve ratings of ACTFL Superior or ILR 3/3+ are: discussing an issue at the abstract level, supporting opinion, hypothesizing, and handling “linguistically unfamiliar situations” (usually indicated by the ability to circumlocute around low-frequency/unrehearsed material).

All three groups of speakers under discussion here are expected to command the language to such a degree that culturally appropriate ways of performing the required tasks are now fully operational. This does not mean that the learner, heritage, and native profiles converge entirely, but all three types of speakers demonstrate that they can be equally effective in performing the required functions successfully in the target culture. Native speakers receiving these ratings obviously perform the functions in a native or native-like manner. If the speaker is a learner or a heritage speaker, however, the way in which he/she performs these functions may range from culturally acceptable (i.e., effective and would not disturb the native interlocutors) to native or native-like. This will depend in large part on the speaker’s experiential contact with the target culture. For some learners, this experience may or may not be greater than it is for a
heritage speaker; in other words, learners may, in fact, perform these functions in culturally more authentic ways than heritage speakers, depending on the level of exposure to the target culture and its norms. It is not unusual for heritage speakers with limited experience inside the target culture to perform the functions in ways deemed successful, but that are clearly influenced by his/her dominant, non-target culture language and environment. In short, a single profile is impossible to establish at this level for natives, learners, or heritage speakers.

To be rated at the very top of the ILR scale (4/4+/5), in addition to being able to discuss issues abstractly, support opinions, and hypothesize – all in highly sophisticated ways – the speaker must also perform the function of spontaneously tailoring the language in two different contexts: formally (often in a presentational mode) and informally (usually in a familiar, interpersonal mode). This function is generally elicited using role plays. Lack of ability to demonstrate this function is probably the single greatest obstacle to attaining levels 4/5 for all speakers.

One interesting observation has emerged: most native speakers do not have difficulty with informal role plays, but many do have difficulty producing spontaneous formal tailoring. On the other hand, some high-level learners who have used the language of the test in professional capacities that require formal tailoring, instead have greater difficulty tailoring language informally. In both instances, performance of these functions depends in large part on a speaker’s experience using the language in the target culture in situations requiring tailoring. Given that the assessment is non-compensatory and all criteria must be met in order to achieve a particular rating, failure to tailor either formally or informally will result in a maximum rating of ILR 3+ for all three types of speakers, regardless of profile.

Attempting to identify indicators that the speaker has a broad knowledge of and ability to use appropriately “native cultural references” (ILR 4/4+, ILR 5 uses “cultural references”) has also become more difficult in an age where access to education has been increasingly democratized and the content of most educational systems continues to diversify, not to mention how to account for differences based on factors such as age, region, background, etc. Defining and then designing assessments that could “test” for such things is extraordinarily challenging in a world where “native” speakers are becoming less and less defined by their belonging to a homogeneous group of speakers who share the same cultural, linguistic and even physical or territorial reality. Currently, the best that we can do when testing at the very top of the ILR scale is to look for
spontaneous production of speech that shows the speaker’s ability to use cultural references in rich and robust ways, rather than attempt to elicit production or evidence of “knowledge” of specific references.

Content/Context
This broad criterion is applied uniformly for all three categories of speakers. Speakers who receive ratings of ACTFL Novice/ILR 0+ are limited essentially to basic courtesy requirements and immediate survival needs that have likely been rehearsed; ACTFL Intermediate/ILR 1/1+ (typically learners or low-level heritage speakers with very little actual experience inside the target culture) are limited in content to topics related to autobiographical information and everyday routines, and contexts that are restricted to everyday interpersonal interactions related to these familiar content areas. The content and context areas manifested in the performance of speakers achieving ACTFL Advanced/ILR 2/2+ increase beyond the familiar and personal to include more high frequency concrete topics such as work and current events. At ACTFL Superior/ILR 3/3+, this criteria expands to include most practical, social and professional (not highly specialized) topics, as well as settings where these topics might be discussed. At ILR 4/4+/5, the speaker is expected to operate effectively in virtually all (4/4+) and all (5) settings and contexts that may be encountered in the target culture.

Comprehensibility/Accuracy
This category includes linguistic features such as delivery – pronunciation, intonational patterns, tones for tonal languages, pace or “fluency,” as well structural control including grammar and syntax. These features are assessed in two primary ways: 1) in relation to the level of accuracy required to perform the function: are these linguistic features adequate to perform successfully the communicative task at the intended level? and 2) in relation to what kind of listener would be able to understand the message: very sympathetic, accustomed to dealing with non-natives, or a neutral, non-sympathetic listener? It is only at the highest level of the ILR scale (ILR 5) that pronunciation is explicitly mentioned. The ILR descriptions of 4 and 4+ say that the speaker, despite high-level language use, would “seldom be perceived as culturally native” (ILR 4) or “not necessarily be perceived as native” (ILR 4+), which are often interpreted as including marked foreign pronunciation, among other things. We discussed the complexity of defining “native” pronunciation in no. 1 above, and I shall address the consternation caused in test-takers that self-identify as “native” in question no. 3 below.
Text Type

At ACTFL Novice/ILR 0+, the text type is essentially the same for all speakers: isolated words and phrases and perhaps some sentences or questions that can be used only in highly contextualized everyday interactions, such as greetings and social courtesies. For ACTFL Intermediate and ILR 1/1+, the primary text type consists of simple sentences and strings of sentences. In order to be rated ACTFL Advanced or ILR2/2+, the speaker must control paragraph discourse in order to perform the functions (narrations, descriptions, for example); often heritage and certainly native speakers rated at this level have more consistent control over this text type than many learners. At ACTFL Superior/ILR 3-5, extended discourse must be controlled in order to perform the functions across a range of topics and tasks.

In terms of text-type/structural control, I would like to offer the following observation. For learners and often for heritage speakers, syntactic patterns—even at high levels of proficiency in the language being tested—tend to be influenced by the dominant language of the speaker. In significant testing of native speakers of English who have also achieved high-levels of proficiency Russian, for example, those who score at Advanced-Mid to Superior in ACTFL and 2+ or above on the ILR scale, dominant English syntactic patterns often stubbornly remain reflected in spoken Russian. Many other features, such as lexical control, grammatical accuracy, content/context domains, socio-cultural appropriateness of language use, and even pronunciation and intonation may approach native-like Russian usage, but target-language syntax appears to be among the most difficult thing for adult foreign language learners to acquire.

For example, English syntactic structure is quite different than Russian, and because of morphology, Russian word order is generally thought of as much more “flexible” than English. Furthermore, impersonal, “subjectless” sentences are quite common in Russian, something that English structure does not tolerate. I must repeat here that this observation is not based on empirical studies, but rather, on experience over nearly 25 years of testing, teaching, and interacting with thousands of learners of Russian as well as “native” speakers, I have come to think that this may be the most difficult part of the language to “acquire” and fully internalize. More recent experience with heritage speakers whose dominant language and language of formal education is English, suggests that the same may be true for many of them, as well.

In other words, learners and heritage speakers who have received little or no formal education in Russian and who consider English to be their “stronger” or “primary” language, often share this feature. Certainly, empirical linguistic
analysis is required to confirm these observations, and perhaps the abovementioned study being conducted by UCLA’s National Heritage Language Resource Center with ACTFL participation (looking at ACTFL-rated heritage speakers of Spanish and Russian) may shed light on this phenomenon. Interestingly, the one most common thing that learners and heritage speakers share is the lack of formal education in the target culture; in other words, speakers of these two profiles become fully literate in English while being educated formally in a primarily English-speaking environment. I would expect interesting findings to emerge from a study examining the relationship between the syntactic structure of the written language of a speaker’s formal education and the syntactic patterns of the target language. Is it possible that in this regard, the syntactic patterns of one’s strongest language (in which one has received formal education, especially post-elementary school and in which one has the highest degree of literacy) has a greater influence on one’s acquisition of target-language syntactic patterns of oral speech than when and where and how the language was learned or acquired?

4. Should there be a different set of assessments or criteria for “native” speakers, especially at the higher levels?

This question continues to be discussed in the ACTFL/ILR testing community, but so far, there is not a significant consensus or even strong sentiment arguing for a new set of assessments designed specifically for “native” speakers. However, there has been some discussion of designing a separate assessment just for the upper ends of the ILR scale that would distinguish between native and non-native performance only in specific cases where there would be a need to distinguish between these two profiles. I suggest that if the focus of proficiency assessments is to remain the functional ability of a speaker to perform certain communicative tasks, then there is no reason to develop a separate set of assessments. An entirely new set of guidelines and assessment tools would need to be designed if our goal were to be the ability to assess how well a given speaker’s performance matches that of a “native” speaker. Such an assessment would have to compare a speaker’s performance with some idealized “native” standard, which begs the questions: How would we define “native” speaker for these purposes? On what basis would we attempt to characterize an “ideal” native-speaker performance of specific tasks? As discussed earlier, the complexities involved in defining “native” are increasing in the 21st century, not diminishing. Hence, it seems that designing valid and reliable assessments that would assess whether a speaker is a “native” or how closely s/he approximates
one, would prove virtually impossible, since setting standards would be impossible.

Since the vast majority of organizations seeking individuals with language proficiency for the workplace usually look for ratings in the ACTFL Advanced-Superior ranges and ILR 2-3+ ranges, the need for such an instrument does not appear to be urgent. However, a reasonable suggestion continues to be discussed that, for cases in which a rating of ILR 4/4+/5 is desirable or required for a specific purpose, it may be both prudent and possible to design a separate assessment for any speaker who has already scored an ACTFL Superior or ILR 3 in order to determine whether s/he can perform the functions of 4/5 as described by the ILR Descriptions. In practical terms, the design and administration of such a test would resemble more the official Test of Russian Language (administered in Russia), or the tests administered by the University of Cambridge ESOL (based on the Common European Framework). These tests (at all levels) are targeted to a specific narrow range of levels; that is, a test-taker applies to take an assessment designed specifically to generate a rating at a given level or at least within a limited range of levels, not across the entire scale. Such an assessment would be used only for speakers who had already achieved a rating of Superior/3+ and only in cases where an ILR 4/4+/5 is the desired target. The potential certainly exists for developing new assessments exclusively aimed at ILR 4/4+/5, but to date, I am unaware of any specific plans to do so. This approach, however, may serve to address concerns of high-level test-takers, particularly those who self-identify as native speakers, about the criteria that are actually being used to assign ratings at these levels currently, an issue I shall briefly address here before moving on to question no. 5.

Test-takers’ understanding of the guidelines and assessment criteria: While it is commonly agreed in the testing community that not all native speakers automatically score Superior (ACTFL) or Levels 3-5 (ILR), pressure is being exerted on the community by those taking the test who consider themselves to be “native” speakers to articulate more clearly what the criteria are for these ratings. When educated native speakers read the guidelines/descriptions, they understandably self-assess at the top of each scale. This is less problematic for the ACTFL Guidelines, since it is not uncommon for native speakers educated in the target culture to score Superior in an ACTFL OPI (where there is no distinction after crossing the Superior border, meaning that an ILR 3-5 would all receive the rating of Superior). But when native speakers read the ILR descriptions for Levels 3-5, they also self-assess at the very top of the scale, and are often
surprised to be awarded a rating of “only” 3 or 3+. While this rating may be based on the speaker’s inability to perform certain functions at L4/5 as tested in practice by looking for the ability to tailor language both formally and informally (as discussed above), as well as the lack of spontaneous production of highly sophisticated language when discussing conceptually complex issues and the use of cultural references with facility, these expectations are not explicit in the ILR 5 description. It is not surprising that educated native speakers find ILR 5 the best match for their abilities, when ILR 4 clearly says “would seldom be perceived as a native”, and 4+ says “the individual would seldom be perceived as culturally native.”

Test-takers who self-identify as “native” speakers clearly need to be given more information about the expectations in practice, not just in the Descriptions/Guidelines, to attain a rating of 4, 4+ or 5 on the ILR scale. Language Testing International, Inc. (the ACTFL Testing Office) has developed supplementary explanatory material about the expectations for attaining a rating of 3+ and higher to be made available to test-takers, explaining that, while native speakers may believe that their language fits the ILR 5 description “speech on all levels is fully accepted by well-educated native speakers in all of its features, including breadth of vocabulary and idiom, colloquialisms and pertinent cultural references”; in practice, the assessment will only award a 3+ unless the speaker can demonstrate the ability to tailor language both formally and informally. In the absence of a new set of high-level assessments aimed at ratings of 4/4+/5, editing these descriptions may also help alleviate much of the distress currently caused when self-identified native speakers do not receive ratings of 4, 4+ or 5.

5. Might a given speaker’s profile matter to the end-user of the ratings? If so, what mechanisms are currently used and/or may be developed to provide this information?

When third-party independent testing is conducted, the testers and raters do not always have any information about the client requesting the test or the purpose for the assessment. In many testing practices, the only information provided to a tester is the candidate’s name (or a test ID number) and often the name of the requesting client. It is generally thought that testers and raters do not need, and in fact, perhaps should not have, any specific background information about the candidate or about the potential uses of the rating. I agree that a tester does not need this information in order to elicit and rate a speech sample using current ACTFL/ILR interview protocols. The only information provided by a tester or rater is a final rating, no matter the profile of the speaker, unless otherwise
requested by the client.

Obviously, end-users may request information about previous language acquisition and use directly from the test-taker, but the testers/raters are perhaps in a unique position to provide observations about the speaker’s linguistic profile, as well. In practical terms, a final rating is indicative of the tasks (limited to the ones elicited depending on the assessment tool) a given speaker can perform using his/her language, while sustaining the other criteria required for that rating. In a number of cases, clients have worked with test administrators to customize oral proficiency assessments based on specific job demands, for example, by developing role-play situations that are specifically job-related. A number of end-users have begun to develop feedback mechanisms that supplement the rating, and that may be used for a variety of purposes, including diagnostic assessment. Such feedback mechanisms may include questions about those features of a speaker’s language that prevent him or her from achieving a desired level (where is the breakdown?), for example, or that ask the tester to indicate the speaker’s acquisition profile (adult foreign language learner, heritage or native) and whether or not the breakdown from the next level was primarily functional, linguistic, or both, or what level of comprehension of the questions was exhibited by the test-taker (a kind of gauge of interpersonal listening).

I would suggest that more end-users reflect on what, besides a final rating, would be useful to know about a particular profile and that the testing organizations work with them to help provide pertinent information. However, an important word of caution is in order here: we must be cognizant of potential unintended consequences of using oral proficiency interviews to glean and record information about a speaker’s background. Official interviews are often used to make high-stakes decisions about employment, job assignments and even pay scales. In official commercial testing, for example, a tester is prohibited by law from asking questions about age, sex, race, color, religion or national origin, sexual preference, marital status, health and political viewpoint. Designing or using assessments that draw conclusions or make assumptions about the ethnic or regional origins of the speaker, for example, could lead to both the perception, and perhaps the reality, of having an individual or entire group disadvantaged. We must be careful not to create a context where discrimination, either real or perceived, based on linguistic profile might evolve.

While I support a standard set of criteria by which to assess the functional

---

4 Under the 1964 Civil Rights Act, prospective employers may not ask age, sex, race, color, religion or national origin. The other categories have been added subsequently by the Equal Employment Opportunity Commission.
proficiency of all speakers regardless of background, such as the ILR Descriptions and ACTFL Guidelines, we must consider the challenges inherent in such an approach now that the testing context has changed significantly from what it was when these guidelines were first written. Inside the ACTFL/ILR testing community, we need more discussion and more empirically based investigations into the complexities of testing in a context that is growing linguistically more diverse. We also need to make greater efforts to educate end-users and, perhaps most importantly, test-takers themselves, about the criteria and practical considerations involved in applying those criteria in currently used assessment tools based on the ACTFL Proficiency Guidelines – Speaking and the ILR Proficiency Descriptions – Speaking.
Taking the “L” out of LCTLs: The STARTALK Experience

Catherine Ingold
Mary Elizabeth Hart

Introduction
The less commonly taught languages in the United States are often those most critical to national security. How, then, can the number of students learning these languages be increased, and how can high-quality instructors be produced to teach these languages? Having determined that foreign language skills are essential to diplomacy, economic competitiveness, and the security interests of the U.S., the Secretaries of State, Education, and Defense, and the Director of National Intelligence coordinated their efforts to expand language education beginning in kindergarten and continuing through elementary, secondary, and postsecondary education and into the workforce (United States Department of Education, 2008). STARTALK is the U.S. government’s most recent effort to address these issues. Sponsored by the Office of the Director of National Intelligence, STARTALK is one outcome of the National Security Language Initiative (NSLI), a presidential initiative begun in January 2006 to increase the number of Americans learning critical need foreign languages. The National Foreign Language Center (NFLC) has been the sole contractor for implementation of this project since its inception.

STARTALK provides summer language learning experiences for students and professional development opportunities for teachers of critical languages in grades K-16. The broad goals of the project are to encourage further study of these languages among students, and also to increase the number and effectiveness of teachers of these languages, thus, consequently improve the quality of the programs in which they teach (“National Security Language Initiative,” n. d.). The ten languages currently taught in STARTALK are Arabic, Chinese, Dari, Hindi, Persian, Portuguese (as of 2011), Russian, Swahili, Turkish, and Urdu. Through 2010, the project has provided professional development opportunities to more than 4,000 teachers and learning opportunities to over 13,000 students of critical languages across the United States, reaching participants from every state and sponsoring programs in all but two states. The planning, implementation, and oversight of the programs have engaged not only
language professionals working in the ten current STARTALK languages, but also many education professionals working in more commonly taught languages, who have shared their expertise in language program planning, curriculum design, instruction, assessment, and teacher development with their colleagues in emerging language fields in which such expertise and resources remain scarce. It seems appropriate in this volume honoring the career of Dr. Richard Brecht, the NFLC’s third director and subsequently the founding director of the Center for Advanced Study of Language, to place the achievements and findings from the STARTALK project within the context of the NFLC’s original and ongoing mission to develop U.S. capacity in languages other than English. In particular, STARTALK addresses the dimension of the NFLC’s mission that focuses on increasing the percentage of Americans who have personally and professionally useful biliteracy in English and another language.

Figure 1. STARTALK Timeline

The STARTALK Experience: Planning for Implementation
The STARTALK contract issued to the NFLC beginning in 2006 mandates the general activities that the NFLC carries out in support of the project, but it has allowed significant latitude in how those activities are carried out. To begin the
planning process in 2006, the NFLC convened an advisory board to review the project’s mission and goals, and to solicit advice and support in executing the project. To ensure maximum inclusion of and support from the field, the NFLC invited leaders in world language teaching and learning, second language acquisition, and less commonly languages to participate in the planning phase. During this meeting, short- and long-term strategies for implementing STARTALK were discussed for both the student and teacher professional development programs. Selection criteria for programs and program sites were identified. The group came to consensus on reasonable expectations for accomplishments in summer programs for high school students.

The initial mandate for summer 2007 was to implement high school programs for 400 students in Arabic and Chinese, and professional development opportunities for 400 current and prospective teachers of these languages in at least five states. The challenges facing STARTALK in its first year appeared daunting. In the *No Child Left Behind* environment, world languages had been de-emphasized in the curriculum and there were no obvious indicators of demand for summer programs in Arabic and Chinese. Although the NFLC was confident that the student programs would reach the enrollment target, there was no evidence that current or prospective Arabic or Chinese teachers would enroll in an unknown summer professional development program. Most Arabic and Chinese teaching in the U.S. was occurring at the postsecondary level, but the 2007 STARTALK contract mandated high school-level programs. In November 2006, the NFLC issued a request for proposals for STARTALK student and teacher programs to be implemented in summer 2007. The solicitation was announced via language listservs, through language association newsletters, and directly to Arabic and Chinese language programs across the U.S. In response to the first RFP issued, the NFLC received 59 proposals and selected 34 for funding.

**Implementing the First Summer Programs**

In February 2007, the NFLC convened a four-day meeting of a task force of language, curriculum, assessment, and program evaluation specialists—to the extent possible, with experience in K-12 education—to design sample curricula for high school student programs and to identify topics for teacher development programs. The NFLC assembled the STARTALK 2007 program directors and their lead instructors in Washington, D.C. in late March 2007 to introduce them to STARTALK and the materials produced by the task force, as well as provide guidance from those who had experience in running successful summer programs, such as Concordia Language Villages. These program directors then began to develop curricula, recruit participants, gather and develop materials,
and implement their programs with far less guidance than STARTALK would provide in subsequent summers.

In collaboration with the staff from the Center for Applied Linguistics (CAL), STARTALK’s core staff (STARTALK Central) developed surveys to determine the demographic and educational background of both students and teachers; ascertain the teachers’ experience; and gather attitudinal and evaluative data. CAL was issued a subcontract to gather, analyze, and report on STARTALK survey data. The NFLC was required to monitor the implementation of summer programs, and this mandate included on-site visits. Observation rubrics and protocols for both student and teacher programs were developed by members of the task force who had experience in K-12 program evaluation. Curriculum and language experts were recruited and trained to conduct the site visits. A site visit report template was developed and used, so that observations could be compared and analyzed.

At the end of the summer, survey results indicated that most participants had a positive experience. Of the 346 students responding to the survey, 324 (94 percent) reported that learning a language during the summer was a good experience. Of the 268 teachers who responded to the survey, 252 (94 percent) felt that the goals of the program had been achieved (Sugarman & Malone, 2007b).

While these affective indicators were encouraging, STARTALK Central was stymied in its efforts to gather data on student language learning. The duration of the programs was generally too short to achieve gains measurable in proficiency terms, and the heterogeneity of program content thwarted the use of standardized instruments. Additional significant findings were that (1) the demand for STARTALK programs, both teacher and student, far exceeded expectations; and (2) STARTALK had provided increased opportunities for national networking among teachers, particularly for sharing materials and experiences.

**Documenting Learning Outcomes**

The major challenges facing STARTALK at the end of the first year were to (1) help programs to specify observable learning objectives so that outcomes could be assessed, possibly by providing a curriculum planning tool; and (2) identify or develop instruments for assessing outcomes of short, intensive language programs. It became apparent that some standardization of STARTALK programs would be necessary to address the first challenge, particularly by requiring more systematic planning of curriculum and instruction. STARTALK Central decided to develop a curriculum-focused planning template that would
impose the use of national standards (ACTFL, 2006) and promote effective instructional strategies identified in programs deemed successful by all participants. This step was taken in recognition of many program directors’ relatively limited experience and expertise in language program planning and in language pedagogy, and also of the need to support the rapid program planning required in the short interval between the notification of funding and the programs’ start dates. In the first year, many program directors complained about the mandated use of the template, but by the following year, its utility was widely recognized.

There was no immediate solution to the lack of assessment tools for short programs, but several efforts began that would lead to significant progress. STARTALK began working more closely with the National Council of State Supervisors of Foreign Language and the University of Oregon to enhance LinguaFolio Online as a means of gathering evidence of student learning. Experiments with the STAMP test in 2007 proved frustrating because so much of the test content had not been covered in the brief STARTALK programs. As a result, CAL undertook, at the request of STARTALK Central, an analysis of curriculum content of the student programs to identify the most commonly occurring topic areas and learning objectives (Sugarman & Malone, 2008c).

The findings from this study have served to define topic areas for an experimental assessment tool, the ACTFL Assessment of Performance toward Proficiency in Languages (AAPPL), developed by ACTFL under a STARTALK subcontract. This computer-based test of speaking proficiency at the ACTFL Novice level was developed in 2009 for Chinese, and piloted in 2010 in a number of the STARTALK Chinese programs. A similar test for Arabic is under development and will be pilot tested in summer 2011 programs.

In addition, the STARTALK project identified a need to add features to LinguaFolio Online to enable teachers to customize can-do statements related to the learning objectives of their programs and to allow students to upload evidence of their learning. These features were implemented in 2009. Subsequent studies by Ross et al. (2010) compared the results of an objective assessment of listening, speaking, reading, and writing via the Computerized Assessment of Proficiency (CAP) against two other data sources: teacher retrospective checklists and student self-reports in LinguaFolio. The first-year findings suggested that students can self-assess with relative accuracy, typically self-rating in can-do statements slightly below their performance on relevant test items. Individual teachers’ ratings of student performance varied quite widely when compared with the objective measure. This preliminary evidence underlines the need for accessible, reliable, and affordable assessment tools for STARTALK languages, as
well as for enhanced teacher professional development in principles and techniques of student assessment. Meaningful assessment of student learning has emerged as one of the key challenges in providing highly effective language learning experiences, and the STARTALK project has continued to devote time and resources to this issue.

**Improving Program Quality**

In addition to efforts to document learning outcomes, the STARTALK project has focused more broadly on describing, assessing, and improving program quality. With data from 2007 and recommendations from the STARTALK advisory board, STARTALK staff began implementation of a continuous quality improvement process that has become a hallmark of the project and, it can be argued, one of the most important factors in the project’s success to date. This process has involved (1) systematic identification of issues that have emerged from extensive data gathering, notably program and site visitor reports and stakeholder surveys; (2) structured brainstorming with a network of experts to find solutions to issues thus identified or, for more complex challenges, to identify strategies to address them over time; (3) careful determination of how best to introduce and implement STARTALK program enhancements—new products, procedures, and/or protocols—including elicitation of feedback on each new element of the program; (4) periodic surveys of stakeholders on every aspect of the program to identify further opportunities for improvement; and (5) consistent, detailed, and timely feedback to stakeholders, including reports on survey findings.

STARTALK stakeholders (Figure 2) include (1) an advisory board whose composition varies from year to year, depending on the nature of topics on which the project leadership needs advice; (2) a task force consisting of teacher educators and expert practitioners that meets annually, examining findings from the previous year and designing training and interventions for the upcoming year; (3) STARTALK program directors; (4) instructors in student and teacher programs; (5) team leaders, who provide mentoring to a group of program directors; (6) site visitors, who accompany team leaders to observe programs; (7) participants in the student and teacher programs (and, in the case of younger students, their parents); (8) subcontractors (CAL, University of Oregon, and ACTFL); and (9) speakers, workshop providers, and invited attendees at the spring and fall STARTALK conferences. According to their experiences with STARTALK, they believe that the STARTALK structure ensures quality through rigorous protocols and extensive resources, and the detailed documentation required provides a clear understanding of all stakeholder roles.
Monitoring Implementation and Mentoring Teachers
In 2008, Hindi, Urdu, and Persian were added as program languages; the age range for student programs expanded to include middle school; and the number of programs increased from 34 to 81. In response to growth in the number and variety of programs, the NFLC decentralized the oversight and mentoring of funded programs. In addition to support from STARTALK Central, a team leader system was established to mentor program directors and instructional staff, and to monitor the implementation of programs in accordance with the approved proposal.

The team leader system was expected to provide expert guidance for program planning and to create synergy among program staff that shared a team leader, with directors who had already run programs in 2007 providing some guidance to new programs in the areas of program administration, meeting STARTALK requirements, and sharing resources. The team leader was to offer pedagogical guidance in the completion of the curriculum template and ensure that all elements of a successful STARTALK program were present in each program's plan for summer 2008.
Data from surveys administered to program directors, site visitors, and team leaders themselves at the end of STARTALK 2008 confirmed that this model was effective and should be continued. Seventy-five percent of the program directors reported that the team leader was of assistance in developing and refining program curricula (Sugarman & Malone, 2008). In addition to providing support and guidance to programs, the team leaders gathered rich information from their interactions with program directors, including observations on what was working well, what programs found difficult to implement, and areas for improvement in the design and implementation of the STARTALK project. Based on their experiences, team leaders have continued to recommend improvements to the STARTALK project and identify professional development priorities; they are an invaluable source of qualitative data for continuous project improvement.

STARTALK has worked closely with a diverse group of team leaders to assure that each program receives effective guidance and that messages on effective practices are clear and consistent. A number of team leaders are curriculum experts and teacher trainers drawn from the more commonly taught languages, many with experience in mentoring LCTL programs, while an increasing number are expert teachers in the STARTALK languages. Team leaders and site visitors work from a rigorously standardized model for visits, using an observational checklist that has been revised over time. Training for consistency among team leaders and site visitors includes observation and evaluation of classroom videos using a previously developed checklist, followed by often lively discussion of how well the videotaped instructors met the criteria on the checklist. Role plays, followed by discussions, help to develop consensus around effective methods for providing formative and constructive feedback to program directors at the conclusion of site visits. These activities take place at every team leader meeting, usually held prior to the spring and fall conferences. In surveys, an overwhelming majority of team leaders report that these activities are useful in making them more reflective about their own teaching, as well as the teaching of others (Sugarman & Malone, 2008).

Sharing Experiences and Building Best Practices
The STARTALK spring pre-program and fall post-program conferences provide an opportunity for STARTALK program directors and their instructional staff to come together for targeted professional development and to share experiences, successes, and challenges. The STARTALK 2008 spring conference enabled STARTALK program directors and their staff to meet with their team leaders to begin work on planning curricula and lessons for the summer. During the spring
conferences, new program directors and instructors receive orientation on STARTALK procedures, requirements, and resources. As programs have gained experience with STARTALK procedures, the time devoted to substantive professional development around key elements of effective language programs has been expanded. The fall conferences are devoted to topics informed by site visit findings on what instructors and programs most need to continue to improve. Results of STARTALK program participant surveys are shared, as well as surveys of program directors, instructors, site visitors, and team leaders, in order to provide an evaluation of STARTALK programs from several perspectives. Again, as the programs gain experience with STARTALK, program directors have devoted more time sharing what works well and discussing unique program features. At the 2008 fall conference, STARTALK Central introduced the STARTALK-endorsed principles for effective teaching and learning, shown in Figure 3.

*Figure 3. STARTALK-Endorsed Principles for Effective Teaching and Learning*

| Implementing a standards-based and thematically organized curriculum |
| Facilitating a learner-centered classroom |
| Using target language and providing comprehensible input for instruction |
| Integrating culture, content, and language in a world language classroom |
| Adapting and using age-appropriate authentic materials |
| Conducting performance-based assessment |

**Expanding to Elementary and Postsecondary Settings**

In 2009, STARTALK expanded to include elementary (K-5) and postsecondary (undergraduate) programs, and added Swahili and Turkish to the list of languages. The number of programs increased from 81 to 116. The addition of elementary students presented challenges to conducting surveys, because most students in this age group would not be able to complete an online survey.
NFLC worked with CAL to develop an interview protocol, which was piloted in 2009 by CAL staff.

**Increasing Opportunities for Teacher Professional Development**

It can be argued that the most important impact of the STARTALK project as of this writing has been in teacher development, both in improving the effectiveness of thousands of teachers, and in identifying and promoting effective practices in teacher professional development in the critical languages. During the first two years of STARTALK’s program implementation, a number of critical issues had surfaced related to world language teacher preparation: the variability of standards and procedures for training and licensure; the scarcity of programs serving the STARTALK languages; and, most concretely, the inability to provide prompt and accurate information to prospective teachers about how to become licensed in their state. Survey data of teacher program participants had, by that time, generated a description of prospective teachers of these languages: educated native speakers of the languages, predominantly women beyond their college years, sometimes with teaching experience in non-traditional settings such as community-based programs, with very different strengths and needs compared with the undergraduate language major for whom many world language licensure programs had been developed.

Supplemental funding enabled STARTALK to host a national conference in December 2009 to address teacher preparation. Conference participants demonstrated a surprising degree of consensus about what was needed and shared information about successful models, resources for teacher assessment, and common problems. One concrete outcome was the publication of *The Teachers We Need: Transforming World Language Education in the United States* (Ingold & Wang, 2010).

End-of-year funds were made available in 2009 to institutions hosting STARTALK programs to create a total of 12 online teacher development workshops. The institutions selected had all exhibited specialized strengths in their face-to-face programs, and were willing to develop online materials on the accelerated timeline necessary to use the federal funds before the end of the fiscal year. The NFLC hosted a planning meeting with representatives of the institutions featuring sessions on instructional design of electronic materials. The resulting collection of online workshops is publicly available at [startalk.umd.edu](http://startalk.umd.edu) and has been widely used, both online and in various conference presentations.

In addition, the NFLC received funds to film classes from selected STARTALK programs, creating a small video collection of teaching in STARTALK languages. For a number of these languages, no such materials had
previously been available. The development of these resources for STARTALK program staff was intended to bridge the gap between well-developed curriculum plans and the implementation of day-to-day instruction in the classroom. Site visit reports from 2009 documented the need for more professional development in executing successful instructional strategies in the student programs. A number of the instructors reported difficulties in adhering to the communicative mode of instruction prescribed in the national standards and often defaulted to a teacher-centered mode of instruction; the videos show how a student-centered mode of instruction can be successfully implemented in all languages.

In 2009, awards to STARTALK institutions were extended to February 28, 2010, allowing programs the opportunity to offer follow-up activities to their summer programs. A number of teacher programs offered late winter meetings to monitor the effectiveness of their summer instruction and experiences, and allow teacher trainees to share the challenges and successes of implementing STARTALK practices in their own classrooms. Some student programs also developed and administered online opportunities for students to continue their language learning, although at a much less intensive pace than during the summer.

Recent Developments
In 2010, STARTALK added Dari and Russian for teacher trainee programs or student programs combined with a teacher program. Evaluations from 2009 recommended providing teachers an opportunity to be trained by STARTALK before implementing student programs. This recommendation allows prospective teachers and program directors the opportunity to introduce the national standards and effective instructional strategies in the student programs. In many of these teacher trainee programs, microteaching or peer-teaching experiences are provided.

