To Advanced Proficiency and Beyond

THEORY AND METHODS FOR DEVELOPING SUPERIOR SECOND LANGUAGE ABILITY

EDITED BY TONY BROWN AND JENNIFER BOWN

To Advanced Proficiency and Beyond

This page intentionally left blank

TO ADVANCED PROFICIENCY AND BEYOND

Theory and Methods for Developing Superior Second Language Ability

Tony Brown and Jennifer Bown, Editors

GEORGETOWN UNIVERSITY PRESS
Washington, D.C.

Georgetown University Press, Washington, D.C. www.press.georgetown.edu

Library of Congress Cataloging-in-Publication Data
To advanced proficiency and beyond : theory and methods for developing superior second language ability / Tony Brown and Jennifer Bown, editors.
pages cm
Includes bibliographical references and index.
Summary: Due to the pressures of globalization, American society increasingly needs citizens who can carry out Superior level functions in languages other than English. Instructors, researchers, and students of second language acquisition seek scholarly resources to help satisfy this demand. In this volume, leading experts in second language acquisition and language planning supply cutting-edge research on working memory and cognition and empirical studies of effective teaching. The theoretical and empirical work in these pages is complemented by descriptions of successful pedagogical practices that take students from the Advanced to the Superior levels and beyond. With examples ranging across a number of languages, including Russian, Chinese, and Arabic, the volume will serve a broad audience. This practical handbook will help seasoned instructors improve outcomes, while it can also be used for training new instructors in methods courses.
ISBN 978-1-62616-173-3 (alk. paper) -- ISBN 978-1-62616-174-0
1. Second language acquisition. I. Brown, Tony (Newel Anthony), editor II. Bown, Jennifer, editor
P118.2.T6 2015
418.0071--dc23
2014023184

This book is printed on acid-free paper meeting the requirements of the American National Standard for Permanence in Paper for Printed Library Materials.

15 14 9 8 7 6 5 4 3 2
First printing

Printed in the United States of America

Cover design by Debra Naylor, Naylor Design, Inc.

Contents

Acknowledgments

WE WISH TO THANK Ray Clifford, Director of the Center for Language Studies, and John Rosenberg, Dean of the College of Humanities, both of Brigham Young University, for sponsoring a symposium in 2012 that precipitated this volume. Agnes Welch of the Center for Language Studies provided invaluable behind-the-scenes assistance in the way of planning and carrying out the symposium. We also wish to thank Mel Thorne of Faculty Editing Services at Brigham Young University for providing critical input and direction in the final stages of format editing.

The contributors to this volume made editing it a true pleasure. We express our sincere thanks to each of them for their thoughtful input, graceful acceptance of feedback, and their careful attention to detail.

We would like to thank the editorial staff at Georgetown University Press for their insightful feedback and attention to detail. In particular, we are indebted to David Nicholls, acquisitions editor, Deborah Weiner, former editorial and production manager, and Glenn Saltzman, editorial, design, and production manager.

Outside readers provided important organizational, compositional, and stylistic comments that enhanced the volume as a whole. We extend our sincere thanks to them for their careful reading of the manuscript and quality feedback.

Lastly, we wish to thank our spouses, Emily Brown and Tom Bown, for their support and encouragement in seeing this project through to completion.

This page intentionally left blank

List of Abbreviations

ACIE	American Councils for International Education
ACTFL	American Council on the Teaching of Foreign Languages
ACTR	American Council of Teachers of Russian
AH	Advanced High
AL	Advanced Low
AL2	Advanced language ability
ALC	Advanced language capability
ALPPS	Advanced Language Performance Portfolio System
ALT	Adaptive Listening Test
AM	Advanced Mid
AOF	Arabic Overseas Flagship
ART	Adaptive Reading Test
BA	Bachelor of Arts
BYU	Brigham Young University
CAL	Center for Applied Linguistics
CALL	Center for the Advancement of Language Learning
CASA	Center for Arabic Study Abroad
CATRC	Computer Adaptive Test for Reading Chinese
CCALT	Chinese Computer Adaptive Listening Test
CDLC	Coalition of Distinguished Language Centers
CE	Continuing Education
CLIL	Content and Language Integrated Learning
CLS	Critical Languages Scholarship
DA	Diagnostic Assessment
DLAB	Defense Language Aptitude Battery
DLIELC	Defense Language Institute English Language Center
DLIFLC	Defense Language Institute Foreign Language Center
DLNSEO	Defense Language and National Security Education Office
DLPT	Defense Language Proficiency Test
DTRA	Defense Threat Reduction Agency
ECA	Educational and Cultural Affairs
EHLS	English for Heritage Language Speakers
ERP	event-related potential
ESP	English for Specific Purposes

ETS	Educational Testing Service
FEHAWENS	Functional Equivalent of a Highly Articulate, Well-Educated Native Speaker
FL	Foreign Language
FLES	Foreign Language Elementary Schools
fMRI	functional Magnetic Resonance Imaging
FSI	Foreign Service Institute
FORTE	Family, Occupation, Recreation, Travel, Education
GAO	Government Accountability Office
GLOSS	Global Language Online Support Systems
GSS	General Social Survey
GUGD	Georgetown University German Department
HBCU	Historically Black Colleges and Universities
HHE	High Hispanic Enrollments
HR	Human Resources
HSK	Hanyu Shuiping Kaoshi
ICC	intercultural competence
IED	improvised explosive device
IH	Intermediate High
IL	Intermediate Low
ILR	Interagency Language Roundtable
IM	Intermediate Mid
IQ	intelligence quotient
KSA	knowledge-skills-abilities statements
L1	first language
L2	second language, target language
LC	listening comprehension
LSP	language for specific purposes
LTI	Language Teaching International
MESA	Middle East Studies/Arabic
MLAT	Modern Language Aptitude Test
NASA	National Aeronautic and Space Administration
NATO	North Atlantic Treaty Organization
NEH	National Endowment for the Humanities
NFLC	National Foreign Language Center
NGO	non-government organization
NH	Novice High
NL	Novice Low
NH	Novice Mid
NMELRC	National Middle Eastern Language Resource Center
NORC	National Opinion Research Center
NSEP	National Security Education Program
NSLI	National Security Language Initiative
NSLI-Y	National Security Language Initiative for Youth
OCP	Overseas Capstone Program

OCQ	Overclaiming Questionnaire
OFC	Overseas Flagship Capstone
OIG	Office of the Inspector General
OPI	Oral Proficiency Interview
OSAP	Open Source Analytical Research Project
OSU	Ohio State University
PDF	Portable Document Format
PLUS	Partnership for Language in the U.S.
Q&A	Question and Answer
RC	Reading Comprehension
ROF	Russian Overseas Flagship
S	Superior
SEALang	Southeast Asian Languages
SLA	Second Language Acquisition
SOPI	Simulated Oral Proficiency Interview
SP	Speaking Proficiency
STEM	Science Technology Engineering Mathematics
TBLT	task-based language teaching
TOEFL	Test of English as a Foreign Language
UNF	University of North Florida
US	United States
WPT	Writing Proficiency Test
ZPD	Zone of Proximal Development

This page intentionally left blank

Introduction

Past Context, Present Focus, Future Directions: Shifting Focus from Intermediate Skills in Classroom Training to Advanced/Superior and Beyond

CYNTHIA L. MARTIN

THE YEAR 2012 MARKED thirty years since the initial development of the American Council on the Teaching of Foreign Languages (ACTFL) provisional proficiency guidelines for speaking (as part of the ACTFL Language Proficiency Projects in 1982), and twenty years since the launch of ACTFL's official testing office, Language Testing International (in 1992).[1] The original framework for the ACTFL Proficiency Guidelines was based on the Interagency Language Roundtable (ILR) Descriptions for Speaking, which remain in widespread use by US government agencies.[2] During the first two decades subsequent to the initial publication of ACTFL's provisional guidelines, a growing proficiency movement transformed the foreign language classroom and curricular planning in the United States in ways many had hoped for, but no one had imagined. This transformation took place, as one might expect, mostly at the lower levels of instruction in what is generally thought of as Novice and Intermediate levels. During the 1980s and well into the 1990s, the focus remained on how to get students to the Intermediate level—the level conceptualized in the ACTFL proficiency community as the realm of the *linguistic survivor*. Much of the research cited and still conducted today focuses on initial acquisition, methods, and pedagogy related to early stages of foreign language learning. The successes have been many, and today there is no doubt that at the center of their planning, programs have placed a notion of what students can *do* with their foreign language, rather than what they *know*.

The proficiency movement has been so successful in transforming the US field of foreign language instruction that university language programs commonly expect to graduate students with a minimum Intermediate level of oral proficiency. However, increasingly the question has become the following: how can instructors get students to the Advanced and Superior levels? Particularly in the years since September 11, 2001, there has been a major shift in focus to produce speakers of foreign languages at the Advanced and Superior levels. Not only has the outside world demanded users

of languages at these levels for purposes of national security, but learners themselves have sought higher proficiency levels in order to take advantage of growing opportunities to use language skills in an era of rapid globalization in ways unimaginable just a few decades ago. This increase in demand for Advanced-level or Superior-level speakers across a wide range of specializations has naturally expanded the need for and use of a reliable and valid assessment instrument that uses a standardized protocol across languages in order to assess how well and in which contexts a speaker is able to function in a language. The most widespread and familiar assessment tool for speaking has become the Oral Proficiency Interview (OPI), rated according to either the ACTFL Proficiency Guidelines 2012—Speaking or the Interagency Language Roundtable Descriptors—Speaking.

Defining Our Terms: What Does a Rating of Advanced (ILR Level 2) or Superior (ILR Level 3) on an Oral Proficiency Interview (OPI) Tell Us?

The chapters in this volume use the ratings of Advanced or Superior according to the ACTFL Guidelines as reference points for discussing higher levels of language use; therefore, it is important that the reader understand (1) how an OPI is administered, (2) what a rating in the Advanced or Superior range says about a speaker, and (3) what rating criteria are used for each level. Note that OPIs conducted by government agencies and rated according to the ILR Skill Level Descriptors are similar to the ACTFL OPI in terms of the assessment criteria and functional expectations, but since there is no standard ILR OPI administered by one central organization, each government agency uses its own testing protocol. Therefore, the explanation below is limited to the ACTFL OPI, as it uses a standardized protocol for elicitation in addition to standardized rating criteria.

An official ACTFL OPI is conducted by a certified ACTFL tester either over the phone or in person. The OPI is recorded and lasts between fifteen and thirty minutes. After the interview, the tester listens to the OPI recording in order to assign a final rating. The OPI is then rated by another certified tester. If the first and second ratings do not agree, the OPI is submitted for further arbitration before a final rating is issued.[3] The OPI itself is a criteria referenced assessment, and the criteria are set forth in the ACTFL Proficiency Guidelines 2012—Speaking. The guidelines focus on the speaker's functional ability to perform certain tasks at each major level of the scale. As a result of this focus on functional language use, the OPI can be used to assess the proficiency of any language user, regardless of his or her profile (a true second language or foreign language learner, a heritage or native speaker) and of how or where the language being assessed was learned or acquired. This focus on assessing a limited set of functions at each major level largely explains the resiliency and flexibility of these guidelines since their conceptualization.[4]

The assessment criterion related to expected functions for a given level is referred to as Global Tasks and Functions. The list of functions is limited at each major level of the scale: Novice speakers can use memorized material, such as formulaic language, to produce memorized lists in highly familiar topic domains.

Intermediate speakers can use language to engage in simple conversation, ask questions, and handle a simple social transaction. Advanced speakers can narrate and describe in major time frames and handle a situation with an unexpected complication. Superior speakers can discuss abstract issues extensively, support opinions, hypothesize, and handle linguistically unfamiliar (i.e., not previously encountered) situations. A more detailed discussion of the Advanced and Superior functions follows, but it should be emphasized again that the lists of functions required at each major level are limited, meaning that an official OPI may not elicit or document many things that a speaker at a given level can or cannot do. All speakers who receive the same major rating on an official ACTFL OPI share a common characteristic: they can perform the required functions for the level assigned and, unless rated Superior, cannot sustain performance of the required tasks at the next level.[5]

In addition to the category of Global Tasks and Functions, there are three other major assessment criteria used to rate a speaker's language proficiency (discussed in detail below): Content/Context, Comprehensibility (including Accuracy), and Text Type. These other assessment criteria are logical extensions of what one's language actually has to be like in order to accomplish the functions of a given level. In order to be rated at a given level, a speaker must maintain all of the criteria for the level all of the time (when asked to do so) throughout the interview. The discussions presented in this volume of teaching, learning, and assessing at the ACTFL Advanced and Superior levels are predicated upon a common understanding of how these criteria are defined and assessed in an actual OPI.

What does a rating in the Advanced range mean?

Speakers rated in the ACTFL Advanced range must be able to perform the following functions: narrating and describing in major time frames and handling situations in everyday social transactions that include an unexpected complication. In terms of the assessment criterion of Content/Context, the speaker must demonstrate the ability to perform Advanced-level functions across a range of concrete topic areas, which include not only the autobiographical and familiar but also more general topic domains related to work, as well as topics of interest to the broader community, like local, national, or even international news. We think of an Advanced speaker as a reporter. The category Comprehensibility/Accuracy includes linguistic features such as pronunciation, intonation patterns, tones for tonal languages, pace or fluency, and structural control, which includes grammar and syntax. These features are assessed in two primary ways in an official OPI: (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 the kind of listener required for successful communication, does the speaker need a very sympathetic listener accustomed to dealing with non-natives, or a neutral, non-sympathetic listener? In order to be rated ACTFL Advanced, the speaker must be able to control the basic structures of the language sufficiently (not perfectly!) so as not to miscommunicate or confuse the listener when performing Advanced-level functions. At Advanced, a speaker should no longer need a sympathetic listener in order to communicate. In terms of the assessment criteria of Text

Type, paragraph-type oral discourse is required for a rating in the Advanced range because that is the text type required by the Advanced functions (e.g., narrations and descriptions) and, therefore, would not apply to a speaker capable of only producing simple sentences.

A word about the sublevel ratings: a speaker rated Advanced Low can perform the functions of Advanced while maintaining the other assessment criteria, albeit minimally. At Advanced Low, a speaker does not usually show much evidence of Superior-level criteria. A speaker rated Advanced Mid is a speaker who performs the functions of Advanced while sustaining the other assessment criteria with appropriate quantity and quality across a range of topics. A speaker rated Advanced Mid is likely to show some evidence of performance or linguistic features associated with the Superior level. An Advanced High speaker is able to perform the functions of the Superior most of the time, but is not able to sustain that performance across a range of functions or topic areas. A speaker with a rating of Advanced High is not just a superstar when performing Advanced functions; rather, an Advanced High speaker should be thought of as an emerging Superior.

What does a rating of Superior mean?

The Global Tasks and Functions criteria that a speaker must fulfill to achieve a rating of ACTFL Superior include discussing an issue at the abstract level, supporting opinions, hypothesizing, and handling linguistically unfamiliar situations (usually indicated by the ability to circumlocute and to deal with any low-frequency or unfamiliar linguistic material). A Superior speaker is one who can think aloud in the target language about a range of general issues. The assessment criterion of Content/Context expands to include most practical, social, and professional (not highly specialized) topics, as well as settings where these topics might be discussed. In terms of the criterion of Comprehensibility/Accuracy, a Superior speaker is able to communicate effectively with speakers not used to dealing with learners; errors may occur, but they are sporadic (not patterned) and neither impede communication nor distract the listener from the message. In terms of the assessment criterion of Text Type, the ability to produce and control extended discourse is required for a rating of Superior because that is the text type required by the functions; less sophisticated text types will prove inadequate when attempting Superior-level functions. The full descriptors for both the ACTFL and ILR proficiency scales are available at http://www.actflproficiencyguidelines2012.org and http://www.govtilr.org, respectively.

How can we help learners achieve an Advanced or Superior rating?

Our goal to help learners achieve OPI ratings of Advanced or Superior necessitates our consideration of what we know about learners, what kinds of methods and approaches can facilitate the attainment of these proficiency levels, and what policy and programmatic considerations are at the institutional and even national levels. Taken as a whole, this volume is largely focused on documenting what we know about successful students, programs, and materials, and the results of the research described herein offer both implicit and explicit recommendations for advancing

our collective goals. Individual chapters in this volume align with these three broad categories:

1. Defining Characteristics of High-level Language Learning and Learners
2. Approaches to Maximizing Language Gain at Home and Abroad
3. Future Directions in Assessment, Program Design, and National Policy

I. Defining Characteristics of High-level Language Learning and Learners

One thing we know for certain is that all language learners are unique and do not comprise a homogenous group. That being said, learners who achieve high levels of proficiency in a second or foreign language they begin studying as adults may share common characteristics. In chapter 1, "Experience with Higher Levels of Proficiency," Betty Lou Leaver and Christine Campbell of the Defense Language Institute present research findings on successful learners in programs within the US government and elsewhere that have helped students achieve Advanced, Superior, and even Distinguished levels of proficiency. Leaver and Campbell present the information through research analysis and comparison of students whose data were gathered through qualitative interviews. There are surprises in the data: some previously unquestioned assumptions are shown to be myth, and some learner characteristics previously not considered at all are shown to be important, if not critical. The study suggests that there are many kinds of learners, differing motivations, and a variety of paths to achievement. That being said, the authors found several commonalities that appear to facilitate successful language acquisition. The authors discuss some of the most successful curricular design elements, such as content-based classes, skill integration, focus on the sociolinguistic and discourse aspects of language, inclusion of sociocultural knowledge and exercises meant to develop the ability to read not only between the lines but also "beyond the text," de-fossilization activities, practice in recognizing genre shift, and special techniques for rapid improvement of speaking skills. These practices can be emulated by other teachers who desire to bring students all the way to ILR Level 4/ACTFL Distinguished. Overall, this research supports a rationale for looking at teaching upper levels of proficiency not as more of the same, but as a fundamentally different endeavor for both learners and teachers.

In chapter 2, "L1, L2, and Cognitive Development: Exploring Relationships," Dan Dewey, Ray Clifford, and Troy Cox of Brigham Young University (BYU) explore relationships between cognitive ability and language proficiency: They align elements of the cognitive process dimension from the revised version of Bloom's taxonomy with the communication tasks and topical domains associated with the ACTFL and ILR Speaking and Writing Proficiency Guidelines. Then they assemble a parallel array of cognitive skills necessary for successful performance of higher-level communication tasks, such as presenting and supporting hypotheses and arguments, explaining complex matters in detail, presenting structured discourse,

and organizing and evaluating ideas. The results from language proficiency tests that included these aligned language proficiency and cognitive skill development expectations, along with the results from an intelligence test, are used to report on the relationships observed between first language (L1), second language (L2), and cognitive abilities. Their research findings indicate that when speakers attempt high-level communication tasks in a second language, either a lack of advanced first-language skills or undeveloped critical thinking skills may serve as blocking factors that inhibit effective communication. Dewey, Clifford, and Cox present findings from the ACTFL OPIs and Writing Proficiency Tests (WPTs) both in participants' dominant language (English) and in their second language (Spanish) in order to distinguish individuals' level of cognitive skill development from their level of second language development. The results of the four assessments were examined for patterns capable of explaining which aspects of an individual's performance reflected cognitive skill development and which reflected second language ability.

It was observed that none of the participants' L2 proficiency levels exceeded their L1 proficiency levels. Not all of the participants had attained ACTFL Superior/ILR Level 3 in their first language, and intelligence scores were somewhat related to breadth of vocabulary. While this study is in many ways preliminary, the authors identify several areas for further research.

II. Approaches to Maximizing Language Gain at Home and Abroad

There is broad consensus among those involved in language learning and assessment that extensive experience with the target language and culture is essential for achieving high levels of proficiency. Generally, an immersion experience inside the target culture is viewed as the optimal way to attain this experience, and, for this reason, stateside instruction has focused much energy on the lower levels of acquisition, counting on the in-country experience to take students to the higher levels. But a growing body of evidence suggests that much can be done in the US academic context to facilitate success prior to and after an in-country immersion experience and that efforts to integrate stateside and overseas approaches warrant improvement.

Three chapters in this section present findings about successful stateside approaches for higher-level learners of English, Russian and Chinese (Kennedy and Hansen; Brown, Bown, and Eggett; Christensen and Bourgerie) and two chapters focus on successful in-country experiences for students of Arabic (Belnap and Abuamsha; and Davidson). The chapters also discuss the increasing numbers of US students learning different critical languages who take part in one or more of the major federally funded overseas immersion programs that were expanded substantially under the National Security Language Initiative of 2006 (Davidson).

In chapter 3, "Developing Superior Language Proficiency and Analytical Skills for the Federal Workforce," Deborah Kennedy of the Center for Applied Linguistics and Christa Hansen from the Center for Language Education and Development at Georgetown University discuss the instructional practices that promote the development of the Superior-level (ILR Level 3 and beyond) language proficiency in

adult learners who plan to pursue professional careers with the US government. The chapter draws on the authors' experiences as directors of the English for Heritage Language Speakers (EHLS) program, which prepares native speakers of critical languages (nonnative speakers of English) to meet the federal government's need for personnel who can conduct research and analysis that is informed by deep cultural and linguistic knowledge. Participants are required to possess ILR Level 2 proficiency in all four modalities in English at program entry; their goal is to achieve ILR Level 3 proficiency by the time they graduate eight months later.

In this chapter, the authors first present the theoretical underpinnings of their instructional approach and then describe the distinguishing features of the program, including principles of backward design, content and language integrated learning (CLIL), and task-based instruction. The authors show how these features are combined with a thorough analysis of the skills included in the descriptors for ILR Level 3 proficiency when constructing the curriculum. The principle of backward design dictates an approach that starts with the functional goals of learning and focuses assessment and instruction on enabling learners to achieve those goals, which are drawn from the characteristics of ILR Level 3 proficiency. These characteristics include the ability to participate effectively in formal and informal interactions on practical, social, and professional topics; to understand and use abstract language and concepts; to state an opinion and support it with detailed factual information; and to hypothesize plausibly on the basis of established information. The chapter challenges readers to reconsider their own teaching and learning practices and to design programs that will yield maximum gain relevant for learners' linguistic and professional goals.

In chapter 4, "Advanced Foreign Language Study through Debate," Tony Brown, Jennifer Bown, and Dennis Eggett of BYU present their findings that demonstrate how innovative curricular design and development at one's home institution can produce measurable gains over a single semester in a university language classroom. In this program, the key to success was a focus on debate—a forum that involves matters of rhetoric, composition, and language functions consistent with Superior-level proficiency. Furthermore, the task-based nature of debate treats language as a means rather than an end. The focus is on arguing for or against a position, and language is the tool for achieving that end. In addition to facilitating dialogue rich in domain-specific lexical items, debates stretch the students' capacity to think critically as they wrestle with challenging social, political, and economic issues. The focus on real-world issues that are relevant to students' lives increases motivation to learn and often spurs students to attain higher proficiency levels than would otherwise be possible. This chapter provides a thorough description of the instructional methodology applied, including scaffolded reading, writing, listening, and vocabulary exercises. Pre- and post-program assessment measures, such as the ACTFL Writing Proficiency Test and OPI, revealed that learners made significant proficiency gains as a result of this approach. Multiple iterations of this pedagogical approach have culminated in the writing of *Mastering Russian through Global Debate* (also described in the chapter), which promotes Advanced- and Superior-level proficiency in writing and speaking through argumentation and debate.

Another approach to facilitating high levels of proficiency by integrating stateside and in-country study is presented by Matthew Christensen and Dana Bourgerie of BYU in chapter 5, "Chinese for Special Purposes: Individualized Instruction as a Bridge to Overseas Direct Enrollment." The authors discuss their experience with an individualized instructional approach used in the Chinese Flagship Program at BYU that prepared students for direct enrollment in content classes in a Chinese university. This approach is predicated upon the assertion that in order for students to maximize in-class instruction abroad, they must be prepared in three ways: (1) they need to be very familiar with their domain of study in the target language, (2) they need to be comfortable with nonstandard accents of Mandarin Chinese, and (3) they must understand Chinese presentational discourse and professional cultural practices. The authors discuss the design of a curriculum that allows students to pursue research within their selected domains of study while simultaneously receiving intensive individual mentoring and tutoring. The approach calls for group meetings in which students are trained in cultural literacy and professional presentational skills. Students also enroll in an additional course in advanced conversation and rhetoric that focuses on professional-level discourse and nonstandard Chinese accents. The results of data collected over the past nine years are presented, indicating significant success when using this approach.

In chapter 6, "Taking on the 'Ceiling Effect' in Arabic," Kirk Belnap of BYU and Khaled Abuamsha of Qasid Institute for Classical and Modern Standard Arabic discuss the importance of content-based instruction as an integral, if not essential, part of BYU's semester-long intensive study abroad program for students of Arabic. No undergraduate Arabic program graduates more Advanced-level speakers and readers of Arabic annually than BYU. At the core of the program is its semester-length intensive study abroad program. Content-based instruction has been integral to the program from its beginning in 1989. During the fall semester of 2011, 52 BYU students with four semesters of prior Arabic experience studied intensively at the Qasid Institute in Amman, Jordan, using a special curriculum developed by BYU and modified in consultation with its Qasid partners. This chapter documents the impact of this experience on students, as seen through student reflections and standardized measures of proficiency, and the impact of a subsequent capstone internship experience. The study also compares this program with others, such as Cornell's Intensive Arabic Program, which also relies heavily on content-based instruction. In addition, the authors provide detailed suggestions for how content-based instruction could be improved in US programs.

In chapter 7, "The Development of L2 Proficiency and Literacy within the Context of the Federally Supported Overseas Language Training Programs for Americans," Dan Davidson of Bryn Mawr College and American Councils for International Education (ACTR/ACCELS)[6] presents new proficiency-based, pre-program and post-program proficiency test data on the immersion learning of a broad range of critical languages by more than 1,700 US students. These students have participated in a distinguished group of federally supported overseas study programs, which are open on a competitive basis to qualified American students: the National Security Language Initiative for Youth (NSLI-Y), Critical Languages

Scholarship program (CLS), and the national Flagships. The reports provide heretofore unreported information on the impact of instruction in the less commonly taught languages undertaken in the overseas immersion context for learners' cohorts ranging from 16-year-old students with no prior training to young adults seeking to advance their linguistic and cultural proficiency at the Intermediate, Advanced, and Superior levels. Davidson's study documents and compares the outcomes of immersion study across different languages and for students of different ages at different stages of their language-learning career. While clearly demonstrating the advantages of an early start in the study of a critical language, exemplified by the results of the NSLI-Y program, the author also provides strong evidence of success at reaching the Superior level for students who begin their study at the college level with access to appropriate levels of overseas immersion training, such as those provided by the Critical Language Scholarship Summer Institutes, Fulbright-Hays, and the Language Flagship Program.

III. Future Directions in Assessment, Program Design, and National Policy

The growing body of evidence about successful learners, programs, and approaches at the advanced levels has promoted thinking about new directions in assessment, teaching, and learning related to professional language use. In chapter 8, "From Proficiency to Expertise: Using HR Evaluation Methods to Assess Advanced Foreign Language and Culture Ability," Patrick McAloon of Ohio State University examines alternative methods for documenting the abilities of high-level users of foreign languages. McAloon discusses the use of approaches that are common in human resource management to document skills and abilities, specifically approaches that would be useful in a foreign language context. Though many language educators now agree that high-level foreign language ability is connected inextricably to its professional applications, along with the ability to interact with native speakers of the target language in a culturally appropriate manner, McAloon suggests that there is still room for discussion about the relationship between language and professional ability at these high levels. Moreover, he argues that the current standards for foreign language assessment in the United States, the ACTFL and ILR scales, while useful for describing and assessing individuals' foreign language ability, were not designed to assess performance in a professional foreign language environment. He suggests that, at high levels of foreign language ability, professional performance must become the object of assessment and that extant tools for assessing performance in the workplace should be used to evaluate foreign language learners whose language abilities are so advanced that their language skills cannot be separated from their professional performance. In this chapter, McAloon explores how two types of performance assessment regularly used by human resource professionals could prove useful for evaluating high-level foreign language speakers: (1) 360-degree evaluation and (2) portfolio assessment. The author asserts that using human resource assessment tools to evaluate foreign language learners could have implications for current assessment practices that focus only on language proficiency and that these

HR-based assessment tools could enhance educators' ability to graduate foreign language speakers who are prepared to work in the target language.

In chapter 9, "Professional Language Skills: Unprecedented Demand and Supply," Richard Brecht of the University of Maryland, William Rivers of the Joint National Committee for Languages and the National Council for Languages and International Studies, John Robinson of the University of Maryland, and Dan Davidson of ACTR/ACCELS and Bryn Mawr College present an overview of the current state of affairs in US language education at the advanced levels, starting with a brief assessment of the demand for language skills driven by the decade-long growth of public- and private-sector demand for language professionals—individuals with the ability to perform as professionals in a second language across a wide range of disciplines. The authors first discuss the traditional supply system for speakers of foreign languages in the United States and the language proficiency output of higher education over the past few decades, which has been a product of two traditional pathways in achieving Advanced-level skills: intensive instruction, with some in-country immersion, and very modest extended sequences built upon a weak K–12 base. The authors then provide evidence for changes that would support current and projected extraordinary enhancements of these pathways derived from current best practice in classroom and immersion settings, as well as major advances in science and technology. While still some distance away from fulfilling the promise of making efficient and effective programming for acquiring professional language skills universally available, this chapter lays out common parameters for program design that can be used in settings such as industry, government, heritage communities, and education to dramatically improve proficiency outcomes.

In chapter 10, "Expanded Understandings and Programmatic Approaches for Achieving Advanced Language Ability," Frederick Jackson of the National Foreign Language Center at the University of Maryland discusses three examples of instructional programs that aim to help learners reach high levels of proficiency (ACTFL Superior or ILR Level 3), and then argues that what we now know about attaining those levels can help us move beyond them to even more sophisticated levels of language use and cultural understanding.

Jackson documents a number of successful programs that have sought to facilitate learning at the higher levels of proficiency and discusses some findings that may inform future efforts. For example, openness to implicit learning, which research has demonstrated to be essential, represents an important variable that appears to affect the level of advanced mastery that a learner is capable of achieving (Doughty 2011; N. Ellis 2004). The control of language that is represented by the ACTFL Distinguished and ILR Level 4 designations is massive in scope and must be understood and produced with automaticity, allowing the speaker to focus entirely on the meaningful content and context. Elements of the language at this level cannot be taught or learned piecemeal; rather, they must be absorbed from experience in rich, authentic contexts. But Jackson pulls together evidence suggesting that feedback and explicit instruction even at this level can help at least some learners attend to aspects of communication that might otherwise not be observed (see Schmidt, 2001; R. Ellis, 2008; N. Ellis, 2005). Preliminary findings indicate that explicit noticing

of such language features leads to accelerated and deeper acquisition, provided that learners are able to immerse themselves concurrently in an authentic lingua-cultural context.

Additionally, Jackson discusses work by Bernhardt (2005), who pointed out that the Foreign Service Institute's (FSI) understanding of ILR Level 3 language proficiency has been deepened and refined in recent years, resulting in improved training. Jackson also discusses work by Byrnes (2006a; 2006b) and several colleagues, who have proposed a model based on multi-literacies. These multi-literacies are, in turn, based on a recognition that various communicative contexts exist and that certain contexts, particularly the public context, require the language user to take part in broad societal and professional arenas. Finally, Jackson discusses a very recent development that will affect our understanding of advanced language competence and thus affect training and assessment—ILR's 2012 publication of skill level descriptions for competence in intercultural communication, descriptions which incorporate both linguistic and extralinguistic elements at each level and refer to the content and form of communication, both verbal and nonverbal. Jackson suggests that these descriptions serve to complement and supplement the long-standing ILR Language Proficiency Descriptions, incorporating both transactional and public communications, and in several ways reflect a natural extension of ACTFL's *Standards for Foreign Language Learning*.

In the conclusion, "To Advanced Proficiency and Beyond: Charting a New Course in the Twenty-First Century," we see that the past thirty years have witnessed a remarkable evolution in the learning of foreign languages—from the early days of the proficiency movement that encouraged academic approaches to learning, to an emphasis on learning for the purpose of real-life language use, and, most recently, to a focus on how to achieve levels required for professional language competence. The final contribution in the volume is provided by coeditors Tony Brown and Jennifer Bown of BYU, who look ahead at challenges we face in terms of developing more socioculturally appropriate forms of assessments that capture much of what is not captured by current standardized instruments, expanding our design and use of immersion experiences and individualized instruction, and rethinking our curricular materials and instructional paradigms.

Notes

1. American Council on the Teaching of Foreign Languages, "ACTFL Language Proficiency Projects," *Modern Language Journal 66*, no. 2 (1982), 179.
2. For the purposes of this introduction, references are made to the ACTFL proficiency ratings in the Advanced and Superior range, which generally correspond to the ILR 2 and 3 ranges.
3. See http://www.actfl.org for studies on the reliability and validity of the OPI as an assessment instrument and for more information about tester training and certification.
4. There have been three editions of the guidelines: 1986, 1999, and the most recent one, 2012. In all versions, the major levels have remained the same, and a Distinguished level was added to the 2012 edition.

5. Though a level (called Distinguished) has been added to the 2012 guidelines above Superior, the current official OPI does not test for Distinguished. A brief discussion of the Distinguished level and considerations for assessment can be found at the end of this introduction.
6. ACTR/ACCELS administers Advanced and Professional-level language instruction at overseas universities and immersion centers that currently support more than 1,500 Americans annually, including the State Department's NSLI-Y program for high school students and the undergraduate overseas CLS summer institutes, the Department of State's Title VIII training program, US Department of Education's Fulbright-Hays Group Project Abroad, the Defense Language and National Security Education Office's (DLNSEO) African Languages Initiative (AFLI), and Overseas Flagship programs, which include direct enrollment and internships in Arabic, Chinese, Persian, Russian, Swahili, and Turkic languages.

References

Byrnes, Heidi, ed. 2006a. *Advanced Language Learning.* London: Continuum Press.

Bynres, Heidi. 2006b. "Locating the Advanced Learner in Theory, Research, and Educational Practice: An Introduction." In *Educating for Advanced Foreign Language Capacities*, edited by Heidi Byrnes, Heather D. Weger-Guntharp, and Katherine Sprang, 1–16. Washington, DC: Georgetown University Press.

Doughty, Cathy. 2011. "Redesigning Language Learning." Plenary Presentation at monthly meeting of the Interagency Language Roundtable. Arlington, Virginia. October 11, 2011.

Dubinsky, Inna, James Bernhardt, and Joan Morley. 2005. "High-Level Proficiency Acquisition: Oral Skills." In *Teaching and Learning to Near-Native Levels of Language Proficiency II: Proceedings of the Fall 2004 Annual Conference of the Coalition of Distinguished Language Centers*, edited by Inna Dubinsky and Richard Robin. 23–32. Salinas, CA: MSI Press.

Ellis, Nick C. 2005. "At the Interface: Dynamic Interactions of Explicit and Implicit Language Knowledge." *Studies in Second Language Acquisition*, 27(2), 305–52.

Ellis, Rod. 2008. "Principles of Instructed Second Language Acquisition." *CAL Digest* December 2008.

Schmidt, Richard. 2001. "Attention." In *Cognition and Second Language Instruction*, edited by Peter Robinson. Cambridge: Cambridge University Press. 3–32.

Part One

Defining Characteristics of High-Level Language Learning and Learners

This page intentionally left blank

Chapter 1

Experience with Higher Levels of Proficiency

BETTY LOU LEAVER AND CHRISTINE CAMPBELL

SINCE 1983, US GOVERNMENT institutions have set about producing language learners at Interagency Language Roundtable scale Level 3 (ILR-3[1]) and higher in several languages, the most common of which has been Russian. Pilot programs, as well as some enduring programs, have been undertaken at the Foreign Service Institute (FSI), the National Aeronautic and Space Administration (NASA), and the Defense Language Institute Foreign Language Center (DLIFLC). Gradually, the understanding that levels as high as ILR-4/4+ can be achieved in a classroom setting, sometimes even without an in-country immersion experience, has increased as one program after another has actually accomplished this goal.

Literature Review and Historical Background

In 2002, Byrnes referred to the "disjuncture" in American postsecondary language programs between the teachers' interest in helping students attain higher levels of proficiency and the realities of the learning environment where students were exposed to a few semesters of language acquisition courses followed by content courses that rarely stated acquisition goals or included "instructional interventions" designed to help them reach the higher levels (henceforth referred to as advanced second language [AL2] and understood to mean levels 2+ and above; 34). The following year, Byrnes and Maxim (2003, xv) edited a tome focused on the AL2: *Advanced Foreign Language Learning: A Challenge to College Programs*, which they saw as a vehicle for expanding the "nascent discussion" on the topic in the profession.

The profession has evolved considerably in the past decade, in part due to the gradual dissemination and adoption of the *National Standards* (1996) in academe and a deliberate—and successful—movement within some organizations to "push the envelope." Standards within the US government, known today as the ILR scale, were initially developed by FSI in 1954–1955.They were refined at intervals: in

1968, in 1986, and together with other ILR members as part of a Center for the Advancement of Language Learning (CALL)[2] initiative, in 1996 to standardize variations on the application of the ILR scale that had begun to appear in the intervening years at various government institutions. As the 1996 collaboration indicates, the remainder of the US government language schools and agencies had adopted the ILR scale over time, and, beginning in 1976, the North Atlantic Treaty Organization (NATO) joined the group accepting the ILR scale as the basis for its own standards, which have also been refined over time. The initial efforts to create national standards for academe were a collaborative effort between FSI and the American Council on Teaching Foreign Languages (ACTFL) in the mid-1980s, which has since come to be known as the ACTFL scale.[3] The standards, whether those of the US government or those of ACTFL, continue to be a critical common metric that facilitates communication across diverse language institutions and organizations about goals and objectives.

As a result of improved teaching through the application of the standards and the integration of pedagogically sound technology, high school students are arriving at postsecondary institutions, whether in academe or at government centers like DLIFLC with a higher level of proficiency than in prestandard times, ready to achieve the required base level of proficiency of ILR-2 for basic course students, ILR-2+ for intermediate students, and ILR-3 for advanced students. Most recently, agencies to which graduates of the various DLIFLC programs are assigned have designated ILR-3 as the minimal acceptable level for a career as a government linguist. To meet the learning needs of these students, which generally exceed both practice and experience of most teachers and program managers, teachers in the "post-method era"[4] have been challenged to adapt their own "theories of practice," using macrostrategies such as those below when planning lessons and developing curricula for the base levels of proficiency:

- maximize learning opportunities;
- minimize perceptual mismatches;
- facilitate negotiated interaction;
- promote learner autonomy;
- foster language awareness;
- activate intuitive heuristics;
- contextualize linguistic input;
- integrate language skills;
- ensure social relevance; and
- raise cultural consciousness (Kumaravadivelu 2003, 1, 2, 39, and 40).

The *National Standards* focused primarily on K–16 needs, which required addressing levels below ILR-3, and did not look toward developing refined differentiation among the levels at and beyond ILR-3, which were generally conflated as "Superior." The push for a refined definition of the upper levels as well as experimentation in teaching at these levels, known informally as "The Level 4 Movement," was born in 2002 with the founding of two important organizations: (1) the Coalition of Distinguished Language Centers (CDLC), a grassroots organization established

to support programs seeking to produce ILR-4 outcomes, and (2) a precursor of the Flagship programs, the Center for the Advancement of Distinguished Language Proficiency at San Diego State University, funded by the National Security Education Program (NSEP). The program was established to train teachers of high-proficiency programs in pragmatic techniques for achieving near-native levels of foreign language skills. With the appearance of these two organizations, publications and conferences of the CDLC, and a growing, yet still nascent body of research, interest and interaction began to coalesce among a number of centers and programs interested in the highest levels of proficiency: some foreign universities, NSEP-funded Flagship programs that began appearing in the United States circa 2004, a few private organizations, and US government agencies. This coalescence formed the core of the Level 4 Movement.

The tenets promulgated by the CDLC and its member institutions were shared in an early volume on teaching to and beyond ILR-3 (Leaver and Shekhtman 2002), focusing on programs and institutions that routinely produced ILR-3, ILR-3+, and ILR-4 speakers, readers, listeners, and writers. These programs included FSI, where the "Beyond Three" Russian and French programs have produced diplomats with Level 3+ and Level 4 skills for twenty-five years—programs and outcomes that appear in few published forums; the American University of Cairo, where foreign students were able to achieve Level 4+ in reading in a one-semester program; NASA, where astronauts from the United States gained Russian-language skills and cosmonauts from a number of countries, including the Soviet Union/Russia, gained English-language skills at professional levels beyond ILR-3 as needed for joint space activities; the Monterey Institute of International Studies, where "bridge" activities were used to prepare students to participate in professional dialogues and do high-level translation and interpretation; Georgetown University, where an immersion was conducted, focused on developing Advanced-level proficiency; Chinese programs that conducted training in-country; an online reading to Level 4 program (LangNet, founded by the National Foreign Language Center [NFLC]) now part of the online Joint Language University; Bryn Mawr College, where teacher training mixed native speakers of English with native speakers of Russian and expected the same outcomes from professional training programs; and, for students of English, the teaching of high-level English-writing skills at the University of Aarhus (Denmark). As with most high-level programs, heritage learners were often among the student body in the programs represented in this volume (and at CDLC annual meetings), and both integration and attention to the specific needs of heritage students as they differ from nonnative learners were required.

The practices used by these programs shared many common features, yet were not based on research. In fact, to this date, few research projects have been undertaken in order to develop a "theory of practice" for achieving the higher levels of proficiency that is rooted in science and used to enhance the probability of success. Among these few was an effort funded in 2001–2002 by NFLC and collaboratively conducted with DLIFLC, in which Leaver and Atwell (2002, 260–278) conducted in-depth interviews with more than one hundred language users who had been tested at ILR-4. Participants came from all walks of life but shared many

motivational, demographic, and learning strategy traits, including an intense desire to share their "story." Data were collected in more than one hundred categories per skill (reading, writing, listening, speaking) and in more than twenty-eight demographic areas. Among the learners were polyglots who shed much light on the variability in language acquisition at higher levels.

The NFLC-DLIFLC study was followed by another and conducted on learners of English at the University of Jordan who had achieved Level 4. That study, funded by the National Teachers of English's Conference on College Composition and Communication, compared Level 4 achievers with Level 3 achievers and looked mainly at demographics (Jasser, Al-Khanji, and Leaver 2005).

These studies distilled eight typical core characteristics of a language learner able to achieve near-native levels of foreign language proficiency, some of which are counter to those exhibited by good language learners at lower levels of proficiency:

- tenacity;
- a good ear;
- desire to do a good job (a job that requires L2), i.e., instrumental motivation;
- upbringing in a multilingual neighborhood;
- background in a multilingual home;
- experience abroad, particularly in going to school with native speakers and being treated as a native speaker or in holding a job abroad in the same capacity as a native speaker;
- a native-speaking spouse or significant other (or equivalent friendships); and
- access to classroom instruction at levels 2+ and 3 (Leaver 2013).

It is important to note that 75 percent of the students had received upper-level instruction in a classroom setting and considered this instruction at ILR-3 and ILR-4 as "essential." Unlike the results found by Gardner and Lambert (1972), in which the best students at lower levels of proficiency in university classrooms exhibited integrative motivation, i.e., the desire to acculturate—especially to the point of being mistaken for a native—the surveyed learners who had achieved near-native proficiency displayed instrumental and intrinsic motivation. They did not show integrative motivation. In some cases, passing for a native speaker would have been impossible, especially for a six-foot-seven, blue-eyed, blond-haired student of Chinese. In the case of students working for the government in the national security field, the lack of interest in acculturating was rooted in the job requirement that they not acculturate. Students understood the value of language as a means of communication, or more typically, as a job tool, and a number of them already spoke another language. A few were polyglots, speaking at least three languages.

Shared characteristics that were anticipated included a "sharper" cognitive style, i.e., the tendency to focus on differences rather than similarities, which allowed learners to pick up on refined differences between their native language and target language without being directly taught those differences, which was helpful in the upper-level task of tailoring learners' languages to the audience. Moreover, in the classroom they were quick to focus on small differences between synonyms and between synonymous expressions. They noticed them, asked about them, and

remembered them. Most respondents reported actively seeking out access to the culture; e.g., they went to public places like restaurants and cultural events where the language was spoken. The study abroad programs they joined provided an experience in which the students could attend class alongside native speakers, not foreigners. They became tenacious and were able to motivate themselves when they fell into a slump.