During the 2010 spring conference, the professional development resources developed in 2009 were introduced to STARTALK program staff. The conference agenda was structured so that all participants had sufficient time to learn how to use the resources for teacher training purposes, or as a means of self-instruction for teachers themselves. For several of the STARTALK languages, these videos provide the first opportunity for teachers to witness instruction in the target language.

A second major focus of the spring 2010 conference was the use of target language for instruction. Although this is one of the six STARTALK-endorsed principles, an analysis of site visit reports and debriefings with team leaders after
summer 2010 revealed that instructors often reverted to English in the classroom. Several interactive sessions were conducted to provide hands-on experience to instructors and program directors in the use of the target language. Participants shared techniques for presenting common thematic lessons and classroom instructions. Subsequent site visit reports indicated that the target languages were being used for language instruction, but noted that more training was needed in checking for student comprehension during the lessons. In the spirit of continuous improvement, the spring 2011 conference provided two two-hour workshops on this strategy.

A third focus in 2010 was the expansion of the protocol used to interview K-5 students who could not complete an online survey to evaluate their learning experience. Once CAL staff had established that the protocol developed and piloted in 2009 elicited needed information, it was ready for use at more STARTALK sites. The goal of the 2010 pilot activity was to train local personnel to conduct these interviews according to the strict protocol and to report data to CAL. Both the cost and efficacy of the K-5 interview process were evaluated at the end of the 2010 pilot. Results of the analysis were favorable, and pointed to a plan for continued data collection from this age group using these methods in 2011.

**Conclusion**

It is difficult to prioritize the preliminary findings that have emerged from the STARTALK project. There is no way to accurately predict how STARTALK and other efforts to expand language education may positively affect U.S. language education and capabilities in the longer term. However, the STARTALK project has had an impact on thousands of students and instructors of critical languages, and the experience of STARTALK to date is suggestive of a number of preliminary conclusions:

1. Summer is a highly useful time for innovation and quality improvement in educational programs. These programs can have a local and even a national impact.
2. Students and their families are interested in and supportive of language programs, even for “difficult” languages, when the programs are of high quality. Students who participate in such programs are willing, even eager, to continue their language learning.
3. The U.S. has an abundant supply of native speakers of languages critical to national interests who are willing to be trained as teachers, and who can develop strong teaching skills over the course of several summers in well-designed, hands-on programs. However, these candidates have
strengths and needs very different from the traditional teacher trainees for whom most world language teacher education programs are designed.

4. School districts can become interested in offering critical languages, especially when they have engaging experiences with STARTALK and the demand for these languages has been demonstrated locally.

5. New language programs, whether for teachers or students, can improve significantly in quality over a relatively short time when provided with clear quality criteria, technical support (professional development, planning tools, and resources), and tailored, prioritized formative feedback.

Overall, the STARTALK project has met its goals of expanding access to critical languages, improving the effectiveness of critical language teachers, and improving the quality of critical language programs. In 2011, 156 STARTALK programs will be offered, as shown in Figure 4.

**Figure 4. STARTALK 2011 Program Offerings by Program Type and Language**

<table>
<thead>
<tr>
<th>Language</th>
<th>Student</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>24</td>
<td>21</td>
</tr>
<tr>
<td>Chinese</td>
<td>63</td>
<td>49</td>
</tr>
<tr>
<td>Dari</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Hindi</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Persian</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Portuguese</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Russian</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Swahili</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Turkish</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Urdu</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

STARTALK programs will be offered in 46 states and the District of Columbia to an estimated 8,500 participants. The majority of participants would not have these language learning and teaching experiences absent the large number and wide variety of STARTALK programs.

In terms of student perspectives, the optional nature of STARTALK programs (most older students self-select for the programs) has contributed to a remarkable level of student enthusiasm and openness to language learning, even
in cases where the program itself is very much a work in progress. The percentage of students who express a desire to continue to study the language of their summer program has consistently hovered around 80 percent (Sugarman & Malone, 2010). While STARTALK Central does not track students longitudinally, an increasing number of students in other language programs (e.g., NSEP intensive college programs and NSLI-Y summer programs abroad) are reporting enrollment of former STARTALK students, and STARTALK programs also report students enrolling for second and third summers. It seems clear that students, when provided with effective and engaging language programs, are interested and willing to persist in language learning. A number of STARTALK students have enrolled in second-level language courses because of the interest developed during a STARTALK program. Anecdotal reports from their teachers note that these students often assume leadership roles and participate actively in language classes. They are able to demonstrate what they can do in the language, are more likely than other students in the class to take risks in producing language, and are confident in their ability to learn.

The impact of STARTALK programs on the availability of school-year language programs has been one of the more surprising outcomes of the project to date: 39 percent of STARTALK student programs report that school-year programs in their language had been started as a result of the STARTALK program, and an additional 29 percent reported start-up of local afterschool or heritage programs. There are a number of possible interpretations of these data. One is that the STARTALK program offered an opportunity to respond to an existing but unrealized local interest. Another is that the interest generated by the STARTALK program catalyzed a critical mass of interest and commitment. In any case, the data suggest that there is more latent interest and demand for such programs than has been thought.

Both survey and anecdotal data confirm that teachers participating in STARTALK have improved their effectiveness in the classroom: 99 percent of teacher program participants surveyed report increased confidence in their teaching. For teachers of these less commonly taught languages, the STARTALK program often offers the first opportunity for professional development targeted to their respective languages where unique features of instructing others, particularly in the communicative-oriented classroom, can be addressed. Teachers report that network building has been the most enduring effect of the STARTALK program. These teachers remain in virtual contact with each other for purposes of support, materials sharing, and online workshops. Teachers also report that after a STARTALK program, they are more willing to implement new
strategies. Several participants refer to the STARTALK experience as “transformative.”

Approximately 50 percent of teachers have reported their intention to become certified language teachers, a long and arduous process, and one that will contribute to the pool of highly qualified teachers of these languages.

It is clear from evaluation data that both STARTALK student programs and teacher programs have been able to break new ground, experiment and innovate, and improve from summer to summer by drawing the lessons learned from participant and site visitor feedback. Over 90 percent of programs have reapplied for STARTALK funding over several years. The policy of the funder and STARTALK Central is to continue programs that are reasonably successful according to program assessment data, particularly if they give evidence of progress and of learning from past experience and feedback.

With the continuation of successful programs and the growth of new programs, each year STARTALK provides innovative and exciting ways for students to learn to communicate in critical languages, thus fulfilling the project’s goal of increasing interest in and the learning of these languages. Moreover, STARTALK teacher programs have created professional communities of teachers in some of the less commonly taught languages and strengthened existing communities by providing professional development to these teachers, pathways to certification, and resources for language instruction and assessment for LCTLs.

STARTALK stakeholders, from program directors to administrators, report that the STARTALK model as implemented is supporting the vision of STARTALK to provide effective language learning experiences to students and professional development opportunities to current and future teachers.

References


Heritage is Not Enough:
The Changing Demographics of the Chinese Language Field in the
United States

Scott McGinnis

Introduction
In a series of articles beginning in 1994, Richard Brecht and the late A. Ronald Walton first identified four sectors in the United States as sources for providing national capacity in the Less Commonly Taught Languages (LCTLs).

1. Federal: “[G]overnment language-training programs designed for employees of government agencies” (Brecht & Walton, 194)

2. Private sector: Composed of at least three components, (a) proprietary schools, “aimed at professionals…for business, travel, and tourism” (194); (b) corporate programs aimed at employees within an organization; and, (c) language services, including both private language instruction, and translation and interpretation services, as well as foreign visitor escort services.

3. Academic: The traditional K-16 (and now more commonly dubbed pre-K-16+) classroom-based language education system, supplemented by various “auxiliary systems” such as summer programs and study abroad.

4. Domestic ethnic language preservation/enhancement: In the intervening years since first proposed, this has come to be known more succinctly as the “heritage sector,” composed of immigrant populations and their descendants, many of whom have established semi-formal instructional programs for their mother-tongue language and cultural preservation. Brecht and Walton make the point that this fourth heritage sector is “unique in that it possesses the potential to supply language capacity without instruction” (195-196).

In two previous works, I have analyzed the development of Chinese as a heritage language (CHL) education within the United States. Two major points characterize the evolution of the CHL sector.

1. While initially maintaining a pedagogical philosophy that could best be described as derivative, modeling itself after first the language
educational system of the home country of Chinese immigrants (i.e., mainland China or Taiwan) and then later the American K-12 system, CHL educational organizations and institutions have become increasingly innovative in serving a number of societal roles not served by the traditional academic sector-based Chinese language programs; e.g., providing Chinese language and cultural education for both Chinese adoptees and their American adoptive parents through the Families with Children from China (FCC or FWCC) programs (McGinnis 2007).

2. However well-intentioned and purportedly pragmatic an attempt to tap in to this potential pipeline of Chinese language expertise for any one of a range of American needs with regard to greater China—economic, security or otherwise—seeing the CHL population as a “silver bullet” is at once insufficient and ill-motivated. “To the contrary, Chinese heritage language learners are but one sector of our collective national student body, and it is imperative that we work to develop a more extended pipeline that will enable non-heritage learners to reach equally high levels of proficiency” (McGinnis 2005, 594).

Put more succinctly, over the course of the last nearly two decades, the heritage sector has become in some aspects every bit as well-developed any of the other sectors. But as a means of addressing all of the critical national language needs delineated by Walton (1996) for Chinese in particular, it is still just a single sector, and only in combination with the other three sectors can all of those national needs be met.

In this paper, I want to build from what remains a tremendously insightful and innovative analysis by Brecht and Walton to portray the continually evolving Chinese language instructional field in the United States. In the 21st century, it is clear that not only is a single sector—heritage—not enough, but a four-sector analysis of the field is not enough. One cannot understand the means for meeting Chinese language capacity needs in America without adding at least two new sectors to Brecht and Walton’s original framework—domestic non-governmental organizations (NGOs) and foreign governmental agencies. In the former category, most prominent (but not the only ones) are the Asia Society and The College Board; in the latter, it is China’s National Office for Teaching Chinese as a Foreign Language, more commonly known by the Chinese term Hanban, taken from the first syllable of two words in the Chinese name of the organization, Hanyu bangongshi (literally “Chinese language office’’). While each one shall be examined separately, it will quickly become very clear that the
interrelationships among the three of them are considerable and essential to the rapidity with which they have emerged as arguably the major shapers of the Chinese language instructional field in America.

**The Asia Society**

Founded in 1956 by John D. Rockefeller III, the Society describes itself as “the leading global and pan-Asian organization working to strengthen relationships and promote understanding among the people, leaders, and institutions of the United States and Asia” (Asia Society 2010c). While the Society does pursue various projects related to areas such as business and economics, as well as operating a museum in its New York City world headquarters, the overarching mission of the Society is professed to be an “educational mandate.”

Late in the year 2001, two major events took place in the life of the Asia Society. The first was the publishing of a report entitled “Asia in the Schools.” Produced by the National Commission on Asia in the Schools, within its “call to action,” it posed the intriguing challenge of increasing “the number of K-12 students studying Asian languages from the current level of less than two percent studying a foreign language to five percent by 2005, and to ten percent by 2010” (Steinemann, Fiske and Sackett, 8). However, arguably the more significant event in the life of the Society at that time was the appointment of Vivien Stewart, who served as Vice President for Education until 2009. She currently serves as Senior Advisor for Education to the Society, with specific responsibilities as Chair of the Society’s Confucius Classrooms Initiative, to be discussed below.

As reflected in the quote above, the initial thrust of the 2001 monograph seemed to encompass the promotion of all Asian languages. Indeed, as stated on the Society’s website under the Frequently Asked Questions, “[t]he Asia Society’s definition of Asia includes the more than 30 countries broadly defined as the Asia-Pacific region—the area from Japan to Iran, and from Central Asia to New Zealand, Australia, and the Pacific Islands” (Asia Society 2011b). However, it very soon became apparent that Stewart’s primary focus of responsibility was to be on all things Chinese. The Society’s first major Chinese language-specific project was the convening of a meeting in April 2005 that ultimately produced a report entitled “Expanding Chinese Language Capacity.” Over two dozen leaders in both language education in general and the Chinese language teaching profession in particular met to discuss a draft and contribute to the final version of that report. It is interesting to note that whereas the 2001 report targeted enrollments in Asian languages of 5 percent by 2010 and 10 percent by 2015, the
2005 report framed the question as such: “What would it take to have five percent of American high school students learning Chinese by 2015?”—an admittedly more modest and realistic goal, particularly given the growth in interest in Chinese language education in the intervening four years.

In the years since then, the following additional reports were produced and published by the Society.

- (with The College Board) “Chinese in 2008: An Expanding Field”—an overview of the current state of the Chinese language field in general.
- “Learning Chinese in American Schools” (2009)—a DVD distributed at no charge other than shipping, with funding being provided by Hanban.
- “Meeting the Challenge: Preparing Chinese Language Teachers for American Schools” (2010).

Besides Stewart, a critical contributor to all of the above reports and products was Dr. Shuhan Wang, who served as Executive Director for Chinese Language Initiatives at the Asia Society between 2006 and 2009.

But while all of the above projects have in their own way had an important impact on the continuing development of the Chinese language educational field, a relatively new one, undertaken jointly with the Hanban, may provide an even more comprehensive developmental mechanism for K-12 Chinese language teaching in the United States.

“Asia Society’s Confucius Classrooms will be a national network of 100 exemplary Chinese language programs over the next three years...Asia Society will identify, through a competitive process, schools that have the potential to grow into exemplary Chinese language programs...Each Confucius Classroom will be linked with a partner school in China to enhance opportunities for language learning...Asia Society’s Confucius Classrooms provides the infrastructure for the development of effective and sustainable Chinese language programs for American students...This Network is also the beginning of an ongoing field-building process (emphasis mine) with the goal of establishing high-quality, sustainable Chinese language programs in all regions of the United States” (Asia Society 2011a).

What is striking is the explicitness with which Asia Society conceives of this project as one that can build the Chinese language field in toto, arguably supplanting the role of the traditional field developmental leaders that come
from the academic sector; namely, professional membership organizations such as the Chinese Language Teachers Association (CLTA) and the Chinese Language Association of Secondary-Elementary Schools (CLASS). And the Confucius Classroom Network represents just one point of collaboration between the American NGO Asia Society and the mainland Chinese government Hanban, as will be discussed shortly.

But as was noted in the introduction to this paper, the Asia Society is not the only NGO involved in Chinese language field development. Indeed, the Society has been assuming its leadership role most frequently in concert with another domestic NGO.

**The College Board**

To all those that work within the academic sector, it would hardly be surprising that The College Board would seek to exert influence on the development of Chinese language education in the United States. In purely quantitative terms, the Board’s influence on secondary and tertiary levels of education is vast: it currently “is made up of more than 5,900 of the world’s leading educational institutions...Each year, the College Board helps more than seven million students prepare for a successful transition to college through programs and services...including the SAT® and the Advanced Placement Program®” (College Board, 2011a). Far older than the Asia Society (the College Board was founded in 1900), it has been involved in Chinese language education for several decades, beginning with the creation of the SAT II Chinese with Listening subject test in the early 1990s. But in the view of many of those involved with the SAT II Chinese (including the author, who served on the SAT II Chinese Test Development Committee for eight years from 1995 until 2003, including three years as Chair), the SAT II never really moved beyond providing primary service to the Chinese heritage language learner (HLL) community—in principle, the very audience that was not meant to be served by such a testing metric, given the stated focus of the SAT II language tests for assessing achievement by second or foreign language learners of the tested language. What was seen to be needed by many people in the Chinese language educational field was a means to more broadly develop K-12 Chinese language programs throughout the United States in a range of demographic settings, both non-heritage and heritage, and to provide both curricular and assessment resources for that process. Certainly one of the best possible and broadly supported means for doing so would be through an Advanced Placement (AP) Course and Exam.
What has been somewhat obscured by the passage of time is that the original impetus for the establishment of the AP came not from within the College Board, but from a number of national organizations and NGOs, including:

- Chinese Language Association of Secondary-Elementary Schools (CLASS).
- Chinese School Association in the United States (CSAUS), the organization of Chinese community schools with primary connections to immigrant populations from mainland China.
- National Council of Associations of Chinese Language Schools (NCACLS), an “organization of organizations,” all regionally based and with primary connections to Taiwan immigrant populations.
- National Council of Organizations of Less Commonly Taught Languages (NCOLCTL), hosted by the National Foreign Language Center (NFLC) until 2003, and which served as the organizational home for a three-year (2001-2004) project funded by the Henry Luce Foundation entitled the Chinese Language Field Initiative (CLFI). All three of the aforementioned Chinese language associations (CLASS, CSAUS and NCACLS) as well as CLTA were represented on the National Chinese Language Commission (NCLC), which provided the advisory oversight to CLFI.

In June of 2003, the Board of Trustees for the College Board approved a plan to implement four new AP courses and examinations: Chinese, Italian, Japanese, and Russian. As part of the initial planning process for the establishment of the AP Chinese, the College Board and Asia Society put together their first joint venture, in the form of an invitational meeting to present and discuss the preliminary plans for the Chinese AP. Prior to that meeting at the Asia Society on September 16, 2003, at the annual meeting of the NCLC in August 2003, a position paper was drafted, revised, and ultimately presented to the invited participants prior to their arrival at the September meeting. In that position paper, the NCLC stated their view that “no issue is of greater importance at this time to all Chinese language teachers and students in the United States than the establishment of an AP for Chinese” (McGinnis et al, 2003). Four primary conditions or rationales supporting its establishment were presented.

- **Potential market.** Based on data gathered by the two national HLL Chinese associations and the Secondary School Chinese Language Center (SSCLC) at Princeton University, a pipeline of potential
participants in a Chinese AP Course and Exam was estimated to be well over 150,000 students.

- **No single existing curricular model.** Given the multiplicity of state and local regulations governing curricula in the public school, the only arguable means of being able to articulate both horizontally (across K-12 instructional settings, including Chinese community schools) and vertically (that is, into higher educational programs) would be a commonly agreed-upon “capstone” sort of course such as the AP.

- **No single existing assessment tool.** Even while a number of domestically- and foreign-produced assessment metrics were in existence already at that time, they were not tied to a particular curriculum, measuring instead more general, global proficiency. The then relatively recent (1999) development of the Chinese National Standards by CLASS provided a framework that could help shape the content of the proposed AP Course as well as the attendant Exam.

- **Lack of broad-based professional development, both pre-service and in-service.** As has been stated for many years by a number of experts in a number of settings, the greatest impediment to the development of any LCTL field is the lack of well-qualified teachers; Chinese most definitely included.

In examining the initiatives undertaken by the College Board since the announcement of the plans to develop an AP for Chinese Language and Culture, it is interesting to examine how the Board, albeit certainly not in direct emulation of the points presented in the NCLC position paper, nonetheless mirrored and responded to the challenges implicit in those stated conditions and rationales.

- **Potential market.** The College Board conducted its own survey in 2004, in the process determining that nearly 2,400 high schools expressed an interest in offering the AP Chinese course in 2006-07 (College Board, 2011c).

- **No single existing curricular model.** Various resources in support of the AP Chinese Language and Culture Course are now available online, either free or for purchase, including course descriptions with audio files, a teacher’s guide, and sample syllabi.

- **No single existing assessment tool.** In addition to the actual AP Chinese Exam that is revised on an annual basis, online resources are available for support of teachers and students for the AP Chinese Exam, including a complete released exam.
Lack of broad-based professional development, both pre-service and in-service. It is probably in this regard that the College Board, in collaboration with the Hanban, has made the greatest efforts towards providing support of all sorts in response to the lack of highly qualified teachers of Chinese in the United States. In a lengthy press release dated April 19, 2006, it was reported (in part):

“A five-year plan to collaborate on a new Chinese Language and Culture Initiative was announced today by China’s National Office for Teaching Chinese as a Foreign Language (Hanban) and the College Board.... The initiative will address the critical shortage of teachers of Chinese in the United States...expand professional development for teachers.... In an AP survey conducted in 2004, nearly 2,400 high schools expressed an interest in offering the AP Chinese course in 2006-07, but for many of these schools, this goal may go unrealized. They either are understaffed or have no teacher of Chinese, and many see no prospect of finding the teachers necessary to build their programs...the plan announced today will temporarily place 150 guest teachers from China in American classrooms over the next three years...the plan will also enable the College Board to support nearly 300 American teacher candidates in their efforts to attain state certification to teach Chinese....These programs will expand to include greater numbers of teachers...in the coming years. In all, nearly 600 American teachers of Chinese will have access to these programs over the next five years” (College Board 2011d).

What was immediately apparent and continues to this day is the multi-prong approach taken by the College Board and Hanban to address the teacher shortage problem via both tactical and strategic means. Clearly the “guest teacher” concept only fills immediate needs; long-term plans have to include professional development for both pre-service and in-service Chinese language teachers, including teacher certification pathways and opportunities for language immersion experience. One can see the continuation of the original 2006 “model” in an examination of current professional development opportunities from the AP Central Chinese website.

“About the College Board’s Chinese Language and Culture Initiatives Chinese Bridge Delegation to China. Join school and district leaders for a one-week educational tour to China...
**Chinese Guest Teacher Program.** Host a qualified teacher from China for up to three years; start or expand Chinese programs efficiently and affordably...

**Summer Institutes in China.** Apply for a scholarship to attend intensive AP and Pre-AP® Summer Institutes in China.” (College Board 2011c).

The addition of the “Chinese Bridge” program is comparable to the “soft diplomacy” approach by the mainland Chinese government reflected in the rapidly proliferating Confucius Institutes continuing to expand throughout the world—as of early December 2010, numbering a total of 322 individual Confucius Institutes in 96 countries and regions (China Post 2010).

**China-based government agency: Hanban**

As is clearly evident from the above discussion, the Hanban is working with both the Asia Society and the College Board in two unique, complementary ways, in the process neither competing for resources nor duplicating effort. With the creation of the Confucius Classrooms program, the focus in their collaboration with Asia Society seems to be more on the material (as in classrooms, curricula and technology) infrastructure development. With the College Board, however, beyond the continuing development of the AP Chinese Course and Exam, the emphasis appears to be more on human resource support and expansion, both for the short- and long-term.

It should be noted that collaborations between a foreign governmental office, either exclusively funded by the foreign government or one with partial governmental support, and an American organization or entity, is not unprecedented. For well over two decades now, the Japan Foundation (primarily through its office in Los Angeles) has provided funding for a range of projects, such as the Standards for Japanese Language Learning, and JOINT (Japanese Online Instruction Network for Teachers, a professional development support system). The Japan Foundation and its offices throughout the world are supported through a number of funding streams, including Japanese government endowment and annual government subsidies. But grants by the Japan Foundation in the United States have been primarily made to the several professional Japanese language teacher organizations—the Association of Teachers of Japanese (ATJ; comparable to CLTA, with most members teachers of Japanese at colleges and universities), the National Council of Japanese Language Teachers (NCJLT; like CLASS, composed primarily of elementary and secondary school Japanese language teachers), and the Alliance of Associations
of Teachers of Japanese (AATJ; the umbrella administrative office supporting both ATJ and NCJLT). In contrast, the various projects discussed above and funded by the Hanban, while ultimately indirectly supporting the goals of both CLTA and CLASS, do not directly financially support the work either Chinese language teachers association.

To its credit, Hanban has been an extremely influential force in bringing together a wide range of stakeholders in the development of the Chinese language field in the United States. No place is this more evident than the annual National Chinese Language Conference (NCLC), which will be held for the fourth time in April 2011 in San Francisco. It is once again being jointly organized by Asia Society and the College Board. In only four years, it has grown to drawing nearly 1,000 participants from an ever-increasing range of 19 collaborating organizations and agencies, as represented below, including national membership associations of language supervisors, principals and administrators, a range of NGOs, and the United States Department of Education (Asia Society 2011d).

The 2011 Conference also represents a watershed between the three previous years and this year. In each of the past three years, a separate national “Chinese Education Conference” was hosted by the Chinese American International School (CAIS) in San Francisco. Consistent with the seemingly continuing emphasis on rationalizing economic and other resources, CAIS, through a new NGO of its own dubbed the Mandarin Institute, has joined Asia Society and the College Board to serve as a “collaborator” along with the other 18 entities listed above for the 2011 conference.

**The future of sector roles in Chinese language field development**

The first decade of the 21st century has seen the beginning of an arguable sea change in the role of the various national capacity sectors in the development of Chinese language education in the United States.

For the federal and private sectors, which are (particularly in the greater Washington, D.C. area) integrally and inextricably linked for many of the critical LCTLs, the role they can and may come to play in the overall Chinese language field development will most likely remain highly circumscribed and comparatively limited in overall field impact.

There is no question that the academic sector and its representative professional language teacher organizations (i.e., CLASS and CLTA) retain a role in the overall progress of field. But economics are a harsh reality, and in the case of CLASS in particular, which does not have sufficient funding to provide a
salary of any sort for its executive director, it is the “kindness of friends” (cf. strangers) such as the Asia Society, College Board and Hanban that are integral to their success, or even survival.

As for the heritage sector, the jury remains out at this point. Even while they have come to assume some unique and genuinely important educational roles, such as the aforementioned FWCC programs, the hope that many of us held with the announcement of the AP Chinese—that is, that the AP Chinese Course and Exam would serve as a means of facilitating greater horizontal articulation between the compulsory K-12 educational system and community schools—has only been realized to a very limited degree.

The extremely influential roles, both qualitatively and quantitatively, that the two new sectors of domestic NGOs and foreign governmental entities have already assumed have proven to be both a blessing and a challenge for the Chinese language educational field. On the positive side, that the College Board and Asia Society have seen fit to prioritize Chinese language and culture education reflects well on what will hopefully be a continuing long-term commitment to the health and growth of the field. But the political repercussions that comes with the perception of accepting foreign government support for American public schools from a nation that is and will most likely remain the number one challenge to the continuation of United States leadership in the world—economically, militarily and otherwise—are not inconsiderable.

It is also interesting to note that in addition to the roles being played by the Asia Society and the College Board, there is a third NGO that has begun to emerge, or more accurately, re-emerge. The National Foreign Language Center (NFLC), housed at the University of Maryland, played a series of important roles in the multi-sector development of the field of Chinese as a foreign language (CFL) throughout the 1990s and early 2000s. In addition to the previously mentioned Chinese Language Field Initiative, projects from the first two decades in the life of NFLC included:

- Providing advisory and consultative support for the creation of the CLASS in 1987 and NCACLS in 1994.
- Funding of a major project that produced the *NFLC Guide for Basic Chinese Language Programs*, first published in 1997, with a 2nd edition released in 2006.
While the first few years of the 21st century saw a fundamental transformation in the mission and infrastructure of NFLC, what is particularly notable is that since approximately 2005, there has been a significant revival of attention to Chinese language field development initiatives in a number of forms, including:

- **Read Chinese!**, with funding by the United States Department of Education, focusing on online e-learning reading lessons for beginning and intermediate students of Chinese at the secondary school level.

- **STARTALK**, with funding by the Office of the Director of National Intelligence, supporting summer programs for both students and teachers of a number of LCTls, Chinese most prominently included, as well as a website with access to online resources including curriculum development programs, classroom video collections, bibliographic lists of language-specific materials, and lesson plan templates.

- The hiring in 2010 of Dr. Shuhan Wang, formerly Executive Director for Chinese Language Initiatives at the Asia Society, as NFLC Deputy Director—the first hiring of a Chinese language specialist as a full-time staff member at NFLC in nearly a decade.

It would not be at all inappropriate for NFLC, where Brecht and Walton first developed their sector analytical model, to rejoin the ranks of the NGOs truly leading the Chinese language field.

**Conclusion**

The last half of the 20th century was marked with a remarkable degree of regularity in the appearance and development of two LCTL fields—Russian in the late 1950s and throughout the 1960s, and Japanese in the 1980s. In large part, what has enabled teachers of these two languages to sustain their fields in the face of major world changes, either political (the breakup of the Soviet Union and the Eastern European Communist Bloc) or economic (the significant economic downturn for Japan in the 1990s), has been in the ways that they have taken control of their field’s enhancement—or sometimes even survival. In the case of Russian, it was made possible by the creation of an entirely new organization, the American Council of Teachers of Russian, separate from the American Association of Teachers of Slavic and East European Languages (AATSEEL), in some ways similar to the genesis of the American Council on the Teaching of Foreign Languages, which was constituted out of a subset of members of the Modern Language Association. For Japanese, the very different teaching
environments and professional demands prompted first the formation of a completely separate organization for K-12 teachers, the National Council of Japanese Language Teachers (NCJLT), which then later joined with the older, primarily higher education-focused Association of Teachers of Japanese (ATJ) to establish the Alliance of Associations of Teachers of Japanese (AATJ)—an initially tactical, but over time strategic approach to rationalize relatively meager resources.

As a veteran (or survivor) of over two decades of seeing Chinese emerge as the Russian or Japanese of the first quarter of the 21st century, I can say with considerable confidence that neither such model—that is, an ACTR for the Chinese language field, or an AATJ-type alliance of the existing Chinese language teacher organizations and teacher-supporting agencies (e.g., the National East Asian Language Resource Center at Ohio State University)—is possible in the near (and very likely far) distant future. The challenge that therefore remains is that lacking the organizational glue to truly help shape the Chinese language pedagogical field in toto, what can the existing professional organizations do to provide some sort of coherence to the field’s development? Otherwise, we as teachers and students of Chinese will be beholden to those with the most resources at their disposal; namely, the Asia Society, College Board, Hanban, and increasingly, NFLC. However good their intentions, Chinese language education is only part of their agenda. As Chinese language educators, it is our agenda.

References


Russian Heritage Language Speakers in the U.S.: A Profile

Olga Kagan

Introduction
Brecht and Ingold (2002) advocate systematic efforts to develop heritage language (HL) pedagogy to remedy U.S. language deficits: “...because of [heritage language learners’ (HLLs’)] existing language and cultural knowledge, they may require substantially less instructional time than other learners to develop these skills. This is especially true for speakers of the less commonly taught languages” (p. 1).

Russian is one of those less commonly taught languages in the U.S. that is critically important for national security and the global economy. Since the early 1970s, when a large wave of Russian-speaking immigrants began to settle in the U.S., American universities have had to adjust their teaching of Russian as a foreign language to accommodate these immigrants’ children. Students who spoke Russian at home and enrolled in Russian programs that mainly catered to learners of Russian as a foreign language have become a familiar sight in Russian programs in the nearly forty years that have since passed.

Nevertheless, HL teaching methodology is still a subject of lively debate, and most programs continue to struggle in their efforts to blend heritage and non-heritage curricula into one coherent whole. Additionally, with an emphasis on high level proficiency (cf. Flagship mission), learning how to teach these students may lead to more Americans speaking Russian at an advanced or superior level, an achievement that currently eludes most students of Russian as a foreign language.

This paper’s main goal is to present a profile of Russian heritage speakers based on data from a survey by the National Heritage Language Resource Center.2

Linguists are primarily interested in HL speakers (e.g. Polinsky, 1997, 2006), while teachers are primarily interested in HL learners (e.g. Kagan, 2005),

---

1 This paper was presented at the 18th Congress of Scandinavian Slavists in Tampere, Finland in August 2010.
2 Richard Brecht has served on the Board of Directors since the Center’s inception.
i.e. HL speakers who study their HLs. There is a considerable body of research analyzing the linguistic lacunae of Russian heritage speakers (see, for example, Heritage Language Journal, 6(1), Spring 2008, special issue on Russian as a heritage language: http://www.heritagelanguages.org/). However, while research into HL speakers’ linguistic features is necessary, it cannot serve as the only determining factor for curriculum design. Polinsky and Kagan (2007) attempted a study of both HL speakers and HL learners in a joint paper, but such collaboration between a theoretical linguist and a teacher is still rare.

A student’s proficiency at the onset of a program as well as his or her potential proficiency may depend on a number of factors: age at immigration, use of language in the family, motivation and affect, etc. By reviewing various factors forming a multifaceted lingua-social portrait of a Russian heritage language learner, we hope to provide instructors and program designers with some background information, which could serve as a backdrop for program development.

**Russian Speakers in the U.S.**

Limiting our discussion to the American context, a heritage speaker of Russian is an individual who grew up in the U.S. speaking Russian at home but was educated mostly or exclusively in English. Such an individual is a bilingual whose weaker language is Russian. “Russian heritage learners’ […] level of competency in Russian is directly tied to the amount of education they received in the former Soviet Union,” which is, in turn, related to the wave of immigration that brought them to the U.S. (Kagan & Dillon, 2006, p. 87).

**Immigration from the Former Soviet Union**

There were four waves of immigration from Russian-speaking countries in the 20th and 21st centuries. The first wave left Russia after the revolutions of 1917. Most of these émigrés’ went to Europe and came to the U.S. in the years preceding or immediately following World War II. A second wave consisted of those who found themselves outside of the Soviet Union after WWII, and did not wish to go back. The third wave began leaving the Soviet Union in the early 1970s and was largely Jewish, settling primarily in Israel or the U.S. This wave lasted until the collapse of the Soviet Union, which brought on a paradigmatic shift in the former Soviet republics’ immigration policies. For more detailed discussion, see Kagan and Dillon (2010). Andrews (1998) wrote that according to the 2000 U.S. Census, most third-wave immigrants came from large cities, 92% of them had high school diplomas, and 51% had received some form of higher
education; he called them “a sophisticated and cosmopolitan group of immigrants, appreciative of their rich cultural heritage who are consciously adapting to life in a radically different society” (p. 55).