Surprisingly, aside from the typical need for an instructor's insights into the less obvious aspects of sociolinguistics, there was no clear path to the top. Students made it in their own way, on their own schedule, and meeting their own interests. Even more surprising, among the polyglots, each foreign language was acquired in a different way (Leaver 2003). In some cases, the difference was due to the nature of the language (e.g., Arabic requires a cognitive style open to metaphoric/analogue processing, given the high context nature of Arabic expression, making Arabic easier for these kinds of learners and more difficult for literal/digital learners, who had to acquire metaphoric processing strategies; likewise, a random cognitive style often helps learners of Russian and creates difficulties for a sequential learner, given that Russian word order is not fixed). In other cases, a background with a related language made a difference. In other cases, the difference was due to the opportunities that were available to the learner, e.g., the timing of (or even access to) study abroad. These variables can best be addressed by teachers who are intuitive, diagnostically oriented, and highly flexible.

With upper levels of proficiency, particularly for those striving to reach ILR-4/4+, direct teaching (explanations, parsing of language structures, etc.) has been found to be highly important. Exposure and practice alone generally will not push students to these levels (there are always exceptions) because once they are past ILR-3, they cannot learn simply by observation; they do not know what it is that they are seeing (or, especially, not seeing) among the implicit cultural (pragmatic and sociolinguistic) norms. The sections below will explore what curricula, assessments, and faculty development allow teachers to optimize the conditions for learning that can enhance the probability of learner success, making accessible the unseen aspects of language and culture.

A Starting Point: The Foreign Service Institute

In 1984, the Russian program at FSI embarked on an ambitious project: Take students from ILR-3 to ILR-4 in six months. Entry requirements were a recent ILR-3 rating and at least three years abroad, the typical assignment for a Foreign Service Officer. For six years, 1983–1989, every student in the course made ILR-4 and some reached ILR-4+. The two instructors who developed the content-based course (and others who taught it) used a variety of linguistic activities, such as islands, embellishments, complication exercises, and restating an utterance or speech by someone in a different register. Sociolinguistics, pragmatics, register, and genre were a staple of the program. Students were paired with immigrants in need of help; in helping them with their legal, cultural, educational, daily living, and other needs, students acquired higher levels of language ability. The course culminated in a conference

conducted in the Russian language, and was open to the Russian-speaking community in the Washington area. Students presented speeches, conducted debates, and held roundtable discussions on typical State Department topics in Russian, with the public asking unanticipated questions that students had to handle in an erudite manner. This conference was open to students in the basic course, as well, and often pushed students over the ILR-2+/3 threshold.

Today, the program has taken a new turn. Called the Beyond Three Course, there is no guarantee that students will reach ILR-4, although many do, and much more responsibility for learning is left to individual students, with emphasis on strong reading of classics in the language. Different approaches with similar goals have seen nearly thirty years of success in teaching upper levels of proficiency at FSI.

Along the Way: Other Programs

There are other programs that have pushed an occasional student to ILR-4. These include astronaut training, study abroad sites, teacher training programs, some Flagship programs, and some university centers.

The Russian and English program developed for the International Space Station in 1998 saw the development of ILR-4 proficiency by several astronauts at NASA and a cosmonaut or two at Russia's Gagarin Cosmonaut Training Center. The training program, like the DLIFLC and FSI programs, is content-based and uses real-life tasks and authentic materials—an approach that can be very challenging for foreign language teachers since none of those hired to date at either institution have been jet pilots. The cornerstone of the NASA program is a crew-based approach in which astronauts and cosmonauts assigned to the same space station crew do all their technical and language studies together. They attend technical training in Houston conducted in English, and technical training in Star City conducted in Russian. In both locations, an interpreter is on call, if needed, since not understanding technical details can be life-threatening once in space. Although most of the language training is in tutorial format, one of the crew-based activities is to conduct a scenario based on real-life possibilities (e.g., fire on the space station), first in Russian, then in English. Teachers attend but do not assist. They take notes on communication failures, and then incorporate that information in individual follow-up lessons. A detailed description of the program can be found in Leaver and Bilstein (2000).

A relatively small number of study abroad programs are able to develop ILR-3+ and ILR-4 speakers. Among these are American University of Cairo in Egypt (Bidawid 2002), which develops highly advanced reading skills with the ambitious goal of ILR-4+, and the American Councils for International Education's (ACIE) programs, in which advanced language students are individually embedded in regular graduate programs at various Russian universities, which gives them the opportunity to experience education the same way native speakers experience it, thereby building proficiency rapidly.

In the 1990s, Bryn Mawr College developed summer institutes funded by the National Endowment for the Humanities (NEH) for teachers of Russian (Dabars and Kagan 2002). These programs were taught at the native level of proficiency

and focused on linguistics, sociolinguistics, culture, literature, and other high-level topics needed to prepare teachers for instructing all levels of proficiency. Although no proficiency testing was used, four forms of feedback were used to assess the success of these programs: native-speaker instructor observations (conclusion: language proficiency noticeably improved in both fluency and accuracy), formal participant feedback to the Institute (conclusion: without exception, participants stated that the Institute met their expectations and that they would significantly change their teaching methods upon returning to their schools), formal participant feedback to the NEH (conclusion: same comments as on feedback to the Institute), and external evaluation by an outside evaluator required by NEH (conclusion: highly positive influence on teachers' proficiency in the language and capability in teaching Russian).

Flagship programs, funded by NSEP, appeared for the first time in 2003–2004. These programs were not meant to develop language skills to near-native levels, but rather to take ILR-2 students to ILR-2+ during a full year in the classroom and ILR-3 through a full year abroad. Some of these programs have also taken some students to higher levels (as measured on the FSI proficiency test, which tends to produce ratings similar to, but slightly higher than the Defense Language Proficiency Test [DLPT][5] [Leaver and Atwell 2002]). Like a number of universities around the world (Caudery 2002), the University of Jordan routinely produces students capable of reading, writing, and speaking at near-native levels of proficiency (Jasser, Al-Khanji, and Leaver 2005). In the United States, San Diego State University's Center for the Advancement of Distinguished Language Proficiency has trained students to ILR-4 in Russian and Arabic (Lyman-Hager and Degueldre 2004), as well as conducted faculty development open to teachers from any organization desiring to learn how to help students reach these high levels. Similarly, Yale University's Language Center has developed Spanish courses for students at high levels of language proficiency (Tortora and Crocetti 2005), and the Monterey Institute of International Studies has run a "bridge" program for those needing to gain near-native skills in order to enroll in translation and interpretation courses (Angelelli and Degueldre 2002).

From 2002 through 2010, the CDLC held annual conferences on the topic of achieving ILR-4 to assist the work of these centers, programs, and organizations. It also published the main body of work in the field of reaching ILR-4, including six issues of the *Journal for Distinguished Language Studies*, six volumes of conference proceedings (*Teaching and Learning to Near-Native Levels of Proficiency: Proceedings of the Conferences of the Coalition of Distinguished Language Centers*), and *What Works: Helping Students Reach Native-Like Second-Language Competence* (Coalition of Distinguished Language Centers 2006), a small but highly respected book, collectively authored by the leaders in the field at the time.

This list could go on. With the passage of another decade, it very likely *will* go on—much longer. The result of the "Level Four Movement" catching on is that there are more teachers who believe that students can be helped to reach ILR-4 and there are more students who want to do it. This volume is a tribute to those beliefs and wants. Thanks to the experienced ILR-3-and-beyond teachers and administrators

who contributed to this volume, the current limited amount of information on how to move forward in terms of theory, practice, and research (every one of these domains being critical) will be substantially supplemented and enhanced.

A Recent Example: The Defense Language Institute Foreign Language Center

The Proficiency Enhancement Program at the DLIFLC has as its ultimate goal to produce linguists with the highest possible levels of proficiency in its basic, intermediate, and advanced courses. Currently, for basic (initial acquisition) courses the goal is ILR-2+ in listening comprehension (LC), ILR-2+ in reading comprehension (RC), and ILR-2 in speaking proficiency (SP) or Advanced High/Advanced High/Advanced, respectively, on the ACTFL scale. The graduation requirement is ILR-2/2/1+, or Advanced/Advanced/Intermediate High, respectively, although with the recent emphasis by the military services on the ILR-3 as the career goal for linguists the goal may at some point become the requirement.

For intermediate courses, as noted above, the graduation requirement is 2+/2+/2, and for advanced courses, the graduation requirement is 3/3/2+. However, for the past five to six years, DLIFLC has nearly continually produced linguists at ILR-3 and above in both the intermediate and advanced courses at DLIFLC's Directorate of Continuing Education (CE), using a strategic approach that promotes teacher autonomy and innovation in the areas of curricula, assessment, and faculty development. The approach is informed by the survey research on learner characteristics and teaching experience described in the historical background and literature review section above. CE empowers its faculty to do more than teaching: it asks them to create their own curricula, to develop their own assessments, and to seek out opportunities to develop professionally and share that knowledge with colleagues as they collaborate to construct knowledge, ultimately forming a "community of understanding and practice" (Bailey and Freeman 1994). This approach has been a key factor in an increasing number of students who improve as much as 1.5, or even two proficiency points in the skills of listening, reading, and speaking (writing is taught but not tested) over the course of eighteen, thirty-six, or forty-seven weeks, depending on the category of the language. Below are the entry requirements, although, given that students who are working military jobs need to improve their proficiency, these requirements are often waived:

- Intermediate: ILR-2 in LC and in one other skill; minimum graduation requirement—ILR-2+/2+/2 (exception: the Defense Threat Reduction Agency [DTRA] course, essentially a course in interpretation and translation, has no entry requirement, although it is considered an intermediate course and ILR-2 is generally expected);
- Advanced: ILR-2+ in LC and ILR-2+ in one other skill; minimum graduation requirement—ILR-3/3/2.

Course lengths vary according to language category.[5] Below, the lengths of courses in weeks are given by language category:

- Intermediate: eighteen weeks for a Category 1 language (Spanish); thirty-six for Category 3 (Russian, Serbo-Croatian, Hebrew, Persian/Farsi); forty-seven for Category 4 (Arabic, Chinese, Korean);
- Advanced: eighteen weeks for a Category 1; thirty-six for Category 3; forty-seven for Category 4.

As the curriculum and approach to upper levels have matured and improved over time, results have improved from a low of 33 percent of the students reaching graduation requirements in 2008 to between 85 percent and 95 percent reaching them in recent years. Currently, about 54 percent of students surpass graduation requirements, achieving levels as high as ILR-4. Figure 1.1 below presents the results of the DLPT, version 5 (DLPT5), beginning in 2008, when the approach emerged to moving beyond specific goals and assisting students to reach the highest levels possible, regardless of course assigned. All students take the DLPT5 Lower Range (up to ILR-3) upon course completion. Those who score ILR-3 are given the opportunity to take the DLPT5 Upper Range, which goes beyond ILR-3. In addition, for

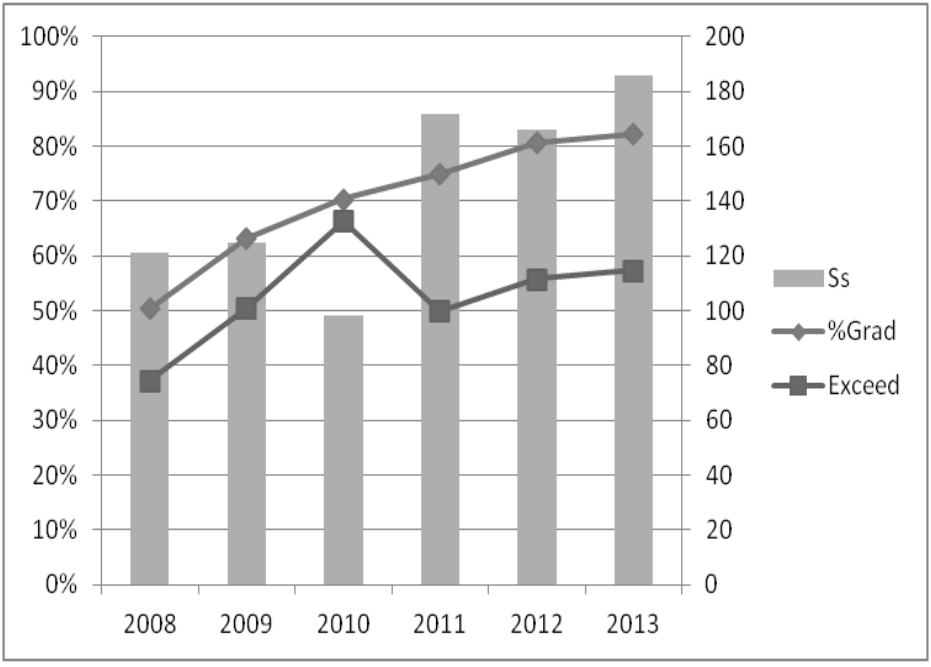


Figure 1.1 CE Overall Graduation Rates

Note a: Total number of students from 2008 through 2013 is 850.

Note b: Reference legend: Ss = students; %Grad = number of students achieving graduation requirements (ILR-2+/2+/2 for Intermediate students and ILR-3/3/2+ for Advanced students); Exceed = number of students with scores higher than graduation standards (ILR-3, 3+, and 4 for Intermediate students and ILR-3+ or 4 for Advanced students). About 75 percent of those who exceed ILR-3 achieve ILR-3+ and 25 percent ILR-4.

speaking scores, students take an oral proficiency interview. Although the DLPT5 is a multiple-choice test of listening and reading skills, CE conducts a diagnostic assessment (DA), which is an interactive, full-range, formative proficiency assessment, on each student. Graduation results have been, for the most part, consistent between the DLPT5 results and the DA results. Figure 1.1 is a representation of the improvements in proficiency that have occurred as the current methodology has been gradually introduced, beginning in 2008.

Curricula

CE curricula are designed to engage the Level 3 and above learner. Emphasis is placed on a balance of process and product where students learn to process language at the higher levels while working on projects, portfolios, presentations, extramural activities in escort interpretation practice (necessary for those who will be accompanying VIPs abroad and interpreting for them), and more.

The CE curricula for both intermediate and advanced courses:

- orient on the individual student: learning style, special interests, Zone of Proximal Development (ZPD), specific linguistic gaps, and the like;
- provide a content-based syllabus;
- emphasize task-based authenticity using authentic language in real-life context;
- focus on language and culture, especially for tailoring language to audience in terms of register and cultural expectation;
- integrate all four skills, as well as interpretation and translation;
- focus on the discourse aspects of language;
- include sociocultural knowledge and exercises meant to develop not only the ability to read and listen "between the lines" (author intent) but also "beyond the lines" (the social milieu in which the piece was written and the audience for which it was intended; the reason for word choice and content inclusion or omission, etc.);
- emphasize relentlessly defossilization activities in pronunciation, grammar, language level, learning (in)dependence, strategy choice, and more;
- teach and practice genre shift; and
- ensure that speech acts and text organization reflect native speaker norms.

The curriculum is individualized in that it is focused on the needs and interests of the students (Higgs 1984), including linguistic weaknesses and processing gaps, and tailored to them (Nunan 1988, 1997, 2006). In some cases, students have developed learning contracts. One case in point was a class that returned from study abroad in Jordan, intrigued by a controversial new novel. The class contracted with the teacher to use the novel as the basis of their language study for several months until the literature had been completely dissected, including author style, the social context, sociolinguistic and genre facets, and more.

Students receive a DA: a face-to-face interview tool that determines a proficiency level in LC, RC, and SP, identifies cognitive styles, and evaluates writing skills. The DA uses a form of dynamic testing that focuses squarely on the ZPD (Vygotsky 1978; Lantolf and Poehner 2004), identifying those elements of language that the

student is most ready to learn next, which will vary by student. Assessors (teachers) provide learners with a learner profile identifying their strengths and weaknesses in the three skills and those linguistic items that fall within their ZPD; this profile will guide their classroom and homework activities, as well as their independent study time. The team of teachers develops a class profile (a picture of the linguistic and cognitive strengths and weaknesses of the group as a whole) in order to determine what lesson plan activities will reach most students and what other ones are needed to reach the remaining few students—what needs to be done in the classroom for individualization and what will best be assigned as homework. Where appropriate, teachers facilitate collaborative learning and assign individualized study. Appropriateness of each approach will depend upon the class profile. In regard to instruction, teachers pay attention to student motivation and other affective variables, utilize information about student learning styles and cognitive preferences to create tailored lesson plans that include authentic materials and tasks, and promote learner autonomy through learning contracts and learning plans.

Content-based instruction is a key feature of the curriculum: the syllabus is based upon an area studies approach. Students learn language as they learn about and analyze social, political, and historical events. (In fact, the American Council on Education recommended awarding of 15 upper-level college credits in these areas, as opposed to foreign language credits, given that proficiency levels of ILR-3 and higher have nearly no counterpart in academe at this time.) The classroom is an immersion environment where students take a language pledge and guest speakers, who are specialists in the subject matter, are invited to interact with students through presentations.

Students hold cultural events. They also go on cultural excursions where, for at least one course, they may gain practice in interpretation. They record and analyze their performance in order to improve before the next excursion.

Task-based authenticity is a priority: Students are given real tasks, such as making formal presentations to distinguished visitors in the foreign language, teaching material to students at the lower levels, preparing reports involving data collected through surveys, blogging, and the like.

Language and culture are blended seamlessly into the curricula. Language exercises focus on repackaging information (i.e., shifting register) for a variety of audiences, simplifying thought, and embellishing expression. Grammar does not come from a textbook but is presented as "grammar in the wild," i.e., grammar that students encounter in their authentic readings.

When students explore how grammar changes across genres, they study a range of genres and learn about sociolinguistics and pragmatics. When they conduct debates, for example, they learn the sociolinguistic practices and expressions associated with debates in the target culture through analyzing debates they watch on television, isolating the ways in which turn-taking is accomplished, and learning expressions used for polite and rude interruption. Thus, when they conduct a debate, it is a close approximation to the way in which native speakers debate. When students conduct interviews, they first observe foreign televised interviews, as well as read interviews from foreign magazines. They identify how these interviews differ from those done

in the United States, specifically in English; then they practice interviewing using the kinds of language, questions, and structure (text organization, discourse, etc.) that parallel the examples they have found. The same holds true for other speech acts: discussion, persuasion, negotiation, and more. Through such activities, they learn to replace the dictionary with real-life discourse models, an approach that most ILR-4 language users report as their preference.

Concerning defossilization techniques, teachers systematically work with students to correct embedded errors in grammar, syntax, phonetics, and register, to broaden the students' language domains, and to improve their communication strategies. Other kinds of fossilizations are also targeted, such as replacing learning strategies that are better used for lower levels (e.g., use of a dictionary or compensation strategies) with more advanced learning strategies (e.g., finding and using authentic texts as models and metacognitive strategies). While not discussed much in the literature, there is also affective fossilization—teacher dependence, for example (Ehrman 2002).The nature of the CE curriculum also addresses these kinds of fossilization.

It is critical to recognize here that teaching to ILR-3 and higher cannot be "more of the same" i.e., it cannot simply be a continuation of what are generally successful practices at the lower levels of proficiency. At the higher levels, the focus is on the text as well as the form, on refined knowledge versus use of context to gain knowledge, and on both extensive and intensive reading, perhaps with the same text. Emphasis is on self-direction versus teacher direction, on associative memory versus rote memory, on in-country immersions that replicate a native speaker's experience rather than attending a language school abroad or simply roaming the streets, and, where something is not understood, in elicitation versus requests for repetition. At these levels, coming close is not good enough—not in expression, morphology, syntax, phonetics, intonation, register, cultural behavior, social reaction, internal emotions, or many other areas that give a person away as a nonnative speaker. In many cases, what is needed to teach upper-level students is counterintuitive, and DLIFLC has a robust faculty development program that helps upper-level teachers adjust their approach to teaching as well as help students adjust their approach to learning.

Immersion Experiences

As stated earlier, the CE curriculum includes immersion (study abroad), although immersion experiences are not routine for all languages and all courses at CE. Languages for which immersion experiences are available are Russian, Chinese, Korean, and Arabic. All students in programs that have immersion opportunities are expected to participate in the immersion programs.

CE immersions last from two to four weeks. The activities of the immersion program differ substantially from both the DLIFLC undergraduate (basic language acquisition) program and typical university immersion programs. Common elements across the CE immersion programs include the following:

- classes are taught by content professors, not language teachers;
- students receive the same lectures as native-speaker students;

- CE teachers accompany students and each evening prepare them for the classes they will attend the next day (see description below);
- learners live the campus experience;
- peer tutors are available;
- excursions and social activities relate to the lesson topics;
- learners have real-life experiences;
- learners write daily reports;
- homestays are arranged; and
- learners make a formal presentation about their immersion experiences upon return.

Having classes taught by content professors (e.g., professors of law, architecture, sociology, etc.) differs considerably from enrolling students in language classes in private sector language schools (or university second language learning programs for foreigners) abroad. First, the professors generally use the same lectures as they use for their own students (in fact, they are asked to do this) so what the language students hear is at ILR-4 and ILR-5, a stretch for most of them, but for which they are prepared the night before by the accompanying DLIFLC teacher. Second, the professors often have not worked with foreign students before and so do not know how to adjust their language; as a result, students learn to adapt to the native speaker. In the process, the students' coping skills develop rapidly, making the students more comfortable with native discourse.

To facilitate student adjustment to native discourse, a teacher is generally needed. Most American students are visual learners (Bennett 1979; Dunn and Dunn 1979); many foreign classrooms are oriented toward auditory learners. For this reason, CE asks that the professors give the reading assignments in advance of class rather than after class, normally ten to twenty pages reflecting highly erudite language (ILR-4 or ILR-5). Obviously, students cannot independently cope with this amount of material in the sixty to ninety minutes available at the end of the day for homework. The accompanying CE teacher assists the students in understanding the material by teaching them different reading strategies, especially those for quickly reading extensive texts, and through preparing and rehearsing questions for the next day's class so that students can ask their questions without slowing down the class or annoying a professor not used to foreign students.

CE immersion programs take place at university campuses. During the time that they are on campus, students are able to interact with native-speaking students, meeting them at snack bars, the cafeteria for lunch, or other campus locations. In most programs, peer tutors from among the student body are hired to help students with their social language; the convenience of the campus location fosters this kind of work.

Lectures generally take place in the morning. (In Arabic immersions, there is generally an additional morning class, taught by a language teacher, introducing students to the local dialect; dialect work begins before students head to the country, and while they are there, having a language teacher to answer specific questions about expressions they hear on the street is very helpful.) After lunch, on many days,

students go on an excursion directly tied to the morning lecture. If, for example, the lecture is about law, the students may visit a law office and, in the case of Arabic countries, ask the lawyers how they integrate government law and *shariya* law. If the lecture is about the military, in some countries it is possible to visit a military post and compare foreign and US military practices. If the lecture is about a social phenomenon, the excursion might be to the office of a journalist writing about the topic. Of course, there are also historical and cultural tours, trips to restaurants, shopping, and other everyday activities. The goal is to provide students with real-life experiences. (On one occasion, a relative of the accompanying teacher was getting married, and the students were invited to the wedding.)

At the end of every class day, students write an analysis of what they have learned, using high-level language. The accompanying teacher has them correct and rewrite any grammatical or lexical errors or Anglicizations, if necessary. The final product is sent back to the teaching team in the United States, where the essays are collected for students' portfolios. (Portfolios are one of the tools used for grading CE students.)

In some countries, it is possible to arrange homestays. Where this is possible, students benefit immensely, especially on the four-week immersions. (Initially, when working with a new country or institution, CE administrators send students for two weeks, building up to three weeks, and then to four weeks when the institutions prove capable of meeting CE student needs.)

Upon return, students are expected to report on their experiences in a formal presentation in the target language with slides, pictures, handouts, and artifacts for the faculty and other students. Sometimes, other members of the community who speak the language are invited to attend these briefings.

Students participating in immersion programs receive a DA prior to the immersion experience and immediately after it. Typical progress in immersion programs is 0.3–0.5 point in proficiency improvement per skill per student for a two- to four-week experience.

Assessments

CE incorporates what practitioners refer to as "alternative assessments" into its language program. The term first began to circulate among teachers in the mid-1990s. Alternative assessment is "classroom assessment of learner language ability using a variety of tools other than tests" (Austin and Campbell 2004, 100). The most widely used alternative assessments are

- portfolios;
- teacher observations;
- peer assessments;
- self-assessments;
- short-term work products;
- short-term projects;
- presentations, exhibitions, skits; and
- conferences (one or more exchanges between teacher and student where the teacher evaluates the student's progress using criteria).

An alternative assessment approach promoted by Wiggins (1998, 12) in the late 1990s, "educative assessment," is also practiced at CE. Educative assessment has two underlying principles:

First, assessment should be deliberately designed to teach (not just measure) by revealing to students what worthy adult work looks like (offering them authentic tasks). Second, assessment should provide rich and useful feedback to all students and to their teachers and it should indeed be designed to assess the use of feedback by both students and teachers.

Two forms of formative assessment are used on a regular basis, as well: DA (as described above) and recall protocols (RP). They are among the tool set available to teachers but are not the only tools teachers use to monitor and improve the progress of their students.

DAs are used at three to five points in the course: initially, for planning the curriculum; midpoint in the course, for adjusting the curriculum; and one month before graduation, for final alterations in the curriculum. Where immersion is part of the course, DAs are used before and after immersion to document progress. A dynamic testing component was added to the DA in 2011 in order to determine the ever-changing ZPD of individual students. Following the final DA, students prepare a learning plan for the post-DLIFLC language sustainment and enhancement activities that will be overseen by their command language program manager wherever they are assigned. This plan takes into account students' learning styles, interests, likely assignments, and linguistic strengths and weaknesses. The students write the plan in the foreign language and design their own activities with faculty correction for linguistic accuracy.

Research programs (RPs) are typically used for conducting research on language processing (reading and listening). However, CE (and DLIFLC) has adapted them to be used as a diagnostic tool to find trends in how students inappropriately process language. Afterward, faculty individualizes instruction to facilitate proper processing. For example, in languages where the verb carries the prominent information, students may mistakenly be focusing on nouns, based on the influence of their native language. Moving students into a more natural way of processing the foreign language typically helps them make rapid gains in comprehension.

At times, students will arrive at CE and expect the normal battery of summative tests, whether those be multiple-choice exams, daily vocabulary quizzes, or some other manner of assigning a grade to a student's progress in learning. At CE, however, testing is not core to the program and, where used, is formative in nature. This approach sometimes creates difficulties (anxieties, even frustration) for students who are convinced that the best way to score well on the DLPT is to take a practice DLPT daily. Based on personal experience, most teachers at CE believe that students who learn the language do well on the DLPT, but those who practice test taking do not necessarily do well on the DLPT. So, several actions are taken to reduce students' "test craving," including explanation, counseling, and training in mental management.

Homework is varied, depending on student needs, and more frequently than not focuses on independent research and preparation of periodic projects. Students

present these projects both orally and in writing and are evaluated by the instructional staff. Daily quizzes and the like are rarely used; the time in the classroom is spent on learning, and the independent time is spent not in test preparation but on project preparation. As stated above, formal presentations are also made upon return from study abroad experiences; outsiders are invited to these presentations, which stretch students' language skills.

Faculty Development

The goal of faculty development in CE is to make each faculty member a "Renaissance person" adept in teaching, curriculum development, assessment, and professional development. Given the general lack of experience in teaching at higher levels of proficiency, CE faculty development includes a strong dose of bootstrapping, in which instructors teach each other new skills through such faculty development activities as sharing of curricula and lesson plans, colloquia, routine analysis of success and failure of students, and peer observation by specialists who are not the faculty members' supervisors. Nearly all teachers in CE have become competent in using tools advantageous to upper-level learning, including periodic DAs, use of recall protocol, applied technology in the classroom (involving online language teaching), accompanying students abroad (and teaching them reading strategies to match their reading speed and comprehension to that required of a native speaker enrolled in a college course), mechanisms to enhance defossilization, and the four pillars of rapid language acquisition: task-based instruction, content-based instruction, learner-centered instruction, and diagnostically oriented instruction.

Conclusion

Byrnes's 2002 definition of the AL2 user is still valid today:

1. someone who is able to draw on a sizable repertoire of fixed or chunked language forms that will ensure the sociocultural appropriateness and acceptability of utterances while, simultaneously, enhancing their fluency, and
2. someone who also has a good command, including a significant level of metalinguistic awareness, of the meaning—form relationships that are encoded with various degrees of fixity and fluidity at the lexicogrammatical level, thereby giving the impression of fluent but also thoughtful, online, situation creation of meaningful language. (51–52)

Whether in academe, government, or industry, AL2 teachers take on the challenge of facilitating learning at the higher levels of proficiency daily.

Keeping in mind that much of teaching for higher levels is counterintuitive and counter to contemporary teaching practices, AL2 teachers can glean ideas for how to get students to the highest levels of proficiency by studying successful language programs like the classroom-based ones at DLIFLC, FSI, and other US government language institutions; university and other academic programs that are presented elsewhere in this volume; and, where study abroad can be a strong component of a program, the recently emerging Flagship programs. While more research is always

desired and a greater range of practices is always welcome, in the past thirty years, since the first introduction of a high-level proficiency course at FSI, much has become known about how teachers can actually "teach" to these higher levels; however, if a greater number of students are to achieve near-native proficiency, more research—both empirical and action based—is needed. Chance and in-country experience alone will not get most students to these high levels of language ability. Quality teaching and quality programs that embrace counterintuitive approaches to learning rather than doing more of the same can achieve that, as demonstrated by a growing number of experiences.

Notes

Disclaimer: The opinions expressed within this article are those of the authors alone and not necessarily those of the DLIFLC or the US government.

1. *ILR* refers to the Interagency Language Roundtable's scale known as the ILR Proficiency Level Descriptions. The latest version of the ILR levels can be found at www.govtilr.org.
2. CALL was established as a collaborative effort by US government language schools to reach out to academe and share experiences. Established around 1989, it continued for more than a decade before closing. By then, routine collaboration was ongoing between government agencies and academic organizations, and academic organizations had begun to participate in the ILR.
3. The ACTFL scale differs from the ILR in that it emphasizes the lower levels of proficiency, which have more divisions (low, mid, high) than the ILR scale which has only two divisions (base, plus—e.g., 1, 1+, 2, 2+ . . .). At the higher levels of proficiency, the ACTFL scale has fewer divisions than the ILR, where the ACTFL scale, until recently, topped out at Level 3 (Superior, including all levels from ILR-3 through functionally equivalent of a highly articulate well-educated native speaker—or listener, reader, writer—[FEHAWENS]), whereas the ILR scale has the same level of gradations throughout the scale (3, 3+, 4, 4+, 5, with 5 being a FEHAWENS). In recent years, ACTFL has looked to reinsert another upper-proficiency level, "Distinguished" (ILR-4), which had appeared in the original iterations of the scale but quickly disappeared for lack of need for that level of granularity at the upper levels among those using the ACTFL scale.
4. In Beyond Methods: Macrostrategies for Language Teaching, Kumaravadivelu (2003) posits that the second and foreign language learning profession is now in a "postmethod era" after recognizing during the last decade of the twentieth century that (a) there is no best method yet to be discovered, (b) the historic distinction between theory and practice has been problematic for teachers, (c) teacher education models that only disseminate knowledge do not produce effective teaching professionals, and (d) teacher beliefs significantly influence the content and character of the daily practice of teaching.
5. The federal government uses a categorization system that divides languages into four categories according to the difficulty they pose to the native speaker of English to learn. For example, Category 1 languages include all the Romance languages and Swahili. Category 4 languages are Arabic, Chinese, Korean, Japanese, and Pashto. By varying the length of instruction by category, students are, in general, fully challenged even if they find themselves in one of the lower categories because the time to achieve the same proficiency level as students in the higher categories is considerably shorter.

References

Angelelli, Claudia, and Christian Degueldre. 2002. "Bridging the Gap between Language for General Purposes and Language for Work: An Intensive Superior-Level Language/Skill Course for

Teachers, Translators, and Interpreters." In *Developing Professional-Level Language Proficiency,* edited by Betty Lou Leaver and Boris Shekhtman, 77–95. Cambridge: Cambridge University Press.
Austin, Teresa, and Christine Campbell. 2004. "Introduction to the New Visions Initiative: Articulations and Assessment." In *New Visions in Foreign and Second Language Education*, edited by Gerd Bräuer and Karen Sanders, 92–106. San Diego: LARC Press.
Bailey, Kathy, and Donald Freeman. 1994. Workshop. DLIFLC, Presidio of Monterey, CA.
Bennett, Christine. 1979. "Teaching Students as They Would Be Taught: The Importance of Cultural Perspective." *Educational Leadership* 36: 259–68.
Bidawid, Elsaid. 2002. "The Quest for the Level 4+ in Arabic: Training Level 2–3 Learners in Independent Reading." In *Developing Professional-Level Language Proficiency,* edited by Betty Lou Leaver and Boris Shekhtman, 56–76. Cambridge: Cambridge University Press.
Byrnes, Heidi. 2002. "Toward Academic-Level Foreign Language Abilities: Reconsidering Foundational Assumptions, Expanding Pedagogical Options." In *Developing Professional-Level Language Proficiency*, edited by Betty Lou Leaver and Boris Shekhtman, 34–58. Cambridge: Cambridge University Press.
Byrnes, Heidi, and Hiram Maxim, eds. 2003. *Advanced Foreign Language Learning: A Challenge to College Programs.* Boston: Thomson and Heinle.
Caudery, Tim. 2002."Teaching High-Level Writing Skills in English at a Danish University." In *Developing Professional-Level Language Proficiency*, edited by Betty Lou Leaver and Boris Shekhtman, 177–96. Cambridge: Cambridge University Press.
Coalition of Distinguished Language Centers. 2006. *What Works: Helping Students Reach Native-Like Second-Language Competence.* Salinas, CA: MSI Press.
Dabars, Zita, and Olga Kagan. 2002. "Teaching Russian Language Teachers in Eight Summer Institutes in Russian Language and Culture." In *Developing Professional-Level Language Proficiency,* edited by Betty Lou Leaver and Boris Shekhtman, 219–42. Cambridge: Cambridge University Press.
Defense Language Proficiency Test, Version 5. 2008. Presidio of Monterey, CA: DLIFLC.
Dunn, Rita, and Kenneth Dunn. 1979. "Learning Styles/Teaching Styles: Should They . . . Can They . . . Be Matched?" *Educational Leadership* 36: 238–44.
Ehrman, Madeline. 2002. "Understanding the Learner at the Superior-Distinguished Threshold." In *Developing Professional-Level Language Proficiency,* edited by Betty Lou Leaver and Boris Shekhtman, 145–59. Cambridge: Cambridge University Press.
Gardner, Robert, and William Lambert. 1972. *Attitudes and Motivation in Second Language Learning.* Rowley, MA: Newbury House Publishers.
Higgs, Theodore, ed. 1984. *Teaching for Proficiency, the Organizing Principle.* Vol. 15 of *The ACTFL Foreign Language Education Series.* Lincolnwood, IL: National Textbook Company.
Jasser, Amal, Rajai Al-Khanji, and Betty Lou Leaver. 2005. "A Demographic Analysis of Arabic-Speaking Students Who Achieve Native-Like Levels in English Writing Skills." *Journal for Distinguished Language Studies* 3: 41–62.
Kumaravadivelu, B. 2003. *Beyond Methods: Macrostrategies for Language Teaching.* New Haven, CT: Yale University Press.
Lantolf, James, and Michael Poehner. 2004. *Dynamic Assessment in the Language Classroom.* University Park, PA: Center for Advanced Language Proficiency Education and Research, Pennsylvania State University.
Leaver, Betty Lou. 2003. *Achieving Native-Like Second Language Proficiency.* Salinas, CA: MSI Press.
———. 2013. Presentation. DLIFLC, Presidio of Monterey, CA.
Leaver, Betty Lou, and Sabine Atwell. 2002. "Preliminary Qualitative Findings from a Study of the Processes Leading to the Advanced Professional Proficiency Level (ILR-4)." In *Developing Professional-Level Language Proficiency*, edited by Betty Lou Leaver and Boris Shekhtman, 260–79. Cambridge: Cambridge University Press.
Leaver, Betty Lou, and Paula Bilstein. 2000. "Content, Language, and Task in Content-Based Instruction." In *Languages across the Curriculum: Interdisciplinary Structures, Intersections of Knowledge, and Internationalized Education,* edited by Maria-Regina Kecht and Katherina von Hammerstein, 79–118. Columbus: Ohio State University Press.
Leaver, Betty Lou, and Boris Shekhtman, eds. 2002. *Developing Professional-Level Foreign Language Proficiency.* Cambridge: Cambridge University Press.

Leaver, Betty Lou, and Boris Shekhtman. 2002. "Principles and Practices in Teaching Superior-Level Language Skills: Not More of the Same." In *Developing Professional-Level Language Proficiency*, edited by Betty Lou Leaver and Boris Shekhtman, 3–33. Cambridge: Cambridge University Press.

Lyman-Hager, Mary Ann, and Christian Degueldre. 2004. "Establishing a Center for Level 4 Study: Issues and Agendas." In *Teaching and Learning to Near-Native Levels of Language Proficiency: Proceedings of the Spring and Fall 2003 Conferences of the Coalition of Distinguished Language Centers,* edited by Betty Lou Leaver and Boris Shekhtman, 115–26. Salinas, CA: MSI Press.

National Standards for Foreign Language Learning: Preparing for the 21st Century. 1996. Yonkers, NY: National Standards in Foreign Language Education Project.

Nunan, David. 1988. *The Learner-Centered Curriculum.* Cambridge: Cambridge University Press.

———. 1997. "Designing and Adapting Materials to Encourage Learner Autonomy." In *Autonomy and Independence in Language Learning,* edited by Phil Benson and Peter Voller, 192–203. London: Longman.

———. 2006. *Go for It!* Boston: Heinle ELT.

Tortora, Margherita, and Maria Crocetti. 2005. "An Interdisciplinary Approach to Advanced through Distinguished Language Proficiency: Bridge Courses." In *Teaching and Learning to Near-Native Levels of Language Proficiency III: Proceedings of the Fall 2005 Annual Conference of the Coalition of Distinguished Language Centers*, edited by Inna Dubinsky and Donna Butler, 59–64. Salinas, CA: MSI Press.

Vygotsky, Lev. 1978. "Interaction between Learning and Development." In *Mind and Society: The Development of Higher Psychological Processes*, edited by Michael Cole, Vera John-Steiner, Sylvia Scribner, and Ellen Souberman, 79–91. Cambridge: Harvard University Press.

Wiggins, Grant. 1998. *Educative Assessment: Designing Assessments to Inform and Improve Student Performance.* San Francisco: Jossey-Bass Publishers.

This page intentionally left blank

Chapter 2

L1, L2, and Cognitive Development: Exploring Relationships

DAN DEWEY, RAY CLIFFORD, AND TROY COX

SECOND LANGUAGE LEARNERS OFTEN express frustration over the reality that their limited second language (L2) skills prevent them from full participation in intellectual discussions, and some have exclaimed, "Hey, I'm smarter than I sound!" Those who have had this experience recognize that a relationship exists between one's language ability and others' perceptions of one's cognitive abilities. In daily life, this relationship between language and cognition is so prevalent that it is often ignored, or the abilities are simply conflated. For example, calls by employers to improve the critical thinking skills of college graduates are typically exemplified by statements about language ability, such as the ability to speak and write accurately, coherently, and persuasively.

Whether one is attempting to accomplish high-level communication tasks in a first language (L1) or L2, the successful completion of those communication tasks likely requires both linguistic and cognitive skills. Therefore, effective communication may be inhibited by the lack of advanced language skills, by the absence of well-developed critical thinking skills, or by a combination of deficiencies in both areas.

For instance, for individuals to successfully accomplish the American Council on the Teaching of Foreign Languages (ACTFL) Superior/Interagency Language Roundtable (ILR) Level 3 communication tasks, they must provide evidence through their spoken or written communications that they possess both the linguistic and cognitive abilities inherent in those communication acts. However, when individuals fail at those communication tasks, questions arise as to what contributed to that failure: 1) limited linguistic skills; 2) undeveloped cognitive abilities; or 3) a combination of limited linguistic and cognitive skills.

Context for Assessing Cognition and Foreign Language Proficiency

To assess language proficiency, we opted to use the ACTFL Oral Proficiency Interview (OPI) and the ACTFL Writing Proficiency Test (WPT), two of the most prominent and widely accepted measures of language proficiency used in the United States. These exams are broadly employed in educational and other settings, and similar instruments are utilized by the US government for screening, promotion, and other high-stakes purposes. In spite of the wide acceptance of these exams, there has been some controversy regarding the exams' validity and use. An issue that has been discussed regularly is the possibility that poor performance on these tests may reflect a combination of limitations in one's L2 linguistic ability and an inability to perform required tasks even in one's L1. As Clark and Lett (1988, 61) point out, a major concern "is that examinees vary widely in the extent to which they are perceived as effective communicators in their native language"—if learners cannot communicate well in their L1, how can they be expected to do so in a L2? To which we might add the question, "Why can't they communicate well in their L1?" Calls for improving the linguistic skills of college graduates usually describe the students' inability to use abstract reasoning and critical thinking skills. Regarding undergraduate students' cognitive abilities, Blanc, DeBuhr, and Martin (1983) note:

> Recent evidence suggests that 50 percent of entering college freshmen have not attained reasoning skills at the formal (abstract) operational level described by Piaget and Inhelder. Students who appear to operate at the concrete (nonabstract) level consistently have difficulty processing unfamiliar information when it is presented through the abstract media of lecture and text. Their questions about materials are often detailed and superficial. Rarely do they ask or answer questions that require inference, synthesis, or application. They can operate at more advanced levels once they have mastered the concept, but they require regular instruction that either anchors the concept directly in the student's previous experience or provides a concrete experience with data from which the concepts may be drawn. (82)

More recently, Conley (2007) states that college readiness for language consists of the ability to reason and to provide solid arguments and proof. Specifically, it is expected that "The student constructs well-reasoned arguments of proofs to explain phenomena or issues; utilizes recognized forms of reasoning to construct an argument and defend a point of view or conclusion; accepts critiques of or challenges to assertions; and addresses critiques and challenges by providing a logical explanation or refutation, or by acknowledging the accuracy of the critique or challenge" (13). Conley further reports that many professors at US universities feel that undergraduates are inconsistently or inadequately equipped with these abilities, and that such skills need to be fostered both prior to and during the undergraduate experience.

Perkins's provocative 1985 study, titled "Post-primary Education Has Little Impact on Informal Reasoning," argues that little progress is typically made in these

areas by undergraduate students during their college careers. Perkins's data are outdated, but the more recent work by Conley (2009) suggests that this lack of formal reasoning skills is still an issue of concern today. Given the limited reasoning abilities described by these authors, it is easy to question whether undergraduate second language learners would have the cognitive and linguistic abilities needed to discuss abstract topics, support and defend opinions, and deal with situations that require higher-level argumentation. Because higher-level reasoning skills are also imbedded in the Superior level performance expectations of the OPI and WPT, the apparent gaps in these skills that regularly exist in undergraduate students raise the question, "Do students have the reasoning skills necessary to perform at the Superior level even in their L1?" Table 2.1 depicts the ACTFL hierarchy of communicative tasks (based on ACTFL Guidelines), the corresponding cognitive abilities listed in Bloom's Revised Taxonomy, and the abilities that undergraduate students are expected to have in order to be successful in their studies (Anderson et al. 2001). Similarities between the ACTFL hierarchy's higher-level skills and the abilities required for successful undergraduate studies suggest that learners ought to bring at least some of these abilities to their undergraduate experiences.

Determining what role one's ability to reason through speaking and writing in one's L1 plays in L2 performance allows for a better understanding of how to approach L2 instruction. Language educators (particularly at the university level) commonly engage learners in class activities that are associated with ACTFL's descriptors of Advanced- and Superior-level proficiency. Two recent articles (Brown 2009; Brown et al. 2009) describe such approaches and exemplify the widespread interest in helping learners achieve Advanced or higher levels of proficiency in undergraduate foreign language programs through the practice of higher-level skills. Advanced- and Superior-level speech involves actions such as narrating in a cohesive and coherent manner, discussing abstract topics, supporting and defending opinions, and dealing with linguistically challenging or unfamiliar situations. By understanding whether learners possess the abilities necessary to perform such higher-level tasks in their native languages, we can determine whether it is necessary to focus attention on these L1 skills in addition to providing learners with a set of L2 linguistic tools (vocabulary, grammar, cultural knowledge, etc.) that can be used when attempting to accomplish these tasks in the L2.

For this exploratory study, we used the Overclaiming Questionnaire (OCQ) to assess cognitive development (Paulhus and Harms 2004). This measure of intelligence was selected largely for practical reasons. As Paulhus and Harms point out, this evaluation can be administered to learners at home (i.e., remotely by web browser) without sacrificing any form of legitimacy. Individuals completing the OCQ view 150 terms or phrases and indicate their degree of familiarity with these items. In the Paulhus and Harms validation study, they found strong and significant correlations between OCQ results and results from multiple standardized IQ measures. They point out that the OCQ is practical (administered in fifteen minutes or less and requiring minimal or no supervision), nonthreatening (no one correct answer and no clinician observing or making judgments as the task is completed), and robust

◆ Table 2.1 ACTFL Hierarchy of Communication Tasks and the Related Development of Cognition and Knowledge

	ACTFL Proficiency Guidelines		Corresponding Elements from Bloom's Revised Taxonomy Cognitive Process Dimension	Cognitive Abilities Expected of US Undergraduates (based on Conley 2009 and Perkins 1985)
Level	**Speaking Tasks**	**Writing Tasks**		
Novice	Can communicate short messages using memorized words and phrases.	Can produce lists, notes, and limited formulaic information on simple forms and documents. Writing is typically limited to words, phrases, and memorized material.	Remember (Retrieve from memory)	
Intermediate	Can recombine learned material to express simple statements/questions about familiar topics of a personal or predictable nature.	Can meet a range of simple and practical writing needs, e.g., simple messages and letters, requests for information, notes, etc. Can communicate simple facts and ideas in a loosely connected series of sentences on topics of personal interest and social needs, primarily in the present. Because of vocabulary limitations and errors in basic structures, writing is comprehensible to those accustomed to the writing of nonnatives.	Understand (Construct meaning)	Memorize and state facts
Advanced	Can describe people, places, and things in the major time frames of past, present, and future. Can provide detailed, concrete, factual narrations about events and activities. Can give coherent directions or instructions.	Can write routine, informal, and some formal correspondence, narratives, descriptions, and summaries of a factual nature in all major time frames in connected discourse of a paragraph in length. Writing is comprehensible to all native speakers due to breadth of generic vocabulary and good control of the most frequently used structures.	Apply (Carry out, use)	Exchange ideas (when more concrete topics and factual narrations are involved) Conduct experiments (largely involves following factual directions, acting questions about procedures, etc.)

Level	ACTFL Proficiency Guidelines		Corresponding Elements from Bloom's Revised Taxonomy Cognitive Process Dimension	Cognitive Abilities Expected of US Undergraduates (based on Conley 2009 and Perkins 1985)
	Speaking Tasks	Writing Tasks		
Superior	Evaluate ideas and provide structured arguments to support one's opinions. Develop hypotheses to explore alternative possibilities. Discuss topics from an abstract perspective.	Can produce informal and formal writing on practical, social, and professional topics treated both abstractly and concretely. Can present well-developed ideas, opinions, arguments, and hypotheses through extended discourse. Can control structures, both general and specialized/professional vocabulary, spelling, punctuation, cohesive devices, and all other aspects of written form and organization with no pattern of error to distract the reader.	Analyze (Distinguish, organize) Evaluate (Judge appropriateness) Create (Hypothesize)	Engage in reasoning, argumentation, and proof Defend a point of view Description of a point of view (abstract topic/perspective) Evaluation of a variety of perspectives Analyze conflicting explanations, supporting arguments, etc. Offer explanations Weigh and synthesize to best judge the soundness of a claim Seek reasons with a formal bearing on a claim in an attempt to resolve the truth of the claim.

(high in reliability and exhibiting strong correlations with a variety of IQ measures under a number of conditions).

Research Questions

In sum, because our aim is to help learners develop Advanced and Superior level L2 proficiency, it is beneficial to know whether those who fall short do so because they merely lack sufficient L2 linguistic abilities (i.e., those described at the higher levels in Table 2.1) or because their lack of observable cognitive/reasoning abilities in their L1 is affecting their ability to demonstrate these skills in their L2. These ideas lead us to the questions of this study:

1. What is the relationship between L1 proficiency and L2 proficiency test scores?
2. What is the relationship between the cognitive difficulty of the task and the amount of language produced in L1 vs. L2?
3. What is the relationship between cognitive development and language proficiency?

Review of Literature

As we explore the cognitive abilities of undergraduate L2 learners, we first review the history of the OPI, the WPT, and the implementation of the current higher-level communicative requirements. We then review literature describing the typical cognitive abilities of US undergraduates.

The Development of the OPI and WPT as Measures of Academic Language Abilities

Concern for L2 speaking proficiency that adequately met the demands of an academic setting appeared as early as 1928. One such evidence is the College Board's recommendation that testing the English proficiency of learners coming from abroad to US universities should include a portion assessing learners' ability "to understand spoken English, and to express his thoughts intelligibly in spoken English" (Spolsky 1995, 56). Furthermore, the Board recommended that these same learners be capable of "speaking well enough to take part in class discussions" (56). In short, they recommended that students be assessed on their ability to communicate and reason through speaking in the college classroom. In spite of these suggestions, the Board's outline for a test of this nature included only an oral test where an examiner was to rate "fluency, responsiveness, rapidity, articulation, enunciation, command of construction, or connectives, usable vocab and the use of idiom." The rater was also to report the degree to which the candidate appeared to be "diffident or shy" (57). Also to be included was a 250- to 300-word composition written on an academic topic that was rated according to similar criteria.

Much progress has been made since these early efforts at assessing speaking and writing. In the years immediately following World War II, the United States shifted from evaluating knowledge of grammar, vocabulary, reading comprehension,

and listening comprehension (principally receptive skills) to evaluating the ability to perform in real-life situations. This desire to see learners' performance in real-life situations eventually resulted in the US Foreign Service Institute (FSI) OPI, an oral test that involved describing pictures, giving a short talk without preparation, and responding to recorded audio prompts. Results were then rated according to a checklist of features that included accent, grammar, vocabulary, comprehension, and fluency. This test ultimately transitioned into a face-to-face interview designed to elicit speech samples according to a hierarchy of increasingly difficult, real-world tasks. One impact of this task hierarchy was that it created a situation in which the cognitive demands placed on the examinees made monitoring language use progressively difficult, resulting in the elicitation of speech samples that reflected the examinee's internalized language system and nonrehearsed material.

As described elsewhere in this volume, the transitions in government testing were eventually reflected in an ACTFL-led initiative to produce a similar global scale for use in academic settings (Martin; Leaver and Campbell; Jackson). These revised scales required learners to demonstrate higher-level cognitive abilities as they progressed up the scale.

To summarize, the ACTFL, OPI, and WPT are the result of several decades of transitioning from largely multiple-choice discrete point tests, to assessments of ability, to integrated tests that range from the simple parroting of learned material to cognitively intense tasks that require examinees, to present thoughts and ideas in unique and extemporaneous ways (for more on the move from discrete point tests to more highly cognitively demanding testing in the United States, see Fulcher 2003 or Spolsky 1995)

US Undergraduates' Cognitive Abilities

Conley's (2009) study of college readiness describes the following core ability expected of US undergraduate students:

> Reasoning, argumentation, proof: The student constructs well-reasoned arguments of proofs to explain phenomena or issues; utilizes recognized forms of reasoning to construct an argument and defend a point of view or conclusion; accepts critiques of or challenges to assertions; and addresses critiques and challenges by providing a logical explanation or refutation, or by acknowledging the accuracy of the critique or challenge. (13)

Conley elaborates that as part of this ability, the student "presents orally or in writing an extended description, summary, and evaluation of varied perspectives and conflicting points of view on a topic or issue" (13). He notes that college instructors expect students to be able to "make inferences, interpret results, analyze conflicting explanations of phenomena, support arguments with evidence, solve complex problems that have no obvious answer, draw conclusions, offer explanations, conduct research, engage in the exchanging of ideas, and generally think deeply about what they are being taught" (6).

Perkins's (1985) earlier study, which addressed the contribution of university education to the types of skills that Conley described, focused on informal reasoning,

which "involves considering a claim and seeking reasons with a nonformal bearing on the claim, pro or con, in an attempt to resolve the truth of the claim" (562). Further, in informal reasoning, "reasons typically occur on both sides of the case, no one line of argument settles the truth of the claim, and no computational procedure assigns the claim a numerical probability. The reasoned must weigh and synthesize to best judge the soundness of the claim" (ibid). Perkins found only minimal effects (i.e., only a marginal significant effect in ability to provide an explanation) of college education on informal reasoning. More substantial growth was seen in high school, and for those who went on to graduate school, greater development occurred during that period. Conley (2009) notes that incoming freshmen typically lack many of these abilities, which suggests that universities may need to be responsible for helping learners develop them if such abilities are necessary to meet the demands of higher-level courses and work beyond the college setting.

Given the demands of the OPI and WPT at the Advanced and Superior levels described previously, learners would need to demonstrate the informal reasoning skills Conley and Perkins described to receive high ratings on these two measures. The question then arises, "Do learners taking the OPI and WPT possess these skills?"

Significant relationships have regularly been found between bilingualism and L1 cognitive abilities. People who are proficient in more than one language consistently outscore monolinguals on measures of verbal and nonverbal intelligence (Bruck et al. 1974; Hakuta 1986; Weatherford 1986). Furthermore, L2 learners tend to have better listening skills and working memory than their monolingual counterparts (Morales et al. 2013; Ratte 1968). Bilingual children have been shown to have great cognitive flexibility, better higher-order thinking skills, and greater problem-solving skills than monolingual children (Hakuta 1986). Both children and adults who study a second language show evidence of greater executive control (i.e., decision-making processes) and creative abilities than their peers (Bialystok et al. 2012; Bamford and Mizokawa 1991; Lee et al. 2012). In fact, adults who spend as little as a semester abroad return showing greater evidence of creative abilities than they did prior to their study abroad experience (Lee et al. 2012). Leung and Chiu (2010) found that exposure to a different culture for a short instructional period could enhance learners' creative conceptual abilities significantly. In short, learning another language and being exposed to another culture can have significant cognitive benefits and affect the learners' abilities to demonstrate intelligent thinking.

The Study

Now that we've mapped out the cognitive and reasoning abilities expected of OPI and WPT test takers and those expected of US undergraduate students, we'll describe a study designed to evaluate and compare learners' L1 and L2 abilities and explore the relationships between

1. The learners' L1 and L2 proficiencies;
2. The cognitive difficulty of the communication tasks presented and the amount of language produced when performing those tasks in the learners' L1 and L2; and
3. The learners' level of cognitive development and the learners' L1 and L2 test performance.

Participants

Participants consisted of 108 learners of Spanish as a second language (88 male, 20 female, mean age 21.7 years, $SD = 1.54$ years) enrolled at Brigham Young University (BYU) in Spanish 321: Third-Year Spanish Reading, Grammar, and Culture. Spanish 321 is a required course for all Spanish majors and minors and the first class that learners take after returning from an abroad experience in a Spanish-speaking country. Taking Spanish 321 is also a requirement for those learners who wish to take challenge exams and receive course credit for their language learning experience abroad.

Seven of these learners had studied abroad for a semester in a Spanish-speaking country and seventy-seven had lived abroad for eighteen months to two years as church volunteers. All learners were born and raised in North American English-speaking communities, but two claimed Spanish as their native language instead of English.

Proficiency Measures

The OPI "is a standardized procedure for the global assessment of functional speaking ability. It is a face-to-face or telephonic interview between a certified ACTFL tester and an examinee that determines how well a person speaks a language by comparing his or her performance of specific communication tasks with the criteria for each of ten proficiency levels described in the *ACTFL Proficiency Guidelines 2012—Speaking*" (ACTFL 2013, paragraph 1). The OPI is a carefully structured interview designed to find a learner's floor (level at which a learner can consistently perform comfortably and confidently) and ceiling (level at which some ability is shown, but breakdown occurs regularly). By considering this floor and ceiling, a tester is able to assign the following ratings: Novice Low, Novice Mid, Novice High, Intermediate Low, Intermediate Mid, Intermediate High, Advanced Low, Advanced Mid, Advanced High, and Superior. The interview typically lasts less than thirty minutes, but it can be much shorter for the lower levels (for a more detailed description of the OPI process and interpretation of score result see ACTFL 2012). Because a rater assigns a holistic rating to the entire interview, analyzing the effect that the cognitive demands of a single communication task have on the examinee responses can pose a challenge and exceed the scope of this study.

The ACTFL WPT is a standardized test for global assessment of functional writing abilities in a second language. The test measures "how well a person spontaneously writes in a language (without access to revisions and/or editing tools) by comparing his or her performance of specific writing tasks with the criteria stated in the *ACTFL Proficiency Guidelines—Writing* (Language Testing International 2013,

"What are the Writing Proficiency Tests and the Business Writing Tests?", paragraph 1). The WPT consists of four or five written prompts in English "dealing with practical, social, and/or professional topics that are encountered in informal and formal contexts" (paragraph 2). Prompts are selected largely based on a learner's self-assessment of his or her abilities and are targeted to elicit the type of writing representative of a specific proficiency level. Thus, an Intermediate prompt would elicit sentence-length language that would be novel recombinations of learned material, whereas a Superior prompt would elicit responses that would use academic vocabulary to make well-reasoned arguments to support an opinion and hypothesize alternate possibilities. The WPT is rated similarly to the OPI. The rater judges the responses to the different leveled prompts and gives a holistic rating based on a floor and ceiling approach. Because each WPT comprises a range of prompts targeted at different language proficiency/cognitive levels, the examinee responses allow a more in-depth inquiry of the relationship between the cognitive demands and language development. As the test is timed, the length of the response can be an overall estimator of the automaticity with which the examinee can complete the task. This relationship will be measured by the number of words in each response.

As described earlier, we administered the OCQ to measure the learners' cognitive ability (Paulhus and Harms 2004), and we utilized the common-sense scoring technique for scoring the questionnaire. Completing the questionnaire was voluntary, and of the 108 participants in the study, 58 took the OCQ.

Procedures

Learners were required to take the Spanish language exams (both OPI and WPT) as part of their course. They were encouraged to take the English versions of the OPI and WPT as practice for their Spanish exams. All exams were proctored by BYU's Center for Language Studies and followed official ACTFL-required procedures. Exams were double- or triple-scored according to ACTFL's rating processes. A demographic survey was administered to all volunteer participants at the time of their OPI and WPT testing. This survey typically took less than ten minutes to complete.

The OCQ was administered to provide data regarding students' cognitive abilities beyond what we were able to gather from the language tests alone. The hope was the language tests alone would yield a range of analytical reasoning abilities; however, given that students tended to score at or near the Superior level on their L1 tests, little variability was seen, which made comparisons of these abilities in the L1 difficult. The OCQ was given six to eight months after the first round of language testing. Personal communication with Paulhus, the primary author of the OCQ, suggests that this measure is likely to change very little over an eight-month period.

Participants were given $10 for completing a demographic questionnaire and a personality test not used in this study. They also were given up to $20 for the English WPT ($10 for completing the test, $5 for an Advanced rating, and $10 for a Superior rating) and up to $20 for the English OPI (also $10 for completing the test, $5 for an Advanced rating, and $10 for a Superior rating). They received no monetary motivation or compensation for completing the OCQ.

Results

In this section, we compare results for L1 and L2 OPI and WPT scores, look at the relationship between the cognitive difficulty of the communication task and the amount of language produced, and investigate connections between the OCQ and L1 and L2 test results.

The Relationship Between L1 and L2 ACTFL OPI and WPT Scores

The ACTFL ratings of Novice Low to Superior were converted to a 10-point scale starting with 1 as Novice Low, 2 as Novice Mid, concluding with 10 representing Superior (for rationale and discussion of this numeric conversion see Rifkin 2005). Though the language gain between the differing levels is more geometric than additive, this numeric scale is still useful for examining trends and variation in the data.

For the OPI, the mean score on the English tests was 9.75 (SD = 0.48), with median and mode at the Superior level (see Figure 2.1). For the Spanish OPIs, the mean was 7.32 (SD = 0.83), with the median and mode at the Advanced Low level. A Wilcoxon Matched-Pairs Signed-Ranks Test was conducted with the OPI results for both languages. The English scores were significantly higher than the Spanish ($Z = -8.810$; $p < 0.001$), with one hundred instances of the subjects scoring higher in English. There were three instances in which subjects had the same rating in

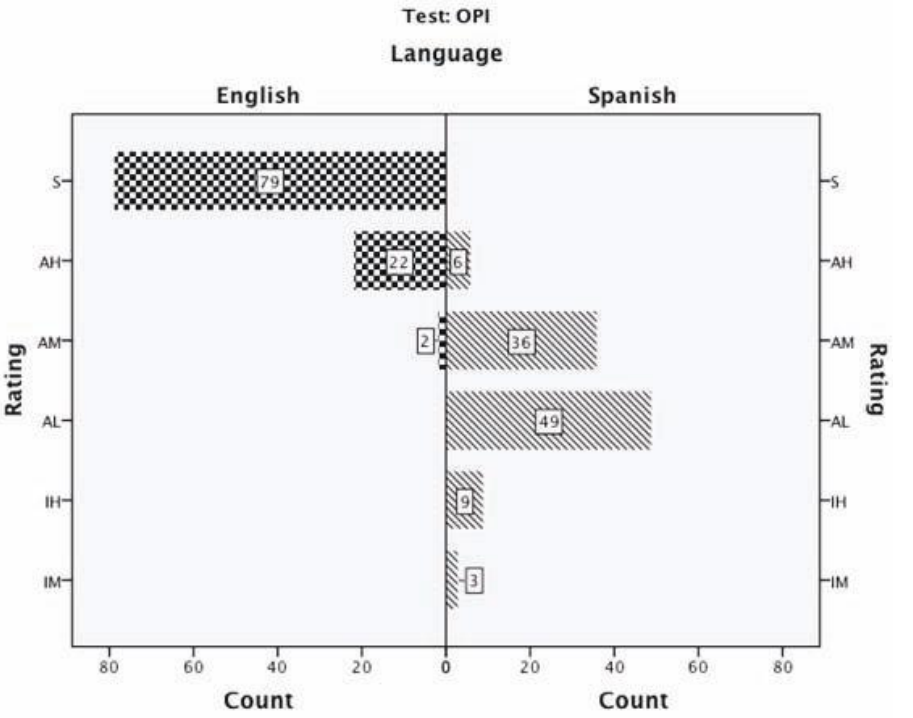


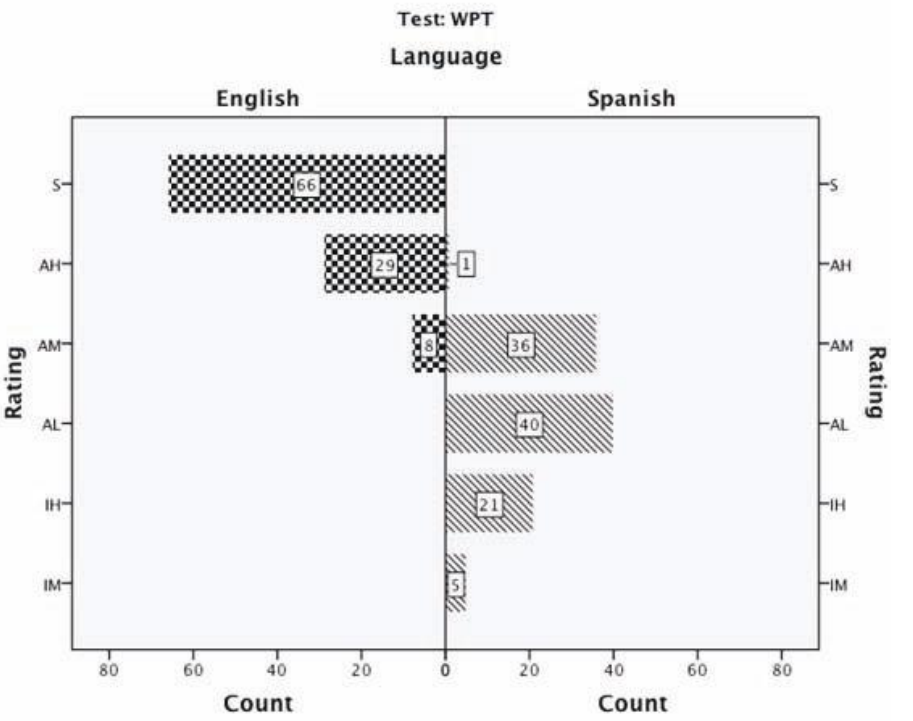
Figure 2.1 Oral Proficiency Ratings for Subjects in English (Native Language) and Spanish (Second Language)

English and Spanish (2 at the Advanced High level and 1 at the Advanced Mid level) and no instances in which someone scored higher in Spanish than in English.

For the WPT, the mean score on the English tests was 9.56 (SD = 0.64), with median and mode at the Superior level (see Figure 2.2). For the Spanish, the mean was 7.07 (SD = 0.89), with the median and mode also at the Advanced Low level. A Wilcoxon Matched-Pairs Signed-Ranks Test was likewise conducted with the WPT for both languages. For the WPTs, the English scores were also significantly higher than the Spanish ($Z = -8.944$; $p < 0.001$), with 102 instances of the subjects scoring higher in English. There was only one instance in which a subject had the same rating in English and Spanish (Advanced Mid) and no instances in which someone scored higher in Spanish than in English.

The Relationship Between Cognitive Difficulty of the Communication Task and the Amount of Language Produced

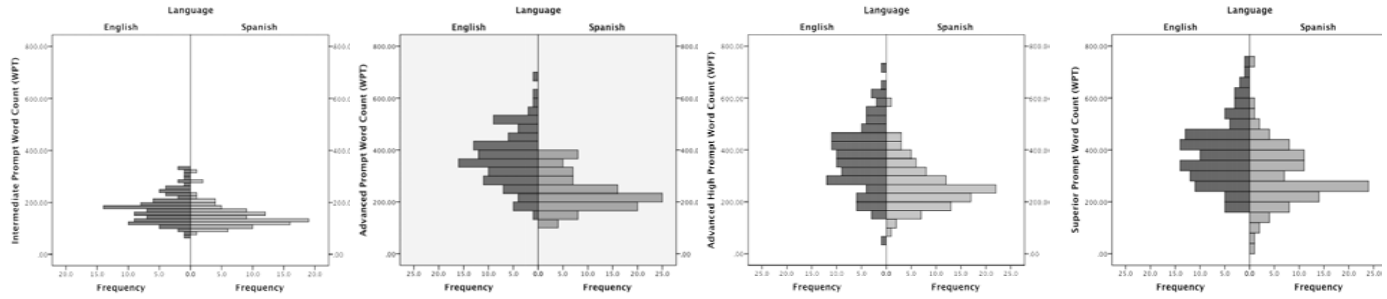
To examine how the cognitive demands of the task affected the amount of language output of L1 compared to L2, we analyzed the number of words produced on the WPT, using the intended prompt difficulty level as our measure. At each prompt difficulty level, participants were able to produce more language in their L1 than in their L2 (see Table 2.2). The Intermediate level prompt had the least amount



◆ Figure 2.2 Writing Proficiency Ratings for Subjects in English (Native Language) and Spanish (Second Language)

◆ Table 2.2 Writing Proficiency Test Word Count by Prompt Type

	Intermediate		Advanced		Advanced High		Superior	
	English	Spanish	English	Spanish	English	Spanish	English	Spanish
N	102	100	103	100	103	100	103	100
Mean	178.04	146.10	367.36	236.12	368.74	262.45	385.96	297.12
SD	56.4	40.57	103.56	69.71	125.25	83.48	120.81	112.2
Mean Difference between English and Spanish	31.94		131.24		106.29		88.84	



of difference, with the subjects writing on average 31.94 more words in English than in Spanish. The largest mean difference (diff) was with the Advanced prompt (diff = 131.24), followed by the Advanced High prompt (diff = 106.29), and then the Superior prompt (diff = 88.84). These data were somewhat surprising because the mean difference between the number of words produced in L1 versus L2 did not follow the increasing pattern observed at the lowest levels and actually declined when subjects were dealing with the Advanced High- and Superior-level prompts.

The Relationship Between OCQ (Intelligence) Results and the Proficiency Test Results

To determine whether IQ and proficiency test results were related, results for the IQ test and each of the proficiency measures were converted to dichotomous variables that indicated high or low performance. For the purposes of this study, IQ scores more than one standard deviation below the mean were considered low. Scores above this were considered relatively high (although we recognize that learners with average IQs were also marked high as a result). For the English ACTFL measures, Superior was considered high and anything below Superior low. For the Spanish exams, Advanced Mid or above was marked as high and anything below as low. Table 2.3 shows the chi-square results for the proficiency and IQ combinations for the English OPI. The results indicate that high- and low-proficiency learners differ significantly in terms of their IQs ($\chi^2 = 3.90$, $df = 1$, $N = 58$, $p = 0.047$). The data suggest that high-proficiency (Superior level) English speakers are more likely to have high IQs than low-proficiency learners. Phi, which indicates strength of association between the two variables, is −0.259; thus, the effect size is considered to be small to medium according to Cohen (1988). None of the other chi-square tests (Tables 2.4, 2.5, and 2.6) showed statistical significance.

In short, the comparisons between the OCG measure of intelligence and the measures of L1 and L2 proficiency showed a significant connection between

◆ Table 2.3 Chi-Square Results for L1 English OPI (Proficiency) and OCQ (IQ)

			Proficiency			
Variable		***N***	**High**	**Low**	χ^2	***p***
IQ	High	32	27	5	3.90	0.047
	Low	26	16	10		
Totals		58	43	15		

◆ Table 2.4 Chi-Square Results for L1 English WPT (Proficiency) and OCQ (IQ) Results

			Proficiency			
Variable		***N***	**High**	**Low**	χ^2	***p***
IQ	High	32	23	9	2.91	0.076
	Low	26	13	13		
Totals		58	36	22		

◆ Table 2.5 Chi-Square Results for L2 Spanish OPI (Proficiency) and OCQ (IQ) Results

Variable		*N*	Proficiency High	Proficiency Low	χ^2	*p*
IQ	High	32	17	15	1.24	0.198
	Low	26	10	16		
Totals		58	27	31		

◆ Table 2.6 Chi-Square Results for L2 Spanish WPT (Proficiency) and OCQ (IQ) Results

Variable		*N*	Proficiency High	Proficiency Low	χ^2	*p*
IQ	High	32	15	17	0.121	0.468
	Low	26	11	15		
Totals		58	26	32		

intelligence/cognitive ability and L1 speaking proficiency. Although the connection between the OCQ and L1 writing was not statistically significant, it was much closer to being significant than the L2 connections.

Discussion

Earlier in this paper, the cognitive abilities required of those achieving various scores on the OPI and WPT were established, along with descriptions of how measurements of these reasoning skills via these two tests developed over time. The abilities undergraduate students in the United States typically possess and the skills desired by undergraduate faculty were also outlined. In analyzing the data, connections between L1 and L2 proficiency levels were sought after. There was clearly a ceiling effect in the measure of L1 proficiency and a dearth of students with Superior-level L2 proficiency, which limited the ability to define relationships between L1 and L2 proficiency levels. However, the data do suggest that the development of L2 proficiency lags behind L1 and that the level of one's L1 proficiency might serve as an upper limit for L2 development.

The high L1 results raised other important questions. In spite of the fact that Conley and Perkins found that undergraduates often lack many of the cognitive abilities (in particular, verbal argumentation and reasoning skills) required by the OPI and WPT, only 24 (23 percent) of the 104 learners who took the English OPI scored lower than Superior. Of these 24, 22 scored Advanced High, indicating they were able to perform many of the Superior-level tasks; however, they had not reached the point at which they could consistently respond at that level. The remaining 2 speakers demonstrated Superior-level reasoning abilities at least some of the time (enough to demonstrate emerging abilities), but they were able only to consistently maintain Advanced-level skills throughout their interviews.

In light of previous findings, it is encouraging to see higher-level reasoning skills consistently present in the L1 for a great majority of these language learners. The fact that learners did so well on the L1 measures contributed to a ceiling effect (i.e., produced little variation in the L1 scores), which prevented an ability to better evaluate relationships between L1 and L2 abilities. However, it also demonstrated that these L2 learners do largely possess the reasoning skills necessary to perform well on the OPI and WPT and might, therefore, be able to transfer these L1 abilities to the L2 tests. Brown's (2009) study demonstrated that L2 learners who were encouraged to practice higher-level reasoning through debate in the L1 were able to benefit and transfer these skills to their L2. At the very least, learners' L1 skills ought to benefit when higher order L1 communication skills are practiced in preparation for addressing those higher-level tasks in the L2.

Conclusions

Despite limited cell sizes in some of the outcome categories, the results of this exploratory study lead to some tentative observations about the relationships among the variables: L1, L2, and Cognitive Development.

1. What is the relationship between L1 proficiency and L2 proficiency test scores? The number of students not demonstrating Superior-level communication skills was very small, but the pattern that emerged might lead to the hypothesis that the level of L1 proficiency one attains might serve as an upper boundary in the attainment of proficiency in one's L2.
2. What is the relationship between the cognitive difficulty of the communication tasks presented and the amount of language produced in L1 and L2? The number of words used by those successfully accomplishing the tasks at each proficiency level indicates that producing successful responses to higher-order tasks requires on average more words than are required for lower-level tasks and that the L1 and L2 responses at the Superior level are closer in length than are the responses at the Advanced level.
3. What is the relationship between cognitive development and language proficiency? A significant correlation was found between higher IQ results and higher L1 proficiency results in speaking, and the correspondence with L1 writing results approached significance.

Implications for Future Research

Several possible research directions are indicated by the results of this study. First, one could use the same instruments to examine the L1 abilities of graduates from a broader range of colleges and universities. Participants in our study were taking a third-year Spanish course, which is typically taken after several years of enrollment at the university or significant time spent abroad as students or church volunteers. The development that occurs during those years following entrance likely contributed to higher OPI scores. In addition, even freshmen at this university may

have relatively high cognitive abilities, given their profile upon entrance. BYU's 2013 acceptance rate was 56.2 percent, and incoming freshmen in that same year had an average grade point of 3.82 on a 4-point scale and an average ACT score of 28.52 (near the 90th percentile). Depending on their admission policies, other universities may have students demonstrating a broader range of cognitive abilities.

In addition to a more varied L1 profile, recruiting subjects with a broader range of L2 ability would allow a more in-depth analysis of the relationship between L1 and L2 development. Secondly, additional research using subjects with more varied L2 abilities will likely provide an understanding of learners' proficiencies. In this study, there was little variation in terms of L2 scores on the OPI, which added to the difficulty of making connections between L1 and L2 abilities. It is possible that our group was rather homogeneous in terms both of L1 and L2 abilities. However, as some (e.g., Luoma 2004; Rifkin 2005) have pointed out, due to the noncompensatory scoring guidelines, learners with the same rating can possess a wide variety of vocabulary knowledge, grammatical competence, pragmatic ability, etc. For this reason, a more fine-tuned analysis of learners' language abilities, such as performance grids, might be necessary in order to better capture and highlight variation in the learners' language development and pinpoint relationships between L1 and L2 proficiency. It may also be necessary to construct other tailored measures of language ability to better tease out relationships that exist among L1 capabilities, L2 proficiency, and higher order cognitive abilities.

The third research thrust might be considered a "byproduct" of the current study. Given the strong reasoning skills that participants demonstrated in their L1 test scores, an important question arises: "Might the Superior L1 reasoning abilities have been fostered or even enhanced by learners' study of the L2?" The research reviewed earlier suggested that US undergraduates lack many of these reasoning skills, yet that did not appear to be the case with L2 learners in our study (Blanc 1983; Conley 2009; Perkins 1985). We do know that L2 learners can benefit cognitively in their L1 from acquiring an L2 (e.g., Lee et al. 2012; Morales et al. 2013; Ratte 1968), but to our knowledge, studies connecting the acquisition of L2 proficiency with the attainment of higher L1 cognitive abilities, such as those demonstrated via an L1 OPI, have not yet been conducted. Following adult learners carefully from the beginning of their L2 study and measuring the possible impact of such study on L1 abilities could prove beneficial. Or given that many typical L2 learners in the United States begin their study at the university level, a comparison of L1 cognitive and linguistic skills of learners who entered a university or college and pursued a L2 versus those who did not would be informative. Would the two groups vary in terms of the development of their L1 cognitive abilities over the course of their undergraduate studies? If it were to emerge that L2 learners demonstrated significantly more development in their L1 written and oral reasoning skills than other students, this finding would make a good argument for the inclusion of L2 study in undergraduate education. Any future studies could also utilize upper levels of the proficiency scales (ILR 4 and 5 and ACTFL Distinguished), pushing learners to higher levels and allowing for greater variability in L1 outcomes as well.

References

American Counsel on the Teaching of Foreign Languages (ACTFL). 2013. "Testing for proficiency." Accessed October 22. http://www.actfl.org/professional-development/certified-proficiency-testing-program/testing-proficiency.

American Counsel on the Teaching of Foreign Languages (ACTFL). 2012. "Oral Proficiency Interview familiarization manual." Accessed October 22. http://www.languagetesting.com/wp-content/uploads/2012/07/OPI.FamiliarizationManual.pdf.

Anderson, Lorin W., David R. Krathwohl, Peter W. Airasian, Kathleen Cruikshank, Richard E. Mayer, Paul R. Pintrich, James Raths, and Merlin C. Wittrock. 2001. *A Taxonomy for Learning, Teaching, and Assessing.* New York: Longman.

Bamford, Kathryn W., and Donald T. Mizokawa. 1991. "Additive–Bilingual (Immersion) Education: Cognitive and Language Development." *Language Learning* 41: 413–429.

Bialystok, Ellen, Fergus I. Craik, and Gigi Luk. 2012. "Bilingualism: Consequences for Mind and Brain." *Trends in Cognitive Sciences* 16 (4): 240–250.

Blanc, Robert A., Larry E. DeBuhr, and Deanna C. Martin. 1983. "Breaking the Attrition Cycle: The Effects of Supplemental Instruction on Undergraduate Performance and Attrition." *The Journal of Higher Education* 54: 80–90.

Brown, N. Anthony. 2009. "Argumentation and Debate in Foreign Language Instruction: A Case for the Traditional Classroom Facilitating Advanced-Level Language Uptake." The *Modern Language Journal* 93: 534–549.

Brown, N. Anthony, Jennifer Bown, and Dennis L. Eggett. 2009. "Making Rapid Gains in Second Language Writing: A Case Study of a Third-Year Russian Language Course." *Foreign Language Annals* 42: 424–452.

Bruck, Margaret, Wallace E. Lambert, and Richard Tucker. 1974. "Bilingual Schooling through the Elementary Grades: The St. Lambert Project at Grade Seven." *Language Learning* 24 (2): 183–204.

Clark, John L., and John Lett. 1988. "A Research Agenda." In *Second Language Proficiency Assessment: Current Issues*, edited by Pardee Lowe Jr. and Charles Stansfield, 53–82. Englewood Cliffs, NJ: Prentice Hall Regents.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

Conley, David T. 2007. *Redefining College Readiness.* Eugene, OR: Educational Policy Improvement Center. Accessed Nov 11, 2013. http://www.aypf.org/documents/RedefiningCollegeReadiness.pdf.

Fulcher, Glenn. 2003. *Testing Second Language Speaking.* New York: Pearson Longman.

Hakuta, Kenji. 1986. *Mirror of Language.* New York: Basic Books.

Language Testing International. 2013. "Writing Proficiency Test." Accessed Nov 11. http://www.languagetesting.com/writing-proficiency-test.

Lee, Christine S., David J. Therriault, and Tracy Linderholm. 2012. "On the Cognitive Benefits of Cultural Experience: Exploring the Relationship between Studying Abroad and Creative Thinking." *Applied Cognitive Psychology* 26 (5): 768–778.

Leung, Angela Ka-yee, and Chi-yue Chiu. 2010. "Multicultural Experience, Idea Receptiveness, and Creativity." *Journal of Cross-Cultural Psychology* 41: 723–741.

Luoma, Sari. 2004. *Assessing Speaking.* Cambridge: Cambridge University Press.

Morales, Julia, Alejandra Calvo, and Ellen Bialystok. 2013. "Working Memory Development in Monolingual and Bilingual Children." *Journal of Experimental Child Psychology* 114: 187–202.

Paulhus, Delroy L., and Peter D. Harms. 2004. "Measuring Cognitive Ability with the Over-Claiming Technique." *Intelligence* 32: 297–314.

Perkins, David N. 1985. "Postprimary Education Has Little Impact on Informal Reasoning." *Journal of Educational Psychology* 77: 562–571.

Ratte, Elizabeth H. 1968. "Foreign Language and the Elementary School Language Arts Program." *The French Review* 42 (1): 80–85.

Rifkin, Benjamin. 2005. "A Ceiling Effect in Traditional Classroom Foreign Language Instruction: Data from Russian." *The Modern Language Journal* 89 (1): 3–18.

Spolsky, Bernard. 1995. *Measured Words: The Development of Objective Language Testing*. Oxford, UK: Oxford University Press.

Weatherford, H. Jerold. 1986. *Personal Benefits from Foreign Language Study*. (ERIC Document Reproduction Services 276305.) Washington, DC: ERIC.

This page intentionally left blank

Part Two

Approaches to Maximizing Language Gain at Home and Abroad

This page intentionally left blank

Chapter 3

Developing Superior Language Proficiency and Analytical Skills for the Federal Workforce

DEBORAH KENNEDY AND CHRISTA HANSEN

THE DEVELOPMENT OF GENERAL professional proficiency (Interagency Language Roundtable [ILR] Level 3 proficiency) in a second or additional language is the goal of many adult language learners, particularly those who intend to use the linguistic and cultural knowledge in their professional careers. Business, industry, and government all have need of personnel with professional-level skills in two or more languages. In particular, many branches of the US government actively seek to employ persons who possess general professional proficiency in both English and another language and thus can conduct research and analysis that is informed by deep cultural and linguistic knowledge.

This chapter describes the English for Heritage Language Speakers (EHLS) program, which was designed to meet the need for individuals with such professional proficiency. Through an annual recruiting and application process, EHLS admits US citizens who are native or heritage speakers of a language identified by the government as critical to national security. The eight-month program gives those admitted the opportunity to develop the professional English language skills needed to obtain and retain federal positions in which they can apply their linguistic and cultural skills.

Background on the Program

EHLS had its origin in legislation authored by the House Permanent Select Committee on Intelligence in 2005. Recognizing that many US citizens possessed native proficiency in languages critical to national security, but lacked opportunities to develop their English skills to the professional level, the legislation provided for

the establishment of an English language instructional program that would provide such skills. The legislation specified that the program would be housed at an institution of higher education and would enable naturalized US citizens with native or Superior-level heritage language proficiency in a language critical to national security to develop the professional English skills needed to fulfill vital positions in federal government agencies.

The overall design for the program was developed by the Center for Applied Linguistics (CAL) in 2005–2006. CAL distributed a survey on language use and language needs to senior personnel in the Department of Defense, the Department of State, and the Intelligence Community and interviewed government and private sector colleagues at meetings of the ILR, the Coalition of Distinguished Language Centers (CDLC), and other conferences. Using the survey and interview results, CAL identified representative task types and the corresponding requirements or expectations for proficiency in both the heritage language and English. The collected information made clear that the program's target should be general professional proficiency (ILR Level 3) in both the native/heritage language and English.

Because the program was not intended to provide instruction in languages other than English, CAL recognized that participants would need to possess ILR Level 3 proficiency in a native or heritage language at entry. They also would need to have resided in the United States for at least five years in order to become naturalized. Therefore, participants would most likely be early- to mid-career professionals who had received education at least through secondary school in their first language and had come to the United States as adults. Accordingly, they would need a program that removed them from the workforce for a relatively short period of time and that provided financial support to enable them to participate.

CAL further recognized that the achievement of general professional proficiency in English in a short-term program requires a level of focus and reinforcement that can only be provided through an immersive style of instruction (Jackson and Kaplan 2001; Long 1983; Malone et al. 2005). Therefore, CAL designed EHLS as a six-month intensive program with approximately thirty hours per week of instruction. The first program iteration took place in 2006; in 2009, the structure was revised to include two additional months of part-time instruction following the six-month intensive portion. The part-time instruction focused on further development of writing skills, career skills, and federal job search skills—areas in which early program participants demonstrated a need for ongoing support. This chapter reports on the instructional curriculum and program outcomes for program years 2009–2012.

The EHLS program is funded and supervised through the National Security Education Program (NSEP), part of the Defense Language and National Security Education Office of the US Department of Defense. CAL provides oversight and management of the program; instruction is designed and provided by the Center for Language Education and Development at Georgetown University.

Foundations of Program Design

The overall design and the instructional curriculum for the EHLS program at Georgetown were developed by faculty at the Georgetown University Center for Language Education and Development with guidance from CAL.[1] The starting point for program design was the concept of language proficiency as understood in the ILR skill level descriptors: what a person is able to do with language in terms of tasks and functions, content and context, accuracy, and text type (Omaggio Hadley 2001), or "the ability to use language as a tool to get things done" (Jackson and Kaplan 2001, 72). Adopting the ILR proficiency guidelines with their focus on language use as the program's framework entailed basing instruction on the theoretical construct of communicative competence (Hymes 1971; Munby 1978), a connection recognized widely in the language teaching community (Chapelle et al. 1997; Leaver and Shekhtman 2002; Spolsky 1989). With the ILR proficiency guidelines as the framework and with communicative competence as the overarching purpose of instruction, CAL and Georgetown then sought to answer three major questions:

1. What does communicative competence at ILR Level 3 (that is, general professional proficiency) look like? What specific abilities (competencies) does a Level 3 listener/speaker/reader/writer possess and demonstrate?
2. What program design and instructional approaches promote the development of communicative competence at ILR Level 3?
3. What assessment methods provide effective formative support and appropriate summative evaluation of an adult language learner's proficiency level?

Profile of Communicative Competence in General Professional Contexts

Project staff at CAL and Georgetown recognized that the answer to the first of these questions would define the goals of instruction and thus dictate the answers to questions two and three. For this reason, an exploration of communicative competence at ILR Level 3 was the first step in program development.

Practitioners whose instruction is informed by the notion of communicative competence have generally focused on the four subareas identified by Michael Canale and Merrill Swain (Canale 1983; Canale and Swain 1980; Savignon 1997):

- grammatical competence—the ability to use a language's grammatical, syntactic, and lexical features
- discourse competence—the ability to understand and structure sequences of language in culturally appropriate ways
- sociolinguistic competence—the ability to understand and meet sociocultural expectations for language interaction
- strategic competence—the ability to negotiate meaning, manage unfamiliar language, and apply communication strategies

In exploring the ILR Level 3 descriptors in relation to these competence areas, project staff noted several critical points (all quotes from Interagency Language Roundtable, n.d.):

- At this level, grammatical (or linguistic) competence is extensive enough to allow an exclusive focus on meaning: "An S-3 can effectively combine structure and vocabulary to convey his/her meaning accurately." In particular, the interlocutor's focus is not diverted from content to form: "An S-3 uses the language acceptably, but with some noticeable imperfections; yet, errors virtually never interfere with understanding and rarely disturb the native speaker."
- Discourse competence includes control of long sequences: "Discourse is cohesive. . . . Use of structural devices is flexible and elaborate." Discourse competence is extensive enough to permit effective participation in "most formal and informal conversations on practical, social, and professional topics" and delivery of "extended, elaborate and informative monologues."
- Sociolinguistic and strategic competence implies participating "effectively," using language "acceptably," and demonstrating the ability to manage situations in which "cultural references, proverbs, and the implications of nuances and idiom may not be fully understood."

These notations indicated that the development of general professional proficiency would require particular attention to refinement of linguistic and discourse competence, as well as to a relatively sophisticated level of sociolinguistic competence. This understanding aligned with the observations of other language professionals working with learners at the ILR 3 and ILR 4 levels. For example, a Generic Learner Profile for reading at Level 4 (developed by surveying Level 4 learners) stressed that discourse competence was "critical for attaining Level 4, especially understanding of text organization and genre differences" (Ingold 2002, 145). Moreover, structural competence and lexical competence, which together comprise grammatical competence, at Level 4 entail internalizing "the complete linguistic system" and "a near-native range of archaic, dialectal, obscure, and invented lexical terms" (ibid). Similarly, in interviews with speakers who had attained high levels of proficiency, Leaver discovered the importance of a focus on form and structure, noting that at these levels, "accuracy and attention to detail is very important" (Leaver 2004, 36). And Leaver and Shekhtman observed that "while all students need most of the components of communicative competence at any given time, there is a changing balance that occurs with proficiency gain" (2002, 10–11) and outlined the "relative balance of componential saliency along the continuum from Levels 0 to 4" (ibid.) with discourse, formal linguistic, and sociolinguistic components having greatest salience at Level 2+ and up.