The fourth wave started after the collapse of the Soviet Union, when it became easy to leave Russia and other former Soviet republics. Its composition was much more diverse in terms of ethnic and geographic origin, linguistic traits, and level of education. Many came from former Soviet republics where they grew up with two languages: Russian and the local language. Even if they did not speak the local language, it may have played a role in the baseline Russian they spoke. This consideration is of importance when we discuss Russian HL learners’ linguistic needs. Some researchers (Zelenin, 2007) identify a “fifth wave,” describing it as a “brain-drain wave” of high-level professionals who find jobs in the U.S.; they may intend to return to their country of origin, but stay so long that their children grow up in the U.S. According to the 2007-08 community survey (U.S. Census), there are over 851,000 Russian speakers in the U.S. The largest populations reside in New York, (29.5 %), followed by 6.3% in California.

A Russian HL Learner
A heritage speaker of Russian who studies Russian at an American educational institution is a heritage language learner (HLL). Russian HLLs today are mostly children of one of the recent immigration waves, i.e., the first generation, who were born in the U.S., or the 1.5 generation, who were born outside of the U.S. but arrived at an early age.

A fairly typical example of such a learner is Igor N. I worked with him individually for ten weeks in Spring 2010. Igor met with me once a week for sixty to ninety minutes. The main task was discussing the text Igor was reading: Aleksandr Pushkin’s The Captain’s Daughter (Kapitanskaia dochka). His experience with Russian is as follows: Igor was born in Moscow. His parents met as students at Moscow State University; his mother was from Krasnodar, and his father from Kirovograd. Igor’s parents moved to the U.S. when he was two years old. He used Russian at home, and his grandmother taught him to read and write a little. In college, he majored in science, and he is currently a graduate student in nanoscience. He decided to take a Russian class in his senior year because he needed FL credit, had some free time, and wanted to improve his Russian. He could not take the first-quarter HL class because he had a class conflict, so his

---

3 UCLA IRB # G071103501. I do not mean to imply that all Russian heritage learners speak exactly like Igor. However, I am sure many of the readers who teach these learners would recognize that the patterns of his speech easily compare to those of their own students.
grandmother worked with him for a few weeks to bring his literacy up to par. He then took the second-quarter HL class (ten weeks in winter), at the end of which he read Pushkin’s *Dubrovsky*.

In the spring quarter, Igor spent ten weeks doing an independent study with me. In addition to working on some grammar from the textbook *Russian for Russians*, he read Pushkin’s *Kapitanskaia dochka*. We spent about one hour a week discussing the story. Igor’s tasks were to recall the events and characters, to ask questions if he didn’t understand, and to answer some of my questions. I followed some recommendations from the recall protocols research (Bernhardt 1983, Carlisle 1999) and recorded all of our meetings.

Instructor: 2-я глава называется «Вожатый» и что в ней было?
The second chapter is called “The Guide” and what happened/happens in it?

- Igor. There goes an interesting story there how they get into / a blizzard and \ they accidentally saw man, a guide, who carried them into / a permanent [postoiannoe] / (forgot what it’s called) so as to spend the night until [the time when] the weather improves (Instructor: - staging inn [postoialyi dvor]; I. repeats after her). And / when the next day Peter gives / his jacket to the man as a gift / \ his \ peasant [the student does not know the correct word for ‘peasant’ in Russian] / \ well, the one with whom \ he everywhere goes, he to him this idea didn’t appeal, but Peter said that he has to give him something, thank him. The guide he was I think also it looks like a drunk and didn’t do anything/ that is the why \ \ Ser, he is only called by... Savel(и)ча

Translation: It’s an interesting story. The characters are caught in a blizzard when they happen to see a man, a guide. He escorts/leads them to a staging inn, where they can spend the night until the weather
clears/improves. The following/next day, when Peter gives the man his jacket as a gift, his servant/peasant, who accompanies him everywhere, objects/doesn’t like it. But Peter insists/says that he has to give (the guide) something as a sign of gratitude. I also believe the guide is a drunk and loafer/ne’er-do-well; this is why people refer to him using only... as simply Savel//ich.

- Instructor: зовут по чему?
- call him by what?
- Igor, по отчестве... по отчеству
- with the patronymic... by the patronymic
- Using only the patronymic.

Igor’s narrative demonstrates the following:

- Good aural comprehension.
- Fairly good pronunciation with some inadequacies.
- Igor has a high degree of fluency, if we understand fluency as a participatory exchange. McCarthy (2006) asserts that fluency cannot to be judged by monologic criteria, but is rather a joint production: conversations (rather than individual speakers) are fluent or non-fluent, with the notion of confluence being central to conversational fluency.
- Good and serviceable strategic competence (забыл, как ... / I forgot how to say...).
- Ability to self-correct (по отчестве/y / using his patronymic)
- Some evidence of complex syntax (до тех пор, как / until [the time when])
- Abundance of lexical and grammatical mistakes.
- Some exceedingly long pauses.

This brief list contains both strengths and weaknesses, but the former outweigh the latter. Igor went from lack of literacy to reading Pushkin’s novella in less than ten weeks. It would take a typical FL student of Russian several years of intensive study to accomplish the same.

This itself points to a difference between HL and FL students, and, consequently, between the optimal curricula for each group. In teaching a student like Igor, what should our focus be? To frame the question more broadly, what kind of curriculum would enable HL students to make the largest gains or, perhaps more importantly, would not hold them back?
A profile of a Russian HLL

Research into the factors important for home-language preservation is still inconclusive. According to Fishman (1978), the loss of the immigrant language typically happens within three generations: the immigrant generation speaks the language, their children are English dominant but continue using the language (what we now call heritage speakers), and their grandchildren are typically monolingual. Lopez (1996) asserts that “[A]sian languages are hardly maintained at all beyond the immigrant generation” (p. 139). Examples of such rapid language loss are evident among Russian immigrants as well. Still, as teaching Russian HLLs shows, some second generation speakers preserve their HL better than others. What factors play a role in one’s maintenance or loss of the home language?

According to Montrul (2008), the earlier a child comes into contact with the dominant language (English) and begins using it more than the HL, the weaker his or her knowledge of the HL is likely to be. Carreira and Kagan (in press) show that the earlier heritage language speakers arrive in the U.S., “the less likely they are to use their HL and the more likely they are to use English to the exclusion of this language.”

These factors may be beyond anyone’s control, as they depend on the family’s immigration history and the parents’ choice of language use at home. Some other factors, though, may be controllable. Au (2008) speculates that “storage strength and retrieval strength of long-ago memory holds at least part of the answer” (p. 339) to language maintenance. She indicates two paths for keeping HL alive: using the language “beyond early childhood” (p. 347) and engaging in relearning the language. She shows that even overhearing the language in childhood helps adults relearn it.

Thus, both age of emigration and use of language at home become crucial factors of language maintenance. Using the language with peers and travel to the home country also help to preserve the language (Hinton, 2001); in fact, travel may promote use of language with peers. Hinton does not believe that attendance at afterschool, weekend, or church schools helps students maintain the language. Still, these programs may motivate (or demotivate) students, and may thus play a role in language maintenance or loss.

In college, many students decide to take courses in their home language in order to improve or relearn it. It may be useful to examine their motivations. An additional factor that may contribute to language maintenance is self-identity, which is always important in a child’s development and learning (Hornberger & Wang, 2008, p. 7). It is easy to see how self-identity would be a crucial factor in
studying one’s home language, which was acquired first and was one’s dominant or only language in childhood.

I will examine the aforementioned factors (age, use of language, travel, attendance at weekend schools, motivation, and self-identity) in order to arrive at a typical profile of a Russian HLL.

The Survey of Russian HLLs

The data comes from a survey conducted by NHLRC in 2007-2009. To date, 1,800 HLLs in 22 languages have taken the online anonymous survey. A total of 219 of the respondents were Russian HLLs. While the latter number may not in itself be impressive, in combination with and comparison to the overall responses, the data provided by Russian HLLs may offer valuable insights.

Most of the Russian respondents were from New York and New Jersey, while the next largest group was from California. This distribution of Russian responses roughly corresponds to the distribution of Russian speakers in the U.S. (see discussion above). The majority of respondents were between 18 and 21, the age of a typical undergraduate. 60.6% percent of the respondents were women. Over 70% report being born in the former Soviet Union. This indicates that the U.S. Russian community is relatively new. 69% of all other HLLs were born in the U.S. Most of the Russian speakers, on the other hand, came at an early age; over 60% never attended school outside the U.S.

57% live at home with their parents while attending college, a somewhat higher percentage than for the other HLLs surveyed (45%). Students report that they began speaking more English than Russian after age five. 43%, however, continued speaking Russian at home. 42% have never traveled to a Russian-speaking country. Fewer than 3% travel to a Russian-speaking country every year. By comparison, 85% of Chinese and 87% of Korean HLLs travel to their home countries regularly. Even Persian-speaking students travel to Iran more frequently than Russian speakers travel to the former Soviet Union: 64% have been at least once.

Many Russian HLLs gain literacy late, in college. Lavretsky et al. (1997) note that Russian families “generally do not insist on speaking Russian to their children and grandchildren. It is quite common that children who came to this country before entering school or elementary school do not speak, read, or write in Russian” (p.337).

Even though families might not compel or encourage their children to speak Russian, 72% of the Russian HLLs surveyed spoke the language at home.
until starting school and, at times, knew no English up to that period. As one student writes, “Russian gave me a tough start with school in America, but after I learned English, it became more of a useful tool to me.”

Once these students start school, however, they seem to acquire English rapidly and may even cease speaking Russian altogether. Another student writes, “My mother would speak Russian and I would reply in English. I rebelled against speaking Russian and Russian culture until about the age of 17/18.”

Even though the use of Russian reportedly diminishes after starting school, students continue using it in some way. 40% report using a “combination of Russian and English,” while 7% speak only Russian and 34% speak only English. 15% claim that they have no preference. Every immigrant population has a name for these combinations: “Konglish” (Korean and English), “Russlish” (Russian and English), etc. “Spanglish” is so widespread that it constitutes a popular discourse, with plays and TV shows written in the language. Carreira and Kagan (in press) report on a pilot study of 36 Spanish HLLs that shows how Spanish and English combine in their daily lives: “with grandparents, nearly all (91.67%) report speaking only Spanish. With their mothers, 25% speak only Spanish and another 33.33% speak mostly Spanish. On the other hand, with siblings, many report using English and Spanish in equal amounts (27.78%) or speaking mostly English (52.78%), and none reports making exclusive use of Spanish with siblings. All respondents reported mixing English and Spanish” when speaking either of the languages. A similar study of Russian speakers is in preparation (Kagan). It appears that the majority of the respondents are still exposed to Russian at some level, even though most of their communication is in English.

In answer to the question, “What did you do in Russian outside of class in the past six months?” students report speaking on the phone (90%), watching TV or video (69%), listening to music (75%), and visiting a website (52%). Between 30% and 40% read a newspaper or a book or a short story, but 18% report never reading in Russian; 26% read in Russian less than fifteen minutes a day, and 20% read fifteen to thirty minutes a day. That differs sharply from time spent reading in English: 70% report reading one to two hours or more outside of school. These numbers show that students’ exposure to Russian continues, mainly in the form of input, but some output (telephone conversations) is also in evidence.

A small percentage (14%) had attended community events. This particular finding underscores some basic differences and similarities between the Russian and some other language communities in the U.S. While, like their Russian
counterparts, only a small percentage (16%) of Chinese respondents had attended community events, a full 50% of Korean and 30% of Persian respondents had done so. Another striking characteristic of Russian HLLs is the fact that less than 15% have attended Russian community or church schools. This finding is very different in regard to Chinese (44% attended more than one year) and Korean (40% attended more than one year), but is similar to Persian (70% never attended weekend or community schools).

How do Russian HLLs assess their language ability? They feel that they are close to native speakers where listening proficiency is concerned, but are mostly at the intermediate level in all other skills, including speaking.\(^5\)

In response to more detailed questions about their perceived proficiency in Russian, respondents felt that they were able to eavesdrop, understand humor (aural ability), use polite language, and be rude (oral proficiency). Their almost daily contact with the language for over 18-20 years certainly justifies such claims. These claims also differentiate HLLs from FLLs.

What do Russian HLLs want to learn in class? Respondents indicate taking Russian classes for three main reasons: (1) communicating better with family and friends in the U.S.; (2) learning about cultural and linguistic roots; and (3) communicating better with family and friends abroad.

When asked what they would most like to learn in class, the majority of respondents identify increasing vocabulary as their primary objective. That is also supported by other HLLs’ responses, with the exception of Chinese students, whose main concern is learning to read and write.

In response to the question about what they want to read in Russian classes, students indicate novels and short stories (84%), followed by poetry (52%), which further evinces their interest in cultural roots. It also indicates that Russian HLLs are close to their families and are aware of the importance of literature and poetry in Russian culture.

The survey also addressed the issue of identity. The question was formulated in the following way: “How do you self-identify (e.g., American, Vietnamese, Vietnamese-American, etc.)?” Responses differed considerably, but most students indicated a dual identity, with a minority saying “Russian” or “American” alone. Here are a few examples: “Russian/American,” “Ukrainian-American,” “American Russian,” etc. In a few cases, identities were more complicated: “Persian-Russian,” “Russian and Serbian and American,” “Russian American Jew.” Even if no additional identities, such as Persian or Serbian, are

\(^5\) These are not ACTFL scale assessments. The scale in the survey is a common sense range, Novice to Native-like.
involved, Russian Jewish immigrant children, for example, must still “struggle with their position in three cultures—Russian, Jewish, and American” (Lavretsky et al., 1997, p. 339). The same may be true for children of families who came from the former Soviet republics. They may understand their identity as “Russian Armenian American,” “Russian Ukrainian American,” etc. It seems that students have a keen sense of their multiple identities and “hybridity” (Bhabha, 1994).

A few examples from the survey’s open-ended responses may help underscore the affective nature of Russian HLLs’ attitude toward Russian culture in their lives. Respondents’ spelling has been preserved.

Example 1. For half of my life, we lived in West Hollywood, so Russian was a big part of that. I guess it was helpful in the sense that it allowed me to communicate with various shopkeepers and neighbors.

Example 2. My knowledge of my heritage language has allowed me to communicate with my grandparents and other family members. It has also allowed me to go to my homeland and better understand the people who live there. At the same time, it has helped me have friends who come from a similar background.

Example 3. Speaking Russian is helpful because there are so many native speakers in New York City. I constantly overhear people speaking in Russian and it often just makes me comfortable to ask a stranger something.

Example 4. My HL is VERY helpful in my church, where 85% of the service is in Russian. And i speak russian with all the people there and sing songs in russian too. at home, i am able to communicate with my relatives, like grandparents who can’t speak any english. I have also been able to talk to strangers in russian in the stores or ask for directions or something in public.

Example 5. i love knowing russian. i have made a great number of friends out of school because of my ability to speak russian especially in the local russian community.

I have tried to find instances of purely negative affect, but it is truly surprising how positively students’ view their Russian heritage. We should, of course, keep in mind that all respondents were taking Russian at the time of the survey, i.e., the audience was, to a degree, self-selected. Below are a few examples that show that some respondents were at times made uncomfortable by their knowledge of Russian. However, some add positive comments to these responses as well.

Example 1. It only affected my experience in school when I was a child because I started kindergarden knowing no english at all. As for right now I
don't think it affects my schooling. My SAT tutor says that it was harder for me to improve on my verbal part of the exam because I was not a native English speaker but i'm not sure how much of an impact that had.

Example 2. Since i have learned Russian first, it made it harder for me to learn English. I was able to learn to read pretty fast because i always loved to read, but i still had difficult time with writing essays. My heritage language also made me more shy and unsure of myself, therefor making it harder to make new friends.

Example 3. I find people's reactions to my heritage language to be sometimes annoying when I live in places where there is little diversity. I very often make friends on the basis of a second common language and a shared culture. Knowing Russian also helped make learning German a little easier. Russian is a good language to have for me because I am majoring in Mathematics, and many very good mathematicians speak Russian better than English, and it is a convenient language in the setting of (for example) an academic conference, or one on one with a professor.

Example 4. The only time that Russian made school difficult for me was when we first moved to the United States. I was young and did not know the language so it was very difficult for me to communicate in school. However, I picked up English very quickly and after that knowing Russian never caused me a problem. Now, I want to learn how to speak better in Russian.

Example 5. I went to religious Jewish schools up until high school, and there, it was highly undesirable to be Russian. So Russian held a sort of stigma for me in that setting, but I liked to talk to people in Russian outside of school in Russian.

Whether positive or negative, their attitude to the language makes them different from learners of Russian as a foreign language.

A General Profile of a Russian HLL
The following is the general profile of a Russian HLL in the U.S. that emerges from the survey:
1. First generation U.S.-born or 1.5 generation (arrived approximately before the age of 10).
2. Sequential bilingual: spoke Russian only before starting school.
3. Continues to use some Russian at home.
4. Retained some proficiency in speaking Russian and is comfortable with aural comprehension. Not infrequently starts speaking more Russian in late adolescence or young adulthood.
5. In college, becomes interested in learning about cultural and linguistic roots and improving language proficiency, particularly in expanding vocabulary.

6. Has a double or triple identity.

Curriculum Development Based on the Profile
The profile above demonstrates that Russian HLLs are different from learners of Russian as a FL in their prior exposure to, experience with, reasons for studying, and emotional attachment to the language, as well as in their self-identification, which determines their motivation. As Au shows (see discussion above), the main difference may be between learning the language by FLLs and relearning it by HLLs. The question is not which group is better at learning Russian, but what curriculum they need in order to learn it well.

Also, considering HLLs’ use of the language throughout their lifetime (no matter how limited or flawed it is), we can expect that they may be able to reach a higher level of proficiency faster than non-HLLs. Brecht and Ingold’s recommendation quoted at the beginning of this paper can only be realized if we teach these groups differently and set different goals for each.

The following principles can be outlined:

1. An understanding of the importance of students’ background, including their families’ immigration histories. The age of arrival and exposure to language make a considerable difference in one’s linguistic and cultural competency and ability to gain higher proficiency. This biographical information can help determine what program would best suit a certain group of students, and can be collected through surveys and interviews.

2. An understanding of the role played by motivation and affective factors stemming from the language used in the family. Since students indicate “communicating with family and friends in the U.S.” among the main three reasons for taking Russian courses, the curriculum should be oriented toward fulfilling this goal. Better communication is impossible without cultural knowledge, which should constitute the core of the curriculum. This has also been recognized by researchers of other HLs. For example, Merino et al. (1993) propose including the home language and culture(s) of Spanish HLLs into curriculum design. They also stress utilizing Vygotskian principles in order to develop learning communities where students would interact not only with their instructors and other students, but also with family and community members [see also Faltis
(1990); Rodriguez Pino (1994)]. The same approach would aid Russian HLLs. Like Spanish-speaking students, Russian HLLs come from many different countries that were part of the Soviet Union before its collapse in 1991. Consequently, their language may contain traces of others (both Slavic and non-Slavic). An effective teacher must understand why students use certain ungrammatical forms, non-standard vocabulary, etc. He or she must also understand these students’ multiple identities.

3. Focusing on what is important for students themselves and rethinking more traditional approaches to the curriculum. Increasing vocabulary emerges as one of Russian HLLs’ main concerns. Targeting curriculum at vocabulary development may yield an increase in proficiency more directly than focusing on grammar and spelling. Research into methodology that foregrounds vocabulary expansion would be most helpful to the profession.

Conclusions and Further Research
To conclude with another quote from Brecht and Ingold (2002), the U.S. has “an unprecedented need for individuals with highly developed language competencies” both in English and in many other languages. Given the importance of Russian in world history, diplomacy, the economy and intellectual development, teaching Russian to HLLs who can gain a high level of proficiency is not only a pedagogical, but also a societal need. I hope that the Russian HLL profile described in this paper shows convincingly that these students need a different curriculum than learners of Russian as a FL. In order to achieve better results in teaching Russian as a heritage language in the U.S., researchers and practitioners would benefit from a corpus of heritage learners’ language, both written and oral, exploration of the attitude of families and communities regarding language preservation, and a database of existing programs of Russian as a heritage language. Such studies may lead us to a comprehensive programmatic approach to heritage language teaching in Russian. Because Russian teaching in this country has a long history and may be more developed than pedagogy in other less commonly taught languages, developments in Russian may also lead to improving teaching of heritage languages in general. Ultimately, the goal of improved pedagogy would be to prepare citizens “who are linguistically and culturally savvy” (Tse 2001, p. 49-51) to pursue their own educational and intellectual goals, to advance the international interests of the U.S., and to play an important role in the global economy.
References


Blackwell Publishing on behalf of the American Association of Teachers of German.


Русификация и казахизация как языковая гомогенизация многоязычного Казахстана

Элеонора Сулейменова

Введение
Русификация и казахизация как процессы целенаправленной гомогенизации в языковом отношении располагаются в одном ряду с белорусификацией, украинизацией, полонизацией, узбекизацией, латышизацией, таджикизациией или исландизацияей, и деятельность политических, образовательных, общественных, экономических, собственно лингвистических институтов по их реализации рассматривается как обычная государственная практика.

Использование названных терминов в разных социально-политических контекстах привнесло в их содержание различные коннотации, как правило, негативно-оценочного характера. Тем не менее, при анализе русификации и казахизации¹ по возможности постараемся избежать каких-либо оценок, в этом, надеюсь, поможет то, что речь пойдет о диаметральных процессах языковой гомогенизации одной и той же страны и даже одного-двух поколений ее жителей.

Русификация – одно из имплицитных направлений языковой политики в дореволюционной России, СССР, Российской Федерации, главной целью которого является языковая гомогенизация многоязычного общества: а) процесс осуществления комплекса мер, нацеленных на распространение русского языка на какой-либо иной (в том числе не исконной) территории (русификация Финляндии, русификация Кыргызской Республики); б) результат такого процесса (графическое, лексико-семантическое, грамматическое сближение какого-либо иного языка с русским языком).

Казахизация – одно из имплицитных направлений языковой политики в Республике Казахстан, главной целью которого является языковая гомогенизация многоязычного общества: а) процесс осуществления комплекса мер, нацеленных на развитие и распространение казахского

¹ Отдельные значения терминов русификация и казахизация, которые реализуются в таких нейтральных контекстах, как русификация или казахизация алфавита (из рекламы мобильных телефонов), русификация или казахизация программного обеспечения (из текстов по информатике), естественно, в статье не рассматриваются.
Русификация и казахизация как языковая гомогенизация многоязычного Казахстана
Элеонора Сулейменова

языка на исконной территории и усиление в соответствии со статусом государственного языка его влияния во власти; б) результат такого процесса (графическое, лексико-семантическое, грамматическое сближение какого-либо иного языка с казахским языком).

Русификация и казахизация осуществлялись и осуществляются в сложных условиях политической интеграции, консолидации многоязычного населения на их пути к гражданской идентичности. Такой путь естественным образом ведет к изменению соотношений между языками, каждый из которых стремится усилить собственные позиции, в первую очередь по отношению к политически доминирующему языку, и получить соответствующие гарантии государства. Государство со своей стороны, учитывая возникающие ограничения в развитии иных языков, кроме политического, может реализовать две возможности: а) реально осуществлять системную поддержку языков в виде политики мультикультурализма (языкового плюрализма), б) ограничиваться декларациями и не предпринимать в отношении этих языков необходимых усилий.

Языковая политика Советского Союза и Республики Казахстан была направлена не только на проведение языковой гомогенизации, но и на решение острых проблем языкового плюрализма. К сожалению, все аспекты реализации устремлений обоих государств к балансу между разными направлениями языковой политики в настоящей статье не могут быть рассмотрены. Основное внимание в ней сосредоточено больше на идеологии и процессе осуществления языковой политики русификации и казахизации как языковой гомогенизации многоязычного Казахстана. Поэтому разделы статьи посвящены соответственно процессам русификации и казахизации в Казахстане, а некоторые результаты русификации казахского языка и казахизации русского языка² представлены лишь фрагментарно.

2 Речь, естественно, идет только о казахизации русского языка в Казахстане как одном из регионов его активного функционирования. Русский язык в Казахстане испытывает все более возрастающее воздействие со стороны казахского языка. Приведем четыре весьма показательных отрывка из казахстанской газеты «Взгляд» (2010): (1) ...аренда шестиканатной юрты обойдется в 30 тысяч тенге, а восьмиканатной – в 40 тысяч тенге... (2) Казахские баксы и кюйши однаково поклонялись Коркуту, и если баксы исполняли сарыны Коркута, то кюйши – его кюи. (3) ...добавьте к этому дополнительные материалы – арканы, кошмы, баскуры... и еще кили, тюкши, тектемет, сырмак, сандык, тосекагаш... (4) Сырмак из кошмы бело-коричневого цвета, баскур из цветной тканой узорной ленты, токым на кошмой или тканевой основе,
Русификация как языковая гомогенизация многоязычного Казахстана

В данном разделе остановлюсь на нескольких общих моментах идеологии русификации советского периода, продолжающих оказывать значительное воздействие на языковую ситуацию в современном Казахстане.

Особая роль в процессах строительства советского государства, как известно, отводилась русскому языку, но русификация в Советском Союзе носила имплицитный характер: даже законов, подтверждающих официальный или государственный статус русского языка, не существовало. Вместе с тем русификация, основанная на латентной идеологии моноязычия (точнее, руссковысокой), вполне откровенно проявлялась в основных Конституциях СССР, Программах КПСС, постановлениях правительства и научных изданиях.

Политика русификации может быть обобщена в виде следующих шести широко использовавшихся в прошлом и несколько модифицировавшихся в настоящем тезисов, которые, конечно же, не исчерпывают собой полностью ее содержание и характер. Более того, такое изложение, возможно, представит политику русификации в несколько схематичном и обостренном виде.

1. Тезис о равноправии всех наций и языков звучал и продолжает звучать многообещающе, однако идеологической базой политики русификации

Вышитый хлопчатобумажными и шелковыми нитями, бостек с оо-орnek, бесик из натурального дерева... Конечно же, приведенные примеры были специально отобраны из текста, заведомо содержащего специфическую национально-культурную информацию и насыщенного безэквивалентной лексикой. Сделано это было также намеренно, чтобы продемонстрировать встречающееся насыщение отдельных текстов казахской безэквивалентной лексикой с непривычными буквенно-звуковыми сочетаниями (сравните сандык при наличии в русском языке успешно освоенного тюркизма сундук), казахско-русскими омонимами, провоцирующими неправильные ассоциации (шестиканатная юрта – это юрта, поддерживающая не шестью канатами, а шестью сегментами основания юрты), непривычными производными (кономный от коным) и так далее.

3 О «Русификации как основном направлении языковой политики русского государства во второй половине XIX века», а также «Национальной и языковой политике советского государства» можно прочитать в [Беликов 2001: 347-405].
4 Типичным воплощением рассматриваемого тезиса может служить следующее высказывание: «Теоретическим фундаментом языкового строительства служит языковая политика КПСС, краеугольным камнем которой является ленинский принцип абсолютного равноправия всех народов и их языков» [Исаев 1978]. Подобные тезисы прямо восходили к статье 34 Конституции СССР: «Граждане СССР равны перед Законом независимо от происхождения, социального и имущественного положения, расовой и
было не всегда эксплицитируемое, но четко проводимое положение о превосходстве русского языка.

Проявлением подобного отношения к иным, чем русский, языкам может служить скрытое принижение степени развитости и историко-культурной ценности тюркских языков. «В советский период ...основывались на непреложной аксиоме: есть древнетюркские алфавиты, но казахи, узбеки, кыргызы к ним никакого отношения не имеют», и это неудивительно, поскольку «развитые тюркские языки с древней историей и не менее древней письменностью противоречили идее превосходства русского языка» [Хасанаев 1997: 4] и его легитимности в роли доминантного языка.

Государственная доктрина со временем стала основываться на идее 'старшего брата' и этнической иерархии народов вплоть до национальных меньшинств (так называемых нацменов): «В созвездии союзных республик первой величиной является Российская Социалистическая Федеративная Советская Республика. И первым среди равных является русский народ» [Правда 1936: 1]. Тезис о равноправии всех наций и языков имел сугубо идеологический характер и во многом был дискредитирован советской языковой политикой. Однако не следует думать, что так было только в советской стране: это происходит во всем мире, какими бы своеобразными ни были условия существования языков. Равноправие языков, несмотря на его популярность и привлекательность, в современных государствах также оборачивается привилегиями по отношению к конкретному языку и его дальнейшей институализацией в качестве официального государственного, стандартного, национального, политического, доминантного языка.

Более того, равноправие языков в существовавшем ранее понимании вообще невозможно [Хобсбаум 2005: 33-43], так как в настоящее время, во-первых, языки поставлены в новые условия выживания и конкуренции; во-вторых, кардинально изменились скорость, способы, формы и каналы получения и передачи информации; в-третьих, более стремительным и

национальной принадлежности, пола, образования, языка, отношения к религии, рода и характера занятий, места жительства а других обстоятельств».

5 Приведем интересное мнение о возможности равноправия языков, иллюстрированное примером, неожиданным для наших рассуждений о русификации: «Заверения в приверженности равноправию всех языков часто служат ширмой для дискриминации, ущемления интересов конкретных групп, например русскоязычного населения в странах постсоветского пространства» [Серебряков 2004].

234
нерегулируемым стало распространение мировых языков (в первую очередь английского). Поэтому сегодня, очевидно, скорее следует говорить не о равноправии языков, а их способности дополнять друг друга в соответствии с коммуникативными и прагматическими потребностями.

2. Марксистский тезис о 'расцвете и сближении наций' в Программе КПСС (1961) звучал в таком виде: «Развернутое коммунистическое строительство означает новый этап в развитии национальных отношений в СССР, характеризующийся дальнейшим сближением наций и достижением их полного единства… Границы между советскими республиками в пределах СССР все более теряют свое былое значение.» «Русский язык фактически стал общим языком общения и сотрудничества всех народов СССР» [Программа 1961: 20, 22].

Видим, что по сути сближение и достижение единства народов без границ и с общим языком – это путь к денационализации и полному переходу на русский язык. Приведем лингвистическое разъяснение данного тезиса советского времени: «При социализме отсутствие национального гнета и утвердившиеся отношения равенства, дружбы народов ускоряют как развитие, так и сближение наций. В капиталистическом обществе политика насильственного навязывания одним народам языка и культуры других народов порождает стремление во что бы то ни стало сохранить национальную специфику со всеми ее и прогрессивными, и реакционными элементами. В обществе, свободном от национального гнета, наоборот, народы охотно отказываются от отживших форм и добровольно воспринимают наиболее прогрессивные элементы культуры других братских народов» [Исаев 1978].

Эффективное управление и контроль над всем населением громадной многоязычной страны могли осуществляться в государстве с единым народом, единым языком и единой графикой. Однако политические и идеологические приоритеты при построении советской страны менялись: то необходимо было предотвратить угрозу панисламизма, то преодолеть сильные пантуркистские настроения, то преградить распространение арабского или турецкого влияния, то противопоставить сепаратистским движениям стремление к интеграции...

Смена приоритетов, например, отразилась на неоднократном переводе письменности многих народов с одного алфавита на другой. Казахский язык, как и все тюркские народы, наследовал руническую (орхон-енисейскую) письменность. Затем арабско-мусульманское влияние привело к тому, что казахи стали пользоваться арабской графикой,
приспособленной А. Байтурсыновым к особенностям казахского сингармонизма. В советское время казахи под мощным административно-партийным напором пережили два перевода письменности: на латинскую и очень быстро — на кириллическую графику. Последствия такой перемены графических систем как в целом для народа, так и для его отдельных представителей ощущаются до сих пор: а) потеряно и до сих не восстановлено культурное наследие на арабской и латинской графике (не всякий литературовед может прочитать рукописи Абая); б) один народ — казахи Казахстана и казахские диаспоры — оказался разделен графически; в) ’графический’ шок — не редкость: казахи-репатрианты (оралманы), вернувшись на этническую родину, оказываются неграмотными; казахи, учившиеся в злополучное время стремительной смены алфавита, начинали обучаться на арабской графике, потом формировали новые латинские буквенно-звуковые представления, а заканчивали школьный курс с навыками чтения и письма на кириллице и так далее.

3. Тезис о русском как ’втором родном’ языке, или тезис о ’двух родных’ языках напрямую связан с предыдущим. Советский народ должен был владеть русским языком, которому отводилась роль ’второго родного’ языка. В статье редактора «Вопросов языкознания», посвященной задачам изучения современных национальных языков в свете Программы КПСС, говорится: «Советские языковеды, исходя из основных задач строительства коммунизма, должны уделять особое внимание изучению …закономерностей взаимодействия языков советских народов в современную эпоху, которая характеризуется …в большинстве случаев бурным развитием и обогащением литературных языков, а также развертыванием процессов добровольной языковой смены у бесписьменных народностей» [От редактора 1962: 3-10]. Смена родного языка в сторону ’второго родного’ русского как результат политики русификации коснулась не только бесписьменных народностей, но и многих народов с богатой национально-культурной письменной традицией (среди них, к сожалению, не последнее место заняли казахи).