Beyond the definition of the competencies that require attention at Level 3, project staff explored the types of skills and tasks expected at that proficiency level. Tasks listed in the ILR descriptors for Speaking 3 include "participat[ing] in formal and informal conversations on practical, social, and professional topics"; "discuss[ing] particular interests and special fields of competence"; "us[ing] the language as part of normal professional duties, such as answering objections, clarifying points, justifying

decisions, understanding the essence of challenges, stating and defending policy, conducting meetings, delivering briefings, or other extended, elaborate and informative monologues"; and "elicit[ing] information and informed opinion from native speakers" (ILR n.d.). Project staff also referred to the American Council on the Teaching of Foreign Languages (ACTFL) descriptors for writing at the Superior level, which reference the ability to "produce most kinds of formal and informal correspondence, complex summaries, précis, reports, and research papers on a variety of practical, social, academic, or professional topics treated both abstractly and concretely" and to "explain complex matters, provide detailed narrations in all time frames and aspects, [and] present and support opinions by developing cogent arguments and hypotheses" (ACTFL n.d.).

Finally, project staff investigated what was known about the nature of learners who attained high levels of proficiency in a second or additional language. Though learners can reach high levels of proficiency through a variety of language teaching methods and learning routes (Jackson and Kaplan 2001; Leaver et al. 2005), Wesche's 1977 study of learners in intensive programs showed that "the most successful learners are those who use their exposure time in the L2 actively and who seek to extend this outside the classroom," demonstrating the central roles of self-motivation and learner autonomy in second language acquisition. Additionally, Leow and Bowles demonstrated the "facilitative effects of attention and awareness on adult L2 learners' subsequent processing, intake, and learning of targeted L2 forms or structures embedded in the L2 data" (2005, 188). Evidence thus suggests that the most successful learners engage actively and attentively in the process of language acquisition and derive their motivation from a sense of self-efficacy as language learners (Ehrman 1998, 2002; Jackson and Kaplan 2001; Leaver et al. 2005; Shekhtman 2003).

Instructional Design for Promoting the Development of General Professional Proficiency

Guided by the ILR framework and the goal of communicative competence, CAL and Georgetown combined a communicative language teaching (CLT) approach with task-based language teaching (TBLT) methods in the design of the EHLS instructional program. CLT builds on the premise that language acquisition takes place through language use. It holds that interaction, with its concomitant need for negotiation of meaning, promotes language acquisition by encouraging learners to draw on the linguistic resources they already possess and extend them to new communication situations (Long 1981). In particular, the process of producing language helps learners recognize and reflect on both their linguistic strengths and holes in their linguistic repertoire (Swain 1985). Thus, CLT provides opportunities for learners to reflect on the language learning process and on their own experience as language learners. It also seeks to link classroom learning with language use outside the classroom (Nunan 1991).

CLT also stresses the use of authentic situations, contexts, and materials in instruction, as is typically done in English for specific purposes (ESP) instruction.

ESP instruction is "designed to meet the specified needs of the learner; related to the content of particular disciplines, occupations, or activities; and centered on the language appropriate to these activities in syntax, lexis, discourse, semantics, and the analysis of this discourse" (Johns and Price-Machado 2001, 44). For these reasons, "efforts to achieve maximum linguistic, strategic, and situational authenticity are made in designing curricula" (47). Here, CLT intersects with TBLT because ESP curricula utilize tasks to generate authenticity.

A task is a language teaching activity that involves some kind of information or opinion gap and has a defined outcome. Completion of a task requires learners to select appropriate (or available) linguistic resources and use them to achieve an expected outcome (Ellis n.d.). Task-based instruction is a "meaning-focused approach that reflects real-life language use," and, therefore, "task-based interaction stimulates natural acquisition processes" (Willis 2004, 8).

For the six-month intensive portion of the EHLS program, instruction is built around tasks drawn from the federal workplace, including participation in meetings, team research and development work, and provision of briefings. Instruction utilizes communicative activities that require the use of ILR Level 3 skills, such as analyzing an argument, distinguishing fact from opinion, stating and supporting a hypothesis, and identifying perspective and bias. Instruction is provided in four main courses, all of which stress improving English skills to complete tasks effectively.

Reading and Writing for Professionals (seven hours per week)

The goal of this course is to improve participants' ability to read, write, conduct research, and interact clearly and effectively in English in professional settings. Instructional activities reflect a mix of individual, independent assignments and joint efforts with colleagues on specific projects. For the capstone activity of the course, participants complete a major independent research project and report. On a weekly basis, participants read authentic materials and demonstrate comprehension by responding to questions in writing and follow-along in-class discussions; participants also have weekly writing assignments that focus on mastery of tasks such as email communications, summaries, memos, business letters, critical reviews, short reports, written briefings, abstracts, and research reports. Over the term of the course, participants

- Read a variety of materials, including government and business documents and authentic articles from journals, magazines, and newspapers;
- Summarize, discuss, and critically evaluate authentic reading materials;
- Practice reading strategies to increase comprehension and speed;
- Analyze discourse in government and business communications;
- Interpret graphic materials such as charts, diagrams, and tables;
- Guess the meaning of new vocabulary from context;
- Write summaries, abstracts, letters, memos, reports, critical reviews, and fact sheets;
- Communicate electronically with instructor, classmates, and agency mentors;
- Organize ideas and information logically and persuasively;

- Develop a systematic editing and revision process;
- Practice paraphrasing, quoting, and citing sources;
- Work as a member of a team to complete a project;
- Use research tools of the Georgetown library and the Internet; and
- Produce a research report with documentation.

Professional Oral Communication (six hours per week)

This course focuses on practicing the kinds of formal and informal presentations and team meeting communications that occur in government work. Instructional activities are a mix of individual, independent assignments—audio journaling, formal and informal speeches—and team-based work with colleagues on specific projects. The capstone activity of the course is a major independent research project and formal presentation. On a weekly basis, participants respond to comprehension questions after listening to authentic materials and prepare for in-class group discussions, team projects, negotiation activities, and formal speeches/presentations. Presentation activities focus on introductions, narratives, supported opinions, exposition, and critical review presentations, whereas team communication activities range from conducting meetings, to addressing problem participants, and negotiating tactics and conflict resolution. Over the term of the course, participants will

- Listen to a variety of materials, including professional presentations and interviews from professional online sources;
- Summarize content and present an opinion based on the listening material;
- Use critical thinking to discern a speaker's purpose, bias, and inferences;
- Participate in presentations by asking relevant questions and offering constructive criticism;
- Practice organizing ideas and information logically and persuasively;
- Practice paraphrasing sources;
- Utilize critical thinking skills to provide support for ideas and opinions and to engage in group problem-solving activities;
- Work as a member of a team to complete projects;
- Interact with classmates in pair and group discussions, meetings, and negotiations; and
- Prepare and deliver presentations on a variety of topics and in a variety of styles.

News Analysis (four hours per week)

In this course, participants focus on improving their analytical thinking, writing, and briefing skills by analyzing authentic news content in terms of objectivity, bias, currency, reliability, and accuracy. The end goal of such exercises is to produce accurate and concise written summaries and oral briefings. Language- and media-analysis tools are utilized to promote the production of increasingly sophisticated reports on the slant or "bias" of news content across sources and topics; use of these tools is extended to other texts, including a participant's own writing. On a weekly basis, participants read authentic news content that serves as the basis for reading

comprehension activities, analytical discussion, oral briefings, written and oral summaries, analytical reports, vocabulary building exercises, and language analysis activities that include vocabulary worksheets and quizzes. Over the term of the course, participants

- Read authentic news media content, as well as commentary and analysis from popular and scholarly sources;
- Summarize, discuss, and critically evaluate authentic reading material;
- Analyze news media discourse by utilizing frame, content, and sentiment analysis tools;
- Build vocabulary through study of high-frequency words, phrases, and idioms through the use of the Corpus of Contemporary American English (http://corpus.byu.edu/coca/);
- Guess from context the meaning of new vocabulary;
- Write summaries and short analytical reports;
- Organize ideas and information logically for summary writing;
- Develop a systematic editing and revision process;
- Practice paraphrasing, quoting, and citing sources;
- Set up personalized news and social media feeds for research purposes;
- Communicate electronically with instructor and classmates;
- Deliver short oral briefings on news content; and
- Deliver short presentations on the analysis of language and graphics used to frame news in different sources.

Career Skills (three hours per week)

This course provides workshop-style classes that focus on the specific skills required to find a position in the US government. Participants develop a professional portfolio and identify appropriate positions within federal agencies. Activities include using online tools for the job search, evaluating positions, researching agencies, preparing targeted documents, and honing interview skills. The companion *Introduction to Federal Government* course (two hours per week) is a weekly speaker series in which representatives from federal agencies and government contractors introduce program participants to their organizations and identify relevant employment opportunities. Over the term of the course, participants

- Speak in settings such as individual, pair, and group work; simulated job interviews; class presentations; and salary negotiations;
- Read a variety of career-related materials in order to understand main ideas and supporting points as well as to acquire specific vocabulary related to federal careers;
- Write effective professional/career search correspondence and professional documents such as résumés, knowledge-skills-abilities statements (KSAs), personal statements, cover letters, and agency research reports with cohesion, coherence, and control of grammar and mechanics;

- Listen in formal and informal settings, such as classroom interactions, simulated job interviews, career-related workshops, and lectures on government agencies, their missions, and their job requirements; and,
- Use discourse appropriate for job search activities and federal employment settings.

In the years since program inception, CAL and Georgetown have built relationships with a number of federal agencies that have provided further guidance on task types and language proficiency requirements through the *Introduction to Federal Government* course. Agency contributions have included presentations on the structure and content of oral and written briefings and similar tasks; provision of guidelines for preparing résumés and KSAs; and the inclusion of program participants in agency training opportunities. These contributions have corroborated the initial identification of ILR Level 3 as the target proficiency level and have enabled instructors to refine the instructional curriculum so that it prepares participants effectively for federal employment.

Instruction takes an integrated-skills approach that combines the four skill modalities of listening, speaking, reading, and writing in every task. In this integrated approach, tasks, activities, and content topics are coordinated across the first three courses using previously selected materials to promote development of "the levels of sophistication found at more formal levels of speech" (Leaver et al. 2005, 29). This integration has been refined over several years of program implementation and now is extensively coordinated. For example, the course instructors might coordinate activities relating to a specific US foreign policy topic as follows:

- *Reading and Writing for Professionals*: Participants read an extended article on a foreign policy topic; analyze the linguistic structures (for example, comparison and contrast) used to present positions and evidence; and summarize their analysis in a written report.
- *Writing Lab*: The written reports form the basis for activities in which participants practice the grammatical tools that are used for expressing comparison.
- *Professional Oral Communication*: In preparation for analyzing and making a group decision about a specific issue under discussion, participants listen to authentic materials on US foreign policy and use the Six Hats meeting decision-making technique (de Bono 1985) to practice looking at a decision from multiple perspectives.
- *News Analysis*: Participants examine news articles about US foreign policy for sources of bias and forms of objectivity, focusing on specific examples of intellectual humility/fair-mindedness and multidimensional thinking/sophistic objectivity (as opposed to "fake objectivity"—considering multiple viewpoints only for the purpose of defending a predetermined choice), and apply the content and frame analysis of designation (noun phrases, naming as framing) to the news articles.

The instructional program also incorporates an extensive amount of individualized instruction to address the specific needs of each participant (CDLC 2008). Individualized instruction occurs in three formats:

- *Oral Communication Tutorial*: A weekly one-hour tutorial with a *Professional Oral Communication* instructor for extensive feedback and individualized instruction on all aspects of oral communication (listening, speaking, intercultural competence) as a complement to the *Professional Oral Communication* course. Feedback is based on a review of participant-produced audio journals and video recordings of their presentations and team meeting interactions. Specific areas of focus include formative feedback on progress in responding coherently and objectively to opinion questions, presenting ideas and information effectively, understanding implications and purpose, inferring meaning, and detecting affective overtones. Comprehensibility and accuracy are addressed as needed.
- *Writing Lab*: The Writing Lab is designed to increase participants' ability to write accurately and appropriately in English for workplace communication. In the lab, program participants apply a systematic editing and revision process to writing assignments from the *Reading/Writing for Professionals* and *News Analysis* courses. Emphasis is placed on attention to grammar and mechanics, use of culturally appropriate content and style, and incorporation of feedback from instructors into the documents that are being rewritten.
- *Career Skills Tutorial*: In this weekly thirty-minute session, program participants receive individual assistance from the Career Skills instructor in identifying career goals, targeting appropriate agencies and positions, and completing application materials.

Formative Assessment of the Development of General Professional Proficiency

Given the high levels of learner autonomy required for the achievement of second language proficiency at the ILR 3 level and beyond, in developing the EHLS program CAL and Georgetown recognized the need to move beyond a basic curriculum design model. A larger instructional design approach that focused on the establishment of a learning-enabling environment was necessary (Merrill et al., 1996). By providing experiences connected with participants' professional backgrounds and goals, this environment would promote the participants' independence as learners and ownership of the learning process, as well as the participants' acquisition of high-level language and cultural skills.

For EHLS, CAL and Georgetown adopted a backward design model of instructional development. This three-step design model begins with identification of the knowledge and skills that learners are expected to develop, the types of tasks that they are expected to be able to complete as a result of participating in the educational program, and the target levels of mastery. The second step involves defining the types and levels of evidence that will demonstrate attainment of the target levels of mastery, and the third step deals with developing instructional materials and

activities that will promote the desired learning outcome. In the backward design model, learning of discrete material is always set in the context of a larger skill or understanding that is being developed, and regular instructor feedback, along with peer- and self-evaluation, keep learners focused and involved in the learning process (Wiggins and McTighe 2005, 2008).

For EHLS, target tasks include both general professional tasks, such as participating in meetings, teamwork, summarizing, synthesizing, and analyzing, and specific tasks that have been outlined by federal agencies, including oral and written briefings and analysis of news reports. Instructors create activities and assignments that involve these types of tasks, conduct ongoing formative assessment, and provide direct feedback to participants on their progress toward expectations and performance goals.

Course-specific language performance goals are used to plan the scope and sequence of instructional and formative assessment activities designed to facilitate attainment of consistent ILR performance within each course. This planning requires a thorough examination and sequencing of the texts, tasks, and performance criteria for each course to ensure full representation of the ILR 3 performance range in the curriculum. Additionally, hierarchies of critical thinking skills, rhetorical argumentation, and claim-making are utilized to identify the subcomponents of the range between ILR 2 and ILR 4 descriptors and thus plan an instructional cycle that moves from the personal, self-referential, concrete, and less complex to the more hypothetical, objective, abstract, and complex. For example, the rhetorical focus in *Reading and Writing for Professionals* becomes increasingly more sophisticated as the program progresses, going from listing and classifying, through cause-effect, definition, chronological sequencing, compare and contrast, and data commentary, to problem solution. As the *Professional Oral Communication* program progresses, participants are challenged to support their opinions with facts, expert sources, definitions, and statistics rather than relying solely on examples and anecdotes. The content analysis focus in the *News Analysis* course begins with gist and summary writing and progresses to applying frame, content, and sentiment analysis to the capstone research project.

Feedback and formative assessment is informed and guided by a language skills performance matrix that is built on the discourse and linguistic/sociolinguistic competences contained in the descriptors for ILR Level 3. Each course has a matrix that is specific to the modalities addressed: reading and writing in *Reading and Writing for Professionals*, listening and speaking in *Professional Oral Communication*, and all four modalities in *Career Skills* and *News Analysis*. The language goals in the *News Analysis* matrix are tailored to a relatively narrow range of text and task types: news articles and broadcasts, oral and written briefings, and summaries.

The thorough analysis and coordination of texts, tasks, and performance criteria allows instructors to provide targeted feedback on assignments and in-class activities and to focus instruction on the skills and strategies a scholar must utilize to consistently demonstrate performance at the ILR 3 level. Instructors meet with participants every seven to eight weeks to discuss their overall progress toward their language goals. For these sessions, instructors gather information from a variety of sources: in-class activities, out-of-class assignments, and end-of-session cumulative

achievement assessments. End-of-session assessments range from timed reading and writing activities in the *Reading and Writing for Professionals* course, to assessment of application of frame analysis concepts in *News Analysis*, to an interactive team project in *Professional Oral Communication*. This design gives the instructor multiple data points to use in determining how well and how often the participant met the ILR 3 criteria for each descriptor. This understanding is particularly important given that attainment of an ILR 3 requires sustained accuracy over multiple events rather than a one-time demonstration of proficiency.

The capstone assessment activity for the six-month intensive portion of the program is an Open Source Analytical Research Project (OSAP) required of each participant. These projects are carried out in collaboration with federal agencies, which provide research topics that draw on participants' language, educational, and professional backgrounds. The agencies also provide mentors who work with participants throughout the research and development process. Ideally, the federal agency mentor provides guidance on relevant, nonclassified materials and sources. The mentor may also review the participant's progress at key points. Work on the OSAP is incorporated into instruction throughout the second half of the intensive component. Participants conduct their research in English and their native language and are encouraged to conduct interviews with knowledgeable individuals in order to augment the information they gain from online and print sources. Each participant develops an analytical report in a format prescribed by a federal agency and gives a presentation at a formal symposium at the end of the six-month intensive portion of the program. The symposium is attended by the participants' federal agency mentors and a variety of other representatives from federal agencies and government contractors. The written report is subsequently given to the agency mentors and published internally on a secure federal government website accessible to the intelligence community and other federal agencies.

OSAP projects conducted in 2012 included the following (abstracts drawn from the symposium program):

- *Online Activities by Islamist Groups in Egypt: Evolution and Trends.* Egyptian Islamists' use of social media testifies to their increased awareness of the value of these tools in spreading their message and polishing their image. Analysis of the content of their online platforms, including social networks, provides insights into their identity and their goals. Research was conducted in Arabic and English.
- *Effects of Political Islam on the Economy and Politics in Turkey.* Turkey's ruling party, the Justice and Development Party (AKP), by combining the values of political Islam, EU reforms, and liberal economic policies, has strongly influenced Turkey's economic growth and domestic and foreign policy. The Anatolian Tigers, a powerful business elite that benefited from AKP's actions, form its core base. Research was conducted in Turkish and English.
- *Iran's National Internet Project.* Analysis of Farsi and English open sources indicates that Iran has the infrastructure for setting up a national Internet and that a permanent disconnection from the Internet is unlikely. This research also discusses possible alternatives to connecting to the Internet in the event of disconnection.

- *Money Laundering Enterprises: Afghanistan.* Money laundering in Afghanistan involves four key groups, with the current government playing a dominant role; findings also identify three primary modes of cash transfers. Research was conducted in English, Dari, Pashto, and Russian using websites, international reports, and Afghan media and government documents.
- *China's Public Position on the South China Sea.* China's public position on the South China Sea, albeit ambiguous, becomes clearer with careful examination of messages underlying the use of key terms. Examination of Chinese news media, official statements, and academic publications indicates that China may be adopting a new approach in order to enhance its bargaining power in the region.
- *Impact of Chinese Economic Competition on Businesses in Kenya and Tanzania.* The increasing Chinese economic competition in Kenya and Tanzania has negatively impacted local businesses. However, positive outcomes include better transportation and telecommunication infrastructure developed by Chinese construction companies. Research in English and Swahili centered on local media sources, interviews, business and academic journals, and websites of key international development agencies.
- *Sudan/South Sudan Border Demarcation.* The complexity of Sudan/South Sudan border dynamics is reflected in the slow progress toward final demarcation and outbreaks of violence. Border demarcation is unlikely to be finalized in the foreseeable future. Increasing international pressure on both states may yield greater progress. Research was conducted in English and Arabic.
- *Somali Conceptions of Religiosity and Governance.* Somalia's progress toward peace and stability depends on the successful integration of religious values and clan affiliations with governance. Analysis of extant organizations, such as Al-Islah, and approaches to governance, such as Somaliland's reconciliation model, can provide useful insights in this direction. Research was conducted with English, Somali, and Arabic sources.
- *Analysis of the Media Environment in Northern Nigeria.* Radio is the preferred and accessible media platform used by youth in northern Nigeria. The research assesses how radio and information technology convergence has allowed the extremist group Boko Haram to gain access to youth in northern Nigeria. Research was conducted using English documents, social media, and interviews in Hausa.

Program Instructors

The five full-time program instructors are Georgetown University Center for Language Education and Development faculty who have developed the curriculum and taught in this program for a minimum of three years. They are seasoned professionals with an MA in Teaching English to Speakers of Other Languages (TESOL) or Applied Linguistics, or an MA in a related field supplemented with a TESOL Certificate. All have extensive post-secondary-level experience in Advanced ESL instruction; three embarked on this project with previous ESP business English experience, two have ACTFL Oral Proficiency Interviewer training in their background, and one is a former federal employee.

At the inception of their tenure in the EHLS instructional program, all instructors received eighty hours of in-house training in the curriculum, programmatic goals, approaches and methodologies, and the needs of this student population. Instructors also participate in annual professional development activities, including federal agency workshops on writing briefings for the Intelligence Community; training on the nature and use of the ILR and ACTFL scales and assessment; and training on applying Backward Design methodology in curriculum development. The Career Skills instructor has also participated in specialized training for certification as a Federal Career Counselor. Faculty research in the following areas has served to inform and enrich the curriculum and instructional design of the program: framing and media analysis; hierarchies of critical thinking skills; rhetorical argumentation patterns; aspects of functional grammar; strategic listening skill development; aspects of intercultural communication, including communication styles, proxemics, high and low context, and negotiation strategies; and vocabulary development through corpus linguistics using the Corpus of Contemporary American English (http://corpus.byu.edu/coca/).

Program Participants

Each year, NSEP provides a list of native/heritage language groups to be recruited for the program; the list is based on input from the Department of Defense and the Intelligence Community regarding the languages of greatest need. The program began in 2006 with scholars from three language communities (Arabic, Mandarin Chinese, and Russian); since then, it has graduated speakers of Arabic, Balochi, Chinese-Cantonese, Chinese-Mandarin, Dari, Hausa, Igbo, Kyrgyz, Pashto, Persian Farsi, Somali, Swahili, Turkish, Urdu, Uzbek, and Yoruba.

Program participants are recruited from across the United States. They are naturalized citizens, with an average length of residency in the United States of sixteen years. For program years 2006–2012, the average age was forty-four, with the youngest scholar being twenty-five at program entry and the oldest being sixty-six. Almost all had bachelor's degrees; 68 percent had a second postsecondary degree, and 15 percent had a third. Approximately two-thirds of the participants in these cohorts had received one degree taught primarily in English.

Program participants are required to possess ILR Level 3 proficiency in the native or heritage language at entry as demonstrated through a telephonic Oral Proficiency Interview (OPI). CAL also uses ILR-aligned tests of listening, reading, and writing, and telephonic OPIs to determine the English language proficiency of applicants who have been provisionally selected for the program. Listening and reading are tested with the English Language Proficiency Test (ELPT), which was developed by the Defense Language Institute English Language Center at Lackland Air Force Base (DLI-ELC) and provides ratings only through ILR Level 3. The writing test was developed by CAL and is scored with an ILR-aligned rubric developed by DLI-ELC. This test also provides ratings only through ILR Level 3. The telephonic OPIs are provided by Language Testing International (LTI), which gives ratings on the ACTFL scale through the Superior level.[2]

To be admitted to the EHLS program, a candidate must receive ratings at or above ILR Level 2/ACTFL Advanced Low in all four modalities in English. This requirement ensures that participants will have the skills needed to participate effectively in all instructional activities, particularly the capstone OSAP. The same tests (in B forms for the listening, reading, and writing) are administered at program exit. The listening, reading, and writing tests each have two parallel forms: Form A, given at entry, and Form B, given at exit.

The total number of participants admitted to the program each year depends on the funding level available. Each participant receives a scholarship that includes full tuition and a living stipend for the duration of the program; receiving a scholarship is conditional on each participant agreeing to work for the federal government for one year in a position with national security responsibilities.

Cumulative Testing Outcomes

Refinement of the instructional curriculum and teaching approach over the years has resulted in progressive improvements in language proficiency outcomes. The goals for the program, as identified by the funder, are for 50 percent of all exit test ratings to be at ILR Level 3 and 75 percent to be at ILR Level 2+ or higher after eight months of instruction (six months intensive and two months part time). The program came close to meeting these goals in the 2011 program year, and in 2012 it exceeded them, as shown in Table 3.1.

In 2011, 34.3 percent of the total entry tests taken ($4 \times 35 = 140$ tests) received ratings of ILR 3 or Superior. At exit, that percentage increased to 41.0 percent. In 2012, 39.1 percent of the total entry tests taken ($4 \times 29 = 116$ tests) received ratings of ILR 3 or Superior. At exit, that percentage had increased to 58.3 percent, greater than the 50 percent target set by the program funder.

In the years prior to 2012, the program had relaxed the ILR 2 entry requirement slightly, as seen in the percentages listed on the "ILR 1+ or lower" row of the

◆ Table 3.1 Cumulative EHLS Entry/Exit Ratings, 2009–2012, All English Modalities (percentages of total tests taken*)

	2009		2010		2011		2012	
	28 participants		37 participants		35 participants		29 participants	
	Entry	Exit	Entry	Exit	Entry	Exit	Entry	Exit
Total—all modalities								
ILR 3	19.6	28.6	31.1	28.4	34.3	41.0	39.1	58.3
ILR 2+ or higher	50.0	61.6	63.5	68.9	63.6	72.7	71.3	81.7
ILR 2 or higher	92.9	91.1	95.9	98.6	93.6	98.6	100.0	100.0
ILR 1+ or lower	7.1	8.9	4.1	1.4	6.4	1.4	0.0	0.0

* Formal reporting of speaking scores used the ACTFL proficiency scale. The corresponding ILR scale ratings are used in this table for purposes of summarization. Correspondence as follows: ACTFL Superior = ILR 3; ACTFL Advanced High = ILR 2+; ACTFL Advanced Mid and Advanced Low = ILR 2; ACTFL Intermediate High = ILR 1+.

◆ Table 3.2 Cumulative EHLS Entry/Exit Ratings, 2009–2012, English Listening and Speaking (percentages)

	2009		2010		2011		2012	
	28 participants		37 participants		35 participants		29 participants	
	Entry	Exit	Entry	Exit	Entry	Exit	Entry	Exit
Listening								
ILR 3	7.1	17.9	10.8	10.8	20.0	34.3	27.6	69.0
ILR 2+ or higher	39.3	64.3	64.9	64.9	80.0	74.3	72.4	90.0
ILR 2 or higher	100.0	96.4	100.0	100.0	100.0	100.0	100.0	100.0
ILR 1+ or lower	0.0	3.6	0.0	0.0	0.0	0.0	0.0	0.0
Speaking								
ACTFL Superior	14.3	25.0	27.0	21.6	34.3	45.7	31.0	58.6
ACTFL Advanced High or higher	32.1	42.9	59.5	73.0	54.3	71.4	65.5	79.3
ACTFL Advanced Low or higher	75.0	75.0	83.8	94.6	88.6	97.1	100.0	100.0
ILR 1+ or lower	25.0	25.0	16.2	5.4	11.4	2.9	0.0	0.0

table. Strict adherence to the minimum entry requirement in 2012 contributed to the program's ability to create a learning environment with a higher level of challenge and thus achieve higher outcome levels in that year.

Oral Skills Outcomes

In listening and speaking, outcomes showed marked improvement in the 2012 program year, particularly with regard to the percent of participants who attained ILR Level 3 (Table 3.2). Adherence to the minimum requirement of ILR Level 2 proficiency at entry contributed to this outcome. Additionally, in preparing for the 2012 instructional program, the instructors revised the listening comprehension and related oral communication activities to ensure that the activities demanded the levels of listening discrimination and speaking finesse that characterize Superior proficiency. The revised activities increased the program's ability to promote the development of the desired levels of proficiency in the oral skill modalities.

In listening, a closer look at participant outcomes illustrates the challenges that the program has faced in addressing this modality, as well as the improvement that resulted from the curricular changes made in 2012 (Table 3.3).

In Table 3.3, the shaded cells indicate the number of participants each year who demonstrated null gain between entry and exit. This number includes those who entered with a rating of 3; because the maximum rating on the test is 3, demonstration of an increase is not possible for an examinee who enters with that rating. The cells to the left of the shaded cells provide numbers of participants whose ratings

decreased, and the cells to the right of the shaded cells provide numbers of participants whose ratings increased.

The majority of participants each year demonstrated no change in rating for listening. Multiple noncurricular factors may explain this outcome:

- The ILR proficiency levels are broad, meaning that a learner must acquire a great deal of language ability before learning is reflected in a level change.
- The test contains only 65 items; one right or wrong answer can mean the difference between a 2+ and a 3, or a 2 and a 2+, and this difference could be attributable to the Standard Error of Measurement of the assessment instrument.*

Table 3.3 EHLS Listening Modality, Entry/Exit Ratings by Number of Participants

		Exit (ratings)			
		ILR 1+	ILR 2	ILR 2+	ILR 3
Entry (ratings)	**Listening 2009**				
	ILR 1+	0	0	0	0
	ILR 2	1	9	5	2
	ILR 2+	0	0	6	3
	ILR 3	0	0	2	0
	Listening 2010				
	ILR 1+	0	0	0	0
	ILR 2	0	7	6	0
	ILR 2+	0	6	12	2
	ILR 3	0	0	2	2
	Listening 2011				
	ILR 1+	0	0	0	0
	ILR 2	0	6	1	0
	ILR 2+	0	3	12	6
	ILR 3	0	0	1	6
	Listening 2012				
	ILR 1+	0	0	0	0
	ILR 2	0	3	3	2
	ILR 2+	0	0	3	10
	ILR 3	0	0	0	8
	Total Listening				
	ILR 1+	0	0	0	0
	ILR 2	1	25	15	4
	ILR 2+	0	9	33	21
	ILR 3	0	0	5	16

- The content of the test is not aligned closely with the focus of EHLS instruction. EHLS activities use topics of relevance to the federal government context, whereas the test contains general interest topics.
- Program participants have varying educational backgrounds, professional backgrounds, and learning goals and motivations.
- *CAL does not have access to the technical manual for the test and therefore is unable to ascertain its Standard Error of Measurement.

However, the outcomes for 2012 also seem to show that the changes made to the curriculum in that year did have an effect. In that year, more than half of the participants improved their listening ratings, a far higher percentage than seen in previous years. In 2012, ten of the thirteen participants who had entered at Level 2+ increased their ratings to Level 3, whereas in 2011, only six of twenty-one did so. In addition, in 2012, for the first time no participants received a lower listening rating at exit than at entry (see Table 3.10 in Appendix 1). These differences can be attributed, in part, to curricular changes that aligned listening practice activities more closely with the specific skills and abilities that characterize the ILR Level 3 listener.

A paired sample *t*-test[3] was conducted on the listening outcomes to determine the significance of rating changes over time. For convenience of analysis, a number from 1 to 4 was assigned to each proficiency level: ILR 3 = 4, ILR 2+ = 3, ILR 2 = 2, ILR 1+ or lower = 1. Then each participant's entry and exit scores were listed. To ascertain the added value of the program, participants with an entry score of 4 were removed from the analysis because they had reached the ceiling prior to starting the program. Table 3.4 provides the results of the *t*-test that reflect a significance value of 0.05.

For the listening modality, the results showed a significant increase between the total entry and total exit scores: total entry scores (M =2.58, SD = 0.50) and total exit scores (M = 2.90, SD = 0.76), t (107) = 4.73, p = 0.000. The results also indicated a significant difference between the two scores in 2009 and 2012; in 2009, entry scores (M = 2.35, SD = 0.49) and exit scores (M = 2.77, SD = 0.82), t (25) = 3.07, p = 0.005; and in 2012, entry scores (M = 2.62, SD = 0.50) and exit scores (M= 3.43, SD = 0.75), t (20) = 6.17, p = 0.000. No significant increases in listening scores were

Table 3.4 Paired Sample T-Test on EHLS Entry/Exit Score, 2009–2012, English Listening

Year		Paired Differences Mean	Std. Deviation	t	df	Sig. (2-tailed)	N	Effect Size
2009	LExit - LEntry	0.423	0.703	3.070	25	0.005	28	0.602
2010	LExit - LEntry	0.061	0.659	0.529	32	0.601	37	0.092
2011	LExit - LEntry	0.143	0.591	1.279	27	0.212	35	0.242
2012	LExit - LEntry	0.810	0.602	6.167	20	0.000	29	1.346
Overall	LTotalExit - LTotalEntry	0.315	0.692	4.725	107	0.000	129	0.455

$\alpha = 0.05$

observed for the years 2010 and 2011. Cohen's effect size values for the listening modality ($d = 0.46$) suggested a moderate practical significance (Cohen 1988).

In speaking, the trend across all years has been for the majority of participants to maintain a proficiency level, with most of the rest achieving a rating increase and a few experiencing a rating decrease (Table 3.5). Because Language Testing International provides the speaking tests, speaking ratings are reported using the ACTFL scale. For the sake of comparison with the ILR scale, in Table 3.5, the ACTFL Mid and Low levels are collapsed.

Table 3.5 Speaking Modality, Entry/Exit Ratings by Number of Participants

		Exit (ratings)			
		ACTFL IH	**ACTFL AM**	**ACTFL AH**	**ACTFL S**
Entry (ratings)	**Speaking 2009**				
	ACTFL IH	5	2	0	0
	ACTFL AM	3	6	2	1
	ACTFL AH	0	1	1	3
	ACTFL S	0	0	1	3
	Speaking 2010				
	ACTFL IH	2	3	1	0
	ACTFL AM	0	4	4	1
	ACTFL AH	1	0	9	2
	ACTFL S	0	0	5	5
	Speaking 2011				
	ACTFL IH	0	2	0	0
	ACTFL AM	0	6	5	1
	ACTFL AH	0	0	4	4
	ACTFL S	0	0	0	12
	Speaking 2012				
	ACTFL IH	0	0	0	0
	ACTFL AM	0	5	4	1
	ACTFL AH	0	1	1	8
	ACTFL S	0	0	1	8
	Total Speaking				
	ACTFL IH	8	7	1	0
	ACTFL AM	3	21	15	4
	ACTFL AH	1	2	15	17
	ACTFL S	0	0	7	28

However, 44.8 percent of participants achieved a rating increase in 2012, compared with 30–35 percent in the three preceding years (see Table 3.11 in Appendix 1). In particular, data for those who entered with a speaking rating of Advanced High show that, whereas only three of five moved from Advanced High to Superior in 2009, as did two of twelve in 2010 and three of seven in 2011, fully eight of ten in 2012 were able to demonstrate such growth in speaking proficiency. This improvement may, in part, reflect changes to the 2012 curriculum and overall stronger cohort in terms of speaking proficiency at entry. Because no participants entered at a proficiency level below Advanced Low, the instructors were able to challenge participants in ways that they could not in earlier years when they had to ensure that lower proficiency participants were able to keep pace.

A paired sample *t*-test was conducted on the speaking outcomes to determine the significance of rating changes over time. For convenience of analysis, a number from 1 to 4 was assigned to each proficiency level: ACTFL Superior = 4, ACTFL Advanced High = 3, ACTFL Advanced Mid and Advanced Low = 2, ACTFL Intermediate High or lower = 1. Then each participant's entry and exit scores were listed. To prove the added value of the program, participants with an entry score of 4 were removed from the analysis because they had reached the ceiling before even starting the program. Table 3.6 provides the results of the *t*-test. Results reflect a significance value of 0.05.

For the speaking modality, a significant increase between the total entry and total exit scores was reported, with total entry scores ($M = 2.18$, $SD = 0.72$) and total exit scores ($M = 2.66$, $SD = 0.93$), $t\ (93) = 6.79$, $p = 0.000$. Significant differences between the two scores were found in the years 2010, 2011, and 2012. In 2010, entry scores ($M = 2.22$, $SD = 0.80$) and exit scores ($M = 2.67$, $SD = 0.78$), $t\ (26) = 3.31$, $p = 0.003$; in 2011, entry scores ($M = 2.13$, $SD = 0.69$) and exit scores ($M = 2.70$, $SD = 0.82$), $t\ (22) = 4.60$, $p = 0.000$; in 2012, entry scores ($M = 2.50$, $SD = 0.51$) and exit scores ($M = 3.15$, $SD = 0.88$), $t\ (19) = 4.33$, $p = 0.000$. No significant increase in speaking scores was observed for the year 2009. For the speaking modality, a Cohen's effect value ($d = 0.70$) suggested a moderate to large practical significance (Cohen 1988).

◆ Table 3.6 Paired Sample T-Test on EHLS Entry/Exit Score, 2009–2012, English Speaking

		Paired Differences						
Year		Mean	Std. Deviation	t	df	Sig. (2-tailed)	N	Effect Size
2009	SExit - SEntry	0.292	0.751	1.904	23	0.070	28	0.389
2010	SExit - SEntry	0.444	0.698	3.309	26	0.003	37	0.637
2011	SExit - SEntry	0.565	0.590	4.596	22	0.000	35	0.958
2012	SExit - SEntry	0.650	0.671	4.333	19	0.000	29	0.969
Overall	STotalExit - STotalEntry	0.479	0.684	6.790	93	0.000	129	0.700

$\alpha = 0.05$

◆ Table 3.7 Cumulative EHLS Entry/Exit Ratings, 2009–2012, English Reading and Writing (percentages)

	2009		2010		2011		2012	
	28 participants		37 participants		35 participants		29 participants	
	Entry	Exit	Entry	Exit	Entry	Exit	Entry	Exit
Reading								
ILR 3	50.0	57.1	73.0	64.9	77.1	71.4	79.3	79.3
ILR 2+ or higher	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
ILR 2 or higher	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
ILR 1+ or lower	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Writing								
ILR 3	7.1	14.3	13.5	16.2	5.7	11.8	17.8	25.0
ILR 2+ or higher	28.6	39.3	29.7	37.8	20.0	44.1	46.4	57.1
ILR 2 or higher	96.4	92.9	100.0	100.0	85.7	97.1	100.0	100.0
ILR 1+ or lower	3.6	7.1	0.0	0.0	14.3	2.9	0.0	0.0

Reading and Writing Outcomes

In reading, obtaining a fully accurate estimate of program participants' ability and progress presents a challenge. As Table 3.7 shows, reading is the strongest skill modality overall for EHLS program participants, all of whom have tested at ILR Level 2+ or higher at both entry and exit in every program year since 2009. Because the ELPT test used for entry and exit testing does not provide ratings above Level 3, it may not provide actual reading proficiency levels for some program participants. The strength of the entry ratings and the limitations of the reading test do not allow for meaningful conclusions to be drawn. Therefore, no further analysis has been carried out on the reading outcomes.

Achievement of the designated exit proficiency goals in the writing modality has presented the biggest challenge to the EHLS program. Not surprisingly, this is the modality in which entry ratings are lowest in all years, with more than half of participants entering at ILR Level 2 or lower each year (Table 3.8; Table 3.12 in Appendix 1). The data provided in Tables 3.9 and 3.10 demonstrate the difficulty of moving participants to ILR Level 3 in this modality. The 2012 cohort was stronger, overall, than those of previous years, with higher percentages of 3 and 2+ ratings at entry. In that year, for the first time, all participants who entered at 2 or 2+ either maintained or increased their rating at exit. However, three of the five who entered with Level 3 ratings saw their proficiency ratings decrease by the end of the program. Table 3.8 provides a graphic representation of the writing outcomes.

A paired sample *t*-test was conducted on the writing outcomes, using the same methodology that was applied in analyzing the listening outcomes. Table 3.9 provides the results of the *t*-test. All results were obtained using a significance value of 0.05.

For the writing modality, a significant increase between the total entry and total exit scores was found with total entry scores ($M = 2.17$, $SD = 0.50$) and total exit

◆ Table 3.8 EHLS Writing Modality, Entry/Exit Ratings by Number of Participants

		Exit (ratings)			
		ILR 1+	ILR 2	ILR 2+	ILR 3
Entry (ratings)	**Writing 2009**				
	ILR 1+	0	1	0	0
	ILR 2	2	13	4	0
	ILR 2+	1	0	3	2
	ILR 3	0	0	0	2
	Writing 2010				
	ILR 1+	0	0	0	0
	ILR 2	0	19	6	0
	ILR 2+	0	1	4	2
	ILR 3	0	1	0	4
	Writing 2011				
	ILR 1+	0	3	2	0
	ILR 2	0	16	5	1
	ILR 2+	0	0	4	1
	ILR 3	0	0	0	2
	Writing 2012				
	ILR 1+	0	0	0	0
	ILR 2	0	11	3	1
	ILR 2+	0	0	6	3
	ILR 3	0	1	2	2
	Total Writing				
	ILR 1+	0	4	2	0
	ILR 2	3	59	18	2
	ILR 2+	1	1	17	8
	ILR 3	0	2	2	10

scores ($M = 2.45$, $SD = 0.69$), t (114) $= 4.85$, $p = 0.000$. The results showed a significant difference in 2011 between the entry scores ($M = 2.00$, $SD = 0.56$) and the exit scores ($M = 2.42$, $SD = 0.66$), t (32) $= 3.44$, $p = 0.002$; the results also showed a significant difference between the two scores in 2012, entry scores ($M = 2.38$, $SD = 0.50$) and exit scores ($M = 2.71$, $SD = 0.75$), t (23) $= 2.89$, $p = 0.008$. No significant increases in writing scores were observed for the years 2009 and 2010. Cohen's effect size values for the writing modality ($d = 0.45$) suggested a moderate practical significance.

◆ Table 3.9 Paired Sample T-Test on EHLS Entry/Exit Score, 2009–2012, English Writing

Year		Paired Differences		t	df	Sig. (2-tailed)	N	Effect Size
		Mean	Std. Deviation					
2009	WExit WEntry	0.154	0.613	1.280	25	0.212	28	0.251
2010	WExit - WEntry	0.188	0.535	1.982	31	0.056	37	0.350
2011	WExit - WEntry	0.424	0.708	3.440	32	0.002	35	0.599
2012	WExit - WEntry	0.333	0.565	2.892	23	0.008	29	0.590
Overall	WTotalExit - WTotalEntry	0.278	0.615	4.853	114	0.000	129	0.453

$\alpha = 0.05$

Though writing represents a major focus of both the six-month intensive portion of the program and the two-month part-time session that follows, the development of professional proficiency in writing evidently may require more time and sustained practice than the EHLS program has been able to provide to date. However, the 2011 and 2012 data suggest that programmatic changes, such as a closer alignment of instructional design, activities, assessment, and performance feedback with ILR Level 3 discourse and linguistic/sociolinguistic competences for writing, and the addition of a Writing Lab designed to increase participants' ability to write accurately and appropriately in English for workplace communication, may have contributed to improved 2011 and 2012 writing outcomes. Program staff continues to explore ways of enhancing and reinforcing instruction to address the challenges that affect positive growth in writing performance.

Other factors that possibly influenced outcomes and that warrant future investigation include

- Educational background, including highest degree received, major field(s) of study, language(s) and country (countries) of instruction, and years elapsed since last participation in academic study;
- Linguistic background, including native language and extent of study of languages beyond the first language;
- Professional background, including professional field(s), language(s) used in professional contexts, and types of language tasks carried out in the workplace; and
- Demographic information, including age, gender, and years in the United States.

Conclusion

EHLS is unique in its nature as an ESP program that prepares participants specifically to obtain and retain professional positions with the federal government. Since the program's inception, its graduates have entered federal service as regional and country analysts, managers, translators, and language instructors with a variety of

federal agencies and government contractors. The program has produced a total of 215 graduates from its inception in 2006 through 2012.

Over the years of program implementation, CAL and Georgetown have confirmed the importance of reflective teaching practice in developing an effective program design. Through ongoing critical reflection on the efficacy of tasks and activities, program instructors have developed a detailed understanding of the curricular elements and instructional techniques that are most effective at promoting the self-awareness and higher-order thinking that underpin the development of professional-level language proficiency. Use of the backward design model of instructional development has played a key role in translating the linguistic skills and cultural understandings that characterize general professional proficiency into a sequenced program of instruction that effectively promotes development of those skills. Weekly in-depth, individualized feedback targeted to the discourse and linguistic/sociolinguistic competences contained in the descriptors for ILR Level 3, combined with whole class task-based instruction, has proven critical in developing participants' language awareness—a necessary element of language acquisition at the more advanced levels (ILR 3, 4, and 5). Perhaps most importantly, the use of the backward design model of instructional development based on the end goals of instruction has enabled the program to produce graduates who meet the needs and requirements of the federal workplace.

Though challenges continue to present themselves—particularly with regard to writing skill development—the program has demonstrated that task-based communicative language teaching, combined with direct attention to the details of structure and mechanics and to the importance of cultural awareness, can move learners to the level of general professional proficiency (ILR Level 3). CAL and Georgetown look forward to carrying out additional formal analysis of learner outcomes and continuing this partnership that enables the federal government to meet its need for high-level speakers of critical languages and English.

Notes

1. A parallel program was developed by CAL and the University of Washington in Seattle and implemented in program years 2006–2008. This program had to be discontinued in 2009. Many of the insights and strengths that characterize the program at Georgetown were gained through interactions with UW colleagues in those first three years.
2. Language Testing International will begin providing ratings on the ILR scale in program year 2014.
3. The authors wish to thank Cathy Cameron and Xin Yu of CAL for their assistance with this analysis.

References

American Council on the Teaching of Foreign Languages. n.d. "Descriptions of the Major ACTFL Proficiency Levels: Superior." Accessed May 9, 2014. http://www.actfltraining.org/actfl_certificate/actfl_superior.cfm.

Canale, Michael. 1983. "From Communicative Competence to Communicative Language Pedagogy." In *Language and Communication*, edited by Jack Richards and Richard Schmidt, 2–27. London: Longman.

Canale, Michael, and Merrill Swain. 1980. "Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing." *Applied Linguistics* 1: 1–47.

Chapelle, Carol, William Grabe, and Margie Berns. 1997. "Communicative Language Proficiency: Definition and Implications for TOEFL 2000. ETS Research Memorandum." Accessed May 9, 2014. http://www.ets.org/research/policy_research_reports/publications/report/1997/hzeu.

Coalition of Distinguished Language Centers (CDLC). 2008. *What Works: Helping Students Reach Native-Like Second-Language Competence*. San Juan Bautista, CA: MSI Press.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

de Bono, Edward. 1985. *Six Thinking Hats: An Essential Approach to Business Management*. Boston: Little, Brown & Co.

Ehrman, Madeline. 1998. "The Learning Alliance: Conscious and Unconscious Aspects of the Second Language Teacher's Role." *System* 26 (1): 93–106.

———. 2002. "The Learner at the Superior-Distinguished Threshold." In *Developing Professional-Level Language Proficiency*, edited by Betty L. Leaver and Boris Shekhtman, 245–259. Cambridge: Cambridge University Press.

Ellis, Rod. n.d. "Professor Rod Ellis on Task-Based Language Teaching." TESOL Alliance video, 8:00. Accessed May 9, 2014. vimeo.com/10621002.

Hymes, Dell H. 1971. *On Communicative Competence*. Philadelphia: University of Pennsylvania Press.

Ingold, Catherine. 2002. "The LangNet 'Reading to the Four' Project: Applied Technology at Higher Levels of Language Learning." In *Developing Professional-Level Language Proficiency*, edited by Betty L. Leaver and Boris Shekhtman, 141–155. Cambridge: Cambridge University Press.

Interagency Language Roundtable. 2014. "Interagency Language Roundtable Language Skill Level Descriptions—Speaking." Accessed May 9. http://www.govtilr.org/Skills/ILRscale2.htm.

Jackson, Frederick, and Marsha Kaplan. 2001. "Lessons Learned from Fifty Years of Theory and Practice in Government Language Teaching." In *Language in Our Time*, edited by James Alatis and Ai-Hui Tan, 71–87. Washington: Georgetown University Press.

Johns, Ann M., and Donna Price-Machado. 2001. "English for Specific Purposes: Tailoring Courses to Student Needs—and to the Outside World." In *Teaching English as a Second or Foreign Language*, 3rd ed., edited by Marianne Celce-Murcia, 43–54. Boston, MA: Heinle & Heinle.

Leaver, Betty L. 2004. "Interviews with High-Level Speakers: Surprises in the Data." *Journal for Distinguished Language Studies* 2: 27–40.

Leaver, Betty L., Madeline Ehrman, and Boris Shekhtman. 2005. *Achieving Success in Second Language Acquisition*. Cambridge: Cambridge University Press.

Leaver, Betty L. and Boris Shekhtman. 2002. "Principles and Practices in Teaching Superior Level Language Skills: Not Just More of the Same." In *Developing Professional-Level Language Proficiency*, edited by Betty L. Leaver and Boris Shekhtman, 3–27. Cambridge: Cambridge University Press.

Leow, Ronald, and Melissa Bowles. 2005. "Attention and Awareness in SLA." In *Adult Second Language Acquisition*, edited by Cristina Sanz, 179–203. Washington: Georgetown University Press.

Long, Michael H. 1981. "Input, Interaction, and Second Language Acquisition." *Annals of the New York Academy of Sciences* 379: 259–278.

Long, Michael H. 1983. "Does Second Language Instruction Make a Difference? A Review of Research." *TESOL Quarterly* 17: 359–382. http://onlinelibrary.wiley.com/doi/10.2307/3586253/abstract.

Malone, Margaret, Benjamin Rifkin, Donna Christian, and Dora Johnson. 2005. "Attaining High Levels of Proficiency: Challenges for Foreign Language Education in the United States." Accessed May 9, 2014. http://www.cal.org/resources/digest/attain.html.

Merrill, M. David, Leston Drake, Mark J. Lacy, and Jean Pratt. 1996. "Reclaiming Instructional Design." *Educational Technology* 36 (5): 5–7.

Munby, John. 1978. *Communicative Syllabus Design*. Cambridge: Cambridge University Press.

Nunan, David. 1991. *Language Teaching Methodology*. London: Prentice Hall.

Omaggio Hadley, Alice. 2001. *Teaching Language in Context*. 3rd ed. Boston: Heinle & Heinle.

Savignon, Sandra. 1997. *Communicative Competence: Theory and Classroom Practice: Texts and Contexts in Second Language Learning*. 2nd ed. New York: McGraw-Hill.

Shekhtman, Boris. 2003. *Working with Advanced Foreign Language Students.* Salinas, CA: MSI Press.
Spolsky, Boris. 1989. "Communicative Competence, Language Proficiency, and Beyond." *Applied Linguistics* 10 (2): 138–156.
Swain, Merrill. 1985. "Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in Its Development." In *Input in Second Language Acquisition*, edited by Susan M. Gass and Carolyn G. Madden, 235–253. Rowley, MA: Newbury House.
Wesche, Marjorie B. 1977. "Learning Behaviors of Successful Adult Students on Intensive Language Training." Paper presented at the Los Angeles Second Language Research Forum, Los Angeles, California, February 11–13. Accessed May 9, 2014. http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED176583&ERICExtSearch_SearchType_0=no&accno=ED176583.
Wiggins, Grant, and Jay McTighe. 2005. *Understanding by Design.* 2nd ed. Upper Saddle River, NJ: Pearson.
———. 2008. "Put Understanding First." *Educational Leadership* 65 (8): 36–41.
Willis, Jane. 2004. "Perspectives on Task-Based Instruction: Understanding Our Practices, Acknowledging Different Practitioners." In *Task-Based Instruction in Foreign Language Education: Practices and Programs*, edited by Betty L. Leaver and Jane R. Willis, 3–44. Washington, DC: Georgetown University Press.

Appendix 1. EHLS Program English Language Testing: Data Tables

◆ Table 3.10 EHLS Listening Outcomes, 2009–2012 (number and percentage of ratings)

	2009		2010		2011		2012	
	#	%	#	%	#	%	#	%
Entry Score = 3	2	7.1	4	10.8	7	20.0	8	27.6
No rating change*	0	0.0	2	5.4	6	17.1	8	27.6
Decrease 3 to 2+	2	7.1	2	5.4	1	2.9	0	0.0
Decrease 3 to 2	0	0.0	0	0.0	0	0.0	0	0.0
Entry Score = 2+	9	32.1	20	54.1	21	60.0	13	44.8
No rating change	6	21.4	12	32.4	12	34.3	3	10.3
Increase 2+ to 3	3	10.7	2	5.4	6	17.1	10	34.5
Decrease 2+ to 2	0	0.0	6	16.2	3	8.6	0	0.0
Entry Score = 2	17	60.7	13	35.1	7	20.0	8	27.6
No rating change	9	32.1	7	18.9	6	17.1	3	10.3
Increase 2 to 3	2	7.1	0	0.0	0	0.0	2	6.9
Increase 2 to 2+	5	17.9	6	16.2	1	2.9	3	10.3
Decrease 2 to 1+	1	3.6	0	0.0	0	0.0	0	0.0
Total Participants	28		37		35		29	
Total no rating change	15	53.6	21	56.8	24	68.6	14	48.3
Total increased rating	10	35.7	8	21.6	7	20.0	15	51.7
Total decreased rating	3	10.7	8	21.6	4	11.4	0	0.0

*The maximum rating on the test is 3, so an increase is not possible for an examinee who receives an entry rating of 3.

◆ Table 3.11 EHLS Speaking Outcomes, 2009–2012 (number and percentage of ratings)

	2009		2010		2011		2012	
	#	%	#	%	#	%	#	%
Entry Rating = Superior	4	14.3	10	27.0	12	34.3	9	31.0
No rating change	3	10.7	5	17.9	12	34.3	8	27.6
Decrease Superior to Advanced High	1	3.6	5	17.9	0	0.0	1	3.4
Entry Rating = Advanced High	5	17.9	12	32.4	7	20.0	10	34.5
No rating change	1	3.6	9	24.3	4	11.4	1	3.4
Increase Advanced High to Superior	3	10.7	2	5.4	3	8.6	8	27.6
Decrease Advanced High to Advanced Mid or Advanced Low	1	3.6	1	2.7	0	0.0	1	3.4
Entry Rating = Advanced Mid or Advanced Low	12	4.3	9	24.3	12	34.3	10	34.5
No rating change*	6	21.4	4	10.8	6	17.1	5	17.2
Increase Advanced Mid or Advanced Low to Superior	1	3.6	1	2.7	1	2.9	1	3.4
Increase Advanced Mid or Advanced Low to Advanced High	3	10.7	4	10.8	5	14.3	4	13.8
Decrease Advanced Mid or Advanced Low to Intermediate High	2	7.1	0	0.0	0	0.0	0	0.0
Entry Rating = Intermediate High	6	21.4	5	17.9	4	11.4	0	0.0
No rating change	4	14.3	1	2.7	1	2.9		
Increase Intermediate High to Advanced High	0	0.0	1	2.7	0	0.0		
Increase Intermediate High to Advanced Mid or Advanced Low	2	7.1	3	8.1	3	8.6		
Entry Rating = Intermediate Mid	1	3.6	1	2.7	0	0.0	0	0.0
No rating change	0	0.0	0	0.0				
Increase Intermediate Mid to Intermediate High	1	3.6	1	2.7				
Total Participants	28		37		35		29	
Total no rating change	14	50.0	19	51.4	23	65.7	14	48.3
Total increased rating	10	35.7	12	32.4	12	34.3	13	44.8
Total decreased rating	4	14.3	6	0.0	0	0.0	2	6.9

*Includes scores that moved from Advanced Mid to Advanced Low and vice versa.

◆ Table 3.12 Title EHLS Writing Outcomes, 2009–2012 (number and percentage of ratings)

	2009		2010		2011		2012	
	#	%	#	%	#	%	#	%
Entry Rating = 3	2	7.1	5	13.5	2	5.7	5	17.2
No rating change	2	7.1	4	10.8	2	5.7	2	6.9
Decrease 3 to 2+	0	0.0	0	0.0	0	0.0	2	6.9
Decrease 3 to 2	0	0.0	1	2.7	0	0.0	1	3.4
Entry Rating = 2+	6	21.4	6	16.2	5	14.3	9	31.0
No rating change	3	10.7	2	5.4	4	11.4	6	20.7
Increase 2+ to 3	2	7.1	2	5.4	1	2.9	3	10.3
Decrease 2+ to 2	1	3.6	2	5.4	0	0.0	0	0.0
Entry Rating = 2	19	67.9	26	70.3	23	65.7	15	51.7
No rating change	13	46.4	20	54.1	16	45.7	11	37.9
Increase 2 to 3	0	0.0	0	0.0	1	2.9	1	3.4
Increase 2 to 2+	4	14.3	6	16.2	5	14.3	3	10.3
Decrease 2 to 1+	2	7.1	0	0.0	1	2.9	0	0.0
Entry Rating = 1+	1	3.6	0	0.0	4	11.4	0	0.0
No rating change	0	0.0			0	0.0		
Increase 1+ to 2+	0	0.0			2	5.7		
Increase 1+ to 2	1	3.6			2	5.7		
Entry Rating = 1	0		0		1	2.9	0	0.0
No rating change					0	0.0		
Increase 1 to 2					1	2.9		
Total Participants	28		37		35		29	
Total no rating change	18	64.3	26	70.3	22	62.9	19	65.5
Total increased rating	7	25.0	8	21.6	12	34.3	7	24.1
Total decreased rating	3	10.7	3	8.1	1	2.9	3	10.3

Chapter 4

Advanced Foreign Language Study through Global Debate

TONY BROWN, JENNIFER BOWN, AND DENNIS L. EGGETT

FOR CENTURIES, DEBATE HAS been a major component of universities throughout Western Europe and the United States. Aside from educating students about significant social and political issues, debate fosters critical thinking and analytical skills, not to mention respect for opposing opinions and an increased capacity to relate to individuals of different persuasions. In a debate, every claim is subject to questioning, thus creating an environment rich in rhetorical strategies and complex linguistic constructions.

Owing to its overt public nature, debate also demands sophisticated public speaking abilities that extend beyond mere narration and description to persuasion and support of a hypothesis or claim. From the standpoint of language proficiency, the very criteria that constitute a well-educated native speaker likewise constitute an articulate debater (see Martin, this volume, for an overview of the American Council on the Teaching of Foreign Languages [ACTFL] proficiency guidelines). Given the striking similarities between the two skills, the question arises as to whether a forum capable of improving one's first language (L1) proficiency can likewise improve second language (L2) proficiency as well.

Our efforts to answer this question began in earnest during the Spring 2007 semester with the creation of a course entitled "Russian Global Debate" that culminated in 2013 with the writing of a task-based foreign language textbook in which exercises that included parliamentary-style debate and scaffolding served as a means for developing Advanced- and Superior-level proficiency. Findings presented in this chapter reflect data from four similar iterations of the course offered during the Spring 2007, Spring 2011, Fall 2011, and Fall 2012 semesters. This chapter likewise provides a detailed description of the pedagogical methods applied in the aforementioned textbook to illustrate an innovative approach towards converting foreign language learning theory into classroom practice. On the whole, this chapter argues for

university foreign language instruction as a critical step toward achieving Advanced- and Superior-level proficiency.

Debate and Superior-Level Functions—Speaking

Supporting and defending opinions is a core task at the Superior level of foreign language learning, and the criteria outlined in the ACTFL speaking performance profile describe qualities emphasized in public speaking and debate (i.e., fluency/integrative abilities, breadth of vocabulary, sociolinguistic/cultural appropriateness, grammatical accuracy, and debate-related communication tasks).

Fluency/Integrative Abilities

In terms of fluency of speech, the Superior-level speaker easily connects extended thoughts together, thus moving beyond sentential-level speech to that of connected, cohesive paragraphs. Fluency of speech likewise characterizes an experienced public speaker, particularly in societies that have a high regard for interpersonal communication skills.

Vocabulary

Superior-level sp eakers can participate in a wide variety of conversations, ranging from professional and/or scholarly topics to informal topics, and tailor their language accordingly. Although circumlocution represents a valuable skill developed by Superior-level speakers in order to engage individuals on an array of topics, frequent usage of the same limited number of vocabulary items often results in "inexpressive and boring speech" (Fessenden et al. 1973, 92). Redding further asserts that "the debater, like any public speaker, should command a precision of word choice that will reflect the most subtle shadings of meaning" (1954, 205). In short, having breadth of vocabulary facilitates fluency and accuracy and expands one's capacity to provide uninterrupted, extended, and in-depth discourse on a topic rather than spotty, fragmented statements that lack transitions, cohesion, and continuity.

Sociolinguistic/Cultural Appropriateness

Parliamentary style debate forums challenge one's ability to demonstrate intellectual prowess, cultural sophistication, and wittiness. Unlike American-style debate, parliamentary-style debate accords greater value to overall presentation than to the defense of a position. As such, manner of presentation in the form of idiomatic statements, reference to cultural artifacts, and wit plays a particularly important role in determining the outcome of a debate. Similarly, Superior-level speakers distinguish themselves with their ability to make references to the culture of the target language and control the range of registers needed to respond appropriately in a variety of situations.

Grammatical Accuracy

Correct grammar usage naturally plays an important role in any public speaking forum. Sentence construction in oral presentation tends to have a looser structure than most forms of writing, but certainly not at the expense of grammatical rules.

Similarly, a Superior-level speaker may make occasional grammatical errors, but one does not typically notice a recurring pattern of errors in speech, especially with regard to basic structures.

Debate and Superior-Level Functions—Writing

The intensely methodical task of crafting a sophisticated written argument differs considerably from the rather extemporaneous and fluid nature of orally debating a topic; however, the success of the latter often depends on the care and attention afforded by the former. Honing the ability to write clearly and persuasively requires considerable time and effort for native speakers and L2 learners alike since academic writing reflects a different set of conventions than spoken language (e.g., information gathering, analysis and organization of information, and presentation of ideas in a manner that communicates effectively to the reader) (Williams 2005). When criteria for persuasive writing are placed side-by-side with ACTFL functions for writing (i.e., tasks, context/content, text type, and accuracy) one finds a direct correlation between well-crafted academic writing and debate rhetoric.

Kaplan (1987) observes that oral and written argumentations are similar in their aim to produce conviction, but they are different in their extension, structures, lexicon, and interrelation of structures. As with speaking, supporting and defending opinions and constructing hypotheses and conjectures is a core task at the Superior-level according to the ACTFL Rating Grid by Sub-Level for Writing Proficiency. The subsequent criteria of context/content, text type, and accuracy likewise dovetail with qualities emphasized in argumentation and debate (e.g., writing in a variety of content areas including practical, professional, and social topics [context/content] cohesive text of multiple paragraphs and pages [text type]; and good control of a firm range of grammatical structures, wide vocabulary, and no patterned errors in basic structures [accuracy]).

Foreign Language Learning through Task-Based Instruction (TBI)

Simply put, a task is "an activity conducted in the foreign language that results in a product with a measurable result such that students can determine for themselves whether or not they have adequately completed the assignment" (Leaver and Kaplan 2004, 47). Debate in the foreign language classroom enables language to become a vehicle for communicating ideas for meaningful purposes instead of functioning solely as an object of study (Coyle et al. 2010; van Lier 2005; Long 2007; Stryker and Leaver 1997; MLA Ad Hoc Committee 2007; Shaw 1997; Hedegaard 2005). Given that competing positions evolve during the course of a debate, careful attention to ongoing exchanges takes on paramount importance. In such a *meaning-focused* task, participants "are not simply displaying their control of particular patterns or structures or phrases, which would be a linguistic objective" (Willis 2004, 13). Rather, they are constructing meaning for the purpose of learning the language. Such a collaborative approach to language learning mirrors what Swain refers to as the "Output

Hypothesis," in that students "push their linguistic competence to its limit as they attempt to express their ideas" (1993, 162) and negotiate meaning. Long (1996) and Gass and Varonis (1994) argue that perhaps *interaction* rather than output more aptly describes the hypothesis in question owing to the inherent role of interaction in negotiating meaning, which in turn contributes to L2 acquisition. Semantics aside, debate promotes dialogue rich in rhetorical functions that are used to convey *meaning* in a logical and persuasive manner.

In discussing the function of rhetoric, Inch and Warnick assert that "the rhetorical function of language aims to direct or influence thoughts and behavior. It is persuasive" (2005, 114). Additional research including that conducted by Massie (2005) and Connor (1987) identifies the task of argumentation and debate as a valuable strategy for improving L2 oral and writing proficiency, especially at the Advanced and Superior levels. However, finding scaffolding materials suitable for this method presents a significant challenge for instructors, especially for third-year instruction and beyond. In order to meet this challenge a course centered on debate was created, and, subsequently, a textbook was designed to facilitate Advanced- and Superior-level foreign language uptake through debate.

Debate Textbook

Mastering Russian through Global Debate and *Mastering English through Global Debate* represent the first two volumes in a series of textbooks that will soon include Chinese and Arabic. The primary objective of these textbooks is to facilitate acquisition of Superior-level proficiency through the forum of oral debates and written position papers.[1] The topics selected for the Russian and English volumes were as follows: Economy vs. Environment, Interventionism vs. Isolationism, Wealth Redistribution vs. Self-Reliance, Cultural Preservation vs. Diversity, Security vs. Freedom, and Education vs. Field Experience. These topics reflect real-world issues that transcend borders and capture the interest of adult language learners.

Each chapter in the textbooks revolves around one of these debate issues. Each chapter begins with exercises that introduce the topic and activate learners' background knowledge, thus preparing them for reading the texts that comprise the heart of each chapter.

At the center of each chapter is a text written by native speakers of the target language. These texts were pedagogically created and, thus, not authentic. Nevertheless, they are not simplified in any way and provide a rich source of input, particularly in terms of contextualized vocabulary. Commissioned texts were selected to ensure that each article presents a balanced overview of the topics, including the major arguments of both sides of the debate. Each reading assignment is followed by comprehension checks.

A major portion of each chapter is dedicated to vocabulary development. Lexical items are introduced not as single words, but rather as collocations—groups of words commonly used together. The choice of vocabulary in each chapter was, in part, governed by frequency of usage as found in linguistic corpora. Students are directed to use online corpora to further their word knowledge so that they learn

words as they are used in the target language, not as they are used in isolation. Open-ended discussion questions offer a starting point for learners to apply new vocabulary introduced in the articles.

Once students have become acquainted with a topic and the vocabulary necessary for discussion, they begin to prepare for the debate. In the "Constructing Critical Discourse" section, learners are introduced to advanced syntactical features of the language, particularly those used to form hypotheses—a function important for performing at the Superior level. Students put their new knowledge to use in a role-play by representing various stakeholders discussing the issue in a concrete way. Listening comprehension skills are developed by having learners listen to brief mock debates that illustrate important turns of phrases used for turn taking, arguing a point, agreeing, and disagreeing.

These previous sections prepare the speaker for the final two sections, "Formatting the Argument: Writing" and "Formatting the Argument: Speaking," which represent the culminating tasks for the textbook. In the writing portion, learners are introduced to the elements of persuasive writing, such as writing a thesis statement, constructing paragraphs, and, finally, revising arguments. These exercises culminate in the form of a persuasive essay, in which learners argue one side of the issue. The speaking section introduces learners to important rhetorical strategies used in debate, such as conjecture, questions of definition, and questions of value. At the end of the speaking section, learners record a monologue of their prepared remarks and then put their skills to the test by debating with other class members.

Russian Global Debate Course

Materials for the debate textbook were piloted on four separate occasions at Brigham Young University (BYU). The course set out to push learners to the Advanced and Superior levels in their respective foreign languages by integrating rhetorical methods into the curriculum. Course outcomes suggest that the course and materials indeed reached their goals and contributed to language gain in both oral and writing proficiency.

Russian Global Debate was open for students completing either their third or fourth year of language study. In order to establish average pre- and post-proficiency levels and to measure gain over the course of a semester, students completed a pre- and post-Oral Proficiency Interview (OPI) and Writing Proficiency Test (WPT), administered by the ACTFL. Given that proficiency levels in any one group could vary considerably, constructing exercises tailored to specific proficiency levels became all the more critical.

Participant Demographic and University Foreign Language Curricula

A total of fifty-five students participated in the course (fourteen in Spring 2007, eighteen in Spring 2011, eleven in Fall 2011, and twelve in Fall 2012). All of the students learned the bulk of their Russian while living in a Russian-speaking country for eighteen months to two years. Although these students had considerable

contact with the target language while abroad, their exposure consisted primarily of speaking and listening. Upon returning to their home institution, each student matriculated directly into a two-semester, third-year advanced grammar course that reviewed formal elements of the language—including verbal conjugation, adjectival endings, and noun declension—and incorporated a modest writing component that required students to summarize assigned short stories.

Classroom Structure and Schedule

Students received credit for two academic hours and met twice a week for fifty minutes. Class periods were spent working on language exercises associated with a given debate topic. At the beginning of each semester, instructors distributed copies of the ACTFL oral and writing proficiency guidelines in order to explain the criteria certified raters use to establish proficiency and to guide students when setting language proficiency goals.

Language exercises culminated in the form of in-class debates with fellow class members conducted in the target language. The class was divided into two teams (proposition, those speaking in favor of a motion, and opposition, those speaking against a motion). The proposition began each debate by framing a motion with the wording "This house believes . . . " or "This house would . . .". For example, if the motion was "This house believes that governments should mandate wealth redistribution," then the proposition (or government) speakers would explain why wealth redistribution is a good idea, and the opposition would demonstrate why it is not. Additionally, the government would propose a course of action and support it with philosophical, practical, and consequential arguments. The burden of proof was on the government, but the opposition also needed to demonstrate the strength of its arguments. Individual speeches alternated between the proposition and opposition and lasted for four minutes. After the first minute and before the last minute of a speech, any member of the opposing team could offer an argument or ask a question, otherwise known as offering "points of information." For example, when any member of the proposition team was speaking, any member of the opposition team could stand for a point of information. In the same way, when any member of the opposition team was speaking, any member of the proposition team could stand. Instructors were responsible for keeping time and signaling to debaters when they could offer points of information and when they could not. Students were allowed to reference notes for talking points, but they were not allowed to read from them, thus underscoring the extemporaneous aspect of public speaking and the inseparable relationship between language and critical thinking skills.

Students received a schedule at the beginning of the semester that indicated both their team for any given debate and the position that they would be responsible for defending (i.e., whether they would be for or against a given issue). Given that students did not have a say in their respective positions, they felt pressed to articulate a coherent argument on a position that they did not necessarily agree with in the first place.

Homework

Students met weekly with a native Russian teaching assistant to review assigned readings, receive individualized feedback on writing assignments, and converse in the target language.

Reading Assignments.

Students received a reading assignment in Russian related to each debate topic. For example, if the topic for the upcoming week's debate addressed the Second Amendment to the United States Constitution, students would read a relevant article (3–4 typed pages) and then respond orally to comprehension questions during class. The purpose of such readings was twofold: 1) to introduce students to topic-related lexical items that, in turn, they could use applied when writing their individual position papers and 2) to deepen their understanding of the relevant issues under consideration. In essence, the researchers sought to create a linguistic environment in which instructors and students simultaneously facilitated learning by using "the discourse to help them produce utterances that they would not be able to produce on their own" (Ellis 1997, 48).

Writing Assignments.

Writing assignments consisted of outlines, summaries, and position papers. After reading the assigned text for a given topic, students carried out an Intermediate-level task of writing a sentence outline in Russian. They then gradually edged their way into the Advanced level by taking their outlines and turning them into summaries, which encouraged descriptive language in all time frames. Such a stepwise approach to writing stemmed largely from scaffolding theory[2] and Lev Vygotsky's zone of proximal development (ZPD), which means the "distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (1978, 86). So as not to exceed students' ZPD, the first half of the semester focused solely on outlines and summaries. Only after students had worked on this task at length did they proceed to develop their summaries into full-fledged position papers during the second half of the semester, which pushed them to produce solid Advanced- or Superior-level opinion pieces.

Students revised and resubmitted their assignments, taking into account feedback from the native-speaking assistants, thus emphasizing a step-by-step approach, rather than a product-oriented approach to each writing assignment (Cohen 1994; Shaw 1997; Connor 1987), or as Kaplan insightfully observed, "recognizing writing (composing) as a process in which a given text is not at all finished output, but merely a waystage" (1987, 17).

Conversation Practice.

Students met weekly with a native-speaking teaching assistant for conversation practice using domain-specific terminology relative to assigned topics. Directing the

conversation in this way focused students' attention on lexical items that were introduced in a given unit and encouraged their usage in responses to questions posed by the native speaker. As with feedback on writing assignments, native speakers provided corrective recasts when discussing topics with students as a way of modeling correct usage of newly introduced items.

Course Outcomes and Implications

As aforementioned, participants completed a pre- and post-OPI[3] and WPT. Figure 4.1 illustrates their pre- and post-OPI ratings and Figure 4.2 their pre- and post-WPT ratings.

Tables 4.1 and 4.2 provide pre- and post-OPI and WPT ratings and indicate participants' overall gain, or lack thereof. For example, of the twenty-one participants who received Advanced Mid (AM) on the pre-OPI, sixteen showed null gain on the post-OPI, as indicated by the quadrant where the AM row and column meet, whereas five moved up to Advanced High (AH).

The data from the pre- and post-tests underscore the premise of the ACTFL scale, according to which "facility increases exponentially within the various global tasks" (Swender 1999, 11) (i.e., progressing from Novice to Intermediate presents less of a challenge than progressing from Intermediate to Advanced, and on to Superior). Penetrating the Advanced High sublevel, in particular, arguably presents the most difficult stage on the ACTFL scale since learners must demonstrate partial control of Superior-level tasks to attain such a rating. Performing a *t*-test on the slope of the parameter indicates that the higher the pre-OPI rating, the lower the gain in oral proficiency over the course of a semester (slope/effect size = −0.39,

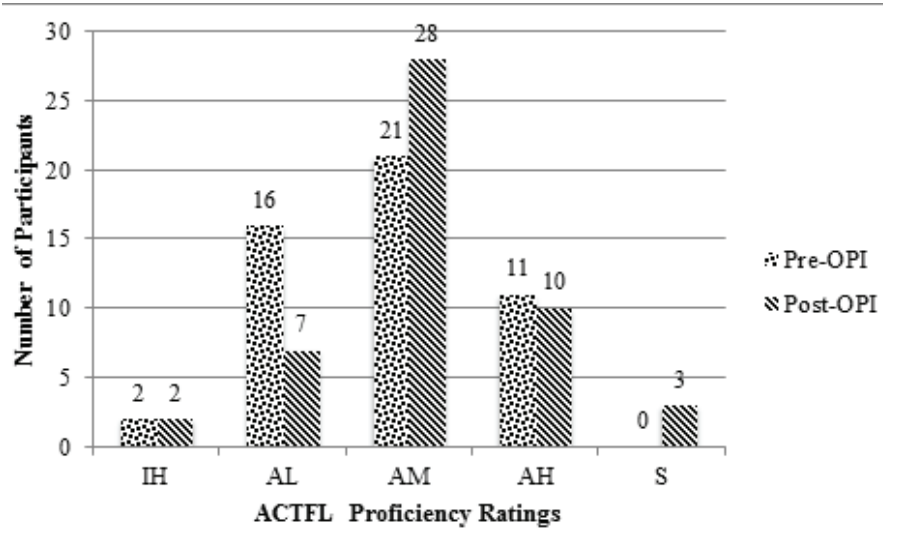
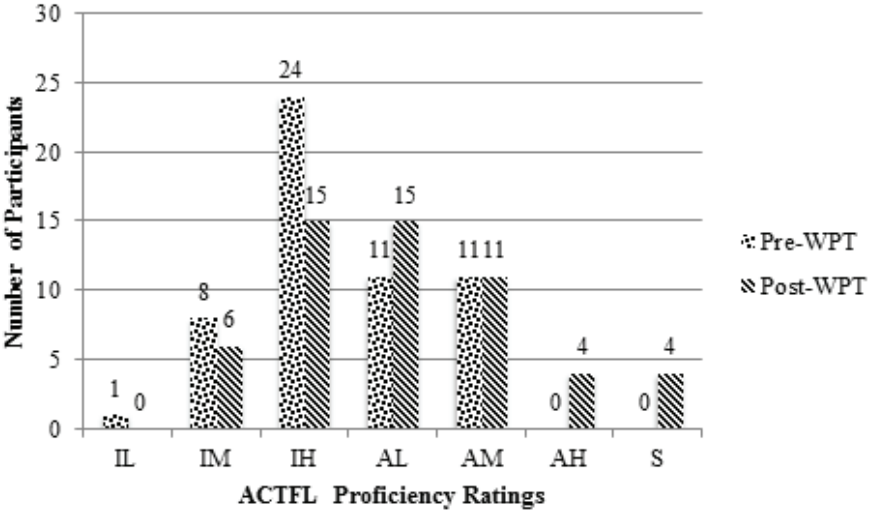


Figure 4.1 Comparison of Pre- and Post-OPI Ratings



◆ Figure 4.2 Comparison of Pre- and Post-WPT Ratings

◆ Table 4.1 Pre–/Post–Oral Proficiency by Number of Participants

		Post-OPI				
	ACTFL rating	**IH**	**AL**	**AM**	**AH**	**S**
Pre-OPI	IH	0	1	1	0	0
	AL	2	6	8	0	0
	AM	0	0	16	5	0
	AH	0	0	3	5	3

◆ Table 4.2 Pre–/Post–Writing Proficiency by Number of Participants

		Post-WPT						
	ACTFL rating	**IL**	**IM**	**IH**	**AL**	**AM**	**AH**	**S**
Pre-OPI	IL	0	1	0	0	0	0	0
	IM	0	4	3	1	0	0	0
	IH	0	1	11	7	4	1	0
	AL	0	0	0	4	4	1	2
	AM	0	0	1	3	3	2	2

$t = -3.78$, $p = .0005$). In terms of writing proficiency, the data likewise suggest that the higher participants rated on the pre-WPT, the less gain they demonstrated during the semester (slope/effect size = −0.19, $t = -1.47$, $p = 0.15$). Results of carrying out a paired t-test for overall gain indicated that participants demonstrated significant gain in both skill sets (OPI: estimated mean = 0.28, $t = 2.95$, $p = 0.0049$; WPT: slope/effect size = 0.65, $t = 4.59$, $p < 0.0001$).

Such rapid language uptake accomplished by meeting for two hours per week over the span of one academic semester speaks to the efficacy of a task-based approach to foreign language instruction in the classroom. Figures 4.1 and 4.2 suggest general upward mobility in oral and writing proficiency, and Tables 4.1 and 4.2 further clarify the number of participants who demonstrated sublevel and/or threshold gain. Of the five participants who rated Advanced High on the pre-OPI (a rating that indicates emerging ability at the Superior level), three crossed over into Superior. Interestingly, in writing, four participants reached Superior; two of those four pretested at Advanced Low and the other two pretested at Advanced Mid, thus demonstrating both sublevel and threshold gain.

The outcomes of this course likewise suggest that improved writing proficiency facilitates improved oral proficiency. Performing a t-test on the slope of the parameter indicates that gain in writing proficiency acts as a strong predictor of gain in oral proficiency (slope/effect size = 0.21, $t = 2.51$, $p = 0.016$). Data from this research strongly accord with this observation and past research (cf., Brecht et al. 1995; Higgs and Clifford 1982; Magnan and Back 2007), in particular, that accuracy, cultivated in the reflective skill of writing, contributes not only to improved writing proficiency but also to overall improved oral proficiency. As Sir Francis Bacon observed, "Reading maketh a full man, conference a ready man, and writing an exact man" (1908, 155).

Conclusion

Rifkin (2003) examined OPI scores in relation to classroom instruction hours of students of Russian. Findings from his research revealed that "students entering at the Intermediate Mid level alone represented a range of 180 to 600 hours of prior classroom instruction (with an average of 350 hours of classroom instruction)" (583). The vast differences in the amount of class instruction suggest that time alone may not be the most critical factor in terms of acquiring proficiency. Instead, Rifkin (2003) and Robin (2000) argue that designing appropriate curricula, tailored to specific proficiency outcomes, is needed. Indeed, the chapters in this volume indicate that focused curricular design can have a major impact on the development of proficiency. Leaver and Campbell (this volume) and Kennedy and Hansen (this volume) illustrate the value of well-designed curricula for at-home study. Similarly, the results of the course described in this chapter demonstrate that students can make appreciable gains in oral and writing proficiency without traveling to the host country. Moreover, course outcomes indicate that students can make significant progress in proficiency with as few as two contact hours weekly within the span of a single semester.

Thus, an approach in which learners are given tasks that are aligned with the desired proficiency outcomes, along with the appropriate scaffolding to complete those tasks, can have a significant impact on learners' linguistic outcomes. Moreover, the data from this study suggest that writing should be given a prominent place in the curriculum. As learners became more proficient in their written Russian, their oral proficiency also improved. The careful attention writers give to tailoring their language for their audience may lead to greater uptake of the vocabulary and grammatical features that are critical at the Advanced and Superior levels.

In this chapter, as with others in this volume, we demonstrate the importance of curricular design for achievement of Advanced and Superior levels of proficiency. While time in the target language certainly plays a valuable and needed role, this research suggests that innovative curricular design and instruction in the university foreign language classroom combined with appropriate scaffolding can equal, if not exceed, uptake that occurs in extended immersion environments. Rather than discounting the value of extended immersion programs or setting up an "either/or" dichotomy, this research suggests that carefully designed learning tasks can lead to overcoming the "ceiling effect" for classroom foreign language learning and facilitate ongoing progress in the target language.

Notes

1. Georgetown University Press is scheduled to release both volumes in fall 2014.
2. Although often attributed to Lev Vygotsky, the concept of scaffolding originated with Jerome Bruner (see Bruner and Sherwood 1975).
3. Five of the fifty-five participants rated Superior on the pre-OPI, which created a ceiling effect in data. This obstacle was overcome in the data analysis by accounting only for those with pre-OPI ratings of AH and lower, hence an N value of 50 for OPI-related data and an N value of 55 for WPT-related data.

References

American Council on the Teaching of Foreign Languages. 2012. *ACTFL Proficiency Guidelines—Speaking*. Alexandria, VA: ACTFL.

Bacon, Francis. 1908. "Of Studies." In *The Essays of Francis Bacon*, edited by Clark Sutherland Northup, 154–156. Boston: Houghton Mifflin.

Brecht, Richard D., Dan E. Davidson, and Ralph B. Ginsberg. 1995. "Predictors of Foreign Language Gain during Study Abroad." In *Second Language Acquisition in a Study Abroad Context*, edited by Barbara F. Freed, 37–66. Amsterdam: John Benjamins Publishing Company.

Bruner, Jerome S., and V. Sherwood. 1975. "Peekaboo and the Learning of Rule Structures." In *Play: Its Role in Development and Evolution*, edited by Jerome S. Bruner, Alison Jolly, and Katherine Sylva, 277–285. Harmondsworth, UK: Penguin Books.

Cohen, Andrew D. 1994. *Assessing Language Ability in the Classroom*. 2nd ed. Boston: Heinle & Heinle.

Connor, Ulla. 1987. "Argumentative Patterns in Student Essays: Cross-Cultural Differences." In *Writing Across Languages: Analysis of L2 Text*, edited by Ulla Connor and Robert Kaplan, 57–71. Reading, MA: Addison-Wesley.

Coyle, Do, Philip Hood, and David Marsh. 2010. *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press.

Ellis, Rod. 1997. *Second Language Acquisition*. Oxford: Oxford University Press.

Fessenden, Seth A., Roy I. Johnson, P. Merville Johnson, and Kaye M. Good. 1973. *Speech for the Creative Teacher*. Dubuque, IA: William C. Brown Company Publishers.

Gass, Susan, and E. Marlos Varonis. 1994. "Input, Interaction, and Second Language Production." *Studies in Second Language Acquisition* 16: 283–302.

Hedegaard, Mariane. 2005. "The Zone of Proximal Development as Basis for Instruction." In *An Introduction to Vygotsky*, edited by Harry Daniels, 227–251. 2nd ed. London: Routledge.

Higgs, Theodore V., and Ray T. Clifford. 1982. "The Push toward Communication." In *Curriculum Competence and the Foreign Language Teacher*, edited by Theodore V. Higgs. Lincolnwood, IL: National Textbook.

Inch, Edward S., and Barbara Warnick. 2005. *Critical Thinking and Communication: The Use of Reason in Argument*. 5th ed. Boston: Allyn and Bacon.

Kaplan, Robert B. 1987. "Cultural Thought Patterns Revisited." In *Writing across Languages: Analysis of L2 Text*, edited by Ulla Connor and Robert B. Kaplan, 9–21. Reading, MA: Addison-Wesley.

Leaver, Betty L., and Marsha A. Kaplan. 2004. "Task-Based Instruction in US Government Slavic Language Programs." In *Task-Based Instruction in Foreign Language Education: Practice and Programs*, edited by Betty L. Leaver and Jane R. Willis, 47–66. Washington, DC: Georgetown University Press.

Long, Michael H. 1996. "The Role of the Linguistic Environment in Second Language Acquisition." In *Handbook of Second Language Acquisition*, edited by William C. Ritchie and Tej K. Bhatia, 413–468. San Diego: Academic Press.

———. 2007. *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Magnan, Sally S., and Michele Back. 2007. "Social Interaction and Linguistic Gain during Study Abroad." *Foreign Language Annals* 40: 43–61.

Massie, J. 2005. "Consideration of Context in the CBI Course Development Process." In *Content, Tasks and Projects in the Language Classroom: 2004 Conference Proceedings*, edited by Renee M. Jourdenais and Sarah E. Springer, 79–91. Monterey, CA: Monterey Institute of International Studies.

MLA Ad Hoc Committee on Foreign Languages. 2007. *Foreign Language and Higher Education: New Structures for a Changed World*. The Modern Language Association of America.

Redding, W. Charles. 1954. "Composition of the Debate Speech." In *Argumentation and Debate: Principles and Practices*, edited by David Potter, 193–213. New York: The Dryden Press.

Rifkin, Benjamin. 2003. "Oral Proficiency Learning Outcomes and Curricular Design." *Foreign Language Annals* 36: 582–588.

Robin, Richard. 2000. "Foreign Languages across the Curriculum and the Proficiency Paradox." In *The Learning and Teaching of Slavic Languages and Cultures*, edited by Olga Kagan and Benjamin Rifkin, 29–43. Bloomington, IN: Slavica.

Shaw, Peter A. 1997. "With One Stone: Models of Instruction and Their Curricular Implications in an Advanced Content-Based Foreign Language Program." In *Content-Based Instruction in Foreign Language Education: Models and Methods*, edited by Stephen B. Stryker and Betty L. Leaver, 261–282. Washington, DC: Georgetown University Press.

Stryker, Stephen B., and Betty L. Leaver. 1997. "Content-Based Instruction: From Theory to Practice." In *Content-Based Instruction in Foreign Language Education: Models and Methods*, edited by Stephen B. Stryker and Betty L. Leaver, 3–28. Washington, DC: Georgetown University Press.

Swain, Merrill. 1993. "The Output Hypothesis: Just Speaking and Writing Aren't Enough." *The Canadian Modern Language Review/La Revue Canadienne des langues vivantes* 50: 158–164.

Swender, Elvira, ed. 1999. *ACTFL Oral Proficiency Tester Training Manual*. Yonkers, NY: American Council on the Teaching of Foreign Languages.

van Lier, Leo. 2005. "The Bellman's Map: Avoiding the 'Perfect and Absolute Blank' in Language Learning." In *Content, Tasks and Projects in the Language Classroom: 2004 Conference Proceedings*, edited by Renee M. Jourdenais and Peter A. Shaw, 13–21. Monterey, CA: Monterey Institute of International Studies.

Vygotsky, Lev S. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Williams, Jessica. 2005. *Teaching Writing in Second and Foreign Language Classrooms*. Boston: McGraw Hill.

Willis, Jane R. 2004. "Perspectives on Task-Based Instruction: Understanding our Practices, Acknowledging Different Practitioners." In *Task-Based Instruction in Foreign Language Education: Practice and Programs*, edited by Betty L. Leaver and Jane R. Willis, 3–46. Washington, DC: Georgetown University Press.