Система образования стала одним из основных инструментов русификации и распространения русского языка как ’второго родного’: а)

6 „С точки зрения советского государств важным следствием утверждения латиницы стала утрата коранической грамотности среди мусульманских народов. Там, где существовали давние письменные традиции, во многом была подорвана и светская культурная преемственность” [Беликов 2001: 386].
существовало и качественно развивалось общее образовательное пространство, основанное на унифицированных типовых программах по русскому языку, учебниках и учебных пособиях, словарях, наглядных пособиях и прочее, создававшихся в Москве и столицах республик;
b) централизованно и масштабно осуществлялось создание мощной научно-методической инфраструктуры: по методике преподавания русского языка в национальной школе защищались диссертации, выпускались монографии, учебники и учебные пособия; в) в разные годы проводились многочисленные реформы образования, итогом которых стало повсеместное введение русского языка как обязательного предмета во все типы учебных заведений всех уровней (это в полной мере коснулось и системы детских дошкольных учреждений в национальных республиках);

4. Тезис о наличии и необходимости дальнейшего увеличения общего лексического фонда языков народов СССР. Появление данного тезиса было напрямую мотивировано планомерно осуществлявшимся и жестко регулируемым процессом сближения (слияния) наций с общей интернациональнаной (советской) культурой. В национальных языках стал целенаправленно формироваться общий словарный фонд, основу которого составляла интернациональная лексика греко-латинского, арабского, итальянского и другого происхождения, так называемые советизмы, обозначающие явления и понятия советской идеологии и деятельности (партия, совет, колхоз, совхоз, социализм, коммунизм, съезд, философия, бригада, звено, совнархоз, книготорг, райком, райисполком, сельсовет, лейтенант, майор, маршал и многие другие), а также слова языков народов СССР, обозначающие главным образом культурно-бытовые реалии (например, тюркизмы и казахизмы разного времени: айран, аксакал, акын, арба, аркан,
Русификация и казахизация как языковая гомогенизация многоязычного Казахстана
Элеонора Сулейменова

архар, арьк, аул, базар, башлык, бешбармак, буран, кавардак, казан, калым, караача, катык, кумган, кумыс, манкурт, сайга, саксаул, тазыр и многие другие).

Русский язык стал практически единственным проводником интернациональной лексики: заимствования контролировались, и их источником, нормирующим произношение и написание, мог быть только русский язык. Подобная централизация процесса заимствования (а значит, и образования общего лексического фонда) сделала влияние русского языка неограниченным. Широкий поток заимствований из русского языка и через русский язык не останавливался даже там, где в языке уже имелись соответствующие эквиваленты.

5. Тезис о разделении функций между русским и национальными языками. Все языки были объявлены равноправными, однако это не означало, что им предписан одинаковый объем общественных функций.

Русский язык для народов Советского Союза был не только средством межнационального общения, он должен был обслуживать наиболее важные сферы: государственную, хозяйственную, правоохранительную, военную, общественно-политическую и в особенности партийную деятельность в союзном, республиканском, областном и районном масштабе; науку (например, диссертации и квалификационные документы к ним должны были быть написаны на русском языке); съезды, конференции, собрания и совещания; официально-деловое общение. Это, впрочем, не исключало возможности отдельных выступлений, обращений, судебного производства и прочее на национальных языках. И все же главные функции национальных языков сосредоточивались в пределах родной культуры: художественная литература, фольклор, театр, радио, телевидение, газеты и журналы, обиходно-бытовое общение и другое [От редактора 1962: 3-10].

Подобное разделение сфер общения между русским и другими языками, ограничивающее развитие ресурсов языков и их функционально-стилистических возможностей, подвергалось критике. Вместо этого предлагалось, например, «параллельное использование русского и

Аргументы, которые использовались при этом, могут быть иллюстрированы следующим отрывком из уже цитированного источника: «Подобно языкам некоторых других мусульманских народов, в таджикском языке преобладала до революции арабизированная терминология, чуждая народной речи. Сама жизнь потребовала ориентации на русскую и европейскую культуру. Поэтому пришлось преодолеть тенденции, исходящие от ислама, и в дальнейшем развитии лексики и терминологии широко использовать ресурсы народной таджикской речи, а также русскую и интернациональную терминологию» [Исаев 1978].
национальных языков в разных сферах, но с учетом конкретной ситуации общения» [Иванов 1982: 3-13].

Однако разделение функций между русским и национальными языками набирало обороты, и к моменту образования постсоветских государств многие национальные языки, сохраняя символические функции, оказались перед необходимостью выполнять больший объем новых функций, предписываемых высоким статусом государственного языка, а их носители, утратившие (или не приобретшие) соответствующие коммуникативные навыки, оказались не готовы сразу выполнять профессиональные обязанности на государственном языке.

6. Тезис о гармоничном и паритетном национально-русском и русско-национальном двуязычии, как правило, поддерживался требованиями максимального охвата двуязычием населения союзных и автономных республик и областей, широкого применения двуязычия во всех сферах народнохозяйственной жизни, высокого уровня владения русским языком. Гармоничность и паритетность обоих видов двуязычия в сочетании с вышеприведенными положениями политики русификации на практике обернулась, например, тем, что казахско-русское двуязычие действительно стало массовым и... не нашло противовеса в единой коммуникативной среде в виде практически отсутствующего русско-казахского двуязычия.

Русификация как языковая гомогенизация Казахстана осуществлялась с помощью разнообразных мероприятий и в значительной степени за счет расширения русской коммуникативной среды, например, массовых переселений славянского (преимущественно русского) населения8. Масштабы и последствия таких переселений в Казахстан были беспрецедентными: трагическая принудительная депортация репрессированных 'народов-предателей', 'подозрительных народов';

8 «Большое значение для процессов сближения народов имеет миграция населения, которая в Советском Союзе носит качественно иной характер, нежели в капиталистических странах. Там это обычно связано с классовым или национальным гнетом. В СССР миграция населения представляет собой плановое перераспределение рабочей силы по территории страны, производимое с целью более рационального размещения производительных сил, с целью широкого освоения природных богатств в слабозаселенных районах» [Исаев 1978]. Интересно по прошествии времени читать объяснения созданной потребности в изучении и владении русским языком: «Русский язык в таком качестве никогда не навязывался сверху в приказном порядке, а выступал в качестве языка межнационального общения в силу объективной потребности совместного сосуществования народов... Эта объективная необходимость определялась большим удельным весом русских среди других народов нашей страны, широкий расселенностью русских по территории страны» [Березин 1997].
Русификация и казахизация как языковая гомогенизация многоязычного Казахстана
Элеонора Сулейменова

Эвакуация в предвоенные и военные годы; исправительно-трудовые лагеря КарЛАГ, СтепЛАГ, Долинка, АЛЖИР (Акмолинский лагерь жен изменников Родины); приезд в страну более двух миллионов целинников; комсомольские ударные стройки и оргтабы специалистов и рабочих силы из разных регионов СССР в промышленность, согласно идеи превращения Казахстана в мощную аграрно-индустриальную республику – так в Казахстане появились малые города и поселки городского типа с преимущественно славянским населением, необходимость владения русским языком становилась жестко обусловленной жизненными потребностями, казахи в собственной стране превратились в этническое меньшинство вплоть до 1966 года, а русский язык стал языком большинства.

Итак, приведено только шесть популярных тезисов, лежавших в основе советской языковой политики и отражавших существовавшие противоречия между их содержанием и фактически осуществлявшейся русификацией. Эти противоречия существовали в разных республиках Советского Союза, и их драматизм в полной мере вскрылся с обретением ими независимости.

Перечисленные особенности советской политики русификации со всей очевидностью демонстрируют реальность и действенность целенаправленного административно-управленческого регулирования языковой ситуацией. Особо следует сказать о планомерности и последовательности осуществления всех мероприятий, сильном воздействии советской идеологии и пропагандистской машины, а также оптимальном использовании всех административных ресурсов советской власти. Более того, по прошествии нескольких десятилетий стала очевидной высокая эффективность, широкие манипулятивные возможности и инерционная сила русификации.

Русский язык стал играть сложную роль в формировании советской идентичности. С одной стороны, успешность русификации как беспримерного распространения русского языка – мощного ресурса мобилизации населения для развития и распространения советской идентичности означала построение гомогенного в языковом отношении общества. С другой, – языковая гомогенизация общества за счет других языков противоречила декларируемым положениям советской национальной и языковой политики.

Успехи советской политики русификации (несмотря на ее противоречивый характер) убеждают, что языковая и культурная гомогенизация как инструмент интеграции всех регионов громадной
страны, сопровождающаяся стремительной индустриализацией, централизацией управления и упрощением коммуникации за счет распространения русского языка, оказывается сегодня, пожалуй, одним из убедительных примеров эффективной языковой политики.

3. Коренизация как поиск баланса между русификацией и языковым плюрализмом

Истоки языковой политики казахизации восходят к коренизации 20-30-х годов, выразившей стремление советской власти строить государство в соответствии с провозглашенными правами, равенством и суверенностью народов (Декларация прав народов России, 1917). В Декларации прав трудящихся Киргизской / Казахской АССР (Оренбург, 1920) было объявлено, что КАССР как автономная единица, имеющая право на национальное самоопределение, входит в свободный Федеративный Союз Советских республик, окончательное национально-государственное оформление КАССР осуществилось в 1925 году.

Вопреки ожиданиям, одним из первых официальных документов Киргизской (Казахской) автономной республики стал Декрет Совета Народных Комиссаров об употреблении киргизского (казахского) и русского языков в государственных учреждениях (1921). В нем (с оговоркой «выбор языка в каждом отдельном случае диктуется соображениями целесообразности и практической пользы») жестко прописаны функциональные сферы обоих языков: в государственных учреждениях губерний и уездов на равных правах употребляются казахский и русский языки; декреты, инструкции, циркуляры и прочие законодательные и правительственные акты, имеющие общий принципиальный характер и исходящие от центральных и губернских органов, мандаты, служебные удостоверения, выдаваемые центральными, губернскими и уездными (районными) учреждениями, публикуются одновременно на оба языка; уездные учреждения ведут делопроизводство и 'сносятся с высшими установлениями' (центральными и губернскими) исключительно на русском языке и так далее.

Необходимость учитывать интересы народов фактически привела к жесткой привязке политики коренизации к административно-территориальному делению Казахской Республики, и коренизация проводилась в административных единицах всех уровней (вплоть до аулсоветов). В 1928-1933 годах неоднократно проводилось 'районирование' республики, район стал основной единицей административного
управления, и общение с вышестоящими государственными органами осуществлялось на русском языке.

Подобная установка безоговорочно выполнялась на всех уровнях, и ее исполнение постоянно контролировалось: так в степи и городах, а главное, в государственных структурах всех уровней крепло убеждение, что советская власть – это русская власть. «Необходимо, – говорил несколько позже И. Сталин в докладе на XII съезде РКП(б) в 1923 году, – чтобы власть пролетариата была столь же родной для инонационального крестьянства, как и для русского… Тогда тогда, и только постольку Советская власть, являвшаяся до последнего времени властью русской, станет властью не только русской, но и интернациональной, когда учреждения и органы власти в республиках этих стран заговорят и заработают на родном языке» [Сталин 1923].

Политика коренизации проводилась с помощью, во-первых, 'приближения' власти к населению с помощью перевода делопроизводства на казахский язык; во-вторых, увеличения представительства казахов во власти; в-третьих, перевода письменности на латинскую, а затем и кириллическую графику; в-четвертых, обучения казахскому языку партийных и хозяйственных (так называемых европейских) работников; в-пятых, многочисленных агитационно-пропагандистских мероприятий.

В этот период выходят многочисленные постановления, декреты, циркуляры (в том числе секретные), письма и положения, по которым можно судить об активности, но не всегда о результативности политики коренизации: о мероприятиях по введению делопроизводства на казахском языке в государственных учреждениях (Совнаркоме, КазЦИК, Наркомпросе, Наркомюсте, Наркомздраве, Наркомсобезе, Верховном суде, прокуратуре и иных организациях) и уездных органах (1923, 1924, 1925, 1927), о порядке 'сношения' с краевыми, областными и районными организациями, а также с низовой сетью социально-бытового обслуживания и торгово-заготовительных организаций исключительно на казахском языке (1933), о порядке увольнения рабочих и служащих государственных учреждений и предприятий в связи с введением делопроизводства на казахском языке и укомплектовании их лицами, знающими казахский язык и письменность: так процентный подход к комплектованию кадров был заменен функциональным (1924), специалистам (за исключением тех, для кого казахский язык является родным или родственным, то есть татар, узбеков и так далее) в зависимости от степени овладения казахским языком устанавливались повышенная (от
пяти до пятнадцати процентов) оплата труда, давались преимущества при поступлении на работу и другие льготы [Сулейменова 1997].

Всем советским учреждениям предписывалась обязательная организация курсов по изучению 'европейскими работниками' казахского языка, при этом строго регламентировались количество курсантов, оплата труда преподавателей, определялись уровни владения языком, рекомендовались конкретные учебные пособия и так далее (1930, 1934); за счет республиканского и местного бюджета открывались курсы делопроизводства, счетоводства и машинописи (1927); предусматривалась выдача вознаграждения за составление на латинской графике учебников на казахском языке и переводы произведений казахских писателей на русский язык (1921, 1923, 1934) и так далее.

Параллельно проводилась колоссальная по объему и напряжению работа по переводу казахской письменности на латинскую графику, по обязательному обозначению на казахском языке наименований учреждений, предприятий и организаций, товаров, знаков и прочее; строго контролировалась работа Государственной терминологической комиссии и всех 'терминологических ячеек'; организовывались и проводились научные языковые экспедиции; разрабатывались теоретическая и практическая грамматика, лексикология, фонетика казахского языка; систематизировался и издавался материал для орфографических, отраслевых терминологических, переводных и иных словарей; повсеместно работали курсы по ликвидации неграмотности (ликбеза), или обучению латинскому алфавиту и письму; создавалась учебно-методическая база для обучения казахскому языку – обязательному предмету в неказахских школах; в вузах, а также на специальных курсах осуществлялась подготовка ученых-филологов и учителей казахского языка.

Языковое строительство периода коренизации, связанное с идеями построения советского государства и одновременно активным ростом национального самосознания, было масштабным, осуществлялось по стратегически важным направлениям с горячим и искренним энтузиазмом.

Однако все это делалось в условиях жесточайших репрессий, уничтоживших талантливых лингвистов-теоретиков и методистов А. Байтурсынова, К. Жубанова, Н. Турекулова, Т. Шонанова, Е. Омарова, Х. Досмухамедова, К. Кеменгерова, писателей и составителей букварей и
пособий по казахскому языку С. Сейфуллина, Б. Майлина, М. Дулатова, М. Жумабаева, Ж. Аймаутова, Т. Жургенова и многих-многих других⁹.

Постепенно политика коренизации обернулась 'буржуазным национализмом', деятельность Центрального комитета 'нового тюркского алфавита' при КазЦИК, Всесоюзного тюркологического съезда, а также работа по тюркской (казахской) научной терминологии, орфографии, методике преподавания тюркских языков оказались 'проявлениями пантюркизма и панисламизма' или 'автономистских и сепаратистских настроений'.

«Национальная интеллигенция ...в результате политических репрессий все больше редела. Новая образованная прослойка ...в особенности ее партийно-административная часть, была двуязычной, при этом использование родного языка ограничивалось бытовыми ситуациями. В следующем поколении русификация большей части образованного слоя оказывалась неизбежной» [Беликов 2001: 391].

Коренизация, несмотря на всю сложность и трагичность, была чрезвычайно плодотворной для формирования и легитимации национальной идентичности и национального самосознания, и главное – для развития теоретической и учебно-методической базы казахского языка и подготовки языковедов.

Распространение казахского языка во власть – партийно-государственный аппарат – оказалось тем не менее сильно затруднено, хотя политика коренизации как реализация потребности национальной идентичности и расширения представительства казахов в государственных структурах продолжилась и особенно активно в годы правления Д. Кунаева (1960-1986-е годы).

Постепенно коренизация практически утратила одну из своих определяющих особенностей, а именно, распространение казахского языка в основные общественные сферы коммуникации. Этому препятствовала мощная советская государственная машина с монополией национально-языковой политики, имплицитно основанной на идеалах моноязычия (русскоязычия), единой строго централизованной административной системой, советской экономикой и армией, общим образовательным пространством, социальная мобильность и успешность в которых достигалась только при владении русским языком.

⁹ К 1937 году оказалось, что «казахское языкознание надо было создавать заново: никого из профессиональных лингвистов не осталось» [Ашин 2002: 211].
4. Казахизация как языковая гомогенизация многоязычного Казахстана

Самыми действенными и популярными инструментами любой национально-языковой политики, которая осуществляется через политические, образовательные, экономические, общественные и собственно лингвистические институты, являются статусное планирование (изменение функционального положения языка), корпусное планирование (регулирование структуры языка), планирование усвоения (создание условий для овладения языком) и распространения языка (создание условий для расширения сфер функционирования языка в пределах / вне пределов государства). Каким же образом осуществляется имплицитная политика казахизации?

Статусное планирование в Казахстане – первый, самый очевидный и сильный инструмент регулирования функционирования казахского языка в важнейших коммуникативных сферах. Более того, оно позволяет осуществлять государственный контроль за соблюдением законодательства и соответствием языкового состояния объявленному статусу. Здесь необходимо иметь в виду два важных момента.

Первый связан с тем, что сегодня все больше распространяются коммуникации новых типов, порождаемые глобализацией и новыми информационными технологиями. Глобализация также стремится к максимальной гомогенизации коммуникации и стимулирует новые способы использования мировых языков, уже не зависящие от государственных / языковых границ и не подчиняющиеся статусному планированию или контролю в какой-либо отдельной стране. Это интернет, новейшие информационные технологии, спутниковое телевидение, разные виды мобильной связи, международный бизнес, мировая наука, международное образование и другие.

Второй момент связан с тем, что реализация решений по языковому планированию в Казахстане, как и государственный контроль за ней, существенно отстает от принимаемых законодательных актов.

Тем не менее при статусном планировании, планировании усвоения и распространения государственного языка с переменным успехом используются разнообразные средства и мероприятия казахизации.

1. Одни из мероприятий вполне традиционны для стран с политикой моноязычия.
Во-первых, это многочисленные постановления, указы и распоряжения о расширении функционирования казахского языка в так называемых регулируемых сферах общения: государственном управлении, делопроизводстве, в том числе единым системой электронного документооборота (это коснулось всех предприятий и учреждений независимо от форм собственности, сами же сроки окончательного внедрения делопроизводства на казахском языке несколько раз официально передвигались); вещании на казахском языке (требование паритетности эфирного времени на радио и телевидении для казахского и русского языков в настоящий момент, по данным Министерства культуры, выполнено, однако, к сожалению, казахстанский эфир по качеству передач на казахском языке (как, впрочем, и на русском языке) не составляет серьезной конкуренции российскому и не способен пока существенно расширить собственную аудиторию); создании во всех крупных учреждениях специальных отделов государственного языка со штатом переводчиков и другие.

Во-вторых, системная модернизация и расширение инфраструктуры образования на казахском языке (это коснулось всех уровней обучения: предшколы, среднего образования, бакалавриата, магистратуры, докторантуры PhD); регулирование часов, отводимых на изучение казахского языка во всех типах и уровнях учебных заведений; улучшение качества учебно-методической базы обучения казахскому языку (унификация типовых программ, внедрение уровневого обучения и системы обязательной сертификации владения казахским языком, лексикографическая поддержка, массовое издание учебной литературы для разных категорий обучающихся и так далее); развитие интернет-ресурсов и дистанционного обучения; внедрение в учебный процесс компьютерных обучающих программ; открытие центров и организации курсов; проведение фестивалей, олимпиад и конкурсов на знание государственного языка среди всех слоев населения и многое другое.

Несмотря на обилие 'правильных' и взвешенных предложений и их разумную реализацию, в стране до сих пор сохраняется недостаточный уровень владения казахским языком и его неравномерное функционирование в регулируемых сферах коммуникации.

2. Другие мероприятия не столь традиционны. Более того, не всякое государство идет на подобные финансово затратные кардинальные преобразования. Речь идет о территориально-административном переструктурировании Казахстана и привязке языковой политики к
конкретным особенностям регионов. Самые значительные реформы были осуществлены в 1997 году: столица Казахстана была перенесена из Алматы в Астану; упразднены Жезказганская, Кокшетауская, Семипалатинская, Талдыкорганская и Тургайская области, как следствие, утратили свой статус областные города (позже только Кокшетау и Толдыкорган стали областными центрами); было существенно изменено внутриобластное деление на районы.

Территориально-административные реформы, в особенности перенос столицы и изменение статуса областных городов, имели значительные последствия для осуществляющейся языковой политики моноязычия (казахоязычия): во-первых, для укрепления территориальной целостности унитарного государства и преодоления сепаратистских настроений; во-вторых, для этнической мобилизации казахов и создания этнодемографического баланса в регионах с преобладанием славянского населения; в-третьих, для перераспределения занятого, в основном, в сельском хозяйстве казахского населения юга страны в промышленное производство северных регионов; в-четвертых, для увеличения общей численности северных областей за счет внутренней миграции казахов. Все вместе изменило, например, этнодемографические характеристики Астаны и в немалой степени способствовало ее казахизации: а) за двадцать межпереписных лет население города выросло более, чем в два раза – если в 1989 году в городе проживало 281 252, то в 2009 – 613 600 человек; б) в 1970 году казахи составляли 12,7% населения города, русские – 57,4%; в 2009 году доля казахов выросла до 69,4%, русских сократилась до 19,9%.

3. Третьи мероприятия получают полярные оценки в обществе.

В соответствии со статусом казахский язык – это «язык государственного управления, законодательства, судопроизводства и делопроизводства, действующий во всех сферах общественных отношений на всей территории государства», и «долгом каждого гражданина Республики Казахстан является овладение государственным языком, являющимся важнейшим фактором консолидации народа Казахстана» (статья 4 Закона «О языках в Республике Казахстан» 1997).

Требования свободного владения казақским языком предъявляются в равной мере ко всем гражданам Казахстана: президенту (статья 41 Конституции), председателям обеих палат парламента – Сената и Мажилиса (статья 58 Конституции), обыкновенным служащим, начальникам и руководителям высшего уровня, список которых составляет 220 наименований (Закон «О перечне профессий, специальностей и
Русификация и казахизация как языковая гомогенизация многоязычного Казахстана
Элеонора Сулейменова

должностей, для которых необходимо знание государственного языка в определенном объеме в соответствии с квалификационными требованиями») и так далее. Владение казахским языком сертифицируется в установленном порядке (система КАЗТЕСТ).

Однако, с одной стороны, несмотря на требования, даже самые высокие посты в Казахстане сегодня могут занимать лица, не владеющие казахским языком в необходимой для исполнения служебных обязанностей степени. Среди них не только представители неказахских этносов, но и европейизированные, обладающие высокими показателями полиязычия и владения русским и английским языками русскоязычные казахи, присутствие которых во власти велико. С другой стороны, требования к владению казахским языком стало восприниматься частью населения как объяснение непропорционального представительства казахов и вытеснения русских из руководящих органов.

4. Четвертые исключительны по масштабу и результативности воздействия на языковую гомогенизацию страны.

Имеется в виду успешно действующая национальная программа «Нұрлы кош», согласно которой в страну прибывают репатрианты-казахи (так называемые оралманы) из Китая, Монголии, Турции, Ирана, России, Узбекистана, Туркменистана, Кыргызстана и других стран.

Языковая адаптация и социализация репатриантов-казахов часто проходит болезненно, особенно в городах и русскоязычных регионах, поэтому государством предпринимаются целенаправленные усилия по поддержке репатриантов-казахов (компенсация расходов на переезд и приобретение жилья, преференции и специальные гранты в учебные заведения, подготовительные курсы в университетах, помощь в трудоустройстве, освобождение от службы в армии, гарантированный объем медицинских услуг и другие).

Заметное участие репатриантов-казахов в росте абсолютной численности и удельного веса казахов в населении страны оценивается сегодня в 24,9%. Число репатриантов-казахов приближается к миллиону и ощутимо (особенно в городах) влияет на казахизацию различных сфер коммуникации.

Увеличение численности казахов за счет репатриантов и их влияние на преимущественный выбор казахского языка для общения справедливо расцениваются как проявление казахизации [О'Бикен 2005], которая,
впрочем, проявляется не столь отчетливо, как в других государствах СНГ. 
Массовая репатриация казахов вместе с территориально-административным переустройством северных областей и переносом столицы 'плавно вписалась' во внутренние процессы, происходящие в этнической и языковой структуре страны и отмеченные в результатах переписей населения (в 1999 и 2009 годах): 
во-первых, а) Казахстан стал страной с преобладающим казахским и сокращающимся русским населением — за десять межпереписных лет численность казахов выросла на 26,0%, составив 63,1% населения страны; численность русских уменьшилась на 15,3% и составила 23,7% населения страны; б) произошло увеличение одних (узбекской, уйгурской, таджикской, кыргызской, азербайджанской, дунганской, курдской и др.) и уменьшение других (русской, украинской, белорусской, немецкой, татарской, польской и др.) диаспор, это привело к перераспределению удельного веса диаспор между собой: 1989 год — русская, украинская, немецкая, татарская, узбекская, белорусская, уйгурская, корейская диаспоры; 1999 год — русская, украинская, узбекская, немецкая, татарская, уйгурская, белорусская, корейская диаспоры; 2009 год — русская, узбекская, украинская, уйгурская, татарская, немецкая, корейская, белорусская диаспоры; в) на 24,0% сократилась численность алтайской, армянской, башкирской, болгарской, бурятской, греческой и многих других диаспор (представляющих референтные группы более 100 языков по спискам 1999 года), в 2009 году их доля в населении страны составила 1,1%; 
во-вторых, произошла решительная перергрупировка языков в сторону казахского языка: а) по росту казахской языковой компетенции — свободное владение казахами устной речью достигло 98,4%, письменной речью — 93,2%; русскими — 25,3% и 6,3% соответственно; узбеками — 95,5% и 61,7%; украинцами — 21,5% и 5,2%; уйгурами — 93,7% и 60,8%; татарами 72,6% и 33,7%; немцами — 24,7% и 7,9%; корейцами — 43,5% и 10,5%; турками — 91,0% и 43,4%, азербайджанцами — 81,3% и 43,2%; белорусами — 19,0% и 4,8%; дунганами — 49,7% и 16,0%; курдами — 77,5% и 37,1%; таджиками — 89,4% и 12,9%; поляками — 20,9% и 6,6%; чеченцами — 56,4% и 21,4%; кыргызами — 92,7% и 62,6%; другими этносами — 42,2% и 15,9%; б) по удельному весу в образовании (например, впервые в высшем образовании обучение на

Доля азербайджанцев в Азербайджане составила 90,6% (1999) против 82,7% (1989); доля белорусов в Белоруссии повысилась до 81,2% (1999) против 77,9% (1989); доля кыргызов в Кыргызстане выросла до 64,9% (1999) против 52,4% (1989) [Топилин 2002].
казахском языке превысило 50%: в этом учебном году в 149 вузах страны обучается 620 422 студента, из них на казахском языке – 319 940 студентов, остальные обучаются на русском, английском, арабском и иных языках); в) по отдельным коммуникативным сферам: они по-прежнему преимущественно остаются в традиционных сферах общения (художественная литература, публицистика, народные промыслы и так далее), но к ним уверенно добавились делопроизводство и сфера обслуживания.

Формирование новой реальности сказывается на политическом доминировании казахов и некотором продвижении казахизации, которая находит устойчивое сопротивление в виде сохраняющихся «многих черт недавней исторической эпохи»: «безусловное господство и высокая престижность русского языка, без которого во многих сферах просто невозможно обойтись; широкое общение между людьми разных национальностей, как правило, требующее знания русского языка; ...распространение рыночных отношений... в большинстве случаев стихийно усиливают роль русского языка» [Алпатов 2005].

Действительно, русификация и ее результаты сегодня сохраняют свою инерционную силу в Казахстане – стране с отчетливым присутствием русского языка и его многочисленных носителей во всех коммуникативно важных сферах. Роль русского языка в Казахстане, как и в некоторых других странах СНГ, оценивается даже «на уровне сил ядерного сдерживания» [Азаттик 2010].

Именно в таких условиях в Казахстане проводится новая языковая политика, с объявленными новыми стратегическими приоритетами, новой целью, новыми задачами и законодательной базой. Оговорены конкретные законодательные положения, направленные на государственное регулирование деликатнейшей сферы национально-языковых отношений и осуществление языкового планирования по трем стратегическим направлениям – расширение и укрепление социально-коммуникативных функций государственного языка, сохранение общекультурных функций русского языка, развитие других языков народов Казахстана. Всем содержанием законодательные акты Казахстана направлены на то, чтобы сгладить центральную дилемму между поддержкой языкового
многообразия и провозглашением казахского языка государственным с обязательными для этого акта последствиями казахизации.

С одной стороны, данная дилемма в теории и практике казахстанской языковой политики решается в форме «социокультурного варианта» политики мультикультуризма и языкового плюрализма: поддерживается все языки, защищаются языковые права их носителей, допускается «создание параллельных институтов (школ, средств массовой информации, гражданских ассоциаций), которым предоставляется равный статус в общественной сфере», поощряется билингвизм или мультилингвизм в образовании и средствах массовой информации [Кёниг 2003: 16]. Убедительным примером того, как осуществляется официальная поддержка языкового многообразия в стране, является созданная в 1995 году Ассамблея народа Казахстана – консультативно-совещательный орган при президенте, объединяющий представителей государственных органов, национально-культурных и других общественных объединений. По информации Секретариата Асамблеи, сегодня действует 820 этнокультурных объединений, работают центры и воскресные школы национального возрождения, комплексы и кружки изучения 30 родных языков и культур. Ежегодно государство увеличивает финансовую поддержку деятельности национально-культурных объединений и малых асамблей. В стране действует также системная государственная поддержка языков этносов в образовании, культуре и информации: в 2010-2011 учебном году из 7516 общеобразовательных школ работали 1524 школы с русским, 58 с узбекским, 14 с уйгурским, 2 с таджикским языками обучения, 2097 школ – смешанные; из 2261 дошкольных учреждений 272 с русским, 3 с другими языками обучения, 808 детское учебное заведение – с двумя языками обучения; из 50 казахстанских театров 9 смешанных, 15 русских, 1 корейский, 1 немецкий, 1 узбекский и 1 уйгурский театр; средства массовой информации работают на 13 языках, из 2,5 тысяч действующих в стране средств массовой информации 469 издаются и вещают на казахском, 856 – на русском, 879 – на казахском и русском языках, на языках этнических групп в Казахстане издается 28 газет и журналов. Приведенные факты

11 Из-за заданного объема статьи не приводятся данные о том, какие языки представлены в Казахстане, сколько референтных носителей они насчитывают и какова динамика численности референтных носителей языков диаспор, как и какие функции им свойственны, наконец, каким образом регулируется их функционирование со стороны различных субъектов языкового планирования (см. подробно в [Сулейменова 2009: 21-36]).
убеждают в том, что Казахстан может служить примером страны, которая стремится реализовать мультикультуралистскую идеологию и поддерживать языковой плюрализм.

С другой стороны, в Казахстане как обычная практика при строительстве любого государства осуществляется политика моноязычия (казахизации). Единый язык – это средство национальной и гражданской консолидации, символ, который отличает одно государство от другого. Поэтому государству так важно было выделить из множества языков казахский язык, наделив его особый статусом и полномочиями. Языковая политика сегодня направлена не просто на закрепление и развитие символической роли казашского языка при определении государственной идентичности, но и на кардинальное изменение языкового поведения всех этнических и возрастных групп, а также социальных слоев страны.

5. Заключение

Ни в одном из законодательных документов Республики Казахстан не говорится о казахизации как направлении языковой политики, точно так же, как не говорилось о русификации в Советском Казахстане.

Точно так же, как и политикой русификации, казахизацией управляют идеалы и идеология моноязычия (но уже казахоязычия), а конечной целью казахизации является языковая гомогенизация многоязычного общества, которая, конечно же, имеет свои пределы и никогда не может стать абсолютной.

Точно так же, как когда-то русификация, казахизация провоцирует столкновение интересов казахской и русской национальных элит.

Точно так же, как русификация в советский период, так и казахизация в современном Казахстане – это мощный ресурс мобилизации населения и инструмент политической консолидации.

Точно так же, как русский язык – символ и носитель этнической (русской) идентичности использовался для построения советской (коллективной, гражданской, государственной) идентичности, так и казахский язык – символ и носитель этнической (казахской) идентичности используется для удовлетворения потребности казахстанской (гражданской, государственной) идентичности.

Точно так же, как русификация советского периода стремилась найти компромисс в виде политики коренизации, так и казахизация в качестве собственного противовеса использует ресурсы мультикультурализма и языкового плюрализма.
Итак, во-первых, сходство русификации и казахизации как процессов языковой гомогенизации на этом вряд ли заканчивается; во-вторых, языковая политика Казахстана характеризуется сегодня как сбалансированная, и ее поиски баланса реализуются в напряженных взаимоотношениях между казахизацией и языковым плюрализмом (поддержкой всех языков, включая русский).