This page intentionally left blank

Chapter 5

Chinese for Special Purposes: Individualized Instruction as a Bridge to Overseas Direct Enrollment

MATTHEW B. CHRISTENSEN AND DANA S. BOURGERIE

STUDYING IN AN OVERSEAS university environment alongside native speakers of the respective target culture is a significant step toward obtaining American Council on the Teaching of Foreign Languages (ACTFL) Superior-level language proficiency. Such an approach to foreign language study differs significantly from many study abroad programs that offer classes in the target language designed specifically for foreigners. Direct enrollment requires a student to have a high level of proficiency in all four modalities, namely, listening, speaking, reading, and writing. Furthermore, students must undergo domain-specific training prior to enrolling in the classes, familiarize themselves with nonstandard accents of the foreign language, and demonstrate cultural literacy. The traditional foreign language classroom falls short of the kind of intense training needed to prepare students for direct enrollment courses and general integration at a foreign university. Ultimately, a direct enrollment experience elevates students' proficiency levels to the point where they can fully participate in the target language environment, regardless of their chosen discipline.

This chapter begins by describing what students need to know in order to succeed in a direct enrollment environment. It then discusses the importance and challenges of domain-specific language training and how universities typically approach this matter. The core of the chapter describes an individualized teaching and learning approach used in the Chinese Flagship[1] Program at Brigham Young University (BYU), which prepares students for direct enrollment and professional opportunities in China. Findings presented herein reflect the learning outcomes of courses taught between 2003 and 2012.

The Need for Superior-Level Skills

For years, the expectation of foreign professionals conducting business in the United States has been that they do so in English. However, Americans conducting business in China typically hire an interpreter or expect their Chinese counterparts to speak English. Such an uneven playing field can lead to a language disconnect between the two parties and possible misunderstanding. Grosse and Voght point out that "most American professionals—whether in business, government, medicine, law, or travel and tourism—lack the basic skills needed to cultivate professional relationships with colleagues in foreign countries (or speakers of other languages in the United States) and do not have easy access to new ideas and developments in their areas of professional activity" (1991, 184).

Cultivating relationships is essential to conducting business in China. Chinese often will spend a couple of days sharing meals and visiting cultural sites before beginning any serious discussions. They want to know and trust their counterparts before discussing business with them. If direct communication is not possible, conducting business may become challenging. Conversely, an American speaking Chinese sends a strong, positive message that s/he is serious, committed, and can be trusted.

Challenges of Direct Enrollment

Students from non-English-speaking countries typically come to the United States to enroll directly in degree granting programs. According to the Institute for International Education's (IIE) *2013 Open Doors Report* ("Leading Places of Origin"), 194,029 Chinese students enrolled in institutions of higher learning in the United States in 2011/12, but only 5.1 percent of those enrollees were in an English as a Second Language program. By contrast, of the nearly 15,000 (14,887 to be precise) Americans who studied in China in 2011/12, virtually all of them were in protected study abroad programs (IIE "Leading Destinations"). There are many complex reasons for this difference.

Historically, few American students have developed the level of language proficiency requisite for direct enrollment in overseas universities. Beyond linguistic impediments, cultural and logistical issues factor in as well. For example, differences between various countries' educational systems have made receiving degrees abroad less attractive for Americans. Moreover, the relative strength and reputation of the American higher education system contributes greatly to comparatively higher international enrollments in the United States. This disparity in direct enrollment largely stems from American students' lack of the necessary linguistic and cultural skills to function in authentic overseas language environments.

Beyond Linguistic Ability

Gaining high-level linguistic skills, e.g., a strong command of grammar and vocabulary, though necessary to fully engage in a direct enrollment course, does not guarantee successful communication of the target culture. Rather, it represents only a part

of the skill set needed to ensure successful communication. In addition to general linguistic training, learners need to understand their chosen domain of study in the target language, have some familiarity with nonstandard varieties of target language, and understand and use presentational discourse and professional cultural practices. Finally, they must demonstrate cultural literacy, i.e., an understanding of the target culture value system, aspects of popular culture, media, history, and societal norms.

Domain Training

To understand what takes place in a foreign university class, learners must first familiarize themselves with the academic culture, the language, and the context of the domain they are studying—be it biology, journalism, business, or the like. Each discipline has its unique jargon and culture that requires specialized training. Such training enables students to participate fully in the class from the outset; students who lack this domain training typically struggle in a direct enrollment class. Moreover, beyond the specialized knowledge of the subject matter, students also need to know something about university culture itself. The Flagship training addresses each of these factors.

Exposure to Nonstandard Language

There are seven main dialect groups in China, which comprise approximately 240 distinct dialects within these groups. Not only are these dialects distinct from each other, they often are mutually unintelligible. Even subdialects of the same main dialect group can be unintelligible, or at least difficult to understand; at times, even native Chinese struggle to understand each other. The university professors that the US students interact with in China come from all over the country and speak a variety of dialects. Although the students are educated Chinese speakers and understand Standard Mandarin Chinese (the national language), they may speak it with varying degrees of accuracy and standard usage. If students are exposed only to Standard Mandarin, they will struggle to understand nonstandard accents. Specific training in understanding nonstandard Chinese pronunciation plays an essential role in preparing students for direct enrollment, even if that training is limited to listening comprehension.

In addition to exposure to different varieties of spoken Chinese, students also can benefit from training in different styles of handwriting. Although the Chinese writing system is standardized and commonly used in nearly all contexts of society in China, handwriting styles can vary substantially from person to person—far more so than it can with Latin-based languages. Accordingly, students benefit from training that exposes them to different varieties of both spoken and written Chinese.

Chinese Presentational Skills

The fact that the Chinese do things differently than Americans should come as no surprise, and classroom practices in China are no exception. Students must learn presentational skills in order to fully participate in their classes. Not only do presentational styles differ, but the very culture of the classroom contrasts with the US

classroom as well. For example, in a traditional Chinese classroom students view the teacher as the ultimate authority and, as such, asking questions can appear disrespectful. Likewise, Chinese education tends to emphasize theory over practice and memorization of detailed facts over critical analysis. Unlike in US classrooms, tests often serve as the sole basis for grades (Hu et al. 2010, 78–79). For these reasons and others, US students need explicit training in giving presentations, writing papers, interacting in the classroom, maintaining teacher-student relationships, and observing in-class protocols within the Chinese university educational context.

Cultural Literacy

Cultural literacy refers to the kinds of information and practices that members of a society value and participate in, and may include ways of interacting, presenting oneself, understanding how the media functions in society, having a general knowledge about society, and properly using idioms, allusions, and historical references. The burden of cultural literacy can be so great that even native members of a particular culture may demonstrate significant gaps in their own cultural knowledge (see Schweizer 2009). Thus, predeparture instruction in matters of cultural literacy facilitates successful integration into a new cultural context.

Learning Languages for Special Purposes

Language for special purposes courses began appearing during and shortly after World War II when there was considerable interest in Russian science and technology and Japanese technology and business. Later, efforts to engage immigrant populations prompted courses like Spanish for social workers or health care professionals. Grosse and Voght report that a new direction in foreign language study can be attributed to "the national trend toward broadening the nature of language study to economic and political pressure, rather than to a renewed interest in humanistic education" (1991, 184). Historically, language for special purposes courses covered a broad range of domains, such as Japanese for tourism, Spanish for health care personnel, Business Chinese, Technical Russian, French for journalism, and the like. Despite the seemingly wide variety of courses taught across multiple domains, the vast majority of them, in fact, deal with business language. Technology language ranks second in terms of popularity for domain-specific courses. According to Grosse and Voght (1991), numerous articles and professional journals have reported on language for special purposes courses, the majority of which are housed in foreign language departments in countries such as the United States, Canada, France, and Germany. Long and Uscinksi (2012) discuss how courses in language for special purposes have evolved in the United States since 1990. They report that 62 percent of foreign language departments that responded to their survey offered courses in language for special purposes, the majority of which were for business language. They also report that the depth and breadth of language for special purposes has grown to address the changing needs of working professionals.

Classes taught in foreign language departments that focus on a single domain, such as Business Chinese or French for Business, work well as they have the advantage

of devoting an entire term to a specific domain. In recent years, there has been a marked increase in the development and publication of Business Chinese materials. Curricula and methodologies also have been designed to meet the specific needs of those interested in learning Business Chinese (see Liu 2004; Yuan 2006); however, domains other than business remain largely undeveloped, putting students with an interest in those other domains at a disadvantage. With the exception of a few select technology-related classes in languages such as Japanese (Dowling and Mitchell 1993), demand typically has not justified the creation of dedicated courses in other disciplines. Language instructors are then faced with the challenge of finding ways to serve students wanting domain-specific training in low-demand languages. Often, students with other professional interests are out of luck or left to pursue training on their own.

Teaching Chinese for special purposes is at the core of the Flagship programs around the country. Some Chinese Flagship programs have addressed this need through offering content courses taught in Chinese. For example, green chemistry and Chinese history have been taught at the University of Oregon's Chinese Flagship Program by native Chinese faculty members in these fields. Likewise, Arizona State University's Chinese Flagship Program has offered courses in Chinese religions and economics. While content courses do provide some training in language for special purposes, these courses do not offer structured, focused attention on language learning and language utilization strategies. In other words, they are not foreign language courses like language for specific purposes (LSP) courses are. For more information on Flagship programs in general, see chapter 7 and chapter 9 (this volume).

Individualized Instruction

Individualized instruction, which was particularly popular in the 1970s and 1980s, has proven to be a viable alternative to the traditional foreign language classroom. Harlow reported that in the 1970s, over 150 articles were published on the topic (1987, 389). She further reported that as of 1986, 111 foreign language departments across the United States offered courses with an individualized option or track. In most cases these individualized tracks provided an alternative to traditional foreign language classes and mostly dealt with either novice or intermediate foreign language learners or with heritage learners who already had prior background in the language. Most recently, Yuasa's (2013) collection addresses individualized instruction in the teaching of East Asian languages. In one chapter from that book, Christensen discusses how individualized instruction can be used effectively for teaching Chinese for special purposes at the advanced level.

Christensen and Wu (1993) have pointed out several advantages to an individualized curriculum, the most notable being the actual amount of student performance time. In a traditional classroom, they estimate that, on average, students have only four to six minutes of actual speaking time, which includes responding to questions posed by the teacher in a typical fifty-minute class period. This figure assumes a class size of twenty or fewer students, thus making the aforementioned figure a conservative estimate. In an individualized approach, students have a great deal more contact

time since during a typical individualized session, which may last from fifteen to fifty minutes, students work one-on-one with a teacher. With respect to speaking, traditional classrooms also present a challenge in that the bulk of communication occurs in a group setting rather than one-on-one, as is typically the case in the real world. Finally, teachers in the traditional classroom deal with multiple student backgrounds and differing strengths and weaknesses. An individualized approach affords the flexibility to spend more time on student weaknesses and less time on their strengths. More importantly, individualized instruction allows teachers to address a wide variety of specific domains, thereby circumventing the need to offer traditional classes in which only a few students may express an interest.

Individualized Instruction Is Not Independent Study

A common misconception about individualized instruction is that students learn independently without interaction with a teacher or other students. In fact, individualized instruction allows learners to move at their own pace under comparatively *more* guidance and direction from a teacher. Additionally, students meet regularly with a teacher to "pass off" their lessons. In some programs, such as the well-established one at Ohio State University, each fifteen-minute individual meeting between a student and a teacher represents a typical fifty-minute class period. Students who pursue individualized instruction meet regularly with teachers and sometimes with other students in group meetings. This method of instruction creates a more intensive and efficient learning environment than a traditional classroom. Students must come fully prepared; there cannot be any passive learners in this kind of a setting.

Individualized Instruction and Language for Special Purposes

There has been a great deal written about individualized instruction, but very little that describes how it applies to teaching foreign languages for special purposes (see Christensen 2013). As early as 1975, Pilkenton proposed the development of individualized materials for domains such as Spanish for medical care, Spanish for child care, Spanish for social workers, Spanish for law enforcement officers, etc. Knorr (1977) similarly proposed using an individualized approach for intermediate and advanced levels of instruction at small colleges where enrollment numbers in upper division courses are often low. Paucity of literature aside, individualized instruction seems a viable option for teaching languages for special purposes and, in turn, preparing students for future professional pursuits. It is this model that BYU has developed for its Flagship Program.

The BYU Flagship Model

The BYU Flagship program includes the following key components:

1. Individualized domain training course
2. Advanced conversation and rhetoric course

3. Personalized tutoring from Chinese native peers
4. Cultural preparation course (for direct enrollment and internships)
5. Direct enrollment at a Chinese university (Nanjing University)
6. Four- to six-month internship in a Chinese organization (business, industry, nongovernmental organizations [NGOs], educational institutions, etc.)

BYU is one of eleven Chinese Flagship programs in the United States. All Flagship programs share a common desired outcome: to produce global professionals with Superior-level (ACTFL scale) linguistic skills and the accompanying cultural skills necessary for professional work in China. Qualified students in all programs participate in the overseas capstone program by taking courses at Nanjing University and completing a four- to six-month professional internship in China (items 5 and 6 above). However, the curricular approach at each of the domestic Flagship programs varies.[2]

This chapter focuses specifically on the aforementioned individualized domain-training course of BYU's Flagship Program. Among its goals, the Chinese Flagship Program at BYU seeks to develop professional-level, domain-specific competence, both linguistically and culturally. This includes the linguistic and cultural skills necessary to read and write in professional contexts, and the ability to speak intelligently about a subject while giving a professional presentation on a specific topic within a given domain. Direct enrollment at a Chinese university represents an essential part of this learning process. Undergraduate students with Advanced-level Chinese skills are carefully selected to participate in the Flagship program. Having already attained ACTFL Advanced-level proficiency through enrollment in regular departmental courses and overseas programs, these students come to the Flagship program prepared to learn about their chosen disciplines in the target language. In order to ensure that students possess sufficient background in their major field of study prior to targeted language training, only students in their junior or senior year participate in the advanced track of the program.

Domain training also includes learning appropriate behavioral skills, such as participating in job interviews, giving professional presentations, and writing a Chinese-style résumé. To this end, the Chinese Section of the Department of Asian and Near Eastern Languages has designed specialized courses for Flagship students that develop the aforementioned skills. Prospective students must demonstrate Advanced proficiency in a battery of standardized tests, including the ACTFL Oral Proficiency Interview (OPI), the Chinese government *Hanyu Shuiping Kaoshi* (HSK), as well as computer adaptive reading and listening tests. Students also demonstrate their language ability through interviews, writing samples, presentations, etc. Table 5.1 outlines the components of the individualized domain training course, which will be discussed below.

Course Goals

The Chinese Flagship individualized course culminates with students writing a term paper in Chinese on their chosen topic and giving a professional oral presentation to a live audience. In order to achieve this objective, learners must develop two areas of

Table 5.1 Outline of the Individualized Training Course

Component	Description
Individual tutorials with faculty member	• Language tutorial (50 minutes); focus on linguistic issues • Content tutorial (50 minutes); focus on content and cultural issues
Weekly group meeting	• Focus on cultural literacy, professional presentation skills, and résumé and interview skills
Portfolio	• Final term paper (in Chinese) • Accumulated domain-specific lexicon • Article reviews
Final oral presentation	• 15-minute oral presentation to faculty and peers • 5-minute Q&A
Individual tutoring	• Peer tutors (5 hours per week)

competence: linguistic competence, which includes knowing the lexicon and communicative styles of the domain, and content competence, or the ability to talk and write about a subject in an intelligent manner, i.e., the way in which professionals are expected to communicate in the domain.

Course Materials

Compared to the quantity and variety of materials available for learning Business Chinese, the number of materials designed for teaching the language of nonbusiness domains pales in comparison—a difference likely stemming from low demand. Accordingly, traditional textbooks do not suit this course; instead, learners identify the bulk of their materials themselves with help from their instructors. One resource available to learners is a BYU-developed wiki that includes links to information on a broad variety of domains, such as engineering, biology, chemistry, journalism, politics, and law.

The wiki format is a collaborative online environment that allows all participants to add and edit useful sources, making them available to future students as well as themselves. This Flagship wiki site includes links to businesses, science and technology organizations, universities, high schools (often with links to textbooks and other information on specific domains), professional journals, and newspapers. It also provides abstracts of specific articles and reports on a variety of topics, thus providing a starting point for learning the language and content of the chosen discipline through the target language. Through the wiki site, students also can reference professional journals available electronically and in print form in the university library. Thus, as students become increasingly familiar with their disciplines, they begin to discover sources on their own, which they then add to the wiki. This process helps students to become experts in their chosen domain.

Course Methodology

Individualized sessions meet twice weekly and group sessions once weekly. While the course focuses on the individual student, the group session allows students to interact with fellow students and practice their oral presentations. In addition, students read about and discuss cultural issues and receive instruction in Chinese-style presentation techniques.

The two individualized sessions consist of fifty-minute one-on-one meetings with instructors. The first session focuses on language issues, i.e., helping a student understand language used in an article, and the second session focuses on content. Instructors conduct both sessions entirely in the target language. In fact, all communication between instructors and students occurs in the target language, inside and outside the classroom. Students receive instruction when learning about their selected domain in the form of two tasks: (1) creating a lexicon of terms related to their domain and (2) interviewing Chinese students who majored in their domain in China.

From the outset of the course and throughout the semester, students compile a lexicon of terms important for them to know about their discipline. As students read articles, they create a list of previously unfamiliar terms that they deem useful. The list expands with each weekly article that they read and takes the form of Chinese-to-Chinese definitions. Using the target language to explain what a word or phrase means gives students valuable language practice and represents an important Advanced-level skill. Every three weeks, students submit these lists to their instructors for input and evaluation.

Since the training received in China often differs from the training students receive at a university in the United States (e.g., educational paradigms, theoretical vs. practical types of knowledge), interviews help prepare students for direct enrollment at a Chinese university. Beyond their language-specific studies, learners also must interview a Chinese student (typically a graduate student) pursuing a similar field of study, in order to learn what majoring in a specific field at a Chinese university entails, e.g., the kinds of classes that are required, the typical amount of coursework, how the major differs in China from the United States, etc.

American students, even those who have been abroad before, are often shocked by differences between the Chinese and American systems. Direct enrollment exposes them to a type of classroom setting few Americans experience. It is distinctly different from a traditional study abroad context: no class syllabi, no office hours, and a general lack of access to the instructors. Communication generally happens through a class leader, who is not always obvious to the uninitiated. Moreover, students must learn the social structure of the classroom. For example, one may not wear hats or eat in class, but one may drink tea. Students may be criticized publicly for raising hands to ask questions because that is often interpreted as an affront to the teacher's authority. Flagship training also includes video footage of classrooms in China and entries from learning journals of previous students to supplement the efforts of the native peer tutor at bridging cultural differences.

At the beginning of the semester, students select a research topic within their domain that both interests them and that will be covered in courses related to their major. Once the topic has been selected, students search the wiki and other sources for articles that relate to their topic. Based on students' individual backgrounds and previous training, teachers and tutors help them select level-appropriate articles. With the help of their instructors, students select one article per week and conduct a careful, close reading of it. Once the lead instructor (content session) approves the article, it is sent to the other instructor (language session) one week in advance of the individualized sessions in order to allow time for instructor preparation.

The purpose of the individualized language sessions is to ensure that students understand the meaning of the articles they read on their own. To that end, teachers focus on grammar, vocabulary, and other structural issues. Students come having read the article and noted areas where they need help. To expedite the comprehension process, they use electronic dictionaries and annotators.[3] Graduate students trained in Chinese language pedagogy usually conduct the individualized language sessions and help students understand the social and cultural implications of the language in the article. Naturally, the effectiveness and efficiency of these sessions depends largely on the degree of student preparation.

In addition to answering questions, the instructor asks leading and probing questions related to grammar and language usage. Every three weeks, students complete a quiz based on the lexicon that they have created and on their understanding of the articles read and discussed. In addition, they undergo evaluation of their ability to respond to probing questions and to speak intelligently about their respective article.

Once students understand the content of an article, they discuss it with an instructor. Similar to the aforementioned individualized session, these sessions also last fifty minutes. The content instructor, who previously has read the article, will discuss the significance of the article with the student, how it relates to his/her research topic, and how this content would differ from a Western perspective. These discussions give the learner a chance to talk about pertinent issues within his/her domain, which represents an important step in becoming a global professional. In these sessions, the instructor not only questions students on what the article says, but more importantly, presses them to express and defend their opinions on the subject. Since instructors are usually not experts in multiple given domains, they rely on the student to explain otherwise unfamiliar concepts and terms. In this manner, the content instructor serves as a pedagogy expert with the skills to elicit discussion and explanation.

In these individualized content sessions, instructors also coach students in their preparation of a professional oral presentation and a written term paper. As the semester progresses, sessions shift from lively discussions of the articles to rehearsals for the final oral presentation. Every three weeks, students are evaluated on their presentations and discussions with the teacher. For students to be successful, they must demonstrate a high level of motivation and confidence as well as work well independently. Students typically work with their tutors in order to prepare for individualized sessions. They may also consult their tutors about general language issues.

Because many of these peer tutors have experience within their Chinese fields—either through undergraduate study or work—they prove invaluable as both cultural guides and linguistic informants.

In addition to the individualized sessions, students meet together in a weekly fifty-minute group session that serves three purposes: (1) to have students read and discuss articles on important cultural literacy topics, (2) to learn about Chinese-style professional practices, and (3) to give students the opportunity to practice their oral presentations in front of their peers. Cultural literacy topics refer to issues of particular relevance to China today, e.g., the People's National Congress, food safety, Mandarin and other Chinese dialects, the economy, the Internet age in China, and "Made in China." Many people in China know about these issues and frequently reference them in speech and in writing. Each week, students are assigned to read articles on these topics and write short reaction essays in Chinese. Class time is then spent discussing these issues. Additionally, students develop Chinese-style résumé writing skills, interviewing strategies, and professional oral presentation techniques. Such skills are essential in order to function in Chinese professional contexts. Finally, students have opportunities to practice their oral presentational skills by responding to questions from their instructor and peers and by expressing and defending their opinions. They are critiqued not only on their language, but also on their adherence to cultural norms.

Approximately every three weeks (three times during the semester), students submit milestone assignments for evaluation, that help keep them motivated and on track. These milestones play a valuable role in assessing learners' progress and are broken down as follows:

- Milestone 1
  a. Accumulated lexicon based on articles read (number of lexical items left to student's discretion)
  b. Final presentation outline/draft (both oral and written reports)
- Milestone 2
  a. Updated accumulated lexicon
  b. Final presentation second draft and at least eight slides for the oral presentation
- Milestone 3
  a. Continued accumulated lexicon
  b. Final presentation final draft and at least eighteen slides of the oral report

At each of these milestones, students are given an individualized language test based on their lexicon and the articles they have read for a given unit. By the end of the course, they have compiled a lengthy lexicon specific to their topic that reflects what they need to know. The lexicon serves as a personal resource for future research and coursework.

Outcomes and Assessment

Although the Flagship assessment routine has evolved over the ten-plus years of the program's existence, measuring student performance has always been part of the program. Accordingly, BYU Flagship has used a range of assessments to gauge student progress and to inform curricular design. Early in the program, data collected largely reflected exit tests (ACTFL-OPI or Interagency Language Roundtable (ILR)-OPI), but assessments have been adjusted according to demand and availability.

During the first three years of the program, students were tested via informal oral assessments and through interviews, but due to funding limitations, no full, rated OPIs were administered. In the last seven years, the testing process has been expanded considerably and students are now selected, in part, according to scores on standardized proficiency testing. To enter the advanced program, potential students must rate in the ACTFL Advanced range. They must also undergo in-house oral interviews to assess their ability to handle the target language in their domain of interest. Lastly, applicants are required to provide a Chinese writing sample and undergo standardized listening and reading tests.

Since the program's inception, the available tests have included the Chinese Computer-Adaptive Listening Comprehension Test (CCALT), the Computer-Adaptive Test for Reading Chinese (CATRC), and the ACTFL-OPI. Both CCALT and CATRC are also rated according to ACTFL proficiency guidelines. In 2012 the BYU Flagship Program replaced the CCALT and CATRC with the Adaptive Test for Listening (ALT) and the Adaptive Test for Reading (ART) respectively, both of which are also ACTFL-referenced assessments.[4] Since 2011, pre- and post-overseas capstone tests for all Chinese Flagship programs nationally have been administered by the American Councils for International Education (ACIE) using their own ILR-rated reading and listening tests. To assess writing, the program now uses the ACTFL Writing Proficiency Test (WPT) as an exit test. The HSK, a multi-skill test, also has been provided. Table 5.2 summarizes the current testing regime in the BYU Flagship Program.

Domestic Program Gain

Outlined in Tables 5.3 and 5.4 are pre- and post-domestic test data for the ACTFL-OPI, CCALT, and HSK, which show strong gains made on the domestic program on the ACTFL-OPI.

The gains for speaking, as demonstrated by the OPI, are striking and statistically significant, with 90 percent of the students tested gaining at least one sublevel. Only three students did not gain and one of the null gainers was already rated at Superior. On listening, the data set is smaller, but likewise confirms a pattern of significant gain, as illustrated in Tables 5.5 and 5.6.

The gains on the listening test are similar to the OPI results in their pattern. Over 73 percent of students gained one sublevel or more, and several gained multiple sublevels. Unlike the ACTFL-OPI data, there are some obvious outliers on both ends of the spectrum. It is worth noting that there were some instances of technical

Table 5.2 Assessments for the BYU Chinese Flagship Program

Entrance Tests	Baseline on Acceptance	Post-Domestic/ Pre-Overseas Capstone	Program Exit Tests
• ART, ALT • Informal interview • Writing sample	• ACTFL-OPI	• ACTFL-OPI • ACIE reading and listening	• ACTFL-OPI • ILR test • ACTFL-WPT ACIE reading and listening • HSK

Table 5.3 Gains Made on ACTFL-OPI after Domestic Program (Two Semesters)

Proficiency level gains	Number of students	Percentage of students
+1 sublevel	16	47
+2 sublevel	14	40
+3 sublevel	1	3
Null gain (Superior)	1	3
Null gain (AM or AH)	1	3
−1 sublevel	1	3
Total students	35	100[a]

[a] Total with rounding.

$p < 0.00000011$

N = 34, excluding a participant who had already attained a Superior rating on entry to the program.

Table 5.4 Cross Tabulation of ACTFL-OPI Gains by Sublevel after Domestic Program (Two Semesters)

		Post-		
		S	AH	AM
Pre-	S	1	1	0
	AH	1	2	0
	AM	5	5	0
	AL	7	3	9
	IH	0	1	0

N = 35

◆ Table 5.5 Gains Made on CCALT after Domestic Program (Two Semesters)

Proficiency level gains	Number of students	Percentage of students
+5 sublevel	1	6.7
+4 sublevels	2	13.3
+3 sublevels	1	6.7
+2 sublevels	4	26.7
+1 sublevels	3	20
Null gain	1	6.7
-1 sublevel	2	13.3
-2 sublevels	1	6.7
Total students	15	100[a]

[a] Total with rounding.
$p < 0.00157$
N = 15

◆ Table 5.6 Cross-Tabulation of CCALT[a] Gains by Sublevel after Domestic Program (Two Semesters)

		Post-				
		AH	AM	AL	IH	IL
Pre-	AH	0	1	0	0	0
	AM	0	0	1	0	0
	AL	0	1	1	0	0
	IH	0	7	1	0	1
	IL	0	1	0	1	0

[a] The Computer-Adaptive Listening Comprehension Test (CCALT) follows the ACTFL rating system: Novice-Intermediate-Advanced-Superior, with low, mid, and high gradations. In this table, we have AH=Advanced High, AM=Advanced Mid, AL=Advanced Low, IH=Intermediate High, and IL=Intermediate Low.

N=15

problems with the CCALT that may have affected the results on both ends, perhaps accounting for these outliers. Nevertheless, the pattern of gain seems as clear for listening as it is for speaking.

Finally, the HSK test data echoe the results for ACTFL-OPI and CCALT. The HSK tests several skills (vocabulary, listening, reading, and grammar), but Tables 5.7 and 5.8 show the global score on the original HSK test, i.e., they show the nature of student gains, with all but one of thirteen students who took a pre- and post-test gaining at least one level on the Intermediate or Advanced HSK (scale of 1–14 covering three subtests).[5]

Despite relatively small samples, outcomes for all three tests demonstrate clear evidence of the effectiveness of the domestic training with respect to core linguistic

Table 5.7 Gains Made on HSK after Domestic Program (Two Semesters)

Proficiency level gains	Number of students	Percentage of students
+3 levels	1	7.7
+2 levels	4	30.7
+1 levels	7	53.8
Null gain	1	7.7
Total students		100[a]

[a] Total after rounding.
$p < 0.00014$
N = 13

Table 5.8 Cross-Tabulation of HSK[a] Gains by Sublevel after Domestic Program (Two Semesters) on Scale of 1–14

		Post-						
		3	4	5	6	7	8	9
Pre-	3	0	1	0	0	0	0	0
	4	0	0	1	0	0	0	0
	5	0	0	1	1	2	0	0
	6	0	0	0	0	3	1	1
	7	0	0	0	0	0	1	1

[a] The original HSK was divided into three tests: Beginner (1–3), Intermediate (3–8), and Advanced (9–11). As a reference, a level 6 had been the requirement for a foreigner to directly enroll in a Chinese university in the social sciences and humanities, whereas level 4 was the requirement for the sciences. In 2010, the Chinese government released a new HSK with six separate tests.

N=13

skills. It should also be noted that the higher learners progress on the scale, the more difficult advancing to the next level becomes. In other words, crossing the Advanced High/Superior threshold represents a far greater jump than, for example, crossing the Intermediate High/Advanced Low threshold or making a sublevel gain from Advanced Low to Advanced Mid.

Beyond the standardized measures, students have portfolios of work that attest to the skills and abilities that they develop. These portfolios include video samples of debate performances and of professional-level presentations to Chinese-speaking audiences. Additionally, students produce a written term paper involving multiple revisions, resulting in highly developed prose. Both the presentation and writing skills become part of a foundation that allows participants to eventually participate in direct enrollment courses in China and to successfully negotiate an internship in a Chinese context—the ultimate goal of the program.

Conclusion

Taking direct enrollment courses at universities in the People's Republic of China presents a daunting task, thus making individualized instruction and training prior to entering the university classroom in China essential for students of the Chinese language in the United States. Traditional language courses and methodologies, e.g., language for special purposes courses, often fall short in this regard, meaning they excessively restrict topic offerings. To succeed in direct enrollment courses, students need training in their specific domain, as well as training in areas of cultural appropriateness.

An individualized approach to teaching Chinese for special purposes has proven successful in preparing students for direct enrollment in Chinese universities. The program at BYU enables instructors to focus on individual students' weaknesses and to direct their learning based on a self-selected domain. It also targets nonlinguistic skills that factor heavily in a student's success within demanding overseas learning programs. The statistics cited above and administrators' and instructors' personal observations clearly illustrate the value of this model; however, more rigorous and systematic proficiency testing in all four skills stands to strengthen these claims.

Notes

1. The Brigham Young University Flagship Program was established in 2002 as part of a larger Flagship consortium, which now includes eleven Chinese programs. There are also Flagship programs for other critical-needs languages: Arabic, Hindi, Korean, Persian, Portuguese, Russian, Swahili, Turkish, and Urdu. Although each domestic model differs, all Chinese Flagship Programs flow to two overseas centers at Nanjing University and at Tianjin Normal University. For a full list of institutional participants, see http://thelanguageflagship.org.
2. Spring (2012) describes one such approach at Arizona State University. Falsgraf and Bourgerie (2008) also discuss the general features of the Flagship approach in relation to Advanced-level instruction.
3. Electronic annotators are similar to electronic dictionaries, but they also can generate romanization from characters, word lists, etc.
4. ART and ALT (http://chineseflagship.byu.edu/node/10) were developed by BYU's Center for Language Studies in cooperation with ACTFL and with support from the Defense Language National Security Education Office (DLNSEO). Early versions of the tests were part of a battery commissioned by the Tri-Service Academy group. These criterion-referenced tests currently assess through the IRL 3 level (ACTFL Superior).
5. In 2010, the Chinese National Office for Teaching Chinese as a Foreign Language, or Hanban, released a new version of the HSK. The revamped format has six subtests that students must take individually based on perceptions of their own language ability.

References

Christensen, Matthew B. 2013. "Chinese for Special Purposes: An Individualized Approach." In *Individualized Instruction in East Asian Languages*, edited by Etsuyo Yuasa, 39–59. Columbus, OH: Foreign Language Publications, Ohio State University.

Christensen, Matthew B., and Xiaoqi Wu. 1993. "An Individualized Approach for Teaching False Beginners." *Journal of the Chinese Language Teachers Association* 28 (2): 91–100.

Dowling, Carol, and Anita Mitchell. 1993. "Reading in a Specific Purpose Foreign Language Course: A Case Study of Technical Japanese." *Modern Language Journal* 7: 443–44.

Falsgraf, Carl, and Dana Scott Bourgerie. 2008. "The Language Flagship: Multiple Approaches to Creating Global Professionals." In *U.S.-China Educational Exchange: Perspectives on a Growing Partnership*, edited by Shepherd Laughlin, 83–97. New York: Institute of International Education

Grosse, Christine Uber, and Geoffrey M. Voght. 1991. "The Evolution of Languages for Specific Purposes in the United States." *Modern Language Journal* 75 (ii): 181–95.

Harlow, Linda. 1987. "Individualized Instruction in Foreign Languages: A Survey of Programs at the College Level." *Modern Language Journal* 71: 338–94.

Hu, Wenzhong, Cornelius N. Grove, and Zhuang Enping. 2010. *Encountering the Chinese: A Modern Country, An Ancient Culture*. 3rd ed. Boston and London: Nicholas Brealey Publishing.

Institute for International Education. *2013 Open Doors Report*. http://www.iie.org/Research-and-Publications/Open-Doors.

Knorr, Walter L. 1977. "Individualized Specialization in Intermediate and Advanced Language Courses in the Small College." *ADFL Bulletin* 8 (iii): 36–37.

Liu, Meiru. 2004. "Designing a Functional Chinese Curriculum for a Graduate International Business Program." *Journal of the Chinese Language Teachers Association* 39 (1): 101–16.

Long, Mary K., and Izabela Uscinski. 2012. "Evolution of Languages for Specific Purposes Programs in the United States: 1990–2001." In "Languages for Specific Purposes in the United States in a Global Context: Update on Grosse and Voght (1991)," special issue, *Modern Language Journal* 96: 173–89.

Pilkenton, W. D. 1975. "A Proposal to Develop Individualized Materials for Limited-Objective Career Spanish Programs." *ADFL Bulletin* 6 (iii): 36–38.

Schweizer, Bernard. 2009. "Cultural Literacy: Is It Time to Revisit the Debate?" *Thought and Action* Fall: 51–56.

Spring, Madeline. 2012. "Languages for Specific Purposes Curriculum in the Context of Chinese Language Flagship Programs." *Modern Language Journal* 96:140–57.

Yuan, Fangyuan. 2006. "Case Study Approach, Task-Based Language Teaching and a Framework for Business Chinese Instruction." *Journal of the Chinese Language Teachers Association* 41 (1): 13–30.

Yuasa, Etsuyo, ed. 2013. *Individualized Instruction in East Asian Languages*. Columbus, OH: Foreign Language Publications, Ohio State University.

This page intentionally left blank

Chapter 6

Taking on the "Ceiling Effect" in Arabic

R. KIRK BELNAP AND KHALED ABUAMSHA

AS OF 2013, WELL over thirty thousand students were pursuing the study of Arabic in US colleges and universities, and extensive surveying indicates that most of those students want to reach a level of proficiency that would allow them to use Arabic comfortably in their professional activities (Belnap and Nassif 2011, 5).[1] However, there is little chance of timely success unless they can increase the amount of time they spend on task to even more than that typically amassed by foreign language majors in four years, and thereby break through the ceiling imposed by a typical undergraduate course of study (Rifkin 2005). This breakthrough is doable; some undergraduates are reaching Advanced High- and even Superior-level proficiency in Arabic (Al-Batal 2007–2008). The key? Extended periods of intensive study. However, not all intensive study programs are created equal.

Time on task is likely to remain the most important factor, but innovative pedagogical practices that result in higher levels of engagement and, therefore, expedited learning warrant careful consideration. For example, Rifkin's analysis of Russian data from on-campus and intensive contexts shows the latter to result in greater gains in language proficiency (2005). In their landmark study of thirteen hundred study abroad students, Vande Berg, Connor-Linton, and Paige (2009) found that students who took content courses during a study abroad program achieved significantly higher gains in speaking and intercultural understanding than those who enrolled in on-campus language courses. Highly engaging intensive experiences presumably contribute to more time *on task* (i.e., to greater focus and actual use of the target language). This chapter documents one institution's efforts to help its students acquire professional-level proficiency and concludes with suggestions for students who are not enrolled at such an institution. Particular emphasis is given to the cohort of Brigham Young University (BYU) students who studied abroad in 2011, given that their experience has been the most thoroughly documented to date.[2]

Pre-Program Preparation

Participation in BYU's study abroad program became a core requirement of its new Middle East Studies/Arabic (MESA) major, launched in 2002, which has become the most popular area studies major on campus, eclipsing Latin American Studies in 2007. Students are told on their first day of Arabic 101 that in as little as twelve months they could be stepping off of an airplane to begin an experience designed to result in the acquisition of Advanced-level proficiency in both speaking and reading.[3] The nearness of this goal is underscored daily, given that most Arabic 101 and 102 students have an instructor or teaching assistant (TA) who recently returned from the study abroad program and is enthusiastic about it. These peer role models play a vital part in helping students sustain their motivation.[4]

While studying in-country plays a central role in facilitating significant gain for most BYU students, overconfidence in its effectiveness can undercut their resolve to adequately prepare for it. Prior to going abroad, some students find that keeping up with the rigorous expectations of the Arabic program is difficult, especially in Arabic 201 and 202. Many are working to save money for their study abroad experience and, pressed for time, fall victim to the myth that in-country study is so superior that it will more than make up for spotty preparation in the United States. When the BYU Arabic faculty found that they were spending a considerable amount of time abroad providing remedial learning opportunities for students who were not adequately prepared for the intensive program, they instituted a requirement that all participants must pass Arabic 202 with at least a B− (an option for all, given that grading in the course is criterion referenced). Raising the bar in this way has resulted in students who are better prepared and the faculty are now able to focus far more of their efforts on students who are ready to benefit from the program.

As a prerequisite for acceptance into the intensive program, all applicants are interviewed individually. Those admitted to the program must complete a one-credit-hour course where they receive training on the history and culture of the country, program rules, and strategies for making the most of their time abroad (which includes having realistic expectations, avoiding comparisons of themselves with others, and cultivating a habit of positive self-talk, and other insights from Project Perseverance).[5] Students are also exposed to multiple case studies of former students, including some specifically directed toward female students in order to help them understand how to deal with sexual harassment and its effects. Male students are advised on how to assist female students in having a safe and meaningful experience. Pellegrino Aveni (2005) stressed the importance of such pre-program preparations as a critical ingredient for successful in-country language study. Vande Berg, Connor-Linton, and Paige (2009) found that students who had the benefit of a predeparture orientation with a cultural component made significantly greater oral proficiency progress and reported significantly higher satisfaction with their study abroad experience.

The students' third- and fourth-semester classroom experience is also specifically designed to prepare students to succeed in the overseas intensive program and acquire Advanced-level proficiency in speaking and reading.[6] For example, given that

they will be reading newspapers in Arabic on a regular basis, midway through Arabic 201, students begin reading daily headlines from the Arab press. Later in 202, short newspaper readings become a part of the learners' daily routine and they learn strategies for reading texts replete with unfamiliar words. Their professors coach them on how to discriminate between words they should look up or write down (whether in a written text or a conversation) and on how to recognize high-frequency and/or essential words. Mastering large amounts of key vocabulary is their central focus, but knowing how and when to let unessential words go is just as important. Their professors regularly help students practice these and other effective study skills and habits that will contribute to their having successful experiences abroad, including making a consistent effort, breaking out of one's comfort zone, recognizing signs of discouragement, and finding constructive coping mechanisms.

The Intensive Overseas Program

September 2011 witnessed fifty-two BYU students with four semesters of prior rigorous Arabic experience beginning their semester of intensive study at the Qasid Institute in Amman, Jordan.[7] Most students lived together in apartments located between the institute and the University of Jordan.[8] Dilworth Parkinson, accompanied by Kirk Belnap and three undergraduate TAs/past participants, led the program. These peer TAs, all non-native speakers of Arabic and BYU undergraduates who had succeeded in reaching at least the Advanced Mid level in speaking, served as role models for the students. In addition to helping with customary program needs, such as preparing readings, they assisted Belnap with implementing and documenting the impact of Project Perseverance interventions. For example, either a faculty member or TA, who had been trained in reflective listening by Madeline Ehrman, interviewed all students on a weekly basis for at least the first half of the program (these interviews were optional during the second half).[9] At the end of each week, students typically spent 15–20 minutes processing the week's experiences together—an exercise that many, but not all, found beneficial. This innovation came about at Ehrman's suggestion following her site visit to the 2010 BYU intensive program.

The students read extensive and intensive selections from the Arab press and discussed current events of particular interest to Jordanians. On a regular basis, students also informally surveyed Jordanians about their views on various social and political issues and then reported the results of those surveys in their classes. Some of the students attended a political science course twice a week at the University of Jordan and were regularly invited by the professor to comment about ongoing developments in the world.[10]

All students were expected to spend at least two hours per day conversing in Arabic with native speakers; they also had weekly speaking appointments with Qasid staff and weekly individual tutoring sessions for the purpose of going over writing assignments. Students kept online learning journals that faculty and TAs monitored; they also accounted online for their reading and speaking time and quality of effort. Some students embraced the opportunity to reflect and set goals and benefitted considerably; others chafed at this requirement and viewed it as busy work. To maximize

their progress, students were asked to reflect on the effectiveness of the methods they were using to help them learn and, as a result of conferring with BYU faculty, were encouraged to experiment with different approaches in order to discover what worked best for them.

Learning to speak and read at the Advanced and Superior levels represents a major goal of BYU's intensive program. Reading serves as a means of amassing a rich vocabulary critical to extended-level discourse. Parkinson has developed an effective system that consists of two hours of daily timed reading workouts (speed-building exercises) and personal study time (students are also encouraged to apply as much rigor in designing and following a plan suited to their learning style and preferences during this time).[11] Each day, students receive four articles for extensive reading along with a list of key vocabulary. These are chosen carefully based on their level of difficulty, and students are to spend eight to ten minutes with each. They are instructed to focus on what they readily understand and deduce the meaning of the rest as best they can in the time remaining.

For the next hour and twenty minutes, students work on articles chosen for intensive reading. A recommended schedule follows (A is the new article for the day, B is the article from the day before, C is the article from two days before, and D is the article from three days before on which students will be tested for mastery during class, typically by being asked to translate selections from it):

- 20 minutes (strictly limited, given that too many students would easily spend hours): Students try to translate as much of article A as they can. They receive a list of vocabulary items that occur in the new article.
- 10 minutes (also strictly limited): Students do their best to read and understand the rest of article A, but without taking the time to write out a translation.
- 20 minutes: With the help of a translation provided for them—which they did not have the day before—students read Article B. First, they review the section that they translated the previous day and make quick corrections based on the translation provided them. They then go through the rest of the article, noting especially those parts that they did not deduce correctly the day before (this they do only in their heads). Students try to do this fast enough that they can then read the article again, one paragraph at a time, referring back to the translation when they are unsure.
- 15 minutes: If possible, they read article C three times with the aid of the translation, as needed, with the goal of relying on it less with each subsequent reading. Students are told that they may refer to the translation after each paragraph during their first reading, but by the final reading, they must read the entire article before checking the translation. This process necessitates that learners use the context and meaning of the text, with which they are now very familiar, to help them deduce the meaning of words and forms that they cannot remember. In addition, they make mental notes of expressions or constructions that give them trouble and ask about these later in class.
- 15 minutes: Students read article D three times or more (they should be quite good at getting through this article quickly at this point). They use the translation

on the few aspects they still do not understand well, but they try to do so rarely. The last time through, they "pass it off" (i.e., they try to read it with full mastery, remembering the meaning of every word and structure). Preferably, they get a roommate or classmate to hold the translation and follow along as they quickly translate the article for them. At the end, the roommate or classmate determines whether they "passed" or not. A couple of minor lapses (words or constructions that are partially correct) still count for a pass, but if the student simply could not translate some words or had any major errors, such as choosing the wrong noun as the subject of a sentence, then the student must try again to pass off the article.

Some students and teachers initially complain about the almost exclusive focus on reading front-page news articles. This is a strategic pedagogical choice for two reasons, which most students soon grasp: 1) Many Arabs talk about politics and current events a great deal, and a student who has little understanding of what is going on in the region will not make a good impression, and 2) focusing on easier-to-read news articles allows students to build confidence and fluency in reading and thereby establish a firm base from which to expand their repertoire. As the semester progresses and students' reading ability improves, the articles they are assigned become progressively longer and the topics increasingly varied.