Литература
Азаттык (2010) //http://rus.azattyq.org/content/ukraine_moscow
_ language/1962920.html
Алпатов, В.М. (2005) Языковая ситуация в регионах современной России
 http://www.strana-oz.ru/?numid=23&article=1043
Ашнин, Ф.Д., Алпатов, В.М., Насилов, Д.М. (2002) Репрессированная
 тюркология. – Москва
Беликов, В.И., Крысин, Л.П. (2001) Социолингвистика. – Москва
Березин, Ф.М. (1997) Место и роль русского языка в постсоветской России
О’Бикен, Д. (2005) «Казахизация» и языковая политика в постсоветском
 Казахстане http://www.espi.ru/Content/Conferences/Papers2004/donaka.pdf
Иванов, В.В., Михайлова, Н.Г. (1982) Русский язык как средство
 межнационального общения: актуальные аспекты и проблемы //
 Вопросы языкознания – Москва, № 6
Исаев, М.И. (1978) О языках народов СССР
 http://genling.ru/books/item/f00/s00/z0000003/st009.shtml).
Кёниг, М. (2003) Культурное разнообразие и языковая политика // Решение
 национально-языковых вопросов в современном мире – Санкт-Петербург
От редактора (1962) Вопросы языкознания – Москва, №1
Передовая статья (1935) Правда – Москва, 1 февраля
Программа (1961) Программа Коммунистической партии Советского Союза
 – Москва
Серебряков, И. (2004) Глобализация и языковая политика
Сталин, И.В. (1923) Вопрос об объединении независимых национальных
 республик http://www.i-u.ru/biblio/archive/stalin_5/01.aspx
 Сборник документов. Научный редактор – Алматы
Сулеймanova, Э.Д. (2009) Языковая политика и динамика языковой
 ситуации в Казахстане // Russian Language Journal – Volume 59
Русификация и казахизация как языковая гомогенизация многоязычного Казахстана
Элеонора Сулейменова


Хобсбаум, Э. (2005) Все ли языки равны? Язык, культура и национальная идентичность // Логос. – Москва, №4(49)
Language Matters: Some New Contributions from Sociology  
(Emanating from Richard Brecht’s Castle)  

John P. Robinson, William P. Rivers, Cynthia Costell, Jennifer Robinson  

ABSTRACT  
As a result of the projects undertaken by the University of Maryland’s Center for the Advanced Study of Language, several new findings and insights have emerged from the highest quality data collections on Americans’ abilities in and attitudes toward foreign languages (FL). These involve the (now annual) language surveys conducted by the US Census Bureau and the bi-annual General Social Survey (GSS) conducted by the University of Chicago since 1972. The Census Bureau has been documenting foreign languages spoken at home and how well individuals can speak English in such households. The 2000-08 GSS has assessed whether adults can speak a foreign language and how well they can speak it, along with their attitudes on various language policy issues, with over 70% response rates.  

INTRODUCTION  
One of the richest sociological measures in the GSS is each respondent’s score on a 10-item SAT-like verbal ability measure, originally developed in 1942. That objective language-ability measure has been highly predictive of several public behaviors and attitudes in the GSS, even after adjustment for the main survey predictor of such variables, years of formal education, is taken into account. Overall, however, scores on this verbal measure have surprisingly stayed rather steady at about six items correct since 1974, which means that they have decreased after regression adjustment for the increased college education in the public in recent years. Indeed the correlation between respondents’ education and their vocabulary scores has decreased steadily from .54 to .44 since the 1970s, indicating perhaps that college education reflects less of a difference in basic knowledge today. Thus, the effects of increased education in the population have not resulted in any significant increase on this key indicator of verbal ability. Education’s main positive influence across time thus continues to lie in the
increased political tolerance of the population, as predicted by Stouffer’s “generation replacement” hypothesis (1955).

The contributions from sociology to the study of foreign languages in the US has tended to focus on generational differences in language acquisition (Fishman, 1966) and its ability to predict how well immigrants fare in terms of economic well-being (Chiswick & Miller, 1995). However, there is now a much richer and broader set of survey data sources on language-related behaviors and attitudes, and the examination of such data was made possible by support from the Center for the Advanced Study of Language (CASL) at the University of Maryland under the direction of Richard Brecht. These involve the now annual surveys conducted by the US Census Bureau and the bi-annual General Social Survey (GSS) conducted by the University of Chicago since 1972. The Census Bureau has been documenting foreign languages spoken at home and the self-rated ability of people to speak English in these homes. On the other hand, the GSS has asked how well individuals can speak foreign languages, along with their policy attitudes toward various language policy issues.

Among the important conclusions that have emerged from secondary analyses these time-series data are:

1. Americans’ ability to speak a foreign language in 2006 seems to have remained rather steady since the first survey conducted on the topic around 25 years previously (Robinson, Rivers & Brecht 2006a), when about a quarter of the public said they could.

2. However, only about 40% of FL speakers said they could speak the FL well, or about 10% of the public overall.

3. Of those who could speak well, almost 90% learned the FL at home – and not in school, by travel or other sources, suggesting that formal education was a relatively ineffective means of improving American’s FL ability.

4. Those who could speak either well or even poorly, however, expressed more positive, progressive or “cosmopolitan” attitudes on a variety of issues – especially that children should be required to take FL courses in high school (Robinson, Rivers & Brecht 2006b). That remained true after regression adjustment for the other main predictor of such attitudes, years of education, was taken into account. (The overall proportion of college graduates among those over age 18 has risen from less than 10% in the late 1960s to more than 25% now).

5. In terms of FL policy issues overall, most of the GSS sample also agreed with requiring FL courses in high school, and that it was as important as
math and science as a subject of study. Most of this national sample also disagreed with “English only” policies. These conclusions were not only found in the year 2000 GSS, but in follow-up studies conducted in 2006 and 2008. In the latter study, it was also found that significant age differences in policy attitudes were apparent, suggesting a significant new “generation gap” had emerged with younger adults expressing more “open” FL policy attitudes (Robinson, Rivers & Harwood, 2011). There was further evidence that these policy attitudes were more likely predictors of voting for President Obama, it was not possible to conclude they played a significant role in his election victory.

The present contribution, also made possible by Brecht’s encouragement, support, and more than occasional participation, takes advantage of another rich language measure available in the GSS -- one that more objectively measures a GSS respondent’s language ability in terms of a specific language task.

DATA SOURCE:
The main GSS indicator examined here is neither an activity nor an attitude, but more reflective of an important national resource, namely the English verbal ability of its citizens. Greater verbal ability presumably accumulated through years of a college education allows one not only to develop language skills to achieve personal goals and be a more productive contributor to today’s “knowledge economy”, but the opportunity to further build on language skills to appreciate mass media content, literature and the arts.. Indeed, it is a direct measure of the important indicator of literacy, usually only crudely measured in terms of years of schooling – no matter what quality of education that one received in school.

The verbal ability measure examined here is a brief 10-item vocabulary quiz that was originally developed by Thorndike and Gallup (1942). It has been shown to have almost as much predictive power in surveys as the standard survey question on years in school – acting as a further indicator of a person’s intellectual skill.

These 10 multiple-choice matching questions to measure verbal ability involve presenting GSS respondents with the word “beast”, and then asking them to match that word with the most appropriate of five words on a card they are presented, like:

a) Disruption
b) Landscape
c) Animal
d) Lumber
e) Police

Respondents are then scored on a 0-10 scale depending on the number of words they correctly match. This vocabulary score measure (called “WORDSUM”) has been used in a variety of analyses over the years, most notably as an empirical rebuttal to the racial conclusions of Murray & Hernstien (1995), as in Hout et al. (2001) and Hauser & Huang (1997).

These items have been regularly asked in the GSS’s national probability survey conducted by the University of Chicago annually or (biennially) since 1974, with samples of about 1500-3000 adults per survey. The overall average word score since 1974 has averaged about six of the ten items (as shown in the top line of Table 3 below).

In this article, multivariate analyses of how these vocabulary scores relate to responses on other GSS questions are first conducted using the multivariate regression program of Multiple Classification Analysis (MCA) of Andrews, Sonquist and Morgan (1973), which is now part of the regression analysis program of SPSS. What MCA provides then is a calculation of the predictive power for each score on the 10-item word score to various GSS media and attitude questions, both and after adjusting for each score for that group’s age or education level. In other words, it shows how well, say the score 5 group uses the media, after adjusting for their age and education level. Put another way, it shows how much that score group uses the media “other factors equal”, that is having the same age and education as the other 10 score groups.

RESULTS

A: Relations with Other Variables; Mass Media Use: Tables 1 and 2 demonstrate the predictive ability of the word score to predict or explain differences in two separate information-related variables asked regularly in the GSS, namely usage of the mass media (in Table 1) and general social attitudes (Table 2). Table 1 first shows how daily newspaper reading regularly increases the higher the number of words identified correctly, first in simple bivariate terms and then after MCA adjustment for differences in the demographic variables of education and age, variables in Table 3 that are notable predictors of the word score itself).

In the case of days of newspaper reading in the first column of Table 1, for example, it can be seen how newspaper reading days regularly climb from around 3 days for those scoring 0 and 1 on the word scale, up to 3.9 days for those scoring 5 on the word scale, and then up to 5.4 days reading the newspaper for those with the (maximum) perfect score of 10, which 2.5 days higher than the
group knowing 0 words. After MCA adjustment for education and age, these
differences are shown in the second column to reduced by about a third (from 2.5
to 1.8, as they still range from 3.1 days for those with a score of 0 up to 4.9 days
for those scoring 10. That is reflected in the reduction in the summary correlation
coefficient (eta) from .19 before MCA to .11 after MCA adjustment for age and
education.

In contrast, that adjusted beta is about the same as for the more familiar
demographic factor of education (.13), for which the difference between those at
the highest educational level score (5.4 days after adjustment) vs. 3.5 days for
those who did not finish high school -- a 1.9 differential which is very close to 1.8
high-low differential for vocabulary score. Moreover, the differential is almost
monotonic (with steady progression for each category) for the vocabulary score,
while there is only minor increases between high school graduates (4.2) and
college graduates (4.6) using education as the predictor. However, it is the third
predictor variable of age that is most predictive of newspaper reading, with a
beta value of .30 after MCA adjustment.

Table 1: Differences in Media Use by Verbal Abilities Score, Education and Age
(Before and after MCA adjustment for the other predictors)

<table>
<thead>
<tr>
<th></th>
<th>Days Read Newspaper</th>
<th>TV Hours/Day</th>
<th>Information Websites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bivariate</td>
<td>MCA</td>
<td>Bivariate</td>
</tr>
<tr>
<td><strong>Wordscore</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2.9</td>
<td>3.1</td>
<td>3.4</td>
</tr>
<tr>
<td>1</td>
<td>3.1</td>
<td>3.5</td>
<td>3.6</td>
</tr>
<tr>
<td>2</td>
<td>3.2</td>
<td>3.5</td>
<td>3.8</td>
</tr>
<tr>
<td>3</td>
<td>3.5</td>
<td>3.8</td>
<td>3.4</td>
</tr>
<tr>
<td>4</td>
<td>3.7</td>
<td>3.9</td>
<td>3.3</td>
</tr>
<tr>
<td>5</td>
<td>3.9</td>
<td>4.0</td>
<td>3.1</td>
</tr>
<tr>
<td>6</td>
<td>4.2</td>
<td>4.3</td>
<td>2.9</td>
</tr>
<tr>
<td>7</td>
<td>4.4</td>
<td>4.4</td>
<td>2.7</td>
</tr>
<tr>
<td>8</td>
<td>4.6</td>
<td>4.4</td>
<td>2.5</td>
</tr>
<tr>
<td>9</td>
<td>5.0</td>
<td>4.6</td>
<td>2.3</td>
</tr>
<tr>
<td>10</td>
<td>5.4</td>
<td>4.9</td>
<td>2.0</td>
</tr>
<tr>
<td>Diff (10-0)</td>
<td>+2.5</td>
<td>+1.8</td>
<td>-1.4</td>
</tr>
<tr>
<td><strong>Beta=</strong></td>
<td>.19</td>
<td>.11</td>
<td>.18</td>
</tr>
</tbody>
</table>

Education
High school inc 3.6 3.5 3.7 3.4 12.4 12.8
High school grad 4.1 4.2 3.0 3.0 13.9 14.1
Some college 4.3 4.4 2.5 2.6 13.7 13.8
Bachelors degree 4.6  4.6  2.1  2.3  17.0  16.9
Graduate degree  9.8  5.4  1.9  2.1  18.2  17.8
Diff (Grad-High)  +6.2  +1.9  +1.8  +1.3  +5.8  +5.0
Beta=  .14  .13  .24  .17  .35  .31

Age
1. 18-34  3.2  3.3  2.9  2.9  15.0  15.5
2. 35-49  4.2  4.1  2.6  2.7  15.5  15.3
3. 50-64  4.9  4.9  2.9  2.9  15.5  15.1
4. 65+  5.5  5.7  3.5  3.4  13.1  12.8
Diff (4-1)  +2.3  +2.4  +0.6  +0.5  -1.9  -2.7
Beta=  .29  .30  .14  .10  .11  .12

The two (middle) columns in Table 1 show differences for the media variable of TV viewing hours per day. Here, the pattern is reversed with higher viewing hours among those with lower vocabulary scores and lower years of education. However in columns 3 and 4, education emerges as the stronger predictor (beta=.17), compared to the beta of .10 for vocabulary score and for age. That reflects the high-low difference of 1.3 hours (3.4-2.1) for education vs. only 0.5 hours (2.5-3.0) for vocabulary score.

Table 2: Differences in Public Attitudes by Verbal Abilities Score, Education and Age
(Before and after MCA adjustment for the other predictors)

<table>
<thead>
<tr>
<th>Wordscore</th>
<th>Political Tolerance</th>
<th>Personal Trust</th>
<th>Bible Beliefs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bivariate MCA</td>
<td>Bivariate MCA</td>
<td>Bivariate MCA</td>
</tr>
<tr>
<td>0</td>
<td>5.8  6.3</td>
<td>.94  1.04</td>
<td>58%  50%</td>
</tr>
<tr>
<td>1</td>
<td>5.8  6.8</td>
<td>.91  1.16</td>
<td>60  52</td>
</tr>
<tr>
<td>2</td>
<td>5.2  6.1</td>
<td>1.00  1.10</td>
<td>58  51</td>
</tr>
<tr>
<td>3</td>
<td>5.6  6.3</td>
<td>1.10  1.13</td>
<td>55  50</td>
</tr>
<tr>
<td>4</td>
<td>6.2  6.7</td>
<td>1.17  1.22</td>
<td>47  44</td>
</tr>
<tr>
<td>5</td>
<td>7.1  7.2</td>
<td>1.32  1.39</td>
<td>40  38</td>
</tr>
<tr>
<td>6</td>
<td>8.1  8.0</td>
<td>1.45  1.56</td>
<td>33  33</td>
</tr>
<tr>
<td>7</td>
<td>8.7  8.6</td>
<td>1.68  1.65</td>
<td>26  28</td>
</tr>
<tr>
<td>8</td>
<td>9.7  9.2</td>
<td>1.86  1.87</td>
<td>20  23</td>
</tr>
<tr>
<td>9</td>
<td>10.3 10.0</td>
<td>2.14  1.98</td>
<td>13  16</td>
</tr>
<tr>
<td>10</td>
<td>11.5 10.8</td>
<td>2.32  2.10</td>
<td>10  17</td>
</tr>
<tr>
<td>Diff 10-0=</td>
<td>+5.7  +4.5</td>
<td>+1.38 +1.06</td>
<td>-48 pts -33 pts</td>
</tr>
<tr>
<td>Beta</td>
<td>.35  .26</td>
<td>.25  .18</td>
<td>.24  .21</td>
</tr>
</tbody>
</table>

Education
High school inc  6.0  6.8  1.10  1.31  55%  46%
High school grad 8.4  8.1  1.52  1.55  35  34
Some college 10.8 9.0 1.63 1.62 27 24
Bachelors degree 11.2 9.4 1.98 1.83 16 22
Graduate degree 10.9 10.8 2.28 1.88 10 18
Diff' Grad-High +4.9 +4.0 +1.18 +.57 -45 pts -28 pts
Beta .35 .16 .19 .12 .27 .17
Age
1.18-34 9.4 8.5 1.39 1.33 31% 29%
2.35-49 8.4 8.3 1.66 1.60 31 33
3.50-64 7.0 7.1 1.79 1.73 35 36
4.65+ 5.1 5.5 1.79 1.87 41 38
Diff (4-1) = -4.3 -3.0 +.40 +.54 +10 pts +9 pts
Beta .31 .27 .15 .15 .08 .07

Nonetheless, the pattern for those scoring 2 and over on their vocabulary score do show steady declines in watching TV as a function of higher word scores after MCA adjustment.

Much the same pattern is found in columns 5 and 6 of Table 1 for the third media variable: monthly visits to information sites on the Internet (for news, politics, science and the like, rather than for games and other “entertainment” sites). The beta for education as a predictor here is .31, double that for vocabulary (.15) or age (.12) -- and there are notable departures from monotonicity (e.g. the 15.1 visits for those scoring 4 on the word scale).

B. Education-related Attitudes: The differences on attitudes in Table 2 tend to be larger than for media use. All three attitude areas are those that have been found to be significantly education related in the GSS and past studies (e.g. Finkel et al 1999; Uslaner 2002; Robinson 1999). Scores on tolerance are shown in columns 1-2, on trust columns 3-4 and on the Bible in columns 5-6.

The tolerance measure in columns 1 and 2 in Table 2 is based on a set of 15 items originally developed by Stouffer (1955) in his classic study of Conformity and Civil Liberties during the “McCarthy era”. These questions involve three behaviors (giving a public speech, teaching at a public college and having one’s books in a public library) allowable for five political minority groups (atheists, racists, militarists, socialists and homosexuals), with one point given for each “allow” response. As can be seen in the first column of Table 2, tolerance scores go up (after 0-1 scores) quite regularly with the vocabulary score, as well as for increased levels of education; but they decrease regularly with age. After MCA adjustment for the other predictors, much the same pattern can be observed, but here the word score emerges as a stronger predictor than education, with a beta value of .26 vs. .16 for education. That reflects the 4.5 differential for those
scoring 10 on the word score (10.8) vs. those scoring 0 (6.3). Word score is about as strong a predictor of tolerance as age (beta=.27).

The next (middle) two columns in Table 2 show differences on a three-item scale of trust in people, measured by such questions like, “Do you think most people can be trusted, or that you can’t be too careful in life”. Scores here thus vary between 0 and 3, and it can be seen in Table 2 that trust rises almost monotonically with each higher vocabulary score – both before and after MCA adjustment. Again the difference of 1.06 between highest (2.10) and lowest vocabulary score (1.04) are again larger than the 0.57 for those with highest 1.88) minus lowest education (1.31) -- and the 0.54 difference for highest minus lowest age. These are reflective of the beta of .18 for wordscore, .12 for education and .15 for age.

Finally, it can be seen that a similar pattern holds as well for literal beliefs in the Bible shown in columns 5 and 6 of Table 2, with 50% of lowest vocabulary scorers believing strictly in the Bible vs. only 17% of highest word scorers after MCA adjustment, reflecting that 33 point differential with a beta of .21. This is slightly higher than the 28 point differential for education and much larger the 9 point differential for age.

Table 3: Differences in Word Score by Demographics
(After MCA adjustment for gender, age, race, education and income)

<table>
<thead>
<tr>
<th>VERBAL ABILITY</th>
<th>BEFORE</th>
<th>AFTER</th>
<th>MCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL SAMPLE</td>
<td>6.0/10</td>
<td>6.0/10</td>
<td></td>
</tr>
<tr>
<td>BIRTH FACTORS:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENDER</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5.9</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>6.1</td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>+0.2</td>
<td>+0.2</td>
<td></td>
</tr>
<tr>
<td>ETA</td>
<td>.04</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>AGE:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-24</td>
<td>5.4</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>25-34</td>
<td>5.9</td>
<td>5.7</td>
<td></td>
</tr>
<tr>
<td>35-44</td>
<td>6.2</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td>45-54</td>
<td>6.2</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>55-64</td>
<td>6.2</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>65+</td>
<td>5.9</td>
<td>6.3</td>
<td></td>
</tr>
</tbody>
</table>
C: Word Score Differences by Demographic Factors: Here, word score becomes the dependent variable to be predicted, and not a predictor variable of behaviors and attitudes as in Tables 1 and 2. The variations in the vocabulary score by five basic demographic variables are shown in Table 3. They are again shown in two columns, the first in simple bivariate terms and then after MCA adjustment for the demographic predictors of gender, age, race, education and income.

The Table 3 differences are shown for four basic clusters of demographic factors namely birth factors (factors determined at birth, namely gender, age and race), the social status factors of education and income, the three role factors of worker, marriage partner and parenthood and the location factors of region and urbanicity. The following Table 4 then shows differences by year of study, before and after MCA adjustment for the Table 3 demographic differences. Table 5 then breaks these trend data out into more basic terms.
As shown in the both columns of Table 3, women score slightly higher (6.1) on their word score measure than men (5.9), both before and after MCA adjustment. Age differences become even more apparent after MCA, with the uniformly higher scores among the elderly being boosted by MCA because of their lower formal education. The lower scores of blacks and other races are also reduced for much the same reason of their lower education, but the overall differences do remain significant after MCA adjustment.

Not surprisingly, education remains the strongest predictor of word scores in Table 3, with its beta value reaching 0.50 after adjustment. Education differences then largely explain the decline of income’s predictive power from 0.28 to 0.13 after MCA adjustment; that indicates that verbal ability may assist workers to achieve higher income. (Although not shown in Table 3, word scores vary little by other demographic factors, although the slightly lower scores among those with 4 or more children and among those working routine 40-hour workweeks are noted – as are the slightly lower scores in the South and in both most urban and rural (but not suburban) areas).

Overall, then, education has been the most important predictor of word scores, so that with the dramatic population increase in college education since the 1970s, one should expect a corresponding increase in word scores.

Trends across Time: Table 4 then brings us to that important trend question, by focusing on the actual changes in the word measure over the last 34 years. Here, it can be that word scores show only an insignificant overall increase from 6.0 to 6.1 across time, despite the presence of these more college educated in the population. Indeed, after MCA adjustment, it has effectively steadily declined given their greater presence from 6.3 to 5.7.

Table 4: Trends in Word Score across Time  
(Before and after MCA adjustment for the demographic factors in Table 3)

<table>
<thead>
<tr>
<th>GSS STUDY YEARS</th>
<th>WORD SCORE Before</th>
<th>After MCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972-1979</td>
<td>6.0</td>
<td>6.3</td>
</tr>
<tr>
<td>1980-1989</td>
<td>5.9</td>
<td>6.0</td>
</tr>
<tr>
<td>1990-1999</td>
<td>6.1</td>
<td>5.9</td>
</tr>
<tr>
<td>2000-2008</td>
<td>6.1</td>
<td>5.6</td>
</tr>
<tr>
<td>Difference</td>
<td>+0.1</td>
<td>+0.7</td>
</tr>
</tbody>
</table>
As shown in closer detail in Table 5, college education is becoming far less important in predicting word scores than it did 35 years ago. As can be clearly seen in the last column of Table 5, word scores overall have decreased for all but the least educated. That steady overall average (of about 6.0 items since the 1970s) thus obscures the declines that have occurred within all but the lowest education level. While those with least education, not having finished high
school, have shown no change or improvement, from 4.7 in the 1970s to 4.7 in the 2000s, high school graduates have declined from 6.0 in the 1970s (to 5.7 in the 1980s) to 5.6 today, those with some college a decline of 0.5 points (from 6.7 to 6.2), college graduates the largest decline of 0.9 points (from 7.9 to 7.0) and those with a graduate degree a decline of 0.4 points (from 8.2 to 7.8).

Put the other way, the gap between the gap between those with least and most education has declined from 3.5 in the 1970s (8.2-4.7) to 3.1 points (7.8-4.7) today, but the biggest difference Of almost a point) is for those with just a college degree.. That is further reflected in the correlation of education and word score, dropping from .54 in the 1970s to .44 today. Education remains the most powerful predictor, but it makes less of a difference today. One can note also how regular and consistent the trend has been across each decade since the 1970s in Table 4. In that MCA analysis, the year effect goes from an insignificant correlation of .03 to a notable beta of .08, showing an adjusted score of 6.3 in the 1970s, 6.0 in the 1980s, 5.9 in the 1990s and 5.8 in the years after 2000. (The continued declines in word score can be seen visually in its accompanying figure in Table 4).

**Spanish (and other FL) vs. English Language Comparisons in the US**

With this GSS ability measure, it also becomes possible to assess objectively the English language skills of those who say they can speak a FL. This is a crucial issue when recruiting candidates for translation occupations, since it can be as important to translate the FL into English as it is to speak the FL by itself. In brief, Robinson and Rivers (in press) show that FL speakers are about as adept in English as non-FL speakers, with Spanish speakers being slightly less adept and other European FL speakers slightly above average. In other words, speaking a FL provides little clue about ability in English.

An intriguing development on this question was provided in GSS 2006 (and 2008) when this 10-item GSS verbal ability measure was extended to respondents who were interviewed (and thus fluent) in Spanish. The words used were exactly the same used in the English interview, thus allowing the possibility that subtle differences in word difficulty across cultures could confound the results. It is also the case that the total number of Spanish speakers in the two surveys combined was less than 150. However, Table 6 indicates that these two liabilities may not be as serious as one might imagine. It can be seen first that the relations with age and education are quite similar in the Spanish language group. Moreover, when they are inserted into an MCA to adjust for the lower age and education of the Spanish-speaking
subsample, they show no difference in the MCA-adjusted average score. In other words, while the entire Spanish sample scored only 5.2 on the scale vs. 6.1 for the English sample, when adjusted for their lower education levels and age, the Spanish speakers scored exactly the same as the English speakers.

Table 6: English-Spanish Differences in Word Score
(Before And After MCA Adjustment For Education And Age)

<table>
<thead>
<tr>
<th>VERBAL ABILITY</th>
<th>BEFORE</th>
<th>AFTER MCA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TOTAL SAMPLE</strong></td>
<td>6.0/10</td>
<td>6.0/10</td>
</tr>
<tr>
<td><strong>BIRTH FACTORS:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENDER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>6.1 (2407)</td>
<td>6.1</td>
</tr>
<tr>
<td>Spanish</td>
<td>5.2 (143)</td>
<td>6.1</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.9</td>
<td>0.0</td>
</tr>
<tr>
<td>Beta</td>
<td>-.11</td>
<td>-.00</td>
</tr>
<tr>
<td>AGE:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-34</td>
<td>5.7</td>
<td>5.8</td>
</tr>
<tr>
<td>35-49</td>
<td>5.9</td>
<td>5.8</td>
</tr>
<tr>
<td>50-64</td>
<td>6.6</td>
<td>6.5</td>
</tr>
<tr>
<td>65+</td>
<td>6.2</td>
<td>6.4</td>
</tr>
<tr>
<td>Difference</td>
<td>+0.5</td>
<td>+0.6</td>
</tr>
<tr>
<td>ETA</td>
<td>.16</td>
<td>.18</td>
</tr>
<tr>
<td>EDUCATION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School Inc.</td>
<td>4.9 (326)</td>
<td>4.7 (93)</td>
</tr>
<tr>
<td>High School Grad</td>
<td>5.5 (626)</td>
<td>5.9 (22)</td>
</tr>
<tr>
<td>Some College</td>
<td>6.3 (728)</td>
<td>6.4 (13)</td>
</tr>
<tr>
<td>College Grad</td>
<td>6.9 (404)</td>
<td>4.9 (7)</td>
</tr>
<tr>
<td>Grad School</td>
<td>7.3 (318)</td>
<td>6.9 (7)</td>
</tr>
<tr>
<td>Difference</td>
<td>+3.4</td>
<td>+3.2</td>
</tr>
<tr>
<td>ETA</td>
<td>.47</td>
<td>.50</td>
</tr>
</tbody>
</table>
Whether this reflects better linguistic instruction for those learning Spanish or poorer instruction in English cannot be determined from this small sample. However, it will be of interest to monitor this difference in 2010 and future GSS studies, as its sample size and diversity continues to expand.

**Summary and Conclusions:**
In addition to providing important new glimpses into the US public’s skills in and attitudes about foreign languages and related policy, the GSS contains important insights into changes in basic language ability, as reflected its brief vocabulary skill measure. Those more fluent by this measure report being more serious information consumers of news media content and more cosmopolitan and progressive in their social attitudes.

At the same time, despite the dramatic increase in college education in recent decades, this article shows little if any progress in scores of the public’s verbal abilities, as reflected in this 10-item vocabulary measure regularly collected since 1974. This is consistent with evidence from several other surveys of cultural awareness, activity participation, public information and other “sophisticated” or cosmopolitan attitudes that would be expected to parallel these education gains.

While the overall public GSS scores on *verbal ability* in the GSS have stayed rather steady at about 6.0 since 1974, they have decreased after regression adjustment for their increased college education. Whether that reflects college education making less of a difference in the public’s basic knowledge advantage today is a topic that deserves more research attention. The evidence for college education’s positive influence across time, for now, continues to lie more in the increased political tolerance of the population (as in Table 2), as predicted by Stouffer’s (1955) “generation replacement” hypothesis in relation to the side benefits of increased education outside of increased income.

Consistent with these trends, one finds expected gains in certain social attitude indicators, like increased political and racial tolerance. While still far short of ideal, direct discrimination against minorities has declined – as has expressed intolerance against political minorities, as Davis (1974) documented. Davis found increased public tolerance of homosexuals, communists, racists and other political minorities, as compared to levels of tolerance in the classic 1950 survey of Stouffer (1955). At that time, Stouffer predicted such increased tolerance at the time based on his “generation replacement” hypothesis, namely that as older and less educated generations died off, they would be replaced by younger generations with more education and hence greater tolerance.
Much the same trends, then, might be expected in other key aspects of life that have shown level of education to be a main predictor, like increasing civic engagement, arts participation, and increased reading (with reduced TV viewing). However, gains in these measures of societal “progress” unfortunately have not been found for many of these educated-related indicators. Perhaps the most well-known and substantial catalog of societal decline was advanced in Putnam’s (2000) “Bowling Alone”, in which virtually all of his indicators of “social productive” in the US showed signs of lower engagement in socially productive encounters over the decades. Among the reasons behind this decline, Putnam argued, were the increased participation of women in the workforce, again a doubled-edged trend (but more generally seen as a sign of social progress) and increased reliance on television (here usually seen as an indicator of declining quality of life, as documented in the research in Kubey and Czikszentmihalyi (1992).

But Putnam is not alone is seeing the glass as becoming more than half-empty. Keeter and della Carpini (1997) noted how a more cognitive indicator -- Americans’ level of political knowledge about politics -- had not increased over the years. Other observers of survey evidence point out how levels of expressed happiness, another attitude predicted by education, have not increased since the 1960s.

The Spanish-English comparison, while tentative and in need of substantial replication, does advance our argument that the educational input to the attitudes and abilities measured in the surveys presented here has, in fact, been on decline for some time, even as educational levels have risen. One might well posit that the resources in the U.S. educational system directed to the assimilation of immigrants through the acquisition of English are not well suited to rapid shifts in the makeup of the immigrant cohort. Reacting to a sudden increase of poorly-educated immigrants, or to an increase in immigration from FLs more distant and harder to learn with respect to English, may have profound consequences on the immigrants themselves, society, and the educational system.

Considerable policy emphasis and resources have been focused on bilingual education, starting with the Bilingual Education act of 1967. The latter was first to provide Federal funds for the development and provision of bilingual education through the legal framework requiring bilingual education for those children of Low English Proficiency (LEP), as described in Lau v Nichols, (1973) through the burgeoning scholarly and pedagogical fields of English as a Second Language. It may be that these efforts have ameliorated a
situation already less than ideal for immigrant children, but nonetheless, it appears that more recent young immigrants are not as well served by the educational system as were their predecessors.

References


Some Thoughts on the Societal Impact of Linguistics

Arto Mustajoki

"Not everything that counts can be counted, and not everything that can be counted counts."
—Albert Einstein (attributed)

Today, scholars are obliged to justify the necessity of research in their own fields. The purpose of this paper is to discuss the channels of societal impact of linguistic research. In other words, I will try to answer the following question: in what ways is the work of scholars dealing with language(s) useful for society? The paper is based on my articles published in Finnish (Mustajoki 2005, 2011) and on oral presentations in various academic circles. I will first discuss some general ideas concerning the societal impact of research and researchers. I will then try to apply these ideas to linguistics.

General Mechanism of Societal Impact of Research

Picture 1 shows the main ways in which the work of researchers makes an impact on society. The width of the arrows in the picture symbolizes the significance of the impact of the channel concerned.

As the picture shows, the most important way in which university researchers can influence society is through their students, who, after graduation, work in various influential capacities in society. A researcher may also have direct influence on society by publishing books for a wider public and by writing newspaper or Wikipedia articles. (S)he may also act as a specialist by giving interviews, by taking part in committee work, and by maintaining personal contact with decision-makers. An active and prominent researcher has a certain influence on the development of global research in the field concerned. This is the very idea of scientific activity: through the joint efforts of scholars, research reaches new results and triumphs which can be used by society.

Picture 2 shows the complexity of societal impact of research. Some issues are important to highlight. First, the results of studies focusing on any of the four main objects of research (Nature, Physical environment, Mental environment, Human) are relevant from the point of view of all three major users: the private and public sectors and individual citizens. Second, all three actors contribute to the welfare of nature, the economy, and the public sector.
Some Thoughts on the Societal Impact of Linguistics
Arto Mustajoki

Picture 1: Channels of societal impact of research(ers)

<table>
<thead>
<tr>
<th>Researchers</th>
<th>Students</th>
<th>Researcher community</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Society

Picture 2: Paths of exploitation of research results

PATHS OF EXPLOITATION OF RESEARCH RESULTS

1. Object of research 2. Users of research 3. Consequence, relevance 4. Final goal

- NATURE
- PRIVATE SECTOR
- PUBLIC SECTOR
- WELFARE OF NATURE
- WELFARE OF ECONOMY
- WELFARE OF PEOPLE
- HUMAN

Third, the final goal is not the welfare of the economy, of the public sector, or even of nature; it is the welfare of people. Influence does not go from left to right automatically, but needs active measures. In distributing the results of
research to their users, schools and other educational institutions are essential, as are the media.