In their current events course, students are encouraged to speak as much as possible in a more formal register, and reading contributes significantly to their ability to do so. A significant part of their Jordanian Arabic class consists of presenting rehearsed monologues on a weekly topic, such as describing their daily routine. One of the first strategies taught in 2011 was the acronym FORTE, which stands for "family," "occupation," "recreation," "travel," and "education." Students were encouraged to remember this acronym as a way of selecting topics to discuss during their daily two hours of speaking. In addition, the program successfully experimented with focusing on building a specific skill each week. For example, a TA would prepare a presentation on describing future plans or narrating in the past tense—skills that would prove useful to students in their daily interactions and when completing the American Council on the Teaching of Foreign Languages (ACTFL) Oral Proficiency Interview (OPI).[12] Students would practice these skills in class, and then use them with friends and new acquaintances outside of the classroom.

Toward the end of the program students were interviewed via telephone by a certified ACTFL OPI tester in the United States and eventually received an official score.[13] Students' reading and listening comprehension were also measured using online tests developed by the National Middle East Language Resource Center (NMELRC), in cooperation with ACTFL and the Arabic Flagship Program at the University of Texas at Austin. Most students began the intensive program with Intermediate Low or Intermediate Mid proficiency in all three skills (a few were higher in speaking, especially those who spent the summer in Jordan doing internships), and most finished the program at either Intermediate High or Advanced Low (see Table 6.1). Given that Intermediate High represents Advanced-level proficiency, albeit lacking full control and consistency, the data indicate that 79 percent of the students were able to operate most of the time at the Advanced level in speaking.[14]

Table 6.1 2011 Oral Proficiency Gains

Pre-Program OPI	Post-Program OPI (change)	Number of Students
Advanced Low	Advanced Mid (+1)	1
	Advanced Low (0)	2
Intermediate High	Advanced Mid (+2)	1
	Advanced Low (+1)	2
	Intermediate High (0)	3
	Intermediate Mid (−1)	1
Intermediate Mid	Advanced Mid (+3)	3
	Advanced Low (+2)	8
	Intermediate High (+1)	10
	Intermediate Mid (0)	1
Intermediate Low	Advanced Low (+3)	3
	Intermediate High (+2)	8
	Intermediate Mid (+1)	8
Novice High	Intermediate Mid (0)	1
TOTAL:		52

Post-Intensive Program Study Opportunities

Most students return to BYU to complete their undergraduate studies. Because BYU canceled third-year courses at its home campus in favor of an intensive semester abroad, the university has been able to expand its advanced-level course offerings on campus significantly (See Table 6.2). These courses have proven to be effective in helping students who were diligent but not outstanding language learners during their study abroad experience to reap considerable benefits from their investment of time and effort and make personal breakthroughs, especially in terms of reading. A number of these students went on to receive prestigious academic awards such as full-year fellowships to study at the American University of Cairo in the Center for Arabic Study Abroad (CASA).[15]

Current Events, the most popular course listed in Table 6.2, represents a content-based course.[16] While other courses may focus on particular skills, there is, nevertheless, a high degree of content emphasis. For example, the debate course focuses on rhetorical skills in a formal register, but the topics debated are of considerable interest to students and the overall experience constitutes a high-quality content course.

The new Arabic major has resulted in some students taking far more courses after their study abroad than they would or could have previously. Winter semester 2012 therefore saw the highest number of advanced-level Arabic enrollments. After subtracting courses exclusive to the major, enrollments (fifty-four) are comparable to preceding years. For example, enrollments in fall semesters for 2008, 2009, and

Table 6.2 Advanced-Level BYU Undergraduate Arabic Courses Offered Winter 2012

Course	Students Enrolled
Current Events (two sections)	18
Advanced Jordanian Arabic	14
Advanced Grammar (Close Reading)	15
Independent Readings	2
Classical Arabic Texts	5
Debate (for Arabic majors)	14
Tutorial (majors' research projects)	14
TOTAL:	82

2010, the semester that immediately followed the study abroad experience during these years, were fifty-one, forty-two, and forty-four, respectively. In short, we have found that BYU students' experiences during the study abroad program coupled with a variety of courses that address their needs and interests have served to motivate them to persist in studying Arabic after returning to campus.

Careful attention paid to improving BYU's Arabic program, including years of objectively measuring progress, contributed to the university's decision to add an Arabic faculty position for the purpose of establishing a major. In time, the decision was made to design a "second major" that, in many respects, paralleled BYU's Chinese Flagship Program (see description in Christensen and Bourgerie, this volume), whose goals and methods seemed like a good fit for BYU's top Arabic students. In order to gain admission to this major, students must declare a primary major (e.g., MESA, linguistics, political science, or economics) and be certified at ACTFL Advanced Low or higher in reading and speaking. Promising students with Intermediate High proficiency are admitted provisionally.

At the core of the second major experience is a debate course modeled after a BYU Russian course (Brown, Bown, and Eggett, this volume). From January through August of 2013, Arabic majors met three hours a week for fifty minutes (the equivalent of two semesters of this course). Like BYU Chinese Flagship students, Arabic majors also developed facility in an area of personal expertise by reading domain-specific articles, meeting one-on-one with an instructor and native-speaking TAs to discuss the readings, and preparing a polished, written document and oral presentation in their area. They also met together as a group once a week from January to August with a faculty member and native Arab instructor to receive training and to present to each other. Majors were also required to complete two two-credit-hour advanced-level electives. Table 6.3 gives the results for pre- and post-OPI scores (telephonic interviews) for the 2012 cohort of Arabic second majors, all of whom had been students in the 2011 intensive program.

The debate, research, and tutorial experience facilitated impressive progress, but eight months on campus of doing essentially the same routine resulted in at least some degree of fatigue for most of the students. As a result, the faculty looked for

Table 6.3 2012 Arabic Second Majors' Oral Proficiency Gains

December 2011	August 2011 OPI (change)	Number of Students
Advanced Mid	Advanced High (+1)	2
	Advanced Mid (0)	1
Advanced Low	Advanced Mid (+1)	5
	Advanced Low (0)	2
Intermediate High	Advanced Mid (+2)	1
TOTAL:		11

ways to add variety to the students' experience and again found inspiration in the Chinese Flagship Program, whose students culminate their studies by participating in an overseas internship wherein they shadow a peer professional. Such an experience has long been the goal of the Arabic program, but quality internships have, until recently, proven elusive.

A fortunate breakthrough occurred early in 2013 in the form of an interdisciplinary partnership that led to significantly better internship opportunities.[17] As a result, five of the 2013 majors opted to complete their Arabic studies by pursuing an internship with the Ministry of Social Development in Amman, Jordan, from May to August. Interning in this capacity gave each student valuable experience in the culture and workings of an Arab professional setting. The three most linguistically proficient students had the opportunity to work extensively on significant English-to-Arabic translation projects. In addition to their internship work, students met twice weekly at the Qasid Institute for a debate class, which included regular interaction with Arab professionals who briefly addressed a topic of interest, such as freedom of the press, and then fielded questions from students and engaged them in a lively discussion for approximately thirty minutes. At the end of the summer, one student certified as Superior, two as Advanced High, and two as Advanced Mid on the computer adaptive Oral Proficiency Interview (OPIc). None of these students had experience with Arabic prior to taking it at BYU.

These students also benefitted from Project Perseverance. In particular, they were coached on learning to "sharpen" or to notice the gap between specific aspects of their own performance and a native model and to work in a targeted manner to close this gap (Leaver and Campbell, this volume).

The Bigger Picture

The combination of instability in the Arab world and federal funding cuts has restricted student access to overseas intensive study, especially to high-quality, advanced-level study opportunities. The Critical Languages Scholarship program has proven to be an excellent option for a fortunate few, but, like other federal programs, it has been affected by budget cuts. Furthermore, being a summer program, it does not provide the optimal amount of learning time for students. CASA has been hit particularly hard, as evidenced by a precipitous decline in the number of

fellowships offered (from forty-six in 2011/12 to thirty in 2013/14). In addition, the unrest in Egypt resulted in CASA's relocating to Jordan in 2013 (see Belnap and Nassif 2011, 7 for a more detailed discussion of CASA's woes). Fewer advanced-level study opportunities means well-prepared students are being turned away and must either find their own way or apply again in another year, thus creating an even worse bottleneck than has existed for the past decade.

Quality advanced-level, semester-length or longer intensive programs are available for those who can pay. Unfortunately, these programs do not publish data on pre- and post-program proficiency scores, making it difficult for students and advisors to know which programs are more effective. Measuring program outcomes and greater transparency would be most welcome from all programs. The costs involved in running quality programs put them out of the reach of many students. More needs-based funding for well-qualified students should be a high priority.

More and more institutions are shopping for quality and beginning to sponsor their own programs abroad; however, they rarely invest in the oversight necessary to transform an average program into a high-quality program. Programs with students largely from a single institution rarely have enough well-qualified students to constitute a solid cohort ready to engage in the type of learning experiences described here. Consortium arrangements, like CASA, arguably represent the key to addressing these two challenges.

The tutorial model may be a better option for individuals or small groups seeking a quality-intensive learning experience. The Qasid Institute and others lean heavily on this approach for training diplomats. Research on the development of expertise by surgeons, chess players, musicians, pilots, and other highly skilled professionals uniformly points to the need for a more individualized experience. Summarizing this research, Ericsson et al. noted that "One of the key issues is how . . . the quality of their performance can be regularly evaluated in order that detailed and reliable feedback can be obtained for use in guiding deliberate practice in designed practice environments" (2009, 18). Other key methods they mention include tutoring, simulation, problem-based learning, discovery learning, self-assessment, and self-directed learning. Indeed, the question arises as to whether we have become too accustomed to the traditional approaches to intensive study that have been with us for decades. Experiential learning, similar to the work carried out by BYU interns, represents a highly effective path and, fortunately, one that is becoming increasingly common. In the future, we hope to see continued growth in the number of upper-challenge language learning experiences.

NMELRC has documented some cases of individual learners achieving "professional-level fluency" on their own. For example, Robert Lovett, chairman of Landmark Advisory in Dubai, chose a direct enrollment experience for his junior year abroad in Jordan, with impressive results (http://nmelrc.org/arabic-success-stories/robert-lovett), and Ryan Gregg's pursuit of proficiency in Hebrew serves as a model for how to leverage the right kind of internship into an extraordinary cultural and language sharpening experience (http://nmelrc.org/success-stories-ryan-gregg). While the accomplishments of individuals such as Ryan Gregg and Robert Lovett

may seem exceptional, we are convinced that many students could have similar experiences, especially with appropriate guidance.

Findings from this research underscore the importance of extended periods of quality, intensive learning experiences as a key to students' achieving a high level of proficiency in a timely manner. With Pellegrino Aveni (2005), Vande Berg, Connor-Linton, and Paige (2009), Trentman (2012), and others, we emphasize that significant pre-program interventions, well-designed overseas learning experiences, and on-site coaching lead to high levels of engagement, self-awareness, and deliberate practice.

Notes

1. This estimate is based on the Modern Language Association's 2009 survey and annual reports from a sample of institutions to the National Middle East Language Resource Center, as discussed in Belnap and Nassif (2011, 1–2).
2. In addition to those mentioned by name in this chapter, the authors must also thank Spencer Scoville and Shereen Salah who oversaw the implementation of BYU's new Arabic major.
3. Students beginning Arabic 101 in September are able to complete the four semesters required before participating in BYU's study abroad program by taking Arabic 102 during winter semester, 201 during spring term, and 202 during summer term. Relatively few students opt for this expedited track; most take Arabic 201 and 202 the following academic year.
4. Bandura (1997) included "vicarious experiences through social models" as one of his four cornerstones for promoting self-efficacy, which is critical to success in language and many other types of learning (Mills, Pajares, and Herron 2007). Other core factors promoting self-efficacy include experiencing success in learning and gaining confidence in one's ability to overcome learning challenges, receiving praise from others for deserving efforts, and lowering one's sense of stress and replacing negative emotions such as self-doubt with positive self-talk (Bandura 1997). All of these have come to be key components of the BYU Arabic program.
5. For details about Project Perseverance see: http://nmelrc.org/Pathways.
6. Not all students want or need advanced-level skills and many beginners are simply not aware of the significant time investment typically required. Students need to know what to reasonably expect to gain as a result of their time investment as well as what the expectations of their programs are. The BYU Arabic 101 syllabus therefore reads:
7. The past two years have witnessed a significant decrease in enrollments in the study abroad program. This appears to be the result of 1) a national trend of decreasing Arabic enrollments in well-established programs (many newly established programs continue to report growth) and 2) the increased cost of BYU's intensive program.
8. The location was convenient but not without its drawbacks. Female students were occasionally mistaken for Eastern European prostitutes living in the area, which made it difficult for them to feel comfortable in their neighborhoods after dark.
9. Ehrman has a PhD in clinical psychology and is a member of the Project Perseverance team. Until her retirement, she was Director of Research Evaluation and Development at the Foreign Service Institute. She founded and continues to consult for the institute's Learning Consultation Service, whose services include coaching for those involved in intensive language training.
10. Cornell's Intensive Arabic Program, which is no longer in operation due to lack of funding, also relied heavily on content-based instruction and provided an option for direct enrollment in content courses for Jordanians. Students from the College of William and Mary likewise benefitted considerably from direct enrollment at the University of Aleppo, where they had the benefit of good on-site supervision (Belnap and Nassif 2011, 12–13).
11. This reading routine and the required two hours of speaking help students study more effectively and keep perfectionism and avoidance behaviors in check.

12. Pre- and post-program OPIs and significant attention during the program devotes what the test measures have helped students focus during and beyond the program and awakened as well as increased their awareness of what is required to reach the Superior level.
13. OPIs represent an important part of the program, but the experience has also been frustrating. Telephone conversations are not ideal, especially when the quality of the connection cannot be guaranteed, as has been the case in Jordan. But, worse yet, ACTFL interviewers have been instructed to interview in "Modern Standard Arabic" (MSA), which makes for an unnatural conversation, especially if the student responds in a more conversationally appropriate variety that they are more accustomed to using. Over the years, we have repeatedly had interviewers instruct such students to speak MSA, even though we have vigorously complained every time this occurs. In recent years, we have opted for the OPIc (a computer-administered OPI); results have been generally good, but prerecorded prompts make for even less of a natural conversation. Now that a number of Qasid teachers are ACTFL certified, we plan to have them do face-to-face OPIs and then send the audio recordings on to ACTFL for a second rating. We hope to see ACTFL OPIs become more naturalistic generally, in light of the recent position announced in the introduction to ACTFL's new "Arabic annotations and samples." In particular, "During ACTFL OPI testing, testers accommodate to the variety of language that the test taker is producing and accept Arabic language produced anywhere along the continuum" (ACTFL 2012).
14. Reading and listening proficiency results were similar, as measured by NMELRC/Flagship tests developed in partnership with ACTFL.
15. All three TAs from the 2011 program spent the 2012/2013 academic year in Cairo as CASA fellows. Reduced funding coupled with a CASA decision to prioritize funding for graduate students likely explains why only one of them received CASA funding. While no official OPI was administered at the end of their year in Cairo, oral proficiency testing done by CASA faculty suggests that the students reached the Superior level of proficiency.
16. Content-based courses, such as "Arabic Voices: Poetry to Rap" and a course focused on Beirut, are core to the on-campus Arabic Flagship Program experience at the University of Texas at Austin.
17. Discussions with Ralph Brown, who oversaw BYU's International Development Minor, led the Arabic faculty to believe that the internship program he had developed in Jordan with the assistance of a key Jordanian partner could be a good fit. However, this was only the beginning. Significant adjustments and training of students were needed. Building on her experience training and supervising graduate-student interns in Morocco from the University of Pennsylvania's Lauder Institute, Maggie Nassif, NMELRC administrative director, helped to train students on campus and accompanied them to Jordan. In addition to coaching the students and arranging for them to meet well-placed Jordanian professionals, she also interviewed and trained Jordanian students or recent graduates to work with the students as research partners, along the lines of interventions recommended by Pellegrino Aveni, who calls for better training of those who interact with students (2005, 149). Belnap provided coaching for the group and for individuals via Skype and email, and spent a week in Jordan halfway through the program observing the students in a variety of settings and interviewing them individually about their experiences.

References

ACTFL. 2012. "Speaking." ACTFL Proficiency Guidelines for Arabic. Accessed May 9, 2014. http://actflproficiencyguidelines2012.org/arabic/index.php?p=speaking.

Al-Batal, Mahmoud. 2007–2008. "Toward Superior Level Proficiency in Arabic: A Study of Successful CASA Full Year Applicants 2002–2006." *Al-Arabiyya* 40–41: 57–70.

Bandura, Albert. 1997. *Self-Efficacy: The Exercise of Control*. New York: W. H. Freeman and Company.

Belnap, R. Kirk, and Maggie N. Nassif. 2011. "Middle East Language Learning in US Higher Education: Ten Years after 9/11." National Middle East Language Resource Center. Accessed May 9, 2014. http://nmelrc.org/middle-east-language-learning-us-higher-education-ten-years-after-911.

Ericsson, K. Anders, Ray S. Perez, David W. Eccles, Laura Lang, Eva L. Baker, John D. Bransford, Kurt VanLehn, and Paul Ward. 2009. "The Measurement and Development of Professional Performance: An Introduction to the Topic and a Background to the Design and Origin of This Book." In *Development of Professional Expertise: Toward Measurement of Expert Performance and Design of Optimal Learning Environments*, edited by K. Anders Ericsson, 1–24. New York: Cambridge University Press.

Mills, Nicole. A., Frank Pajares, and Carol Herron. 2007. "Self-Efficacy of College Intermediate French Students: Relation to Achievement and Motivation." *Language Learning* 57 (3): 417–442.

Pellegrino Aveni, Valerie. 2005. *Study Abroad and Second Language Use: Constructing the Self*. Cambridge: Cambridge University Press.

Rifkin, Benjamin. 2005. "A Ceiling Effect in Traditional Classroom Foreign Language Instruction: Data from Russian." *The Modern Language Journal* 89 (1): 3–18.

Trentman, Emma Gale. 2012. "Study Abroad in Egypt: Identity, Access, and Arabic Language Learning." PhD diss., Michigan State University. Accessed May 9, 2014. http://etd.lib.msu.edu/islandora/object/etd%3A1904/datastream/OBJ/view.

Vande Berg, Michael, Jeffrey Connor-Linton, and R. Michael Paige. 2009. "The Georgetown Consortium Project: Interventions for Student Learning Abroad." *Frontiers: The Interdisciplinary Journal of Study Abroad Language Learning and Study Abroad* 18: 1–75.

Chapter 7

The Development of L2 Proficiency and Literacy within the Context of the Federally Supported Overseas Language Training Programs for Americans

DAN E. DAVIDSON

A CONSIDERABLE BODY OF research over the past two decades has demonstrated that second language (L2) acquisition in the overseas instructed/immersion environment may contribute greatly to the development of linguistic and cultural competence, provided appropriate program designs and student support conditions are met (Davidson 2010; Kinginger 2008). As the number of US students participating in study abroad learning continues to grow (Open Doors 2013), the community of researchers, educators, and policymakers interested in overseas study has understandably increased attention on the specification of learning outcomes from study abroad, which are understood to affect the development of target language and cultural literacy, intercultural communicative competence, and regional knowledge. Second language proficiency is also closely linked to the development of cross-cultural adaptability, language socialization, identity construction, enhanced cognitive functioning, and critical thinking, to name only the most common themes (Watson et al. 2013; Kinginger 2013; Byrnes 2013).

In 2006, the White House announced a major multiagency initiative designed to "expand dramatically the number of Americans studying and mastering critical foreign languages," known as the National Security Language Initiative (NSLI). The initiative focuses on senior secondary and university-level students. Recipients of the new federal scholarships represent all disciplines and institutional types and are expected to continue their language study beyond the overseas scholarship period. The students are therefore encouraged to apply their critical language skills in their future professional careers. Other components of NSLI, which are not discussed in the present study, address domestic foreign language training and the support of

teachers of the critical languages. None of the new programs include a US government service requirement.

The new overseas language study programs are open on a competitive basis to qualified US secondary school students, undergraduate students, and graduate students engaged in the study of Arabic, Bahasa Indonesia, Bangla, Chinese, Hindi, Japanese, Persian, Portuguese, Punjabi, Russian, Swahili, Turkish, and Urdu.[1] The present study reports and examines proficiency-based outcomes of the participants in each major program: the National Security Language Initiative for Youth (NSLI-Y), the Critical Language Scholarship Summer Institutes (CLS), and the Language Flagship Overseas Capstone Programs (OCP). Taken together, the new programs represent one of the most significant investments in foreign language study for American students in US history. These programs support high school and college-level learners of critical languages at a stage in their education when an intensive summer or year-long overseas immersion program is more readily accommodated and costs far less than comparable training of midcareer or in-service adult learners.

The present study provides the first detailed review of the impact these programs have on the language proficiency of their US participants. It makes use of large-scale (N = 1,457) multi-language, multi-institutional data on learning outcomes, which make it possible to consider the effects on L2 gains of a group of factors (independent variables) previously identified as strongly associated with language growth: program durations (summer—eight weeks; academic year—thirty-two weeks), initial L2 levels of (pre-program) proficiency (from Novice Low [NL]/Interagency Language Roundtable [ILR]-0) to Advanced High [AH]/[ILR 2+]), and post-program proficiencies (from Novice High [NH]/[ILR-0+] to Distinguished [ILR-4]). Three distinct participant cohorts are of both research and policy interest: high school learners in the summer and year-long NSLI-Y program (N = 523), university-level participants in the CLS immersion program (N = 620), and university-level students in the Flagship Capstone programs (N = 314). Descriptions of the major curricular and cocurricular interventions utilized at the school-, undergraduate-, and postgraduate-level overseas programs are also included. Finally, the study addresses the issue of language gains within the context of the eight-week summer program model, where program outcomes in the past have been challenging to demonstrate using the existing proficiency scales (Dewey 2004; Dwyer 2004). Reading and listening gains from short-term study, when tested, have focused primarily on students at more advanced levels of study (Davidson 2010). The new summer (short-term) immersion data presented here should contribute to the understanding of the efficacy of this popular format for language and cultural learning.

Research Questions

1. What measureable gains in L2 proficiency are demonstrated by participants in a group of large-scale, federally funded programs for overseas/immersion

language study? How do language gains in reading, listening, and speaking vary across summer and academic-year program durations for participating students?
2. To what extent does *initial level* of proficiency affect L2 gain in the overseas immersion setting? How does the impact of a summer or year of study for a student with pre-program proficiencies in the Advanced range (ILR skill level 2 in Speaking [S-2]/Reading [R-2]/Listening [L-2]) compare with the same duration of immersion for a student with proficiency in the Novice or Intermediate range (ILR S-1/R-1/L-1)?
3. To what extent do age and the choice of language affect gain in the study abroad setting?

Subjects, Variables, and Methods

The present analysis is based on data relating to 1,457 students who studied under the previously mentioned federally sponsored summer or academic-year immersion programs, particularly between 2010 and 2012. Participants represented more than 1,500 US schools, community colleges, and universities located in all fifty states, the District of Columbia, and US territories. Participants came from public and private institutions, community colleges, and minority-servicing institutions such as Historically Black Colleges and Universities (HBCUs), High Hispanic Enrollments (HHE institutions), and Tribal Colleges. Approximately 10 percent of the university students represented in the current study are science, technology, engineering, and mathematics (STEM) majors (See Table 7.1).

The programs recruit and select participants on a competitive basis and in keeping with specific eligibility standards that are outlined in appendices 1, 2, and 3. Evidence of the applicant's motivation, overall academic achievement, cultural

Table 7.1 Student Characteristics, 2010–2012 (Total Student Population [N = 1,457])

NSLI-Y Summer and Academic-Year (AY) Programs (N = 523)			
Length of Program	**Age**	**Gender**	**Prior Language Study**
Summer N = 479	Mean Age 16.6	Female 61.0%	Yes 58.9%
AY N = 44	Mean Age 17.5	Male 39.0%	No 41.1%

CLS/Summer Program (N = 620)			
Educational Level	**Age**	**Gender**	**Prior Language Study**
2/4-year colleges: 70%	Mean Age 21.1	Female 60%	Yes 61%
Graduate: 30%	Mean Age 23.3	Male 40%	No 39%

Flagship Capstone Program (N = 314)			
Educational Level	**Age**	**Gender**	**Prior Language Study**
Undergraduate: 89%	22.9	Female: 52%	Yes 100%
Post-BA: 11%	24.1	Male 48%	No 0%

adaptability, and commitment to future study of the language is taken into consideration by external selection committees within the review process. Except for Flagship Capstone, which requires demonstrated Advanced-level proficiency (ILR-2) in at least two language modalities, selection for the overseas programs is not based on the attainment of a specified level of proficiency in the target language at the time of application. Programs do not make use of foreign language aptitude measures, such as the Modern Language Aptitude Test (MLAT) and the Defense Language Aptitude Battery (DLAB), in assessing candidate qualifications to participate in the program. Some of the programs require academic prerequisites that are expressed in terms of one or two years of prior classroom study of the target language.

The conclusions of the present study, therefore, cannot be generalized with full confidence to every foreign language student who studies abroad. Nevertheless, to the extent that selection criteria are controlled in the analyses, candidates are not prescreened for language aptitude, and all three programs accept students from all fields and from a broad range of institutions. The results may be seen to apply more generally to the population of US students who study abroad.

Primary Curricular and Cocurricular Intervention Strategies

The American Councils for International Education, henceforth referred to as American Councils, serves as the primary (but not sole) administrator of the overseas components of the three federal programs, which are implemented, in close cooperation with local partner universities, schools, or centers in each of the host countries. All three federal programs provide resources for local faculty, classroom, and curriculum-development needs, as well as for the administration and regular evaluation of cocurricular components, including homestays, cultural programs, peer tutoring, and service learning. In reviewing these design components, it is worth noting that, unlike study abroad in the nations of Western Europe, where student mobility and cross-border language training at all levels are well understood and widely institutionalized, the experience of the countries where critical languages are spoken tends to be much more limited. The hosting of foreign students with coordinated homestays, internships, cooperative learning, and volunteer service work is relatively new in these societies. Many of the countries also face considerable physical, cultural, and religious challenges in finding local families and businesses willing and able to play a role in hosting US students.

The structure of the overseas programs administered by the American Councils is summarized here in terms of academic, cocurricular, and student-support components.[2] Program designs take into consideration the limitations as well as the learning opportunities provided by the host country environment; the challenges of developing the necessary instructional, curricular, and cocurricular mechanisms (and technologies) to optimize L2 learning; and the critical on-program student-support systems for the participants themselves.[3] For many participants, the NSLI-Y and CLS programs represent their first sojourn outside the United States. Flagship Capstone participants are expected to have undertaken previous academic study

(minimum summer-length) in the host country prior to applying for the year-long program.

Academic Components (Age Appropriate)

- Predeparture participant orientation
- Intensive language training in small groups (12–15 hours per week)
- Regular twice weekly or daily meetings with peer tutor (4 hours per week)
- Overseas centers based in local schools, centers, or universities
- Direct-enrollment courses or classes

Cocurricular (Age Appropriate)

- Integrated homestay or residential component
- Internships and/or service learning
- Optional discussion groups with native speakers (5–6 times per term)
- Ongoing evaluation (testing, site visits, teacher/tutor reports, portfolio development, regular student self-evaluations)
- Weekly (or Biweekly) Online Language Utilization Reports (where available)

Given the range of proficiency levels represented in the federal programs, immersion learning designs must strike a workable balance between (a) sheltered and unsheltered forms of study, interaction types, and community outreach; (b) maximization of target language to include obligatory L2 use in public places, homes, and local institutions; (c) level-appropriate scaffolding and pedagogical support in the form of small-group language training classes and tutorials (average of fifteen hours per week) led by professional native teachers; (d) full-time American resident director supervision; (e) recurrent diagnostic and formative assessments, including self-assessment; (f) peer tutoring; and (g) service learning and/or internship programs.

Beyond the challenges of learning a language and studying a different culture, for many US students overseas immersion study of this type also represents one of the most sustained encounters with self-managed learning, self-conscious strategy selection, and formative self-diagnosis in their learning careers—skills essential for lifelong learning.

Language Measures

Oral proficiency (pre- and post-program) testing using the standardized Oral Proficiency Interview (OPI) was performed exclusively by the American Council on the Teaching of Foreign Languages (ACTFL)—certified OPI testers. Testers are language specialists that are not otherwise involved in the training of the participants. Testers administer telephonic or face-to-face pre-program OPIs to all participants in the days prior to their departure overseas and post-program OPIs at the close of each summer or academic-year program. NSLI-Y and CLS participants with no prior study or family background in the target language are normally

exempted from the pretest. All students are required to take the post-program test. OPIs are recorded for subsequent verification and analysis. Preliminary OPI ratings are then cross-checked by ACTFL and submitted for blind double ratings by other certified testers as a regular part of the assessment process. The overall number of unratable samples remained low.

Pre- and post-program online reading and listening proficiency-based testing was initiated in the overseas Flagship Capstone programs for Arabic, Persian, Russian, and Swahili in 2009 and expanded to the remaining languages between 2011 and 2013. Limited reading and listening proficiency test results for NSLI-Y and CLS are also presented here, making use of the online reading and listening proficiency-referenced test instruments produced and administered at the request of the Defense Language and National Security Education Office (DLNSEO) by the Language Assessment Division of American Councils. The new-generation online examinations were initially cross-calibrated and equated with the older government-sponsored Educational Testing Service (ETS) reading and listening proficiency tests, used by the American Council of Teachers of Russian (ACTR) to measure listening and reading gains for students of Russian for more than fifteen years. In addition, participants in the overseas Portuguese, Russian, and Turkish Flagship Capstone programs are systematically cross-tested for speaking, reading, listening, and writing using the Common European Framework tests. Finally, all Boren Fellows taking part in any Flagship program are additionally tested by the Foreign Service Institute (FSI) in speaking, reading, and listening using the ILR proficiency scale. These results, in turn, have been analyzed and compared in the final calibration of American Councils testing of the same students. These ongoing collaborations with other testing agencies have helped to ensure the quality and construct validity of the American Councils tests, as well as the reliability and alignment of all the language measures of language proficiency used in this analysis with the ILR scale. Where available, reading and listening proficiency test outcomes are included in the present study in order to provide a fuller sense of participant L2 functional literacy in the target languages and cultures (Bazarova, Lekic, Marshall 2009).

Beginning in 2011, at the request of DLNSEO, final post-program Flagship test results were provided using both the ACTFL and ILR proficiency scales in order to bring these scores into alignment with internal US government reporting requirements. The State Department programs, whose participants are matriculated in high school or college, require score reporting on the ACTFL scale, which is more widely used by US educational institutions.

Note on Examining Proficiency-Based Speaking, Reading, and Listening Scores

For speaking, listening, and reading skills, *gain* is defined as the difference between pre-program and post-program scores. Given the nature of the proficiency scale, the ratings are treated as ordinal variables, with each level thought of as a grouping of scores on an underlying unobserved scale of proficiency on which variation is more continuous. As first noted in Brecht, Davidson, and Ginsburg (1995), such grouping

actually loses information, in effect introducing a measurement error, in that students with quite different (unobserved) proficiency levels may be assigned the same speaking, reading, or listening proficiency scores.

Novice Low is the default proficiency assigned to all students entering the NSLI-Y program with no prior home experience or study of the target language. Dark bars indicate pre-program oral proficiency test results, while light bars indicate post-program oral proficiency test results for the same population. Both pre- and post-program scores show notable variation typical of most study abroad programs. Figure 7.1, however, indicates that 64 percent of the subject population (N = 479) entered the summer NSLI-Y program at the Novice level; approximately half of this group were absolute beginners. By the conclusion of the program, 66 percent of the population tested in the Intermediate range, and an additional 18 percent tested at Novice High or (ILR 0+), which is a threshold Intermediate score on the ILR/ACTFL scale, not simply a "strong" Novice performance. These results compare favorably with scores typically obtained by university-level students of the same critical languages after two (or more) semesters of regular academic study.

The standard cross-tabulations presented in Table 7.2 provide a more detailed analysis of the impact of the NSLI-Y summer program: post-program outcomes can be viewed in comparison to initial levels of proficiency.

The main diagonal—the cells enclosed in boxes in Table 7.2—represents no gain. Moving to the left of the main diagonal in any row indicates a loss, and moving to

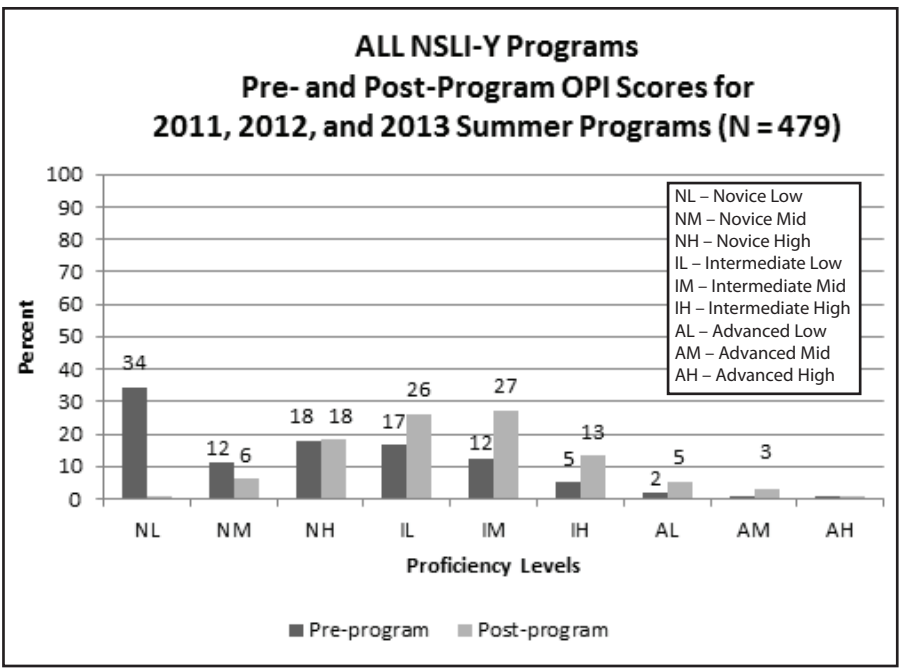


Figure 7.1 All NSLI-Y Programs Pre- and Post-Program OPI Scores for 2011, 2012, and 2013 Summer Programs

the right one column (or two, or three) indicates a gain of a sublevel or threshold on the proficiency scale. Linear models, such as those used here, do not fully capture the three-dimensional nature of the inverted pyramid used by specialists to model the proficiency scale and the reality that L2 gains in the upper ranges of the scale are considerably more difficult to achieve than those at the lower range.

This NSLI-Y data set comprises testing results from seven different language programs hosted in ten nations. Arabic language programs are still rare in US high schools, and, not surprisingly, most NSLI-Y Arabic participants (98 percent) undertook the summer program without prior knowledge of the language (Figure 7.2). The results of their study overseas are comparable to beginning students in the NSLI-Y group as a whole: approximately half (48 percent) of the group reached the Intermediate level during the course of the eight-week program, while another 25 percent attained the Intermediate threshold level (NH). The rest of the group

◆ Table 7.2 Speaking Proficiency Scores for ALL 2011, 2012, 2013 Summer NSLI-Y Learners (N=479), ALL LANGUAGES, Pre- and Post-Program (Count/Row Percent)

Pre-Program Speaking Proficiency Level	Post-Program Speaking Proficiency Level									
	NL	NM	NH	IL	IM	IH	AL	AM	AH	Total
NL	1	27	60	49	18	7	1	0	0	163
%	0.60	16.60	36.80	30.10	11.00	4.30	0.60	0	0	100.00
NM	0	3	22	19	11	0	0	0	0	55
%	0	5.50	40.00	34.50	20.00	0	0	0	0	100.00
NH	0	0	5	39	36	6	0	0	0	86
%	0	0	5.80	45.30	41.90	7.00	0	0	0	100.00
IL	0	0	0	19	41	15	4	1	0	80
%	0	0	0	23.80	51.30	18.80	5.00	1.30	0	100.00
IM	0	0	0	0	25	26	5	3	0	59
%	0	0	0	0	42.40	44.10	8.50	5.10	0	100.00
IH	0	0	0	0	0	9	13	3	0	25
%	0	0	0	0	0	36.00	52.00	12.00	0	100.00
AL	0	0	0	0	0	0	3	4	1	8
%	0	0	0	0	0	0	37.50	50.00	12.50	100.00
AM	0	0	0	0	0	0	0	2	0	2
%	0	0	0	0	0	0	0	100	0	100.00
AH	0	0	0	0	0	0	0	1	0	1
%	0	0	0	0	0	0	0	100	0	100.00
Total	1	30	87	126	131	63	26	14	1	479
%	0.20	6.30	18.20	26.30	27.30	13.20	5.40	2.90	0.20	100.00

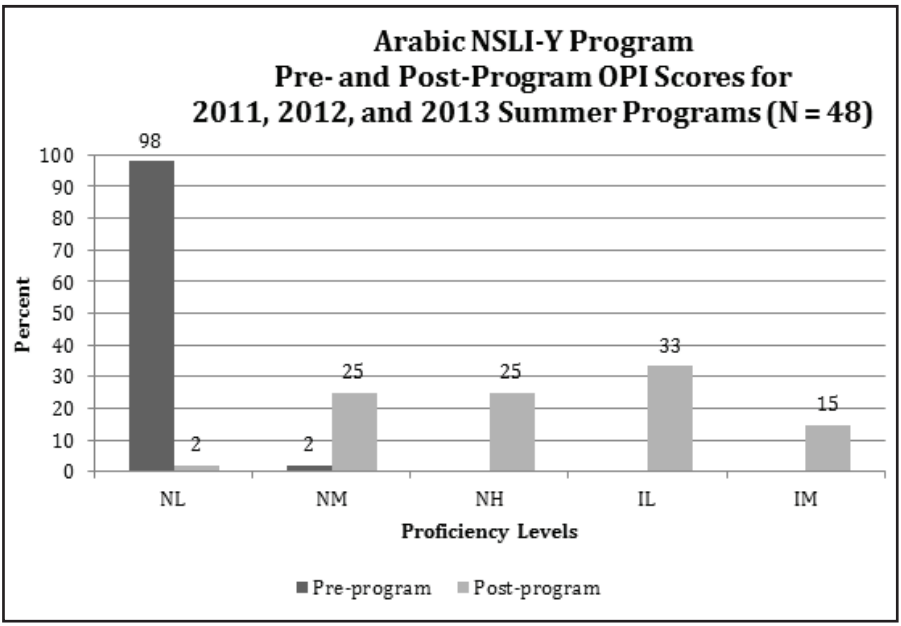


Figure 7.2 Oral Proficiency Outcomes in the Arabic NSLI-Y Summer Programs

remained in the Novice Mid (NM) range, with only one student failing to register a measureable gain in speaking.

By comparison, fewer than 10 percent of participants in the Chinese NSLI-Y summer-program sample entered the program without prior formal study, with most testing at NH at the outset of the program, while 89 percent completed the program well within the Intermediate range or higher (Figure 7.3).

The Chinese NSLI-Y cohort provides a good case for comparing the impact of an eight-week intensive immersion program on language gains for learners progressing through the Intermediate proficiency range (ILR-1). While 77 percent of the group registered either sublevel or threshold gains over the course of the semester, Table 7.3 indicates how the likelihood of achieving measureable gains in proficiency declines somewhat for students as their baseline proficiency increases. As noted, this numerical trend does not necessarily indicate that the more advanced students learned less but is a reflection of the measurement artifact underlying the proficiency scale itself.

Existing and ongoing research regarding time-on-task calculations for attaining advanced levels of proficiency in the critical languages — which are usually languages typologically different from English — indicates progressively longer training and acquisition times required for progression to Levels 2, 3, and above on the ILR scale.

Data on the acquisition of Persian/Tajiki in the NSLI-Y summer program are presented in Figure 7.4. As was the case with Arabic, participants in the Persian/Tajiki program arrived in the host country (Tajikistan) for the most part (85 percent) without prior study of the language. Nonetheless, their gain patterns are consistent with

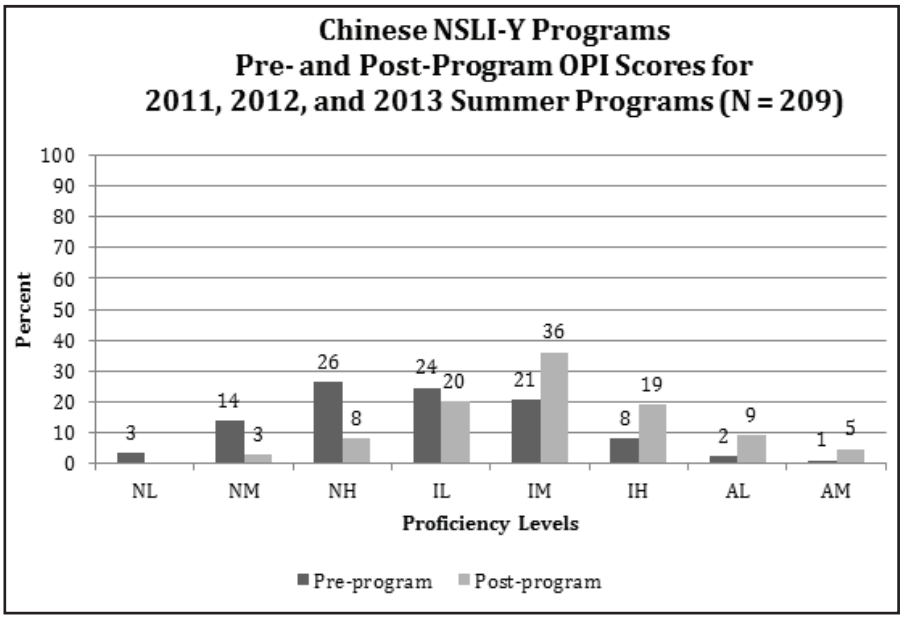


Figure 7.3 Oral Proficiency Outcomes in the Chinese NSLI-Y Summer Programs

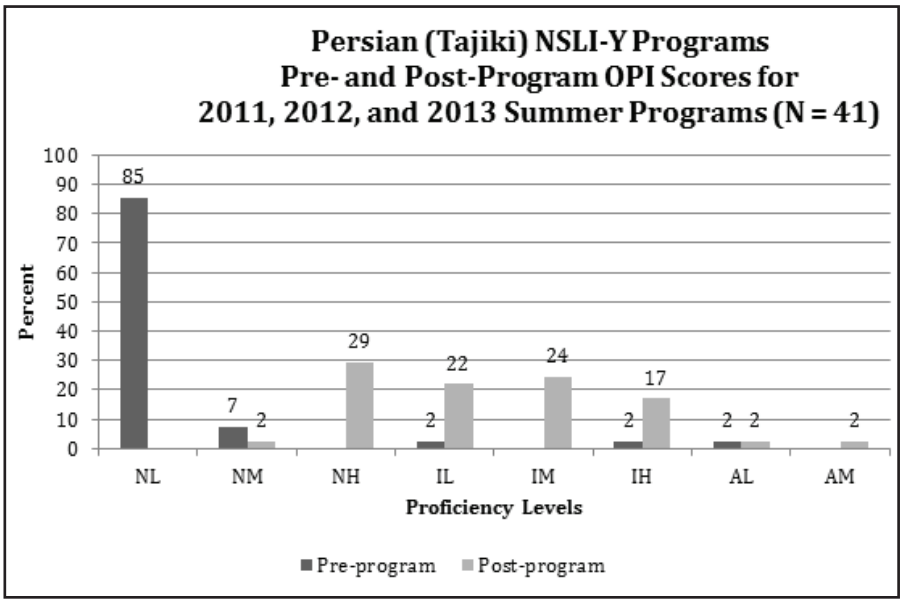


Figure 7.4 Oral Proficiency Outcomes in the Persian (Tajiki) NSLI-Y Summer Programs

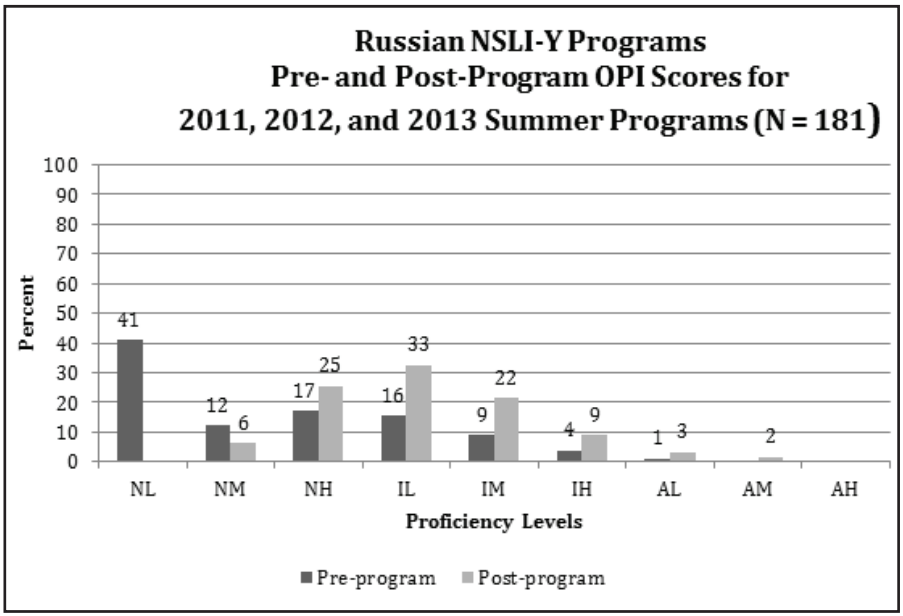
◆ Table 7.3 Speaking Proficiency Scores for ALL 2011, 2012, 2013 Summer NSLI-Y Learners (N=209), CHINESE, Pre- and Post-Program (Count/Row Percent)

Pre-Program Speaking Proficiency Level	Post-Program Speaking Proficiency Level							
	NM	NH	IL	IM	IH	AL	AM	Total
NL	3	3	0	0	1	0	0	7
%	42.90	42.90	0	0	14.30	0	0	100.00
NM	3	11	8	7	0	0	0	29
%	10.30	37.90	27.60	24.10	0	0	0	100.00
NH	0	3	22	24	6	0	0	55
%	0	5.50	40.00	43.60	10.90	0	0	100.00
IL	0	0	12	25	11	3	0	51
%	0	0	23.50	49.00	21.60	5.90	0.00	100.00
IM	0	0	0	19	17	4	3	43
%	0	0	0	44.20	39.50	9.30	7.00	100.00
IH	0	0	0	0	5	9	3	17
%	0	0	0	0	29.40	52.90	17.60	100.00
AL	0	0	0	0	0	3	2	5
%	0	0	0	0	0	0.60	40.00	100.00
AM	0	0	0	0	0	0	2	2
%	0	0	0	0	0	0	100.00	100.00
Total	6	17	42	75	40	19	10	209
%	2.90	8.10	20.10	35.90	19.10	9.10	4.80	100.00

those of the NSLI-Y population as a whole, and prepared participants to continue their study of the language at the second- or third-year level. Two-thirds of the group completed the summer program at the Intermediate level or higher, and 29 percent tested at NH—the Intermediate threshold level. Only one student remained in the novice range at the conclusion of the program.