I will now try to describe different categories of innovations and findings based on research. I will divide them into four types; in each case, some characteristics and examples are given.

1. Production-oriented findings and innovations
Definition: findings are made near to production; they are stimulated by observation of defects in the production process, or by the needs of end-users
Examples: a new model of bucket for an excavator; a new method of indexation; a new service concept for hotels; an updated teaching material distributed by Internet; a computer game
Societal impact: economic prosperity which enables users to “buy” welfare; direct influence on the quality of life
Other relevant factors and conditions: user-friendliness, attitudes of users, design, present trends, compatibility with other products, investment costs, market situation, adjustment to legislation
Time span to benefits: short (from months to some years)
Profits: from some thousands to millions (dollars or euros)
Possibility of measuring societal impact of research: partly possible (patents, earnings of the company, recruitment of employees)

2. Research-oriented findings and innovations
Definition: findings are made in research institutes or universities as a by-product of fundamental research
Examples: a new functional foodstuff; chemical which prolongs the use of an accumulator; a new method of disease diagnosis; a new method in speech recognition
Societal impact: as in 1
Other relevant factors and conditions: as in 1 + knowledge transfer from researchers to producers
Time span to benefits: usually rather long (several years)
Profits: may exceed tens and hundreds of millions
Possibility of measuring societal impact of research: as in 1, but due to the long way from research findings to production, it is difficult to show the significance of research in the final product

3. Impact of research on the work of the public sector
Definition: findings and research results are used in legislation and the everyday work of authorities at different levels of administration and the education system
Some Thoughts on the Societal Impact of Linguistics
Arto Mustajoki

Examples: prohibition of use of asbestos; restrictions on smoking; public campaign for physical exercise; a new method of teaching math; regulative measures influencing the market economy

Societal impact: better life for people

Other relevant factors and conditions: knowledge transfer from researchers to decision makers + public opinion, ”political will”, media, people’s conservatism, rules of economy (may hinder business)

Time span to benefits: usually decades

Profits: huge indirect impact by savings in societal costs (may exceed billions) or substantial increase in people’s knowledge and skills (in the case of education), which has enormous future impact on the way people act as citizens and employees

Possibility of measuring societal impact of research: scarcely possible in any reasonable way, due to the long time span and the fact that research can never be the only condition for implementing new findings and ideas

4. Research improves people’s understanding of the world

Definition: new scientific knowledge expands and deepens people’s world view: how nature works, how their own behaviour affects their health and well-being, how to identify the real causes and effects of different phenomena

Examples: understanding the basic mechanisms of the universe; knowledge on the impact of genes and the environment on the phenotype of living organisms (including humans); knowledge of basic natural phenomena and impact of different materials and regular habits on human well-being; objective knowledge of other cultures; understanding of the course of history

Societal impact: more objective and complete understanding of the world, which affects people’s behaviour as individuals, decision-makers, and employees

Other relevant factors and conditions: as in 3

Time span to benefits: usually decades or generations

Profits: economic impact, as a rule, occurs as very small shifts in different places; as a whole, benefits are enormous

Possibility of measuring societal impact of research: the societal impact is not measurable in concrete terms, but this does not make it less important than other forms of impact

How Research Can Have an Impact on Mechanisms, and be Applied to Linguistics

Linguistics is quite a large branch of scientific research, and languages are one of the biggest educational fields. So we are dealing with a topic that is very important from the point of view of the needs of society. If we look at the
operation of society, the significance of languages becomes even more evident. Both our social life and work are based on interaction between people. Thus, even minor improvements in knowledge about language or in communication skills have a potentially enormous impact on the way society works.

The wide scope of “language research” may also be demonstrated by considering various disciplines within the humanities. We linguists tend to see ourselves as specialists in the study of texts, the realizations of language use. As a matter of fact, we are not the only researchers fishing in these waters: theologians study holy texts; specialists in jurisprudence compile and then interpret legal texts; scholars in literary studies analyse texts created by famous writers; and for historians, various texts are the main source for research. But in comparison with other scholars working in the humanities, linguists can be described as universalists: we study all kinds of texts.

As for the impact of linguists on society, the main channel is, as noted at the beginning of the article, their work as teachers. Usually linguists have a lot of students, who then work in different sectors of society: as teachers, as translators, or as specialists on a particular region in various organizations and enterprises. If they have gained good practical skills, if they understand the mechanism of the language concerned, and if they have a deep, multifaceted knowledge of that culture and society, they are of great benefit to their employers. If they work in the private sector, their work is an important element in profit-making; if in the public sector, they can help the authorities to make decisions that are beneficial to society.

Linguists who do not dedicate themselves exclusively to “pure” research see the possible uses of linguistic knowledge. Bearing that in mind, they take part in compiling dictionaries and textbooks for a wider public and thus make a direct impact on society (Category 1: Production-oriented findings and innovations). In this work they have to be aware of the needs of users; at the same time, they use the linguistic knowledge gathered by the scientific community. From the point view of end users, it is irrelevant whether the author of a textbook or dictionary uses the results of her/his own research or those obtained by other linguists.

Prominent linguists expand our knowledge on languages, on their structure and use, and on the universal features of languages. One cannot see any immediate practical benefit arising from research on, say, deixis, voice, and morphosyntax in Russian. However, this is exactly the way that fundamental research moves forward. Through the collective efforts of linguists, our knowledge of languages increases. Citations are an indicator of the use of a publication by other scholars. So, according to the Publish or Perish database based on Google Scholar, the article on Deixis in embedded structures by
Richard Brecht is cited 51 times. One of those is in RJU Boogaart’s dissertation *Aspect and temporal ordering: A contrastive analysis of Dutch and English*, which has been cited 67 times. The chain continues: one of the citations is made by O. Boric in the book *Aspect and reference time* with 128 citations, etc. The scientific community evaluates and selects new data and observations on language, and on that basis, we can make more complete descriptions of languages in more concrete terms: more comprehensive grammars and dictionaries. They in turn can be used in compiling textbooks for language learners, both as a mother tongue and as a foreign language.

During the last few decades, scientific descriptions of languages have also found other applications in the automatic handling of texts: in creating spell checkers and translation engines, in data mining, in compiling semantic webs. During the last twenty years, we have seen fast development of the field but, as a whole, we are still taking only the first steps in turning out new products of this kind. Especially in speech recognition, the big success stories are still ahead. This is an example of Category 2 impact on society (Research oriented findings and innovations). One can only imagine the economic benefits of all these outcomes of research. Sometimes people in computer science think that they can produce effective language tools without the help of linguists, but as a rule, these trials fail.

As for the third category, the impact of research on the work of the public sector, practically oriented linguists play an important role in planning language policy and strategies of language education for the state or minority communities. Of course, a linguist will only take part in these processes if (s)he is interested in the sociolinguistic role of languages. But this is not enough in order to have an effect on these issues; the linguist also needs to have access to those rooms where these questions are discussed and solved. The final decisions are made by politicians, but the role of a researcher may also be significant. Here again the economic benefits are extensive, though they are very difficult to verify and measure.

How, then, can linguists influence people’s understanding of the essence and use of languages? According to my observations, we linguists still have a lot of work to do here. Laymen’s impressions of language are still rather blinkered and one-sided. Most people lack understanding of variation in language use and language command; many mechanisms of language use are unfamiliar to them.

The question which, to my mind, needs much more attention from linguists is that of communication problems in interaction between people. Insufficient mutual understanding is arguably one of the most serious problems in the world. Given that most human activities are somehow based on communication, success in interaction is of paramount importance. I think
that studying this problem in a more systematic way could multiply the societal relevance of linguistic research. There are, however, some obstacles in gaining that position. I will comment on them briefly.

There are linguists who think that communication and problems in it are not a topic for linguistic research. They are partially right: in studying human interaction, we must also take into account issues such as the psychological and sociological attributes of the communicants (abilities and motivation; behavioural and ideological features). Yet it would be unwise to leave research of (mis)communication only to psychologists and sociologists. To my mind, in investigating this very complex phenomenon, we can reach substantial results only through the joint efforts of researchers from different disciplines.

In discussing the problems of communication with representatives of other fields, I have met the following counter-argument: interaction between people is a unique event and it is therefore impossible to find any systematic features in it. This is a relevant claim. However, it is untenable when we take a closer look. For comparison, let us take another severe problem: cancer. People die from cancer, which is a good reason to make every possible effort to discover the mechanism behind this disease. Thousands of studies have shown that a great variety of factors can lead to cancer and that these factors function differently, depending on the individual. However, by careful and patient research, we have gradually acquired a better understanding of the nature of cancer. By using this new knowledge we are better able to prevent and cure some cancers. Cancer and communication failures are naturally problems of a different order of severity. But there are also similarities between them: in both cases we are dealing with a regrettable, but inherent and inescapable part of human life. With the help of research we are able to reduce the negative effects.

There is a great obstacle to the progress of studies on human interaction: due to the multidisciplinary nature of the problem, researchers work in total isolation. The lack of awareness among researchers can be shown by a single example: a substantial review with more than 200 references made by Berger (2005) is based on a completely different set of studies than another review of interpersonal communication with an equal number of references by Stamp (1999), and it is even more striking that the works of psycholinguists are missing in both of these reviews. In order to solve one of the most serious problems of the human race, troubles in human interaction, we need more awareness of studies conducted by other researchers. It is only in this way that research can be cumulative – a feature that is essential for a scientific approach to any topic of research.
References
Mustajoki, Arto. ”Tutkimuksen vaikuttavuus: mitä se on ja voidaanko sitä mitata?” (Relevance of research: What is it and can it be measured) Tieteessä tapahtuu 6 (2005): 33-37.
Assessment Practices in STARTALK Language Programs:
A View of Current Language Assessment Literacy

Margaret E. Malone, Megan J. Montee, Francesca DiSilvio

1. Introduction
Assessment is essential to education, because it provides information on students’ progress toward learning goals. Reliable and valid assessment can provide not only important summative information, but also formative information to instructors and learners on both what has been learned and what remains to be learned. However, in order for assessment to be used effectively, instructors must understand the components of a reliable and valid assessment system and how to incorporate such a system into classroom testing. Many language instructors in the United States may lack basic knowledge of assessment and measurement (Popham, 2009).

Because assessment of world language learning is not currently required by No Child Left Behind regulations, there is no national system for K-12 world language assessment. Moreover, there is no clear documentation of the actual language assessment processes and practices in world language classrooms in the U.S. These issues affect the teaching of all languages; among the less commonly taught languages (LCTLs), the challenges are magnified. There is a shortage of nationally available, research-based LCTL assessments across language modalities, which are identified by the 2006 Standards for Foreign Language Learning in the 21st Century as Interpersonal Speaking, Interpretive Reading and Listening, and Presentational Writing and Speaking (National Standards in Foreign Language). Furthermore, many LCTL instructors lack the background needed to interpret those tests that are available, or to develop their own reliable and valid classroom tests (Nier, Donovan, and Malone, 2009). This dual lack of materials and of assessment knowledge presents a serious obstacle to documenting language attainment in LCTLs.

Since 2007, the U.S. Department of Defense has funded STARTALK, a program that provides summer learning opportunities in languages deemed critical to U.S. security. In 2010, over 5,000 students in grades K-16 participated in summer language programs in Arabic, Chinese (Mandarin), Dari, Hindi, Persian, Swahili, Turkish and Urdu. STARTALK programs are short-term, varying in
length from two weeks to two months, although many have emerged from existing academic-year LCTL programs. Each potential program applies for STARTALK funding individually. Funded programs develop their own curricula and assessment plans based on STARTALK requirements, and all programs report their assessment plans through required curriculum documents. Given STARTALK’s centralized reporting requirements, an examination of these curricula and their included assessment plans provides an opportunity to investigate the assessment practices of a large group of LCTL programs.

This paper provides a systematic analysis of the formative and summative assessment practices of STARTALK programs and the extent to which these practices support stated program goals. Furthermore, it also discusses how increased language assessment literacy could help STARTALK and academic-year programs improve their assessment practices. The paper begins with an explanation of the two main types of assessment used in STARTALK, formative and summative assessment. It then provides background on language assessment literacy, followed by a description of STARTALK’s relevance to assessment practices in world language programs nationwide. Next, the paper describes the methodology and results of a study conducted to analyze assessment practices across 2010 STARTALK student programs. Finally, it suggests ways that assessment literacy can be improved in STARTALK programs in particular, and U.S. world language programs in general.

2. Formative and Summative Assessment

Most assessment is categorized according to how the results will be used. For example, language assessments can be used to make placement decisions, to certify instructors, or to evaluate the outcomes of a program, among other purposes. Assessments can also be classified on a continuum from high stakes (summative assessments) to low stakes (formative assessments) (Stoynoff & Chapelle, 2005). High stakes or summative language assessments are generally used to make decisions, including whether a prospective teacher has sufficient language proficiency to receive certification to teach that language in a specific state. Other uses for summative assessments include assigning grades, determining whether graduation requirements have been met, or allowing admission to a specific language program for a course of study. On the other end of the spectrum, low stakes or formative assessments are generally used to provide information to students and instructors on progress toward specific learning goals in a course (Stoynoff & Chapelle, 2005). However, the dichotomy presented by these definitions is somewhat misleading; clearly, formative
assessment should be aligned to the goals of summative assessment. Similarly, summative assessments should reflect the learning goals of formative assessments. Such an alignment can improve the likelihood that formative assessments reflect progress toward the outcomes ultimately attained on summative assessments.

Although summative assessments can refer to standardized, nationally-available tests, many language instructors develop their own summative assessments to provide the basis for grades and to demonstrate the attainment of course outcomes. Many STARTALK programs provide students with grades or course credits, but these results are not the only reason that summative assessment is important for STARTALK programs. It can also provide important feedback for program staff and local stakeholders. Furthermore, STARTALK summative assessments in the aggregate help to reveal what can be taught and learned in this type of short-term language program.

Developing formative and summative assessments can be a challenge for short-term language programs. While most of us are familiar with the multiple-choice approach used on most standardized tests in the U.S., STARTALK programs, like many year-long programs, often rely on performance assessments to gauge student outcomes. In contrast to multiple-choice tests, performance assessments require and expect students to use language in real-life situations; demand that students develop an understanding of what real-life situations are like in the language; and are based on performance or the active construction of language, rather than a demonstration of passive understanding only (Norris et al., 1998; Sandrock, 2010). Examples of performance assessments are numerous; for example, students might conduct role plays, in which they work with another student or students to simulate real-life situations. Such role plays are largely spontaneous, unlike a formal student presentation, which, as a performance assessment, generally requires planning. Other examples of performance assessments include interviews, in which a student participates in an interview with the teacher or another student. All performance assessments should be developed and rated according to a specific set of criteria that reflect the goals and purposes of the course.

3. Language Assessment Literacy

Assessment is essential to determining and documenting the outcomes of any language learning experience (Brown, 2004; Hughes, 2003). Furthermore, assessment results can provide valuable information for all participants in the language learning process (Shepard, 2000). Despite its acknowledged importance,
however, many administrators, instructors, and students do not understand the basics of assessment (Popham, 2009). As the central role of testing in the U.S. educational system has only expanded under No Child Left Behind, the gap between what stakeholders need to know about assessment and what they do know must be closed.

The term “assessment literacy” refers to familiarity with assessment processes (Stiggins, 2001). Robust language assessment literacy combines an understanding of testing concepts with knowledge of both language acquisition and language teaching methodologies. Recently, a number of researchers (McNamara and Roever, 2006; Inbar-Lourie, 2008; Malone, 2008; Taylor, 2009) have discussed the importance of language assessment literacy as well as strategies for improving it among stakeholders. Two recent studies (Nier, Donovan and Malone, 2009; Riestenberg et al., in press) have focused specifically on language assessment literacy for instructors of less commonly taught languages. Such publications suggest that researchers largely appreciate the importance of promoting assessment literacy among world language educators. However, there is still little documentation or analysis of current assessment knowledge and practice among world language instructors. Additionally, little research has yet been conducted on how exactly to promote assessment literacy in order to increase positive washback (Hughes, 2003) for world language teaching and learning.

One challenge to defining and addressing language assessment literacy is that, while there is nearly universal acceptance of the need for assessment literacy among language testing experts, and efforts to examine and recommend improvements to existing formal language testing courses, there is not yet any agreed-upon process for attaining different levels of language assessment literacy. The remainder of this paper describes a study that addresses one way to examine current levels of assessment literacy in a national language program.

4. The STARTALK Initiative
STARTALK is a U.S. presidential initiative directed toward teaching languages deemed critical to national security. STARTALK programs provide summer language instruction for students in grades K-16. STARTALK’s stated mission is …to increase the number of Americans learning, speaking, and teaching critical need foreign languages by offering students (K-16) and teachers of these languages creative and engaging summer experiences that strive to exemplify best practices in language education and in language teacher development, forming an
extensive community of practice that seeks continuous improvement in such criteria as outcomes-driven program design, standards-based curriculum planning, learner-centered approaches, excellence in selection and development of materials, and meaningful assessment of outcomes (National Foreign Language Center, 2011).

The STARTALK programs, which began in summer 2007, have added new languages and student grade levels each year. In 2011, STARTALK languages will include Arabic, Chinese (Mandarin), Dari, Hindi, Persian, Portuguese, Russian, Swahili, Turkish, and Urdu. There are two types of STARTALK programs: language learning programs for students, and professional development programs for teachers of these critical languages. Since 2007, STARTALK has served over 7,000 students learning critical languages, and 3,000 teachers and prospective teachers of these languages (National Foreign Language Center, 2011).

In addition to providing language learning and professional development experiences in critical languages, the STARTALK initiative has developed consistent standards and procedures by which participating programs must operate, beginning with their initial applications and ranging all the way through the final reporting process (see also Ingold, this volume). As stated in its mission, STARTALK supports the development and execution of excellent summer language experiences that embody the best of current research and practice. To this end, the standard documents and procedures have been developed based on both government reporting requirements and current best practices for language education. All funded programs are required to develop, submit, and, if necessary, revise a curriculum template. This curriculum template mandates that each program describe in detail its learning goals and objectives and how they will be assessed using both formative and summative tools.

The STARTALK initiative therefore presents a rare grouping of a large, diverse collection of less commonly taught language programs that utilize common standards and forms to describe their curricula and assessment practices. It therefore provides an optimal opportunity to investigate how assessment practices are described and used in LCTL programs across the U.S. The next sections of this paper describe the results of such an investigation, including research questions and a methodology for identifying assessment practices in STARTALK programs.

5. Methods
In 2010, researchers conducted a descriptive and qualitative study of the self-reported assessment practices of the 2010 STARTALK student programs. By both describing the types of assessment practices reported in STARTALK curricula and examining the relationship between assessment practices and instructional goals, the study provided a picture of assessment practices across STARTALK programs, while also identifying areas where additional training and resources may be needed. The study had two phases: the first examined assessment activities and patterns across STARTALK programs, and the second provided a qualitative analysis of the ways in which formative and summative assessment was used to measure curriculum goals. Research questions for the study were:

1. What types of assessment activities do STARTALK programs report using?
2. What patterns emerge across languages, proficiency levels, and grade-level clusters?
3. To what extent are STARTALK programs assessing stated curriculum goals?
4. How are STARTALK programs using formative and summative assessment and how do these reflect current levels of assessment literacy in STARTALK programs?

The researchers answered questions 1 and 2 through a content analysis of the Student Program Curriculum Template and Guide (referred to in this paper as simply “curriculum template”) submitted by the 156 STARTALK student programs that operated in 2010. Each curriculum document contained information on the assessment activities that students would be expected to perform as evidence of learning. Section F of the curriculum template, “End of Program Performance Tasks,” asked programs to list the culminating performance tasks that students would complete to demonstrate their achievement of the program learning objectives. Section G asked programs to list “Other Types of Assessment and Evidence of Learning.” Items in sections F and G were first coded into emerging categories, each representing a different type of assessment activity; all coding was checked by two additional researchers. In case of disagreement, the coders discussed and reached consensus. In all, 34 assessment activity codes were created; each curriculum document contained 3-19 codes, depending on the amount of detail included. Then, to address research question 2, coded data were broken down by program language, proficiency level, and grade level for further analysis.
In the second phase of the study, the curriculum documents were reviewed as part of a qualitative analysis. A researcher reading the documents identified a set of preliminary themes for research questions 3 and 4. These preliminary themes were added to and revised during the analysis process, and curriculum documents were iteratively compared to the emerging themes to ensure that the descriptions accurately captured the data. The curriculum documents were then reviewed again during analysis to ensure the trustworthiness of the data. Finally, two additional members of the research team familiar with the curriculum documents checked the validity of the findings and recommendations. This cyclical method of analysis resulted in a final set of major themes related to each research question, as well as recommendations based on the findings.

6. Results

Research Question 1

The first research question asked: What types of assessment activities do STARTALK programs report using? To address this question, the researchers coded and quantified the different types of assessment practices reported in the curriculum templates, including classroom-based assessments and standardized tests. Table 1 shows the frequency with which various types of assessment practices were reported across the 156 STARTALK programs by raw number and percentage of total programs. As Table 1 below shows, the assessment activity reported by the greatest number of programs was role play, followed by presentation, and then reading comprehension. Overall, oral assessments, including presentations, role plays, interviews, and assessments of oral comprehension, were among the most frequently reported types of assessment. Programs also reported the use of standardized assessments. A small number of programs (17 total, or 11%) reported using nationally available standardized tests; in some cases, programs reported using more than one standardized test. Table 2 shows these tests and the number of programs that reported using each. Furthermore, a total of 26 programs (17%) reported general use of tests or exams, most likely indicating that some program administrators or instructors developed tests internally.

Research Question 2

The second research question asked: What patterns emerge across languages, proficiency levels, and grade-level clusters? To address this question, assessment
activities were analyzed for patterns within program target language, proficiency level, and grade-level clusters.

Table 1. Assessment practices across programs

<table>
<thead>
<tr>
<th>Activity</th>
<th>Total</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role play</td>
<td>135</td>
<td>86.5</td>
</tr>
<tr>
<td>Presentation</td>
<td>128</td>
<td>82.1</td>
</tr>
<tr>
<td>Reading comprehension</td>
<td>115</td>
<td>73.7</td>
</tr>
<tr>
<td>Poster/graphic project</td>
<td>101</td>
<td>64.7</td>
</tr>
<tr>
<td>Interview</td>
<td>97</td>
<td>62.2</td>
</tr>
<tr>
<td>Oral comprehension</td>
<td>94</td>
<td>60.3</td>
</tr>
<tr>
<td>LinguaFolio</td>
<td>82</td>
<td>52.6</td>
</tr>
<tr>
<td>Skit/performance</td>
<td>75</td>
<td>48.1</td>
</tr>
<tr>
<td>Other project</td>
<td>70</td>
<td>44.9</td>
</tr>
<tr>
<td>Writing/composition</td>
<td>68</td>
<td>43.6</td>
</tr>
<tr>
<td>Podcast/video</td>
<td>61</td>
<td>39.1</td>
</tr>
<tr>
<td>Information gap</td>
<td>60</td>
<td>38.5</td>
</tr>
<tr>
<td>Song/rhyme</td>
<td>57</td>
<td>36.5</td>
</tr>
<tr>
<td>Journal</td>
<td>54</td>
<td>34.6</td>
</tr>
<tr>
<td>Venn diagram</td>
<td>54</td>
<td>34.6</td>
</tr>
<tr>
<td>Informal observation</td>
<td>39</td>
<td>25.0</td>
</tr>
<tr>
<td>Research project</td>
<td>36</td>
<td>23.1</td>
</tr>
<tr>
<td>Games</td>
<td>34</td>
<td>21.8</td>
</tr>
<tr>
<td>Field trip</td>
<td>32</td>
<td>20.5</td>
</tr>
<tr>
<td>Quiz</td>
<td>31</td>
<td>19.9</td>
</tr>
<tr>
<td>Worksheet/assignment</td>
<td>29</td>
<td>18.6</td>
</tr>
<tr>
<td>Test/exam</td>
<td>26</td>
<td>16.7</td>
</tr>
<tr>
<td>Portfolio</td>
<td>22</td>
<td>14.1</td>
</tr>
<tr>
<td>Survey</td>
<td>20</td>
<td>12.8</td>
</tr>
<tr>
<td>Dialogue</td>
<td>19</td>
<td>12.2</td>
</tr>
<tr>
<td>Nationally available standardized tests</td>
<td>17</td>
<td>11.0</td>
</tr>
<tr>
<td>Food/cooking project</td>
<td>15</td>
<td>9.6</td>
</tr>
<tr>
<td>E-portfolio</td>
<td>14</td>
<td>9.0</td>
</tr>
<tr>
<td>Oral proficiency interview</td>
<td>14</td>
<td>9.0</td>
</tr>
<tr>
<td>Self-assessment</td>
<td>11</td>
<td>7.1</td>
</tr>
<tr>
<td>Computer/technology</td>
<td>8</td>
<td>5.1</td>
</tr>
<tr>
<td>Student participation</td>
<td>4</td>
<td>2.6</td>
</tr>
</tbody>
</table>
Table 2. Nationally available standardized tests used by STARTALK programs

<table>
<thead>
<tr>
<th>Test Description</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computerized Assessment of Proficiency (CAP)</td>
<td>5</td>
</tr>
<tr>
<td>Standards-based Measurement of Proficiency (STAMP)</td>
<td>5</td>
</tr>
<tr>
<td>Assessments provided by the Center for Advanced Study of Language (CASL)</td>
<td>4</td>
</tr>
<tr>
<td>ACTFL Assessment of Performance toward Proficiency in Languages (AAPPL)</td>
<td>3</td>
</tr>
<tr>
<td>Computerized Oral Proficiency Instrument (COPI)</td>
<td>1</td>
</tr>
<tr>
<td>ALOFA (Arabic Language Oral Fluency Assessment)</td>
<td>1</td>
</tr>
</tbody>
</table>

Assessment Practices by Target Language

Assessment practices within language groups were analyzed for Arabic, Chinese, and Hindi programs. Persian, Swahili, Turkish, and Urdu programs were not included in this part of the analysis, as there were fewer than ten programs in these languages, making consistent patterns difficult to identify. Table 3 lists the top five assessment practices for each language, along with the percentage of programs within that language that reported using this practice.

Consistent with the most frequently reported assessment practices for all languages noted above, the top assessment activities by target language included role play, followed by presentation (equal in frequency to role play within Chinese programs and to oral comprehension within Hindi programs), and reading comprehension for Arabic and Chinese (equal in frequency to oral comprehension within Arabic programs). For Arabic and Hindi, all or almost all programs reported using role plays and presentations. There was less uniformity in reported assessment practices across Chinese programs, perhaps due to the larger number of Chinese programs represented in the data. Notably, Arabic diverged from the other programs in its relatively higher frequency of quiz, test/exam, and worksheet assessment activities. More Arabic programs also reported the use of standardized tests (23%). By contrast, no Hindi program reported using a standardized test (possibly because few standardized tests are available in Hindi). Furthermore, although there were fewer Hindi programs represented in the data, it may also be worth noting that 75% of Hindi programs reported the use of journals compared with 40% of Arabic programs and 24% of Chinese programs.

Assessment Practices by Proficiency Level

Assessment practices were also analyzed by proficiency level. In its initial funding application, each program was required to indicate the proficiency level of the students at which it would aim its courses. Most, but not all programs, reported this information in terms of the ACTFL proficiency guidelines.
(American Council for the Teaching of Foreign Languages, 1999). For analysis, programs were grouped in the four categories listed in the table below.

Table 3. Top five assessment practices by target language

<table>
<thead>
<tr>
<th>Arabic (N=30)</th>
<th>Chinese (N=91)</th>
<th>Hindi (N=16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role play (93.3%)</td>
<td>Role play (79.1%)</td>
<td>Role play (100%)</td>
</tr>
<tr>
<td>Presentation (83.3%)</td>
<td>Presentation (79.1%)</td>
<td>Presentation (87.5%)</td>
</tr>
<tr>
<td>Oral comprehension (70%)</td>
<td>Reading comprehension (75.8%)</td>
<td>Oral comprehension (87.5%)</td>
</tr>
<tr>
<td>Reading comprehension (70%)</td>
<td>Poster/graphic project (65.9%)</td>
<td>Interview (81.3%)</td>
</tr>
<tr>
<td>LinguaFolio (66.7%)</td>
<td>Interview (57.1%)</td>
<td>Poster/graphic project (81.3%)</td>
</tr>
</tbody>
</table>

Table 4 below shows the top five most frequently reported assessment practices for each proficiency level cluster. Table 4 demonstrates that assessment practices do not seem to vary greatly according to proficiency level. Role plays, presentations, and reading comprehension assessments were reported at all levels. However, posters and graphic projects were more common at the Intermediate and Advanced levels, while oral comprehension and interview assessments were more common for the Novice learners.

Assessment Practices by Grade Level

Finally, assessment practices were analyzed by program grade level; all programs listed at least one grade level and some programs fit into more than one grade level cluster. Table 5 shows the top five assessment practices used with each grade level cluster. As Table 5 below shows, all grade level clusters reported frequent usage of role plays. Within the early elementary cluster, oral comprehension was the most frequently reported assessment activity; it ranked third most frequent among upper elementary programs. Interviews were among the most frequent assessment activities reported by middle school, high school, and college programs, but not elementary programs. As might be expected, reading comprehension was not frequently reported as an assessment activity by programs serving early elementary students. Assessment of reading comprehension played a relatively steady role in the four higher grade clusters, though, ranging from use by 64% to 78% of programs at these levels. Writing assessments were reported by 20% of early and upper elementary school programs, and this number increased with each successive grade cluster; at the college level, writing assessments were reported by 73% of programs. The elementary clusters reported little usage of standardized tests or of internally
developed, classroom-based exams. The use of testing, both standardized and classroom-based, was most frequently reported at the middle and high school levels.

Table 4. Top five assessment practices by proficiency level

<table>
<thead>
<tr>
<th>Novice (N=94)</th>
<th>Both Novice and Intermediate (N=32)</th>
<th>Intermediate (N=20)</th>
<th>Both Intermediate and Advanced (N=10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role play (84.0%)</td>
<td>Role play (93.8%)</td>
<td>Presentation (95.0%)</td>
<td>Role play (90.0%)</td>
</tr>
<tr>
<td>Presentation (78.7%)</td>
<td>Presentation (84.4%)</td>
<td>Reading comprehension (85.0%)</td>
<td>Poster/graphic project (90.0%)</td>
</tr>
<tr>
<td>Reading comprehension (68.1%)</td>
<td>Reading comprehension (81.3%)</td>
<td>Role play (85.0%)</td>
<td>Presentation (80.0%)</td>
</tr>
<tr>
<td>Oral comprehension (63.8%)</td>
<td>Interviews (68.8%)</td>
<td>Poster/graphic project (75.0%)</td>
<td>Reading comprehension (80.0%)</td>
</tr>
<tr>
<td>Interviews (60.6%)</td>
<td>Oral comprehension (62.5%)</td>
<td>LinguaFolio (70.0%)</td>
<td>Interviews (60.0%)</td>
</tr>
</tbody>
</table>

Research Question 3
The third research question asked: To what extent are STARTALK programs assessing their stated curriculum goals? As part of their initial funding application, each program was required to identify curriculum goals related to each of the five Standards for Foreign Language Learning (National Standards in Foreign Language, 2006): Communication, Cultures, Connections, Comparisons, and Communities. To address the third research question, each program’s stated goals were compared with its assessment practices. Additionally, the researchers also considered each program’s declared Standards-based goals in relation to its entire curriculum document in order to glean other relevant contextual
information about instructional materials and practices, program themes, and the specific knowledge and skills covered by the program. As described in the methods section above, data for research questions three and four were analyzed qualitatively and are presented descriptively by theme.

Table 5. Top five assessment practices by grade level

<table>
<thead>
<tr>
<th></th>
<th>Early Elementary (N=18)</th>
<th>Upper Elementary (N=36)</th>
<th>Middle School (N=51)</th>
<th>High School (N=122)</th>
<th>College (N=11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral comprehension</td>
<td>Role play (80.6%)</td>
<td>Role play (82.4%)</td>
<td>Role play (88.5%)</td>
<td>Role Play (100.0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(94.4%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Role play</td>
<td>Presentation (72.2%)</td>
<td>Poster/graphic project (78.4%)</td>
<td>Presentation (86.1%)</td>
<td>Presentation (90.9%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(72.2%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other project</td>
<td>Oral comprehension (69.4%)</td>
<td>Presentation (78.4%)</td>
<td>Reading comprehension (76.2%)</td>
<td>Interview (81.8%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(66.7%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poster/graphic project</td>
<td>Reading comprehension (69.4%)</td>
<td>Reading comprehension (78.4%)</td>
<td>Interview (63.9%)</td>
<td>Skit/performance (81.8%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(66.7%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song/rhyme</td>
<td>Poster/graphic project (63.9%)</td>
<td>Interview (56.9%)</td>
<td>Poster/graphic project (63.1%)</td>
<td>Information gap (72.7%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(66.7%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Assessment and the Standards

Programs varied in how they reported goals and expected outcomes. While some programs included multiple detailed outcomes for each Standard, other programs reported general outcomes across Standards. Furthermore, when reporting their assessment practices, many programs only discussed the Communication Standard, making no mention of how they might assess the other four. This indicates that programs may not be planning assessments for areas that are more difficult to measure, such as culture or interest in future language study.
While programs varied in the specificity of their responses, most programs included some detail related to specific program content or instructional activities with mention of general proficiency-based outcomes rather than program-specific outcomes aligned with proficiency levels.

**Level of Planning**

The second theme that emerged from the analysis relates to the level of planning in which programs engaged relative to their assessment tasks. Programs typically described their assessment plans with less detail than their expected outcomes and planned instructional activities. For example, in Section E of the template, “Specific Knowledge and Skills,” an expected outcome listed by one Arabic language program indicated that students would be able to “introduce themselves and others and give appropriate greetings and responses,” and then provided a list of the specific vocabulary students would use for this function. This information was explicitly requested in the curriculum template and most programs provided such detailed responses. This same program, however, reported a planned assessment task (Section F) in which students would “produce recordings of short conversations and dialogues about Greetings/Introductions.” Though this planned assessment relates directly to the course material, the program provided no information about the quantity and quality of language that the students would be expected to produce, or how student responses would be scored on this task. The template, in fact, does not require programs to list any explicit assessment criteria, and most programs did not volunteer such information. Overall, few programs explained how students would be expected to respond to assessment tasks, the criteria for successful task completion, or how the tasks would be scored or rated. For example, only a small number of programs mentioned the use of rubrics to score performance assessments. This indicates that programs may not yet be defining explicit assessment criteria or systematically scoring tasks.

Programs’ general lack of firm assessment plans suggests that they may not be implementing backward design (Wiggins and McTighe, 2005), a curriculum planning method in which a final summative assessment is designed first, followed by plans for the lessons that will directly lead up to this assessment. CAL’s experience in teaching assessment training courses to STARTALK instructors supports the observation that few instructors employ the backward design strategy. The level of assessment planning shown in the curriculum templates also demonstrates a need for increased assessment literacy.
Proficiency Levels

The third theme that emerges from the analysis relates to the degree to which programs accounted for their students’ level of language proficiency when designing assessment practices and predicting outcomes. Analysis showed that outcomes and assessment plans were often less focused on proficiency levels than on the *Standard* being taught. (Of course, the alignment to the *Standards* is likely related to the structure of the curriculum template, which emphasizes themes and *Standards.*) In some cases, programs’ reported assessment tasks were clearly inappropriate for their students’ proficiency level, such as an essay writing task for a program with Novice-level students. In most cases, however, programs simply did not report assessment information.

The curriculum template also gives comparatively little attention to the role of proficiency in assessment design. This lack of focus on proficiency level points to larger issues within STARTALK programs related to proficiency, assessment, and the proficiency outcomes that can be expected for short-term language programs. Several ongoing projects are exploring standardized proficiency measures that could be used across programs. Still, data from the curriculum templates indicate that programs are not currently focused on aligning assessment tools and program outcomes with proficiency levels.

In summary, the researchers’ analysis revealed that STARTALK programs’ assessment plans were generally less developed and less detailed than their instructional plans. This may reflect both the limitations of the curriculum template as well as a generalized lack of assessment planning among programs. Finally, while assessment plans did reflect *Standards* and program themes, they were often disconnected from their students’ proficiency levels and desired proficiency outcomes.

Research Question 4

The fourth research question asked: How are STARTALK programs using formative and summative assessments? While the results of Research Questions 1 and 2 provided a detailed list of the types of assessments programs reported using, this data did not show the extent to which the reported assessment activities were being used summatively or formatively. In the curriculum documents, programs were asked to list formative and summative practices separately in different sections (Sections F and G, respectively). However, analysis during the initial phase of the study indicated that reporting practices were inconsistent, rendering straightforward coding into formative and summative categories impossible. Given this limitation, the researchers instead
reviewed assessment practices for themes related to formative and summative purposes.

This qualitative analysis showed that programs reported formative assessments to a greater extent than summative assessments. In Section F of the curriculum template, programs were instructed to report summative performance tasks, defined in the template as “culminating performance tasks [that] will provide evidence that students have achieved the program learning objectives.” However, programs often used this section to describe activities generally categorized as formative rather than summative. For example, programs often reported daily assessment tasks in Section F, though these are generally considered formative in nature, as they inform the day-to-day planning of classroom activities. Many programs also used Section F to report project-based assessments, often conducted in groups. While such activities between students provide excellent language practice and solid formative assessment information, the programs’ reports generally do not indicate that they were developed to be particularly summative. Overall, very few programs used Section F to describe summative performance tasks intended as final assessments of students’ proficiency or achievement.

Additional analysis suggests that many programs used neither formative nor summative performance task-based assessments. Performance tasks are authentic and contextualized assessments of language use (Norris et al, 1998); most programs’ assessments did not require students to use the target language in any such authentic context. Thus, while programs are moving beyond traditional assessment practices such as multiple-choice testing, it is unlikely that such programs are implementing a task-based approach to assessment.

Finally, programs generally did not report the criteria they used to conduct formative or summative assessments. Rubrics and scoring procedures were not frequently mentioned in the curriculum template, and no clear system for integrating formative and summative assessment emerged. This suggests that, although activities such as role plays were occurring between students and sometimes between students and instructors, clear criteria for evaluating student performance on such activities were not developed. Without such criteria, these activities cannot provide helpful and valid formative or summative assessment.

7. Discussion
The STARTALK curriculum templates provide a useful source of information about the range of assessment activities used by STARTALK programs. Because this information is reported on a standardized form that includes guidelines and
suggestions, data were uniform and may not represent the full range of assessment practices included in STARTALK programs. Additionally, information on how these practices are being implemented, rather than simply how they are being planned, is not consistently captured in the template, and is therefore not reported here. Despite these limitations, however, this analysis revealed several trends in the self-reported data. The data provide important information about current understanding of language assessment literacy in STARTALK programs as reflected in the STARTALK curricula documents.

Focus on Oral Assessment and Communication
Assessments of oral skills including presentations, role plays, interviews, and assessments of oral comprehension were among the most frequently reported assessment practices across STARTALK programs. This remained true when the data were analyzed by language, proficiency level, and grade-level cluster. These assessments cover both Interpersonal and Presentational Speaking, indicating that many programs are assessing both Communicative Modes. Interviews were used less frequently to assess oral skills, possibly because this method may be more time-consuming for teachers to conduct and more challenging to assess than presentation or role play tasks. The overall focus on oral proficiency may indicate that programs need additional resources for assessing literacy skills, particularly writing. While reading comprehension was assessed frequently across programs, writing assessment was limited in the lower grade level clusters. In addition, few programs reported how the assessments were being rated and how the results were being reported, which are two crucial aspects of assessment.

Patterns by Grade Rather than Proficiency Level
Consistent patterns in assessment practices emerged by grade level, while patterns across proficiency levels seemed less systematic. This may indicate that the target grade level(s) of a program play a greater role in determining their choice of assessment practices than do the target proficiency level(s). For example, while assessment activities such as games and songs were frequently reported by lower and upper elementary programs, these activities decreased in frequency with increases in grade level. The data also show that different assessment practices are being used across a variety of proficiency levels. For example, interviews were reported at relatively high rates across all proficiency levels. Additional resources may be needed to help programs to understand how to adapt these practices for target proficiency levels in order to make them level-
appropriate. These results also indicate that professional development for STARTALK program directors and instructors by target grade level and proficiency level may be useful.

**Limited Use of Standardized Tests**

Programs reported limited use of standardized testing. In many STARTALK languages or for certain grade levels, there may be limited standardized materials available. Programs may also need guidance in how to use standardized tests appropriately in short-term programs, as these types of tests may have limited uses in such contexts. This finding highlights the need for projects such as the present study that examine the usefulness and efficacy of testing in STARTALK programs.

The three main findings indicate that STARTALK programs report using a variety of assessments, yet explain little about how these assessments are rated, outcomes reported, and results used for program improvement and reflection. Therefore, professional development and increased awareness of language assessment fundamentals are likely needed for many STARTALK language programs.

8. **Recommendations to Improve Assessment Literacy**

Based on the analysis of the curriculum documents and assessments, recommendations can be made for improving assessment literacy among stakeholders in STARTALK programs in particular and U.S. world language programs in general.

**Facilitate Alignment between Outcomes, Instruction, and Assessment**

The current STARTALK curriculum template includes sections for programs to describe the outcomes they expect to achieve, their instructional practices, and the assessments they plan to use. For STARTALK, the authors recommend that the curriculum template be modified so that for each Standard, programs must list one or more expected outcomes as well as what assessment tools will be used to measure students’ progress toward these outcomes.

It is also recommended that year-long and STARTALK programs alike investigate ways to include professional development training on the alignment between curriculum and assessment. The researchers recommend that all world language teachers participate in relevant training on developing, administering, reporting, and explaining the results of contextualized, authentic tasks that measure stated program outcomes.
Emphasize proficiency levels and understanding
Currently, the STARTALK programs’ reported assessment practices reflect the Standards and incorporation of program themes. However, these programs have not yet aligned these practices with the ACTFL Proficiency Guidelines. A similarly limited relationship exists between the Guidelines and the programs’ expected learning outcomes. While no data exist to confirm a similar gap between expected outcomes and assessment practices and the ACTFL Proficiency Guidelines in school-year programs, the researchers believe that additional resources and professional development on proficiency levels may also be useful for these world language instructors. Developing prototypical tasks and benchmark samples of can-do tasks could greatly help world language instructors better understand the proficiency levels.

Focus on assessment criteria, scoring, and feedback
In order for assessment tasks to be meaningful and useful, methods for scoring and providing feedback should be planned out as part of the initial assessment planning process. Very few STARTALK programs reported the criteria by which they score performance tasks. STARTALK programs and academic-year world language programs alike would benefit from performance task rubrics that could be adapted and shared across programs in the U.S. Stakeholders could also benefit from professional development on criteria, scoring, and feedback that includes benchmark samples; instruction on how to score performance tasks; and tips for providing feedback to students, instructors, and other stakeholders.

STARTALK provides language learning opportunities for increasing numbers of students every year, and its centralized structure has provided a window into its programs’ current, self-reported assessment practices. The current study suggests that STARTALK instructors, many of whom are language instructors during the school year, would benefit from proficiency-oriented, developmentally appropriate professional development on language assessment. While such measures will certainly assist the STARTALK program, the same activities also have the potential to improve world language learning across the U.S. For world language programs in general, the recommendations suggested in this paper provide steps to improved assessment literacy that will close the gap between what world language instructors need to know about assessment and what they do know.
References


A Concurrent Validity Study of Self-Assessments and the Federal Interagency Language Roundtable Oral Proficiency Interview

Charles W. Stansfield, Jing Gao, William P. Rivers

Introduction
The National Language Service Corps (NLSC) was established in 2006 as a pilot program under the auspices of the National Security Education Program (NSEP). This new organization was tasked with providing and maintaining a standing civilian corps of certified bilinguals who would be available for service to federal government agencies as they are needed, and to state and local government agencies in time of emergency. The intent of the NLSC is to fill the gap between full-time language services professionals and individuals who wish to volunteer for temporary services for short- or medium-term assignments. The NLSC recruits a readily-available pool of individuals who have expertise in languages that are important to the security and welfare of the nation. The NLSC must qualify applicants as part of its enrollment process. The NLSC uses the Federal Interagency Language Roundtable Language Proficiency Skill Level Descriptions (the ILR scale) in speaking, reading, and listening as a basis for determining eligibility for Charter membership. The NLSC requires candidates to have proficiency in English and the foreign language at ILR level 3 in the domains of speaking, reading, and listening. All NLSC applicants are asked to complete a series of language self-assessment questions in the four communicative skills and English as an initial screen of language proficiency. These self-assessments provide an indication of applicants’ levels on the ILR scale, and minimize the overall formal testing requirements by eliminating the need to test individuals who self-assess at levels lower than those required by the NLSC. Formal assessment of English language skills is waived for applicants who attended an accredited high school or college in the U.S. for at least three years and graduated.

All NLSC applicants complete a basic application form, respond to a language-background questionnaire and complete a two-part self-assessment form. The first self-assessment is a series of “Can-Do statements,” which are commonly referred to as “Can-Do scales” in the language testing literature. Can-Do scales require the individual to accept or reject an affirmative statement such
as “I can comprehend an oral presentation at a conference on a complex topic in my profession, and I can also comprehend the question-and-answer session immediately following the main part of the talk” (DD Form 2933, Version 4, Sep 2009, National Language Service Corps (NLSC) Detailed Skills Self-Assessment). The Can-Do scales utilized by the NLSC are grouped according to the skill and level on the ILR scale with which they are associated. These Can-Do scales, once completed, provide a useful inventory of the language skills the candidate claims to have. The candidate’s responses can be analyzed in order to produce a predicted language proficiency level on the ILR scale for each skill (listening, speaking, reading, and writing). The second self-assessment is a simplified set of ILR skill level descriptions. The candidate reads the description for each skill and selects the one that best describes his or her language proficiency in that skill. An example of such a simplified ILR description follows for level 4 listening, i.e., advanced professional proficiency level, (DD Form 2934, Sep 2009, National Language Service Corps (NLSC) Pilot Global Language Self-Assessment).

I can understand all forms and styles of speech pertinent to my social and professional needs. This includes speech involving extensive and precise vocabulary, subtleties and nuances in standard dialects of the language, and technical discussion on professional topics within the range of my knowledge. I can understand language tailored to different audiences and purposes, including persuasion, representation, counseling, and negotiating. I can readily infer meanings and implications. I can easily understand all social conversations, radio broadcasts, and phone calls. I may experience some difficulty understanding speech heard under unfavorable conditions, such as through a poor quality loudspeaker or radio or in a noisy room.

The selected skill level description can serve to confirm or disconfirm the predicted language skill level obtained from the analysis of the responses to the Can-Do statements. The self-ranking on the ILR scale is then conjoined with the predicted score on the Can-Do scales to produce a composite score and the predicted language proficiency level.

If the candidate demonstrates proficiency at ILR level 3 or higher on the predicted language proficiency rating, he or she will undergo formal testing of language skills.
The self-assessments were used as an acceptance tool during the piloting of the NLSC program. The NLSC member’s language skills are further tested using a direct measure of language skills prior to any assignment. If the Can-Do self-assessments can be shown to predict scores on direct assessment of language proficiency, then the use of self-assessment could continue for some categories of languages in the permanent program and provide a significant cost reduction for the NLSC.

The validity of this kind of self-assessment instruments is well documented in the literature (ALTE, 2002; Clark, 1981; Davidson & Henning, 1985; Heilenman, 1990; LeBalanc & Painchaud, 1985; Roever & Powers, 2005; Ross, 1998; Shrauget & Osberg, 1981; Tannenbaum, Rosenfeld, Breyer, & Wilson, 1999; Wilson, 2000). In the context of the NLSC, Reed and Stansfield (2006) pointed out the potential advantages of using self-assessment techniques to meet the screening needs of the organization.

Research Design
In 2009, approximately 1,800 NLSC membership applicants across the U.S. filled out the self-assessments online. The self-assessments consist of two parts: the Can-Do statements and the global assessments. Both instruments assess four skills: listening, speaking, reading and writing. However, in the dataset, data corresponding to only three subsets of the Can-Do statements (listening, speaking, and reading) are available. The data for the writing subset of Can-Do statements (hereinafter Can-Dos) were not available for this analysis, since the Government is not currently using it. The following analyses were carried out on 323 admitted candidates to the NLSC, because only these candidates have been formally tested in the target language.

The 158 Can-Dos presented to candidates describe concrete tasks. Some of the tasks are general; and some relate to the work setting. The numbers of statements by domain are: 40 listening, 48 speaking, 32 reading, and 38 writing. The 158 Can-Dos take only a short amount of time to administer and complete. These statements are appropriate for a broad range of settings. For scoring purposes, each Can-Do statement was assigned a skill level between 1 and 5 on the ILR scale. Although there were about four Can-Do statements at each level, an examinee had to answer all statements affirmatively in order to be rated at that level.

For the global self-assessment, if the candidate’s proficiency is substantially better than one level but not consistently as good as the next higher level, they are directed to select the appropriate “plus” level. There is no
description of the plus levels on the self-assessment forms. With this format, the candidate can read and understand the scale quickly, and can make a fairly accurate self-placement at one of the nine points on the scale. For purposes of statistical analysis, the plus level is interpreted as 0.6 levels higher than the base level. For example, level 3+ is converted to a numeric score of 3.6.

**Characteristics of the Sample**

The released dataset consists of 323 observations, including 295 OPIs for which there are corresponding self-assessments. The 323 candidates were tested for their skills in a total of eight languages: 129 for Chinese-Mandarin (39.9%), two for Hausa (0.6%), 18 for Hindi (5.6%), 18 for Indonesian (5.6%), eight for Marshallese (2.5%), 88 for Russian (27.2%), 38 for Thai (11.8%), and 22 for Vietnamese (6.8%). Of these candidates, 228 (70.6%) candidates grew up abroad; 120 (37.2%) candidates received their high school education in the U.S. and 173 (53.6%) candidates attended primary college/university in the U.S. Regarding educational background, 196 (28.5%) candidates hold bachelor’s degrees, 29 (9%) hold master’s degrees, and six (1.9%) hold doctoral degrees.

**Results of the Self-assessments and OPIs**

The self-assessments required candidates to come up with subjective estimates as to how well they could perform various tasks related to overall competence in the target language based on whether they could perform various tasks. In general, as the literature shows, respondents tend to overestimate their abilities in performing discrete tasks, in part because making accurate judgments about one’s own ability relevant to something as complex as language is not easy. It is also not easy to internalize and apply an unfamiliar scale without training. Score distributions for the self-assessment variables are skewed, while score distributions for the OPI scores are closer to normal distribution. Table 1 below gives a breakdown of the self-assessments and the OPI.

The mean ratings of these NLSC candidates were 4.6 (4+) on Can-Do listening, 4.3 (a mid-range 4) on Can-Do speaking, and 4.5 (a high 4) on Can-Do reading. Their mean rating on the global self-assessments were 4.6 on Global listening, 4.2 on Global writing, 4.5 on Global reading, and 4.5 on Global speaking. These are high means, compared with a mean of 3.0 on the OPI test. At first glance, these numbers might seem unusually high. If we consider the fact these candidates are admitted members, these means do not seem as odd. The candidates generally perceived themselves to be quite competent in the target language.
### Table 1: Score Distributions of Self-assessments and the OPI

<table>
<thead>
<tr>
<th>ILR Level</th>
<th>OPI</th>
<th>N</th>
<th>%</th>
<th>1</th>
<th>1+</th>
<th>2</th>
<th>2+</th>
<th>3</th>
<th>3+</th>
<th>4</th>
<th>4+</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Can-Do: listening**

|        |     |   |    |    |    |    |    |    |    |    |    |    |
|        | 0   | 0 |    |    |    |    |    |    |    |    |    |    |
| %      | 0.30% |    | 2.50% | 8.40% | 14.20% | 32.50% | 30.00% | 3.10% | 0.30% |    |    |    |

**Can-Do: speaking**

|        |     |   |    |    |    |    |    |    |    |    |    |    |
|        | 0   | 0 |    |    |    |    |    |    |    |    |    |    |
| %      | 2.20% |    | 2.50% | 13.90% | 1.90% | 16.40% | 1.50% | 74.30% |    |    |    |    |

**Can-Do: reading**

|        |     |   |    |    |    |    |    |    |    |    |    |    |
|        | 0   | 1 |    |    |    |    |    |    |    |    |    |    |
| %      | 0.30% |    | 1.50% | 2.20% | 14.90% | 1.20% | 9.90% | 0.90% | 68.40% |    |    |    |

**Global: listening**

|        |     |   |    |    |    |    |    |    |    |    |    |    |
|        | 0   | 0 |    |    |    |    |    |    |    |    |    |    |
| %      | 0.90% |    | 5.90% | 9.30% | 7.70% | 9.00% | 9.60% | 65.30% |    |    |    |    |

**Global: writing**

|        |     |   |    |    |    |    |    |    |    |    |    |    |
|        | 4   | 3 |    |    |    |    |    |    |    |    |    |    |
| %      | 1.20% |    | 0.90% | 2.50% | 5.90% | 9.30% | 9.00% | 6.80% | 10.50% | 48.60% |    |    |

**Global: reading**

|        |     |   |    |    |    |    |    |    |    |    |    |    |
|        | 0   | 1 |    |    |    |    |    |    |    |    |    |    |
| %      | 0.30% |    | 0.90% | 3.40% | 6.80% | 7.10% | 8.40% | 9.90% | 62.20% |    |    |    |

**Global: speaking**

|        |     |   |    |    |    |    |    |    |    |    |    |    |
|        | 0   | 0 |    |    |    |    |    |    |    |    |    |    |
| %      | 0.30% |    | 2.80% | 6.20% | 7.40% | 10.20% | 9.30% | 62.80% |    |    |    |    |

**Research Design: Predictive Validity Study**

Predictive validity is the extent to which a score on a scale or test predicts scores on some other measure, i.e., the criterion. The most common use of predictive validity is in selecting students for university admission. A high correlation between the admissions test and the criterion variable (grades in the year following admission) indicates that the selection procedure worked well, while a low correlation signifies that something is wrong with the selection method. For NLSC self-assessments to have predictive validity, the correlation between the self-assessment scores and formal language proficiency tests needs to be statistically significant and of at least moderate effect size. It was decided to have the Oral Proficiency Interview (OPI) score serves as the criterion measure for evaluating the validity of the self-assessments. The OPI is a standardized procedure for eliciting and rating functional speaking proficiency. It is a criterion-referenced, direct interview. The OPI measures how well a person speaks a language by comparing their performance of specific language tasks,
not with other candidates, but with the criteria for each level of the ILR scale for speaking. The OPI is a carefully structured conversation between a certified interviewer and the candidate. The validity of OPI and the skill level descriptions for speaking have been documented in research studies (Dandonoli & Henning, 1990; Kenyon & Stansfield, 1992). In this study, the OPIs were administered by Language Testing International, under contract with the NLSC and were scored using the ILR scale.

Research Questions
This validity study addresses four primary research questions:

**Research Question 1**: Among Can-Dos and global self-assessments, which generated higher self-ratings and which generated lower self-ratings?

**Research Question 2**: Are there statistically significant correlations between self-assessment scores and the direct measures of language proficiency? What is the relationship among scores on the two types of self-assessment instruments?

**Research Question 3**: What is the effect size and practical utility of the correlations? How do the correlations compare with those found in predictive validity studies of high stakes tests such as the GRE and the SAT?

**Research Question 4**: What is the predictive validity of the Global self-assessments and the Can-Dos respectively in predicting an OPI score?

Results

**Research Question 1**
Among Can-Dos and Global self-assessments, which generated higher self-ratings and which generated lower self-ratings?

The first research question of this study concerns whether scores from two types of self-rating instruments are alike on the ILR scale. Statistically equivalent scores on the two instruments support the utility of these measures as comparable screening tools in the NLSC certification process. The Can-Do and the Global self-rating scale assess almost the same constructs. Construct refers to the knowledge, skill, or ability that is being tested. In other words, the two types of instruments are different operationalizations of the same concepts. Both the global assessments and the Can-Dos assess the candidates’ listening, speaking and reading ability.
This research question was addressed by examining the means on the Can-Do and the global self-rating scales (see Table 2). Paired sample t-tests have been conducted to compare the means of two comparable variables. For example, a paired-sample t-test was conducted to compare the self-rating of speaking scores obtained by Can-Dos and Global assessments. A paired sample t-tests computes the difference between the two variables for each case, and tests to see if the mean difference is significantly different from zero.

**Table 2: Means of Can-Dos vs. Global self-assessments**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>CAN-DO LISTENING</td>
<td>4.66</td>
<td>.66</td>
</tr>
<tr>
<td></td>
<td>GLOBAL LISTENING</td>
<td>4.61</td>
<td>.65</td>
</tr>
<tr>
<td>Pair 2</td>
<td>CAN-DO READING</td>
<td>4.47</td>
<td>.87</td>
</tr>
<tr>
<td></td>
<td>GLOBAL READING</td>
<td>4.52</td>
<td>.78</td>
</tr>
<tr>
<td>Pair 3</td>
<td>CAN-DO SPEAKING</td>
<td>4.39</td>
<td>.86</td>
</tr>
<tr>
<td></td>
<td>GLOBAL SPEAKING</td>
<td>4.56</td>
<td>.71</td>
</tr>
</tbody>
</table>

Table 3 gives the results of the paired sample t-tests, including the differences in means by skill. There is no significant difference in listening scores obtained by the Can-Dos (M=4.66, SD=.66) and the Global assessments (M=4.61, SD=.65); t (317) =1.78, p=.08. Similarly, there is no significant difference in reading scores obtained by the Can-Dos (M=4.47, SD=.87) and the Global assessments (M=4.52, SD=.78); t (318) =-1.67, p=0.10. However, there is a significant difference in the speaking scores obtained by the Can-Dos (M=4.39, SD=.86) and the Global assessments (M=4.56, SD=.71); t (318) =-0.16, p=0.00.

**Table 3: Paired sample t-tests**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Error Mean</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>CAN-DO: LISTENING-GLOBAL: LISTENING</td>
<td>0.05</td>
<td>0.03</td>
<td>1.78</td>
<td>317</td>
</tr>
<tr>
<td>Pair 2</td>
<td>CAN-DO: READING-GLOBAL: READING</td>
<td>-0.05</td>
<td>0.03</td>
<td>-1.67</td>
<td>318</td>
</tr>
<tr>
<td>Pair 3</td>
<td>CAN-DO: SPEAKING-GLOBAL: SPEAKING</td>
<td>-0.16</td>
<td>0.03</td>
<td>-5.102</td>
<td>318</td>
</tr>
</tbody>
</table>
These results suggest that for listening and reading, the type of self-assessment instrument used does not have an effect on the ILR-scale scores. However, for speaking, the type of self-assessment does have an effect on the scores. Specifically, our results suggest that the Global self-assessment gives a higher speaking score than the Can-Dos. The NLSC could confidently state that self-assessment scores are generally comparable across the self-assessment instruments that assess listening and reading skills.

**Research Question 2**

Are there statistically significant correlations between self-assessment scores and the direct measures of language proficiency? What is the relationship among scores on the two types of self-assessment instruments?

One of the most important goals of the self-assessment is to provide information on the candidates’ true ability in the target language. In this study, the official OPI score is treated as the indicator of the candidates’ true foreign language proficiency. The predictive validity is quantified by the correlation coefficient between the two sets of measurements obtained for the same sample—the measurement performed by the self-assessments and by the face-to-face OPI. Correlation coefficients, which fall in the range of -1.00 to +1.00, reflect the strength of the linear relationship between scores on different tests. A high positive correlation between two tests indicates that candidates who obtain a high score on one test are likely to obtain a high score on the other test.

The Pearson product moment correlation coefficients between the predictor variables and the criterion variables are shown in Table 4. Three of the constructs (listening, speaking, and reading) measured by the Can-Do statement scales, are compared to the four constructs (listening, writing, reading, and speaking) measured by the global assessment scales.

Analysis of the data indicates that the highest correlation is between the criterion (OPI scores) and the global assessment of listening ability ($r=0.54$). The global self-assessment of speaking ability also yields a high correlation (0.49) with the criterion measure (OPI). Next in order of magnitude is the global assessment of reading ability ($r=0.47$). Both reading and listening subsets of the Can-Dos yielded a correlation of 0.45 with the criterion variable.

At the time of this study, self-assessment data of non-admitted candidates was not available to the researchers. Upon receiving the data of the non-admitted applicants, we will carry out a correction of the correlation coefficients for the restriction of the range of the predictors.
Table 4: Pairwise Pearson correlation coefficients: intercorrelations of predictor variables and criterion variables

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OPI</td>
<td>Pearson Correlation</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>295</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAN-DO: LISTENING</td>
<td>Pearson Correlation</td>
<td>0.45**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>293</td>
<td>321</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAN-DO: SPEAKING</td>
<td>Pearson Correlation</td>
<td>0.41**</td>
<td>0.74**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>293</td>
<td>321</td>
<td>321</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAN-DO: READING</td>
<td>Pearson Correlation</td>
<td>0.45**</td>
<td>0.70**</td>
<td>0.80**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>293</td>
<td>321</td>
<td>321</td>
<td>321</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLOBAL: LISTENING</td>
<td>Pearson Correlation</td>
<td>0.54**</td>
<td>0.73**</td>
<td>0.71**</td>
<td>0.75**</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>291</td>
<td>318</td>
<td>318</td>
<td>318</td>
<td>319</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLOBAL: WRITING</td>
<td>Pearson Correlation</td>
<td>0.43**</td>
<td>0.61**</td>
<td>0.70**</td>
<td>0.79**</td>
<td>0.72**</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>280</td>
<td>305</td>
<td>305</td>
<td>305</td>
<td>305</td>
<td>306</td>
<td></td>
</tr>
<tr>
<td>GLOBAL: READING</td>
<td>Pearson Correlation</td>
<td>0.47**</td>
<td>0.70**</td>
<td>0.68**</td>
<td>0.80**</td>
<td>0.80**</td>
<td>0.85**</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>292</td>
<td>319</td>
<td>319</td>
<td>319</td>
<td>319</td>
<td>306</td>
<td>320</td>
</tr>
<tr>
<td>GLOBAL: SPEAKING</td>
<td>Pearson Correlation</td>
<td>0.49**</td>
<td>0.71**</td>
<td>0.76**</td>
<td>0.73**</td>
<td>0.89**</td>
<td>0.75**</td>
<td>0.78**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>292</td>
<td>319</td>
<td>319</td>
<td>319</td>
<td>319</td>
<td>306</td>
<td>320</td>
</tr>
</tbody>
</table>

** Correlation is significant at 0.01 level (2-tailed).

Equation (1) from Lord and Novick (1968, p. 143) will be utilized. This correction uses the ratio of the observed variance of the predictor in the sample to the observed variance of the predictor in the population. The equation is...
where \( s_x^2 \) is the observed variance in the sample and \( \sigma_x^2 \) is the observed variance in the population. When statistical corrections were made to counteract the dampening effects of selection, the correlations of the predictor and the criterion measure usually increase substantially (Raju & Brand, 2003).

The correlations between OPI scores and the self-assessment scores are all statistically significant. In other words, the self-assessments correlate well with a direct measure whose validity is widely accepted. Some of the self-assessments measure the same construct as the OPI, and some measure different constructs, but the four language skills are presumably interrelated.

The relationship among the scores on the two types of self-assessment instruments is addressed by examining patterns of correlations to look for evidence of convergent and discriminant validity. Convergent validity is the degree to which scores on a variable are similar to scores on other variables that are presumed to measure similar skills. Discriminant validity is the degree to which the scores on a variable are different from scores on other variables that it theoretically measures different skills. As expected, in Table 4 the highest correlations appear in the “same construct, same instrument” column, and lowest correlations tend to appear in the “different construct, different instrument” column. For example, the correlation between two reading assessments (Can-Do and global) is higher (0.80) than the correlations between the same two reading assessments and the global speaking self-assessment (0.68 and 0.73). This supports the convergent and discriminant validity of the self-assessments.

Research Question 3
What is the effect size and practical utility of the correlations? How do the correlations compare with those found in predictive validity studies of tests such as GRE and SAT?

In the context of NLSC, it is useful to know not only whether the correlations are statistically significant, but also the size of the observed relationship. In practical situations, effect size is helpful in decision-making,
since a highly significant relationship may be uninteresting if its effect size is small. Therefore, reporting effect size is considered good practice when presenting empirical research findings. By convention, correlation coefficients of 0.10, 0.30, and 0.50 are termed small, moderate, and large respectively in terms of their effect size (Cohen, 1988). In this study, the correlation coefficients between the criterion variable (OPIs) and the self-assessments were in the range of 0.41 to 0.54 (see Table 4), which indicates the effect sizes of the correlation coefficients are from moderate to large. In interpreting these results, it should be kept in mind that the correlation coefficient results are not corrected for the restriction of range.

Clark and Swinton (1979) reported a validity coefficient (0.48), which represented the correlation between a single oral proficient interview (OPI) rating and an examinee’s self-rating. Heilenman (1990) reported a correlation of 0.33 between course grades and undergraduate students’ self-assessments of their French language skills (grammar, vocabulary, accuracy, and fluency). Ross (1998) conducted a meta-analysis of studies dealing with self-assessment in second and foreign languages. For reading, he located 23 correlations, with an average r=0.61. In another summary of research on second-language self-assessments, Oscarson (1997) concluded the accuracy of the assessment depends to a considerable degree on the purpose of the assessment. Compared with correlation coefficients reported in other high risk test validity studies, such as GRE and SAT (Angronow & Studley, 2007; Geiser & Studley, 2002), the NLSC self-assessment instruments have relatively high correlation coefficients with the criterion measure. This provides evidence supporting the validity of the self-assessments.

**Research Question 4**

What is the predictive power of the global assessments and the Can-Dos respectively?

Although the correlations between individual predictor variables and the criterion variables are interesting, they do not provide the full picture of how well a combination of predictors predicts official language test scores. To evaluate the utility of a combination of predictors, subsets of the predictor variables were entered into regression equation.

In the first step of the modeling process, all seven predictors are entered in the regression equation. Table 5 displays results of the regression of OPI scores on seven self-assessment variables. Unfortunately, only one predictor, the listening global, is significant (alpha > 0.05), which seems theoretically
questionable. One possible reason for this outcome is the high correlation between the seven predictors. For example, the correlation between global speaking self-assessment and the listening Can-Dos is as high as 0.70. In general, the correlation between the predictors is in the range of 0.60 and 0.80. When predictors correlate highly among themselves, it increases the possibility of multicollinearity. The greater the multicollinearity, the greater the standard error and the smaller the t-statistics.

Table 5: Regression of OPI scores on Can-Dos and global assessments

<table>
<thead>
<tr>
<th>Model</th>
<th>Coefficients(a)</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>0.69</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAN-DO: LISTENING</td>
<td>0.09</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>CAN-DO: SPEAKING</td>
<td>-0.04</td>
<td>0.07</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>CAN-DO: READING</td>
<td>0.04</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>GLOBAL: LISTENING</td>
<td>0.35</td>
<td>0.11</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>GLOBAL: WRITING</td>
<td>0.01</td>
<td>0.06</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>GLOBAL: READING</td>
<td>0.05</td>
<td>0.09</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>GLOBAL: SPEAKING</td>
<td>0.01</td>
<td>0.10</td>
<td>0.01</td>
</tr>
</tbody>
</table>

a. Dependent Variable: OPI

In predictive validity studies, it is desirable to find large and statistically significant multiple correlation coefficients that account for a substantial amount of variation in the predictor. For example, Linn and Hastings (1984) found a multiple correlation of 0.46 for Law School Admission Test (LSAT) scores combined with undergraduate grades in predicting first-year law school grades, which indicates LSAT and undergraduate grades account for 21.16% of the variance in first-year law school grades. In addition to gauging the strength of the predictors considered simultaneously, multiple regression can also be used to evaluate the utility of each predictor. The regression slopes provide one indicator of the relationship between a predictor and the criterion. If the regression coefficient for a particular variable is statistically significant, it can be concluded
that the variable is important for accounting for variation in the criterion. However, it is difficult to isolate the relative contributions of each predictor from the regression coefficients.

**Table 6: Regression Models**

<table>
<thead>
<tr>
<th>PREDICTED VARIABLE</th>
<th>MODEL 1</th>
<th>MODEL 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPI</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PREDICTOR VARIABLES</th>
<th>MODEL 1</th>
<th>MODEL 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAN-DO: LISTENING</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>CAN-DO: SPEAKING</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>CAN-DO: READING</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>GLOBAL: LISTENING</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>GLOBAL: WRITING</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>GLOBAL: READING</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>GLOBAL: SPEAKING</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Since the Can-Dos and the global self-assessments are different operationalizations of the same constructs, it was decided to perform two sets of regression analyses. Model one had OPI scores regressed on the Can-Dos subsets (listening, speaking, and reading), and model two had OPI scores regressed on the global self-assessment subsets (listening, writing, reading, and speaking) (see Table 6).

**Regression Model One.** Three predictor variables included in the first regression model were: Can-Do listening, Can-Do speaking, and Can-Do reading. As can be seen from Table 7, two predictor variables (Can-Do listening
and Can-Do reading) were significant in their contribution to the prediction of the OPI scores. The three Can-Do statement variables altogether accounted for 23.9% (R-Square=0.239) of the variation in OPI scores.

**Regression Model Two.** The four predictor variables included in the second regression model were: global listening, global writing, global reading, and global speaking. These four variables altogether accounted for 30.5% (R-Square =0.305) of the variation in OPI scores. As can be seen from Table 8, the only significant predictor is global listening. The R-squared values from both model one and model two are slightly higher than R-squared (0.213) obtained in GMAT validity study, using GMAT scores to predict the first year GPA (Sireci & Talento, 2006).

**Table 7: Regression of OPI scores on Can-Dos by skill**

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>1.17</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>CAN-DO: LISTENING</td>
<td>0.23</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>CAN-DO: SPEAKING</td>
<td>-0.002</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>CAN-DO: READING</td>
<td>0.18</td>
<td>0.06</td>
</tr>
</tbody>
</table>

a. Dependent Variable: OPI

**Table 8: Regression of OPI scores on global assessments by skill**

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>0.79</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>GLOBAL: LISTENING</td>
<td>0.39</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>GLOBAL: WRITING</td>
<td>0.003</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>GLOBAL: READING</td>
<td>0.10</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>GLOBAL: SPEAKING</td>
<td>0.01</td>
<td>0.10</td>
</tr>
</tbody>
</table>
Conclusions and Discussion

The self-assessment scores are most useful to NLSC if they can provide information that allows NLSC to make important decisions concerning applicants’ screening. The Can- Dos and Global self-assessments give a clear profile of target-language strengths and weakness. NLSC hopes to use the self-assessment scores to identify individuals with an adequate level of target language competency to perform their jobs. One major finding of this study was that there were no differences in the listening and reading global assessment ratings and the ratings produced by the Can-Dos, while global assessment of speaking produced higher speaking proficiency ratings than the Can-Dos did.

Some researchers have argued that people tend to overestimate their language skills on self-assessments (Davidson & Henning, 1985). While this study also found that to be the case, the results provide substantial empirical evidence that NLSC applicants can make reasonably effective judgments about their own language skills. As anticipated, the highly selective nature of admission to membership produced a limited range of self-assessment scores. Even with the limited range, both types of Reading/Listening/Speaking self-assessment scores (Can-Do and Global) exhibited significant positive relationships with OPI scores. The effect size of the correlation coefficients is between moderate and large. Compared with correlation coefficients reported in other high stakes test validity studies, such as those of the GRE, SAT and GMAT, the NLSC self-assessment instruments have relatively high correlation coefficients with the criterion measure. This provides evidence supporting the validity of the self-assessments.

In regression analysis, two distinct regression models were fit to the data. Model one had OPI scores regressed on the Can-Do statement skill subsets (listening, speaking, and reading), and model two had OPI scores regressed on the global self-assessment (listening, speaking, reading, and writing). The R-squared values from both model one (R-square=0.239) and model two (R-square=0.305) are slightly higher than R-squared (0.213) obtained in GMAT validity study, using GMAT scores to predict the first year grade point average (Sireci & Talento, 2006).

Overall, the implications of this study are that the Can-Dos and the global self-assessments are reasonably valid measures of language skills in NLSC target languages, and should remain as part of the NLSC screening process.

It should be noted that this validity study is based on a sample of applicants in eight languages, which may not be representative of the whole
population of candidates for the NLSC. Thus, decision-makers should use caution when applying these results to applicants with other foreign languages.

References


Examining the Validity of the 2010 Prototype AP Russian Exam through a College Comparability Study¹

Camelot Marshall

Introduction

Since its inception twelve years ago, the Prototype AP® Russian Language and Culture Examination has developed into an assessment instrument that has increasingly become the culminating focus and a hallmark of high school Russian language study in select schools across the United States.² Even more so, through the years of field-testing, piloting, and making the tests operational, the design, content, development, administration, and analyses of the exam have evolved into the model for American Councils’ online assessments of language proficiency not only for Russian, but also for Flagship programs. These tests are already being developed in Arabic, Chinese, Persian, Russian, Swahili, and Yoruba.³

The Prototype AP® Russian Program provides high school students with advanced Russian-language instruction, comparable to that of a college-level language class. The course, instead of being curriculum-based, reflects intermediate level proficiency skills expected and described in the course description of the Prototype AP® Russian Course and Exam. The target proficiency level of the exam for the intermediate level was determined from analyses of language performance of students with four and six semesters of college Russian language study and no Russian language study prior to

¹ The author would like to extend a special thank-you to the Russian Language Specialists and Test Development Members, the American Councils Language Assessment Support System (ACLASS), and to Saodat I. Bazarova for a concise analysis of the data.
² Throughout its development and administrations, the Prototype AP® Russian Exam has held different names reflecting the various stages of existence. Since December 2006, the exam has been known as the Prototype AP® Russian Language and Culture Examination. Initial funding for the project was provided by the U.S. Government through International Research Study Title VI and the National Security Education Program (NSEP).
starting in college. Over the years of its administration, the number of test takers continues to grow, as well as the number of colleges and universities that grant placement and/or credit based on results on the Prototype AP® Russian Exam. In order to assess how well the Prototype AP® Russian Program and Exam prepare students to continue Russian language study at the college level and to assure that the Prototype AP® final grade placement suggestions meet the current performance outcomes of university-level Russian language programs, a short comparability study was conducted with college students in summer 2010. That study is the focus of this report.

**Overview of College Board’s AP® Foreign Language Comparability Studies**

Every five to seven years, the College Board conducts college-grade comparability studies to assure that the final AP grades earned by test takers reflect the standards and practices of colleges and universities and the expectations of the parallel course they offer. The methodology of the College Board’s comparability study for the foreign language exams starts with the administration of a subtest of the exam, which includes the multiple-choice sections and the free response section, to students with at least two years of college language study. For the newer language exams, such as Chinese or Japanese, students are expected to have completed approximately 300 hours of college-level foreign language study. In addition to the administration of a subtest, the college professor is asked to provide a grade for the subtest taken by each student and a course grade for each student. Course grades are requested, because they serve as the criterion used for determining the cut scores on the exams to earn college credit. To motivate the students to take the subtest seriously, they were told that their performance on the exam will be part of a final course grade. Although the multiple-choice subtests were graded like full AP exams, they were not full-length exams; thus, the mean scores and composite scores were calculated from equipercentile scaling of the subtest scores to the scores of the AP examinees. The free-response section was graded

---


by AP readers in the subject area, using the same scoring rubric as the AP examinees. The results on each section were combined to create a composite score, which was used in combination with course grades to determine cut scores for the different AP final grade levels 1-5.\(^7\)

For the college comparability study, data were categorized into three groups: 1) all college students; 2) all AP students, and 3) a standard group of AP students. The standard group of AP students were those who responded “no” to both background questions on their foreign language exam: 1) “Have you lived or studied for one month or more in a country where the language of the exam you are now taking is spoken?; 2) Do you regularly speak or hear the language at home?”\(^8\) Based on the results of the comparability studies, the grading standards are adjusted in order to maintain the performance levels expected of AP students. That is, when comparing the results of the college students with those of the high school AP students, the College Board expects that high school AP students’ composite mean scores on the exams for each AP final grade should be higher than the average course grade of college students in the corresponding college level.

While these comparability studies continue to validate AP examinations today, they remain subjective to the extent that they use final course grades as determined by grading rubrics, which vary from professor to professor. The compatibility study presented here looks at student performance on the multiple-choice sections of the Prototype AP® Russian Exam, which are proficiency-based, using the American Council on the Teaching of Foreign Languages (ACTFL) Proficiency Guidelines, generally accepted and applied as a measurement of functional language.

**Test Design, Content, and Measurement**

The testing objective of the Prototype AP® Russian Language and Culture Exam is to provide a summative, proficiency-based measurement of high school students’ language proficiency in the four skills: reading, listening, writing, and speaking. Students taking the Prototype AP® Russian Exam have completed at least two to three years of high school Russian language study (or the equivalent) with the final year of study, typically the fourth year, an advanced-level course, comparable to that of a college Russian language course. Heritage speakers may present a different course of language study, depending on the extent to which

---

\(^7\) Bischof *et al.*, “College Comparability Studies,” 618-619.

the language is used and/or taught at home or outside a regular school setting.

Test specifications reflect “The College Board World Languages Framework” and the proficiency guidelines developed by the ACTFL. Test items are based on the tasks, skill domains, and performance indicators published in *Standards for Foreign Language Learning in the 21st Century*. The Prototype AP® Russian Examination consists of four subtests, all of which are aimed at the direct assessment of functional proficiency within interpersonal, interpretive, and presentational modes of communication. The key to the test design and content is that it is proficiency-based, so there is not a prescribed program of study or target language units to master. Exam questions range in difficulty from the Novice level to Advanced, with the majority of questions focusing on the intermediate level of proficiency. The intermediate level of proficiency was set as the target level for the AP exam, following analyses of student performance data in the American Councils or Teachers of Russian (ACTR) database. College students who began their Russian language study in college and who had two or three years of study performed at the Intermediate-Mid level of proficiency. The content of the exam, like the content of the course, is not prescribed but rather assesses the student’s ability to perform in all communicative aspects of the language (reading, listening, writing, and speaking). The exam includes authentic stimulus materials that are of interest to high school students and that could be encountered in real-life situations.

The Prototype AP® Russian Exam is approximately three hours, and includes the online administration of three sections—reading comprehension, listening comprehension, and integrated written communication—and a separate oral proficiency interview (OPI), which is administered by telephone. The reading and listening comprehension sections are multiple-choice sections. All of the stimulus material is in Russian, while instructions and multiple-choice questions are in English. Test questions are posed in English, so as to assess students’ comprehension of the reading or listening stimulus materials, rather than assess the added task of understanding the comprehension question. Reading texts may be brief articles, notes, advertisements, or various other types of authentic Russian sources. Likewise, audio clips may be short conversations or interviews, announcements, and advertisements; all of which are self-contained messages. Each text or audio may be followed by one to three multiple-choice questions.

---


items. Items are developed by experienced item writers and undergo a lengthy review process before they are finalized and ready for an exam form. Items are regularly reviewed for proficiency level, topic appropriateness, equity, and fairness, following *ETS Standards for Quality and Fairness*, 2002.

The reading and listening comprehension sections of the Prototype AP® Russian Exam are assembled, adhering to test development processes to ensure quality control and statistical reliability. The reading comprehension section consisted of 34 items, ranging in proficiency levels from Novice-Mid through Advanced, and the listening comprehension section consisted of 36 items, also ranging in proficiency levels from Novice-Mid through Advanced. Prior to appearing on a test form, all items are calibrated by piloting among college students studying Russian. Items selected for inclusion on an exam form must meet the selection criteria of item difficulty between 0.30 and 0.90 and point biserial (discrimination index) of no less than 0.25. Each Prototype AP® Russian Exam form is equated across administrations from year to year, using a set of common items, which represent the content and item characteristics of the earlier form. Common items with a previous test form make up at least 40% of the items in the reading and listening comprehension sections.

A development committee composed of Russian language specialists at the high school and college levels reviews the exam forms and approves them. Following the exam administration, the equating provides raw score-to-raw score conversions for the old and new forms. This procedure produces raw score equivalents on the new form to the anchor points set on the original form based on the standard setting study. American Councils assembled a panel of experts to determine the cut scores that mark the boundaries between proficiency levels for the reading comprehension and listening comprehension sections for the Prototype AP® Russian Exam.¹¹ Using the cut scores from the standard setting, the raw scores of the multiple-choice reading and listening comprehension exams can be placed on the ACTFL scale.

The integrated written communication, composed of a dialogue completion and integrated tasks, requires that test takers produce language within a meaningful, real-life context. For the dialogue completion, test takers supply missing lines of a dialogue, the topic of which is highly plausible in a real-life situation. For the integrated tasks, students respond in writing to reading and listening prompts to create a realistic scenario of incorporating all three language skills. The oral proficiency interview is administered by a certified ACTFL tester and is personalized to each test taker, building on the test

takers’ response to the opening question: Tell me about yourself (Расскажите о себе). Although ACTFL certified testers rate the students’ spoken and written performance in the language, the proficiency ratings assigned should not be considered formal ACTFL ratings, because of the nature of the exam.

Composite scores are provided on the College Board’s AP Grade scale of one to five, and since the total grade is linked to proficiency, sub-scores for each of the four language modalities are also reported in proficiency levels. The sub-score report provides a brief summary of the exam specifications and proficiency orientation of the exam, as well as a brief descriptive summary of the proficiency level achieved on each section of the exam. The Prototype AP® Russian Language and Culture Examination is designed to provide a set of measures of functional proficiency in Russian for use as a predictive assessment for continued language study, and as a placement tool by American colleges and universities for entering freshmen with prior study of Russian. In order to maintain the comparability between the AP final grade, section proficiency levels, and performance in a comparable college course, the reading and listening multiple-choice sections of the Prototype AP® Russian Exam were administered online to college students prior to their summer Russian language study.

For thirty-five years, American Councils for International Education: ACTR/ACCELS has administered and overseen summer, semester-long, and nine-month immersion programs in Russia. The database population in this study includes participants on summer study-abroad programs administered through American Councils and on an intensive stateside summer language program. The students represent a broad demographic of colleges and universities, both large and small, and public and private. Before taking the online exam, students answered a short background questionnaire focusing on Russian language study and use.

**Student Population and General Descriptives**
The entire Prototype AP® Russian Exam was administered online to 266 high school students across the United States, after completing an advanced program of high school Russian language study. Within the high school population, 47% were male and 53% were female. A total of 52% reported they were non-heritage speakers of Russian, and 48% reported they were heritage speakers. Whereas self-reports of factual information are reliable, students’ responses to the question of whether or not they were heritage speakers of Russian were double-checked with their teachers’ reports of heritage status, and further checked with additional questions about language spoken in the home and parents’ native
language. A total of 73% reported that they did not study Russian prior to high school, while 27% reported that they studied Russian prior to their high school Russian language study. The years of reported high school Russian language study ranged from zero to eight years. While high school Russian language programs typically offer four years of study, it is likely that for some of the reported years of language study, students combined their middle and high school years of Russian language study, or attended a school that did not differentiate between middle and high school language study. Within the high school population, 3% of the students were freshmen, 11% were sophomores, 33% were juniors, and 53% were seniors. It is likely that students taking the Prototype AP® Russian Exam who have had fewer than four years of high school language study do so because either they are heritage speakers, they started Russian language study before entering high school, or their high school offers fewer than four years of Russian language study.

Item analyses of the exam data show the Prototype AP® Russian Exam administered to the high school student to have good psychometric characteristics. Looking at the test taker data, the reading comprehension section shows an internal consistency reliability of .85, and the listening comprehension shows a reliability of .88. Reliability provides a level of statistical certainty that repeated administrations of the exam sections would yield similar results of a test-taker’s abilities. The overall group mean score for reading comprehension (M = 25.73, SD = 6.187) was similar to the mean score for listening comprehension (M = 25.23, SD = 7.676). A review of graphical representation of the distribution of scores show a relatively normal distribution. The skewness and kurtosis values are within acceptable range. Looking at the means of just the non-heritage test takers, the reading comprehension mean is lower (M = 21.77, SD = 5.420), and the mean for the listening comprehension is slightly lower (M = 18.78, SD = 5.762). The differences in means, particularly that of the listening comprehension, may be attributed to the heritage students being more proficient in their listening skills.

For the comparability study, only the multiple-choice reading and listening comprehension sections of the Prototype AP® Russian Exam were administered to 133 college students at the beginning of their summer Russian language study program. A total of 83% of the college students were beginning a study abroad summer language program, while 17% were beginning a stateside summer language program. Among the college students, 42% were male and 58% were female. A total of 98% reported they were non-heritage speakers of Russian, and 2% reported they were heritage speakers of Russian. Within this
population, 83% reported that they did not study Russian prior to college, and 17% indicated that they studied Russian before college. The college students reported between zero and twelve semesters of Russian language study. The academic levels of the college students consisted of 1% freshmen, 20% sophomores, 32% juniors, 20% seniors, and 27% graduate students and other.

Item analyses of the reading and listening sections administered to the college students show good psychometric characteristics. The reading section has a reliability of .87, and the listening comprehension shows a reliability of .79. Among the college students, the reading comprehension mean (M = 25.91, SD = 6.546) is slightly higher than the mean of the listening comprehension section (M = 21.66, SD = 5.762). Tests showed a relatively normal distribution of scores. Removing graduate and other students from the analyses, as well as those students who reported they were heritage speakers, reveals little change in the reading mean (N = 95, M = 25.13, SD = 6.622) or the listening mean (N = 95, M = 21.09, SD = 5.670).

Comparison of Means
The standard for the Prototype AP® Russian Exam was set based on measured proficiency levels of university-level students of Russian at different points in their language study, and the target proficiency range is the intermediate range. Entering freshmen who had taken the Prototype AP® Russian Exam would have their results aligned with the typical performance levels of university students at the third, fourth, and fifth semester levels of study. In order to check on whether or not the empirically set target levels for placement correspond to high school students’ performance on the Prototype AP® Exam, a sample each of both high school and college level students was selected for a comparison of their means on the reading and listening comprehension sections.

High school students who were non-heritage speakers of Russian and who reported at least four years of high school Russian language study were selected for the comparison of means. The selected high school sample consisted of 108 students. Since 87% of the non-heritage high school students received final overall AP results with suggested placement, as determined by analyses of student language performance prior to study abroad into either second semester of first-year Russian or into second-year Russian (first or second semester), the selected college student sample for comparison consisted of non-heritage undergraduate students with either one or two years of Russian language study, and who began their Russian language study at the university level. The selected university sample consisted of 57 students, which is a
sufficient sample size for providing reliable statistical results; typically, 30 is accepted as providing reliable results.

Using an independent t-test, the means for the reading and listening comprehension sections were compared for the two groups. Frequency tables of the raw scores for both groups showed the data to be evenly distributed. Normality tests showed a relatively normal distribution, and skewness and kurtosis were within acceptable range. A comparison of the reading and listening comprehension means shows higher means for the undergraduates reading comprehension (M = 23.93, SD = 6.769) and listening comprehension (M = 20.07, SD = 5.480) with one or two years of college Russian than for the high school group’s reading comprehension (M = 22.16, SD = 5.197) and listening comprehension (M = 18.91, SD = 4.179). Given the unequal sample sizes, Levene’s test for homogeneity of variances was examined. The results of this analysis revealed that there was no difference between the groups on the two sections of the exam: Reading comprehension \([t (-1.726) = 91.630, p < .088]\) and for Listening comprehension \([t (-1.522) = 163, p < .130]\). That is, the high school student performance on the reading and listening sections of the Prototype AP® Russian Exam is comparable to the college students’ performance on the reading and listening comprehension sections, and confirms the target university levels for placement. While the comparison of means shows there is no difference between the groups, a look at the proficiency level distributions, based on the set cut scores, shows variation in intermediate levels of proficiency that are also comparable for the target placement levels into college Russian language programs (see Table 1 and Table 2).

<table>
<thead>
<tr>
<th>Table 1 High School Non-heritage Students (N = 108)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proficiency Level</strong></td>
</tr>
<tr>
<td>NM</td>
</tr>
<tr>
<td>NH</td>
</tr>
<tr>
<td>IL</td>
</tr>
<tr>
<td>IM</td>
</tr>
<tr>
<td>IH</td>
</tr>
<tr>
<td>IH*</td>
</tr>
</tbody>
</table>

IH* Signifies that students exhibit proficiency higher than the Intermediate High level, but the test specifications do not allow for a specific rating to be established.
Table 2: Undergraduate Students (N = 57)

<table>
<thead>
<tr>
<th>Proficiency Level</th>
<th>Reading Comprehension</th>
<th>Listening Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM</td>
<td>2%</td>
<td>5%</td>
</tr>
<tr>
<td>NH</td>
<td>21%</td>
<td>5%</td>
</tr>
<tr>
<td>IL</td>
<td>26%</td>
<td>54%</td>
</tr>
<tr>
<td>IM</td>
<td>7%</td>
<td>2%</td>
</tr>
<tr>
<td>IH</td>
<td>16%</td>
<td>30%</td>
</tr>
<tr>
<td>IH*</td>
<td>28%</td>
<td>4%</td>
</tr>
</tbody>
</table>

IH* Signifies that students exhibit proficiency higher than the Intermediate High level, but the test specifications do not allow for a specific rating to be established.

Discussion

A general comparison of the mean scores of all AP high school test takers and the college test takers shows comparable mean scores, which suggests that the groups perform similarly on the multiple-choice reading and listening comprehension sections. Under further review, however, the close mean scores are most likely due to the large number of heritage speakers in the high school group. When this sub-group is removed from the analyses, the reading comprehension mean score for the total high school group decreases, and the listening comprehension mean score decreases significantly more than the reading comprehension mean score. This is not surprising, in that listening comprehension skills of heritage speakers are stronger than those of the non-heritage speakers, and reading comprehension of the non-heritage speakers is closer to the reading comprehension of heritage speakers. Since there were only two heritage speakers in the college student population, the mean scores for the reading and listening comprehension sections show negligible change, when the heritage speakers are removed.

Looking more closely at student performance for placement target levels, the analyses demonstrate that while the college level test takers, who started their Russian language study in college and have at least one or two years of college level study, have a slightly higher mean score for the reading and listening comprehension sections than the non-heritage high school population. The t-test, accounting for unequal sample size, indicates that there is no difference in the two groups’ performance. These findings purport to validate the Prototype AP® Russian Exam as a measurement instrument for reading and listening proficiency, as well as for target placement into a college-level Russian
language course.

This comparability study is slightly different than the methodology used for the College Board’s comparability studies, primarily in that the college students were given the full-length reading and listening comprehension sections of the Prototype AP® Russian Exam, and the results on the sections were based solely on students’ performance on the multiple-choice sections. In the College Board study, college students’ results on a shortened test form are estimated to scores on a full AP Foreign Language Exam. In order to obtain a composite score, the estimated scores are combined with final course grades as assigned by the participating professor or foreign language instructor. The added factor of the final course grade is based on students’ performance in the course, but it is also subjective: in order to determine the final course grade, the instructor created a grading rubric in line with course expectations. With this added variable into the comparability study, the results are not fully on student performance in the language, but also on the language instructors’ expectations for student performance in the course. The comparability study of the Prototype AP® Russian Exam presented here remains objective, in that it is based on student performance on the exam form and objective in the expectations for language performance, which are based on the exam’s target proficiency levels.

Turning to students’ proficiency levels for the reading and listening comprehension sections, the distribution for the high school group and the college-level group are similar. The high school students primarily perform at intermediate-low and -mid levels of proficiency, the target levels of the exam. The college students also perform at intermediate levels of proficiency, and more college students perform at the intermediate-high or above levels than the high school students. This can be accounted for by the inclusion of students with three or more years of college-level Russian language study. The proficiency levels assigned to student performance on the exam sections result from the application of cut scores, as decided during a standard setting exercise to the students’ raw scores on the exam. Because proficiency ratings do not easily lend to statistical analyses, student raw scores were used for the comparison of means.

Tangential to the application of the results on the Prototype AP® Russian Exam for entering freshmen’s college credit and/or placement is the application of the exam design, construct, administration, and scoring to Russian language proficiency tests for other purposes. Currently, some study-abroad programs require that students’ language ability be assessed before and after formal study in the target culture. Thus, the model of the Prototype AP® Russian Exam can be applied to such tests as needed by the study abroad programs. Likewise, with
Examining the Validity of the 2010 Prototype AP Russian Exam

Camelot Marshall

increased expectations for evidence of language learning, college and university departments seek their own entry and exit tests for their students. Although each Russian department has its own core curriculum expectations for language learning, a Russian language test that is based on proficiency, albeit perhaps not targeting all of the programs’ expectations and outcomes, offers a uniform platform for measuring and reporting student language proficiency. With a consistency in measurement and reporting of language proficiency, those in the field of Russian language teaching will have a better means to assess and understand language performance, which in turn will inform curriculum development and language training.

Conclusion
This study attests to the reliability of the online Prototype AP® Russian Exam as an instrument to measure language proficiency in terms of ACTFL Proficiency ratings for Reading and Listening, in addition to providing a final AP Grade on the College Board’s 1-5 scale. Student results can be used for college placement and/or credit, depending on college or university AP grade policies. The exam content and administration to high school and university-level students showed relatively high reliabilities for both multiple-choice reading and listening sections of the exam. This is significant in that the same exam form, given in similar conditions and to a comparable student population, would yield similar results.

This study further confirms that high school students who have taken the Prototype AP® Russian Exam demonstrate comparable mean scores on the reading and listening comprehension sections to mean scores of the university-level students. These results provide evidence to the Russian language field regarding the reliability, validity, and performance outcomes of the exam, and it should be accepted with confidence that the Prototype AP® Russian Exam indeed captures students’ language performance. As such, placement decisions of entering freshmen can therefore be better informed as to what the results and score reports mean, and how they match up with the expected level of college student performance and the appropriate level of language instruction at the receiving institution. The increased use of the exam for placement purposes as colleges and universities gain increasing institutional experience with the AP results, especially given the additional information of proficiency subscores, will hopefully lead to an increasing institutional use of the exams for college credit, and in some cases, even help freshmen gain sophomore standing.
References


The American Council of Teachers of Russian
INVITES YOUR MEMBERSHIP ENROLLMENT FOR THE CURRENT YEAR AND
SOLICITS YOUR SUBSCRIPTION TO THE

Russian Language Journal

CALENDAR-YEAR DUES FOR INDIVIDUALS (ALL EXCEPT STUDENT MEMBERSHIPS INCLUDE RLJ)

- $35 FOR ASSOCIATE AND FULL PROFESSORS AND NON-ACADEMICS
- $30 FOR ASSISTANT PROFESSORS, LECTURERS, PRE-COLLEGE TEACHERS, AND INDEPENDENT SCHOLARS
- $25 FOR RETIRED PERSONS  
- $10 FOR STUDENTS
- $500 FOR LIFE MEMBERSHIP

- $10 NEWSLETTER POSTAGE SURCHARGE FOR ADDRESSES OUTSIDE OF NORTH AMERICA

SUBSCRIPTION TO RUSSIAN LANGUAGE JOURNAL (DOES NOT INCLUDE MEMBERSHIP)

- $10 FOR CURRENT STUDENT MEMBERS WITHIN NORTH AMERICA
- $30 FOR INDIVIDUALS WITHIN NORTH AMERICA
- $50 FOR INSTITUTIONS WITHIN NORTH AMERICA
- $40 FOR INDIVIDUALS OUTSIDE NORTH AMERICA
- $65 FOR INSTITUTIONS/LIBRARIES OUTSIDE NORTH AMERICA

PLEASE CHECK:  
- NEW MEMBERSHIP 
- RENEWAL MEMBERSHIP 
- NEW SUBSCRIPTION 
- RENEWAL SUBSCRIPTION

FULL NAME: ............................................................................................................

INSTITUTION: .............................................................................................................

SHIPPING ADDRESS  
..........................................................  BILLING ADDRESS  
(IF DIFFERENT)  
...........................................................................................

CITY:  
..........................................................  CITY:  
...........................................................................................


COUNTRY:  ..................................................  COUNTRY:  ..................................................

TOTAL DUE: ....................... (PAYMENT DUE IN ADVANCE IN US DOLLARS ONLY)

- CHECK ENCLOSED  
- SEND INVOICE  
- VISA  
- AMEX  
- DISCOVER  
- MASTER CARD

CREDIT CARD NO. ________________________________  EXPIRATION DATE (MM/YY) .................

Return to ACTR Member Services, 3109 Yale Blvd., St. Charles, MO 63301-0462
Editorial Policy

Scope. RLJ publishes scholarly articles related to the study and teaching of Russian language and culture, as well as comparative and interdisciplinary research in Russian language, culture, and the acquisition of Russian as a second language. RLJ also publishes evaluations of teaching/learning materials and book reviews. All articles submitted to RLJ should include adequate documentation, providing credit to primary sources and relevance to current research. Pedagogical articles reporting on experimental research results and/or empirically-based evaluations are also encouraged and invited.

External Review Process. Journal submissions should be crafted without revealing the author’s identity in the body of the work or the bibliographic references. Each submission which meets the overall eligibility criteria for RLJ publication will be reviewed anonymously by at least two external evaluators, who make the final determination (using specific criteria) as to whether a submission is accepted or rejected.

Recommended Length. The limit on length for each article submission is seven thousand words. Authors interested in submitting articles are encouraged to contact the Editor to discuss article length and subject matter prior to submission.

Language. RLJ is a refereed bilingual annual publication. Contributions should be written in either English or Russian.

Citations and References. Citations and references should be kept to a minimum. A list of references follows each manuscript, alphabetized by the last name of the authors; citations are linked to this list. The Chicago Manual of Style, 15th edition, 2003, provides guidelines on bibliographic forms. In order to assist in the editing process, authors should provide all necessary bibliographic information at the time of submission.

Manuscript Preparation. RLJ will accept manuscripts that are neatly typed with one-inch margins on all sides, double-spaced (including notes, block quotes, and references), and follow the body of the paper. RLJ observes The Chicago Manual of Style and the simplified U.S. Library of Congress system of transliteration from Cyrillic, when necessary. RLJ recommends that potential contributors consult the SEEJ Style Sheet at http://aatseel.org.

Submission. Manuscripts should be submitted electronically to the editor at the e-mail address rlj@actr.org as an MS Word attachment. All manuscripts will be acknowledged; authors are notified within six months whether/when their manuscripts will be published. It is the author’s responsibility to obtain and document permission to include copyrighted illustrations in the work. Any illustrations (including tables and charts) must be submitted in camera-ready format. If this is not possible, the costs involved for any graphics preparation will be the author’s responsibility. The author will cover the costs associated with any requested changes made to submissions already typeset.

Originality. RLJ considers only original work for publication. In submitting an article to RLJ, it is understood that neither the manuscript nor any substantially similar version of the work is currently being considered or has been published elsewhere.

Reprints. Two off prints of the published manuscript and one complete copy of the RLJ issue in which the work appears is provided to the author of articles, notes, and review articles.

Back Issues. Back issues of Volumes I to LIV (Nos. 1-179) are sold out and no longer available; microfilm copies of these back issues may be purchased from: XEROX UNIVERSITY MICROFILMS, 300 North Zeeb Road, Ann Arbor, Michigan 48106. All subsequent issues can be obtained from the editor's office.

Advertising. Please contact the RLJ editorial office for current information regarding advertising deadlines, technical requirements, and rates.
The American Council of Teachers of Russian (ACTR) is a national professional membership association of teachers and specialists in the field of Russian language and culture. Founded in 1974, ACTR is dedicated to strengthening the study and teaching of Russian through the design and implementation of academic exchange, professional training, distance learning, curriculum and test development, research and evaluation, advocacy, and programs of public support and recognition for the Russian field. ACTR is a division of American Councils for International Education.

American Councils for International Education:
ACTR/ACCELS
1828 L Street, NW
Suite 1200
Washington, DC 20036
Tel: (202) 833-7522
Fax: (202) 822-1689
www.americancouncils.org
https://membership.actr.org