The Russian NSLI-Y summer programs (Figure 7.5) reflect a mixed population of absolute beginners (41 percent) and students with prior formal study of the language, a pattern consistent with the distributions of high school Russian study programs across the United States (Davidson and Garas 2009).

The impact of the eight-week NSLI-Y Russian summer programs on both the beginning-level students and those with prior study is consistent with the pattern noted previously for the Chinese program and for the NSLI-Y program overall. Two-thirds of the cohort completed their respective programs in the Intermediate range or higher, while 25 percent tested at NH, the Intermediate threshold. Eleven participants out of 181 (6.1 percent) reached the NM level. Movement out of the



◆ Figure 7.5 Oral Proficiency Outcomes in the Russian NSLI-Y Summer Programs

Intermediate range into the Advanced range requires considerable time, as is evident from the rising values of null-gain figures in Table 7.4 for IL, IM, and IH. Nonetheless, a majority of students at these levels (75 percent, 67 percent, and 57.1 percent, respectively) did register measureable sublevel or threshold gains.

Reading and Listening Comprehension Results for NSLI-Y Summer

Reading and listening data have not yet been collected systematically across all NSLI-Y programs, even though reading, listening, and writing are included in the overseas curricula for all seven NSLI-Y languages. Initial piloting of reading and listening proficiency outcomes for Russian was undertaken beginning in 2011 and has now reached a reportable level. Those results are reproduced in Figures 7.6 and 7.7.

Group-level gains in reading are evident, with 79 percent of the entering group testing in the Novice range and 70 percent of the same group testing in the Intermediate range by the conclusion of the program. Final reading scores do not differ substantially from final speaking test scores, although it may be noted that students with prior formal training in Russian in the United States entered the program with somewhat higher reading scores than speaking scores. This is consistent with previous findings on the comparison of speaking and reading scores of US-trained students of Russian (Davidson 2007).

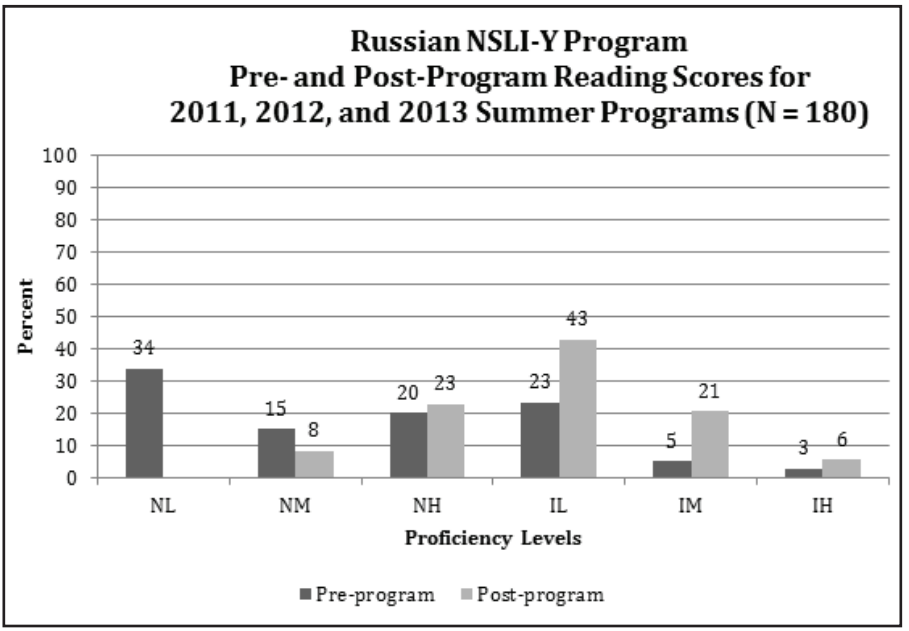
Listening-comprehension testing is distinct from the integrated testing of listening comprehension that takes place in the course of the OPI in that it is not

Table 7.4 Speaking Proficiency Scores for ALL 2011, 2012, 2013 Summer NSLI-Y Learners (N=181), RUSSIAN, Pre- and Post-Program (Count/Row Percent)

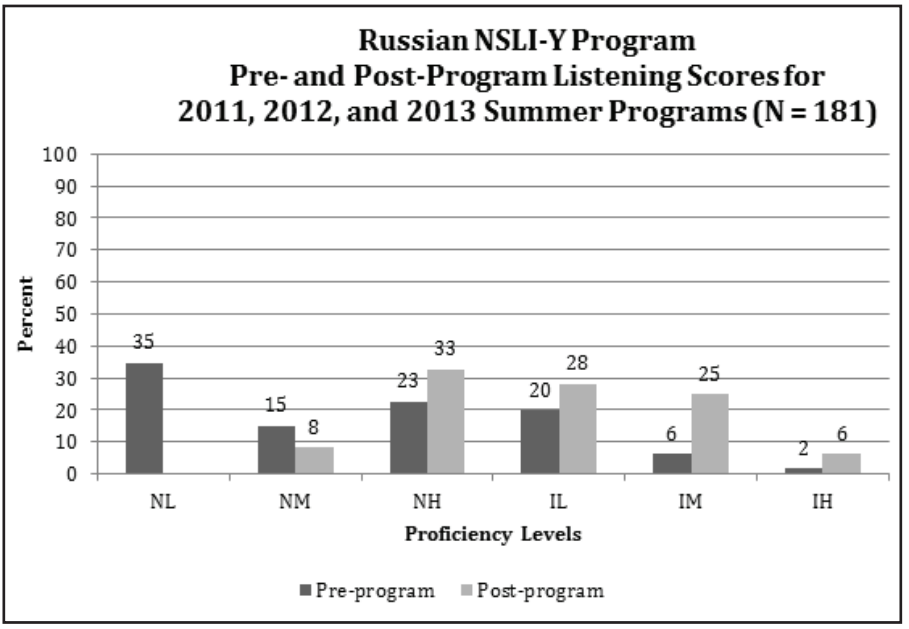
Pre-Program Speaking Proficiency Level	Post-Program Speaking Proficiency Level								
	NM	NH	IL	IM	IH	AL	AM	AH	Total
NL	11	35	26	1	1	0	0	0	74
%	14.90	47.30	35.10	1.40	1.40	0	0	0	100.00
NM	0	9	9	4	0	0	0	0	22
%	0	40.90	40.90	18.20	0	0	0	0	100.00
NH	0	2	17	12	0	0	0	0	31
%	0	6.50	54.80	38.70	0	0	0	0	100.00
IL	0	0	7	16	3	1	1	0	28
%	0	0	25.00	57.10	10.70	3.60	3.60	0	100.00
IM	0	0	0	6	9	1	0	0	16
%	0	0	0	37.50	56.30	6.30	0	0	100.00
IH	0	0	0	0	3	4	0	0	7
%	0	0	0	0	42.90	57.10	0	0	100.00
AL	0	0	0	0	0	0	1	1	2
%	0	0	0	0	0	0	50.00	50.00	100.00
AM	0	0	0	0	0	0	0	0	0
%	0	0	0	0	0	0	0	0	0.00
AH	0	0	0	0	0	0	1	0	1
%	0	0	0	0	0	0	100	0	100.00
Total	11	46	59	39	16	6	3	1	181
%	6.10	25.40	32.60	21.50	8.80	3.30	1.70	0.60	100.00

interactive and requires the processing of both interpersonal as well as interpretive texts on the part of the candidate. The higher the level of testing, the greater the emphasis on interpretive modes of listening and texts of greater length and complexity, such as radio broadcasts, talk show interviews, or public lectures. For that reason, listening comprehension is tested separately and reported here.

Listening gains for the Russian NSLI-Y groups present a pattern similar to those of reading and speaking; however, the alignment with speaking and reading scores is slightly weaker: 59 percent of participants have a post-program score in the Intermediate range, while 33 percent test at NH, the Intermediate threshold. Null gain was observed in 20.7 percent of the cohort in listening.



◆ Figure 7.6 Pre- and Post-Program Reading Scores for NSLI-Y, Summer 2011, 2012, and 2013



◆ Figure 7.7 Pre- and Post-Program Listening Scores for NSLI-Y, Summer 2011, 2012, and 2013

Proficiency Outcomes in the NSLI-Y Academic-Year Programs

The NSLI-Y academic-year programs induct a much smaller number of students annually than do the summer programs (see Appendix 1 for enumeration, eligibility, and conditions). Most students in this program enter with prior training and with proficiencies in the NM/NH range. Typically, only one-third of the participants begin the year of overseas study without any prior study of the language.

The typical post-program-measured oral proficiency is Advanced Low (AL) (48 percent), Intermediate High (IH) (27 percent), and Advanced Mid (AM) (16 percent). Because of the still-small N-sizes within this data set, results for Chinese and Russian are presented in figure 7.8. Similar results were also observed for smaller groups in Arabic and Korean. Of the 44 students participating in the above two programs, 28 students completed the year with oral proficiency scores in the Advanced range, while an additional 12 students were rated IH, the threshold of the Advanced range. Four students completed the program at the IM level.

Reading and listening proficiency testing was piloted in the NSLI-Y academic-year Russian programs beginning in 2012–13 and will be expanded to other languages in the near future. Preliminary results from the pilot year reflect a strong degree of correlation between both pre- and post-program outcomes for reading and listening. Cross-skill correlation of reading and listening with speaking is also notable, although aggregate outcomes in reading and listening are lower by one sublevel than those for speaking: 57 percent of the cohort scored at the Advanced threshold (IH) at the end of the program, 14 percent at AL for reading, and 21 percent at AL for listening. Further research will be required to confirm whether these differences can be accounted for in terms of the emphasis on interpretive reading and listening of the Advanced-level test or by other factors.

Measured proficiencies in the IH/AL levels across modalities are typical of the language levels attained by college juniors and seniors in the less commonly taught languages at many US universities. Alumni of the NSLI-Y academic-year program should have little difficulty placing into advanced-level courses taught in the target languages upon their return to the United States and should be strongly encouraged to seek out colleges and universities where their high levels of language and cultural skills can be fully utilized in their education.

Pre- and Post-Program Proficiency Outcomes: The CLS Summer Institutes

American education has witnessed some broadening of foreign languages offered for the K–12 system over the past half century with the introduction of Russian in the immediate post-Sputnik era in the 1960s, the introduction of Japanese in the 1980s, the introduction of Chinese in the 1990s, and some expansion of Arabic and other major world languages in the post-9/11 decade. For the majority of American students, however, access to non-European languages has been possible only at the college level. The CLS program is intended to provide college students as well as

◆ Figure 7.8 Pre- and Post-Program Proficiency Test Results for Chinese and Russian

graduate students with the opportunity to undertake or continue study of one of thirteen critical languages (see Appendix 2 for enumeration, eligibility, and conditions) at a structured overseas language and cultural immersion program or summer institute. The program, as noted previously, is competitive, admitting 620 students annually, or approximately one in six applicants. The competition rates for certain high-demand languages may be considerably higher. The report on program outcomes across all languages is presented in figure 7.9.

The dark bars reflect the percentage of the total number of students entering the program at this level of speaking proficiency, while the light bars reflect the percentage of students completing the program at the stated score. Over one-third of the total cohort began the CLS program in the Novice range, including 87 students without any prior exposure to the target languages. Close to half (46 percent) of the total cohort entered the CLS program in the Intermediate range of proficiency, a level typical for US college students with two to six prior semesters of formal study of a critical language. The remaining 10 percent of the 2012 CLS population tested at the AL or AM levels at the outset of the program. By the end of the CLS program, 34 percent of the total population had reached the Advanced level of proficiency in speaking, while 60 percent scored in the Intermediate range, and 4 percent were measured at NH, the threshold level for Intermediate.

As has been noted previously in connection with the analysis of the NSLI-Y summer program, movement across the Intermediate-level "plateau" presupposes sustained and effortful L2 learning and several hundred hours of time-on-task.[4] Comparison of pre- and post-program OPI recordings of students whose measured

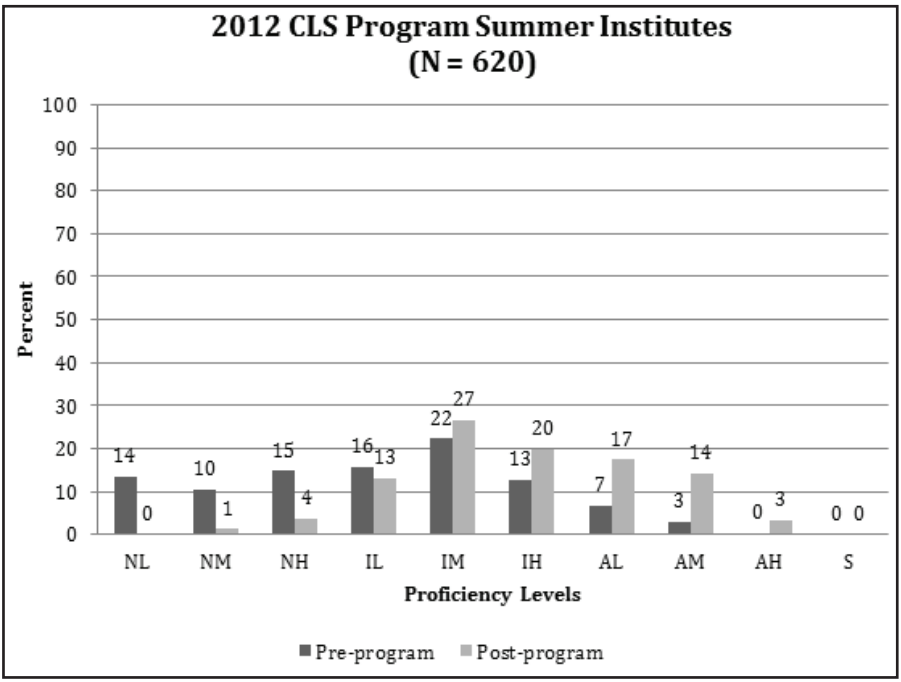


Figure 7.9 CLS Pre- and Post-Program Oral Proficiency Score Comparisons

proficiency scores remained at the Intermediate level throughout the program typically reveal a remarkable range of differences in the latter, including more native-like fluency, sociopragmatic skills, cultural referencing, and a considerably expanded vocabulary evident in the speech production in the post-program learner (Davidson 1982; Magnan and Back 2007; Martinsen 2010).

Table 7.5 presents a cross-tabulation of language-gain data comparing pre- and post-program OPI scores by participant numbers and percentages (N = 620). Measured gain for each entering level of L2 proficiency can be gauged by the distribution of scores to the right (positive gain) or left (negative gain) of the diagonal.

Measureable language gain in the overseas immersion setting is never a guarantee for the L2 leaner. Moreover, the more advanced the learner, the more effort (and time) is required to reach the next threshold. CLS 2012 language-gain reports are summarized in aggregate form in Table 7.6 and Figure 7.10, with sublevel changes included.

The change reports presented here reflect a substantial percentage (57.4 percent) of threshold-level gains produced by the CLS summer program across all thirteen critical languages: 212 of 241 novices achieved Intermediate-level proficiency over the course of the summer program, while an additional 199 students with prior study of the language achieved the Advanced level (see also Figure 7.5, page 128). One Advanced-level student of Chinese reached the Superior level (ILR-3) over the course of the summer.

◆ Table 7.5 Speaking Proficiency Scores for 2012 CLS Program Summer Institutes (N = 620), ALL LANGUAGES, Pre- and Post-Program (Count/Row Percent)

Pre-Program Speaking Proficiency Level	Post-Program Speaking Proficiency Level										
	NL	NM	NH	IL	IM	IH	AL	AM	AH	S	Total
NL	0	7	16	39	19	3	0	0	0	0	84
%	0	8.3	19.0	46.4	22.6	3.6	0	0	0	0	100.00
NM	0	1	6	20	29	8	0	1	0	0	65
%	0	1.5	9.2	30.8	44.6	12.3	0	1.5	0	0	100.00
NH	0	0	1	17	46	18	9	1	0	0	92
%	0	0	1.1	18.5	50.0	19.6	9.8	1.1	0	0	100.00
IL	0	0	0	6	41	28	18	5	0	0	98
%	0	0	0	6.1	41.8	28.6	18.4	5.1	0	0	100.00
IM	0	0	0	0	30	47	39	22	1	0	139
%	0	0	0	0	21.6	33.8	28.1	15.8	0.7	0	100.00
IH	0	0	0	0	0	19	32	22	6	0	79
%	0	0	0	0	0	24.1	40.5	27.8	7.6	0	100.00
AL	0	0	0	0	0	0	8	28	6	0	42
%	0	0	0	0	0	0	19.0	66.7	14.3	0	100.00
AM	0	0	0	0	0	0	2	10	6	1	19
%	0	0	0	0	0	0	10.5	52.6	31.6	5.3	100.00
AH	0	0	0	0	0	0	0	0	2	0	2
%	0	0	0	0	0	0	0	0	100.0	0	100.00
Total	0	8	23	82	165	123	108	89	21	1	620
%	0	1.3	3.7	13.2	26.6	19.8	17.4	14.4	3.4	0.2	100.00

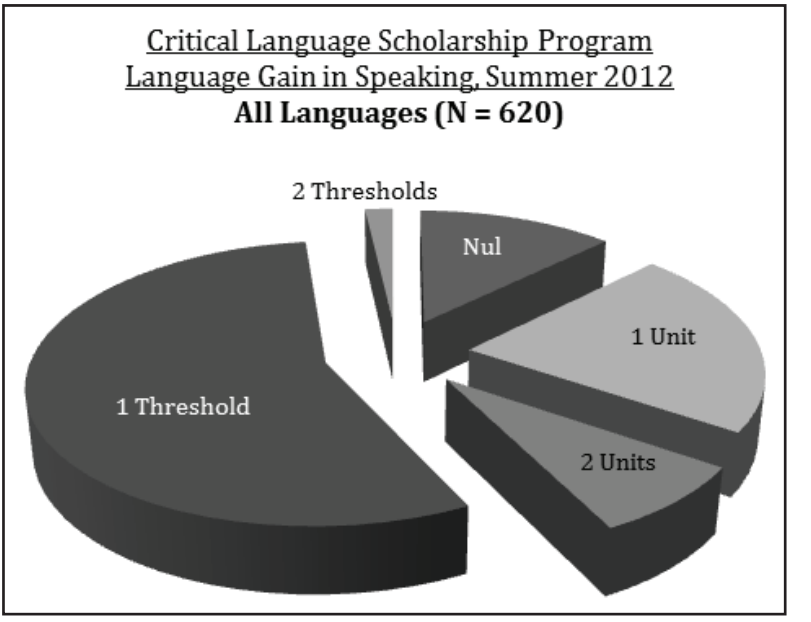
Pre- and Post-Testing Results: The Overseas Flagship Capstone Program

Within the triumvirate of new federal programs supporting the development of US understanding and expertise in major world languages and cultures, the Overseas Flagship Capstone programs (OFC) occupy a special position. Their task is to produce Level-3 speakers in ten critical languages who are well-grounded in language, culture, and regional knowledge of the host country and can function at a professional level in business, government, research, academia, or other sectors of the economy upon completion of the overseas year. The OFC programs accept only those students who have achieved a demonstrated Advanced level (ILR-2) in at least two skills at the time of application and provide a diversified in-country training

experience that draws in varying degrees, depending on the host country, on the full range of intervention strategies and support systems outlined previously. The OFCs emphasize small-group language training and individual tutorials but place increased stress on direct enrollment in courses at the overseas university as well as professional-level internships. The OFC places a special emphasis on in-class and public presentational or project work.

◆ Table 7.6 2012 CLS Test Scores

All Languages	Frequency	Percent	Valid Percent	Cumulative Percent
Loss	2	0.3	0.3	0.3
Null	77	12.4	12.4	12.7
Null	79	12.7	12.7	12.7
1 unit	135	21.8	21.8	34.5
2 units	50	8.1	8.1	42.6
1 threshold	345	55.6	55.6	98.2
2 thresholds	11	1.8	1.8	100.0
Total	620	100.0	100.0	



◆ Figure 7.10 CLS Gain Categories

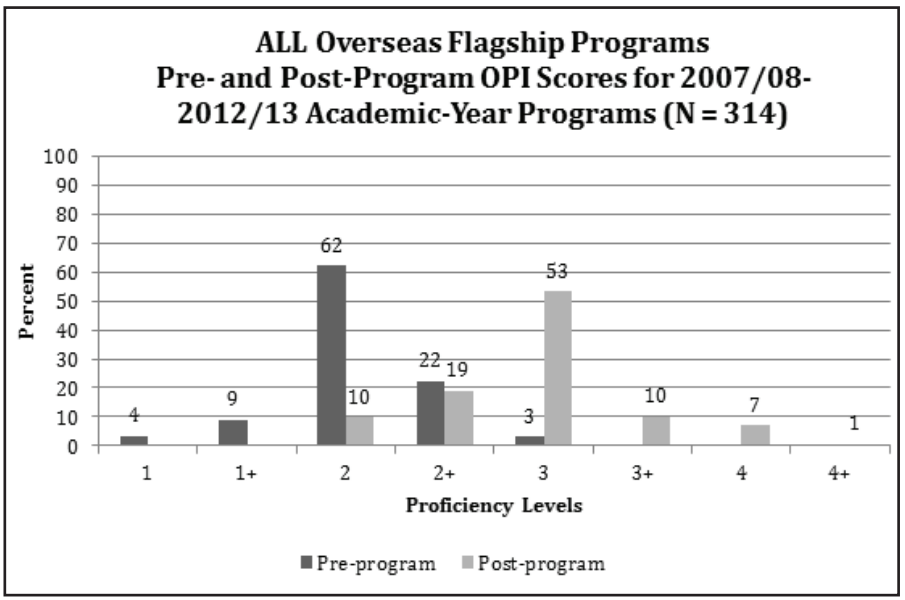
Figure 7.11 presents the pre- and post-program OPI score reports for the several multi-institutional OFC programs. The time period reflected in these data is six years: 2007–2012.

The dark bars reflect measured entering oral proficiency of Flagship students, while the light bars reflect the distribution of post-program OPI scores. It should be noted that in the earlier years of the program, the requirement of 2-level OPI proficiency was not uniformly enforced across all programs. As with other forms of performance-based testing, candidates who have previously scored at target level on an OPI may score lower (or higher) on the same test on a given day, due to fatigue, stress, or other factors.

Overall, 71 percent of all OFC students completed the year-long program at the Superior level (ILR-3) or higher, while an additional 19 percent of the population finished the program at AH (ILR 2+), the threshold of Level 3.

The Advanced-level plateau is well described in the literature, giving rise in the early days of the proficiency movement in the United States to the concept of the "terminal 2" (Higgs and Clifford 1983). Previous research has characterized the likelihood of a US student of Russian moving from Level 2 to Level 3 in the course of a nine-month academic-year program in-country as approximately one in three (Davidson 2010). The likelihood of reaching 2+ or remaining a 2 was also one in three. Taking the OFC model as a whole, the odds of a US student reaching the Superior level (ILR-3) in speaking have improved considerably (see Table 7.7). Currently, 71 percent of all OFC students attain that level or higher.

Reading proficiency outcomes for the same population are presented in Table 7.8.



◆ Figure 7.11 Pre- and Post-Program OPI Score Comparisons: Arabic, Chinese, Persian, and Russian

◆ Table 7.7 Pre- and Post-Program Speaking Scores for NSLI-Y, Summer 2011, 2012, and 2013

Pre-Program Speaking Proficiency Level	Post-Program Speaking Proficiency Level						
	2	2+	3	3+	4	4+	Total
1	7	3	1	0	0	0	11
%	63.60	27.30	9	0	0	0	100.00
1+	9	7	12	0	0	0	28
%	32.10	25.00	42.90	0	0	0	100.00
2	12	41	106	22	13	1	195
%	6.20	21.00	54.40	11.30	6.70	0.50	100.00
2+	3	8	41	9	8	1	70
%	4.30	11.40	58.60	12.90	11.40	1.40	100.00
3	0	1	7	1	1	0	10
%	0	10.00	70.00	10.00	10.00	0	100.00
Total	31	60	167	32	22	2	314
%	9.90	19.10	53.20	10.20	7.00	0.60	100.00

◆ Table 7.8 Overseas Flagship Programs: Reading Proficiency Scores for All Learners (N = 238), Pre- and Post-Program (Count/Row Percent)

Pre-Program Reading Proficiency Level	Post-Program Reading Proficiency Level						
	1+	2	2+	3	3+	4	Total
1	1	2	1	0	0	0	4
%	25.00	50.00	25.00	0	0	0	100.00
1+	0	10	16	8	0	0	34
%	0	29.40	47.10	23.50	0	0	100.00
2	0	8	47	50	18	3	126
%	0	6.30	37.30	39.70	14.30	2.40	100.00
2+	0	0	11	41	6	2	60
%	0	0	18.30	68.30	10.00	3.30	100.00
3	0	0	1	4	9	0	14
%	0	0	7.10	28.60	64.30	0	100.00
Total	1	20	76	103	33	5	238
%	0.40	8.40	31.90	43.30	13.90	2.10	100.00

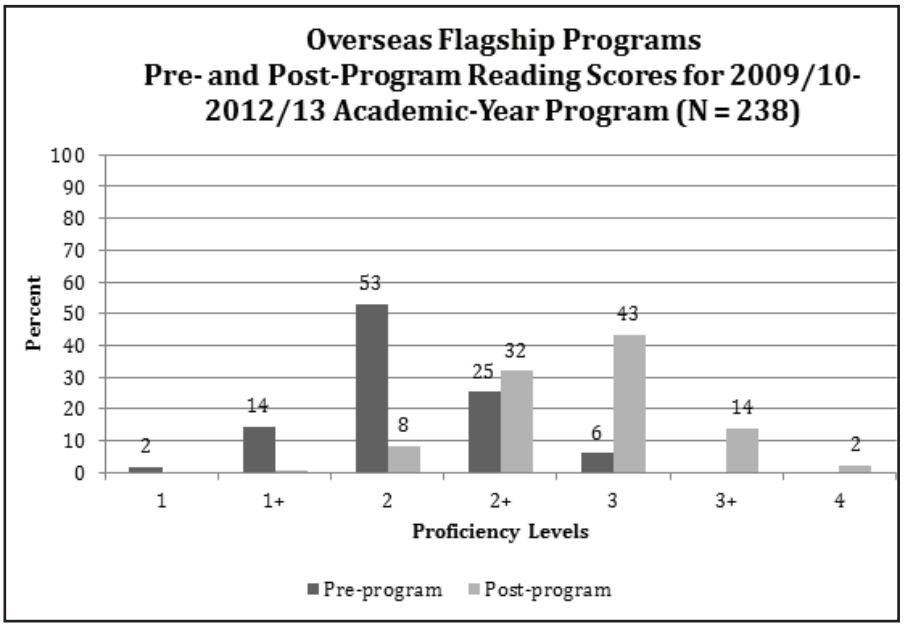
As noted previously, reading and listening proficiency testing of OFC was introduced beginning in 2009, so the N-value in Figure 7.12 is lower than in the OPI report. The pre- and post-testing reading proficiency results present a pattern similar to the OPI speaking inputs and outcomes. A slightly larger concentration (32 percent) of scores at the 2+ threshold level is noted for this modality in comparison to speaking (19 percent). The percentage of OFC students with scores of 3+ or higher in speaking and reading is 18 percent and 16 percent, respectively.

A total of 59.9 percent of all OFC students achieved Level-3 or higher scores in reading proficiency, including a group of 11 (18.3 percent) students who entered the program with Level 2+-proficiency in reading.

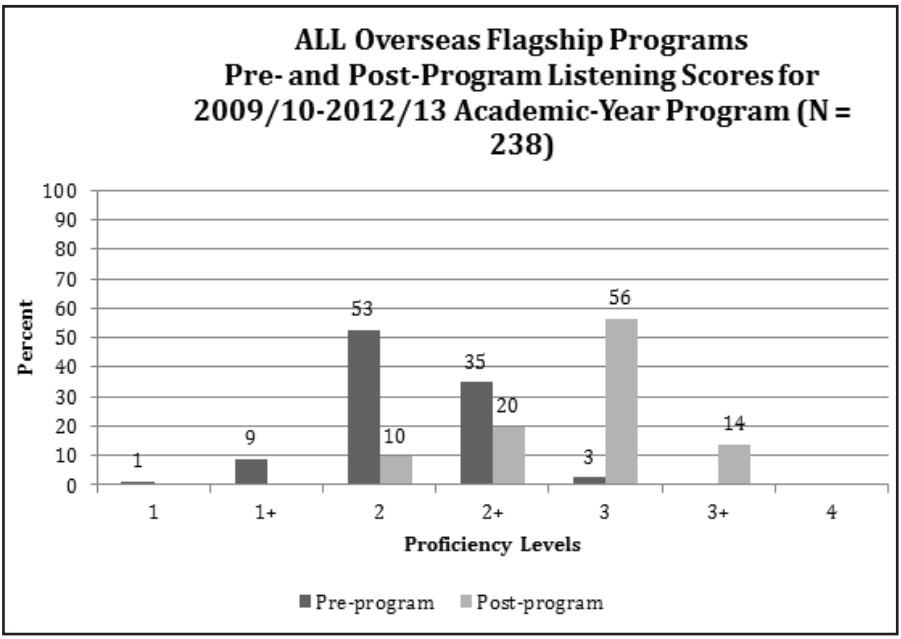
Listening proficiency outcomes for all OFC programs are presented in Figure 7.13.

Listening proficiency outcomes for the OFC groups as a whole parallel OPI outcomes to a very high degree—much more closely, in fact, than was the case with reading. A total of 70.6 percent of OFC participants completed the program at Level 3 or higher in listening (compared to 71.0 percent in speaking and 59.9 percent in reading). As reported previously (Davidson 2010), listening comprehension is strongly correlated with oral-proficiency gain at Level 2 and above.

Approximately 8 percent of all students entering the OFC program with 2- or 2+-level listening skills failed to register measureable improvement in their proficiency over the course of the programs. A total of 34 participants (14.3 percent) achieved 3+-level proficiency or higher.



◆ Figure 7.12 Reading Proficiency Outcomes: Overseas Flagship Capstone Programs



◆ Figure 7.13 Pre- and Post-Program Listening Proficiency Outcomes: OFC

Given the considerable differences in local linguistic and cultural conditions across the major OFCs, breakout data for the Russian OFC alone are provided in Figures 7.14–7.16.

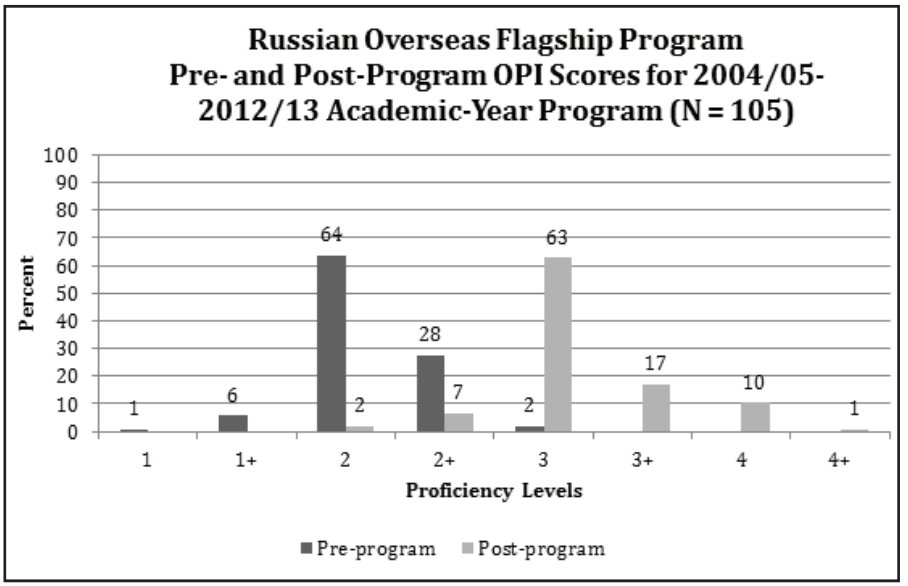
The cross-tabulations here present the input and output rates of the Russian OFC since its inception. The program currently has a null-gain rate of 8.57 percent, with production rates of 62.86 percent for Level 3 and 27.63 percent for 3+ and above (see Table 7.9).

Reading and listening proficiency results in the Russian OFC generally exceed those for speaking, with 38 percent of the group achieving 3+ or higher results in listening and 54 percent achieving 3+ or higher in reading.

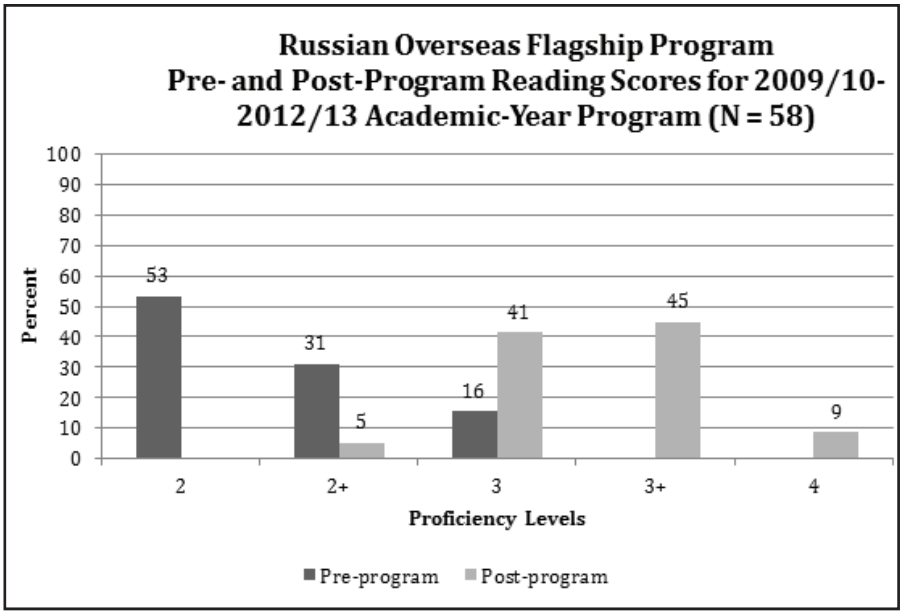
The Arabic OFC pre-/post-program oral proficiency test results are presented in Figure 7.17. Over the past six years, 84 percent of students have entered the program with OPIs in the Advanced-level (ILR-2) range, while 71 percent of all participants have completed the Arabic OFC at Level 3 or higher in speaking. An additional 19 percent of the group has completed the program at 2+ in speaking, the threshold of Level 3.

Null gains for students entering with the requisite 2-levels of speaking proficiency represent 8.7 percent of the total participants. On the other end of the scale, 18 percent of participants completed the program at the 3+ level or higher.

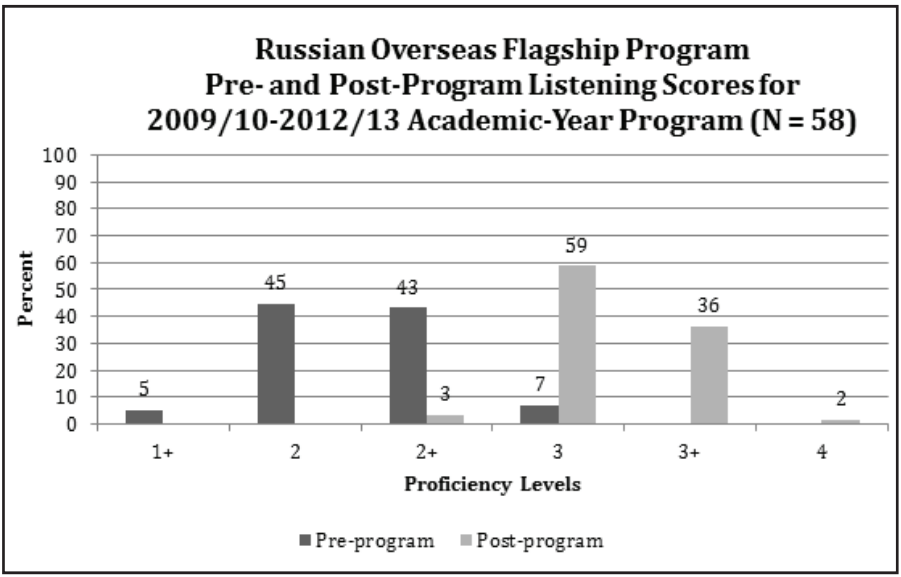
An asymmetry in outcomes is noted between the Arabic OFC reading scores and the speaking and listening scores (compare Figures 7.17–7.19). Speaking and listening outcomes at Level 3 or higher in the Arabic programs are essentially identical:



◆ Figure 7.14 Russian Overseas Flagship, 2004–2012



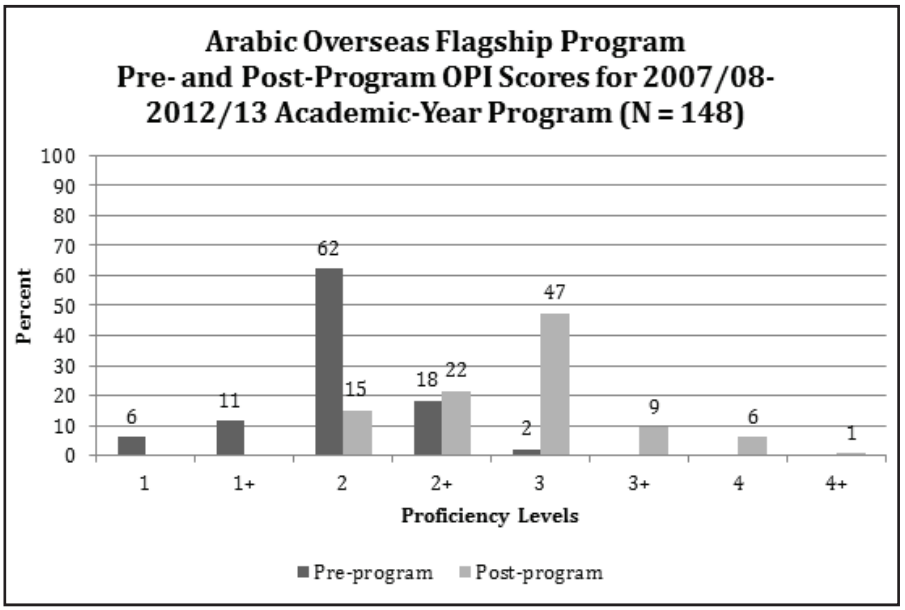
◆ Figure 7.15 Pre- and Post-Program Reading Proficiencies: Russian OFC



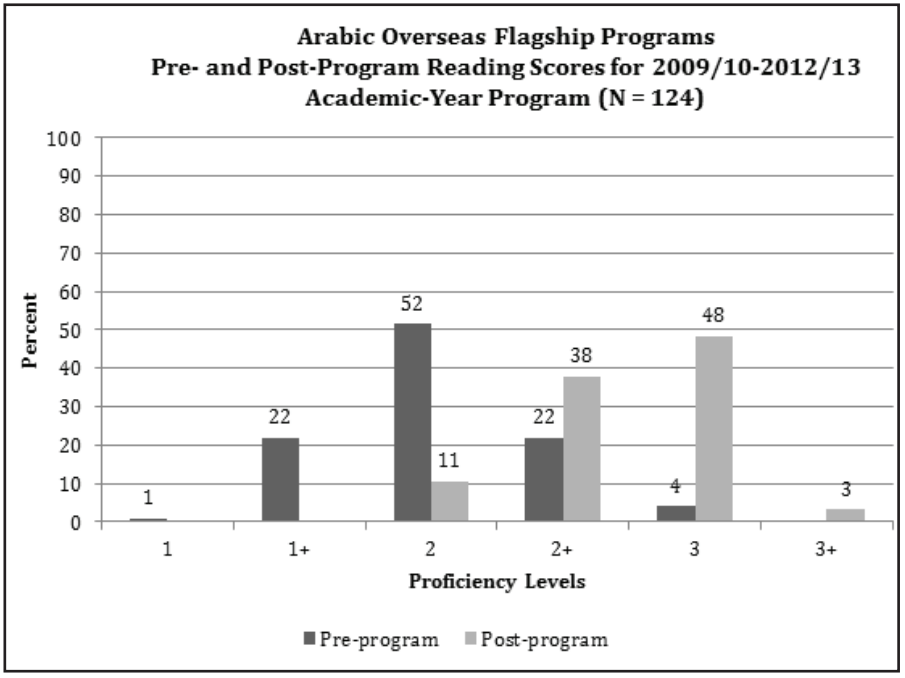
◆ Figure 7.16 Pre-and Post-Program Listening Proficiencies: Russian OFC

◆ Table 7.9 Russian Overseas Flagship Speaking Proficiency Scores for All Learners (N = 105), Pre- and Post-Program (Count/Row Percent)

Pre-Program Speaking Proficiency Level	Post-Program Speaking Proficiency Level						
	2	2+	3	3+	4	4+	Total
1	0	1	0	0	0	0	1
%	0	1.00	0	0	0	0	100.00
1+	2	1	3	0	0	0	6
%	33.33	16.67	50.00	0	0	0	100.00
2	0	5	43	12	6	1	67
%	0	7.46	64.18	17.91	8.96	1.49	100.00
2+	0	0	19	5	5	0	29
%	0	0	65.52	17.24	17.24	0	100.00
3	0	0	1	1	0	0	2
%	0	0	50.00	50.00	0	0	100.00
Total	2	7	66	18	11	1	105
%	1.90	6.67	62.86	17.14	10.48	0.01	100.00



◆ Figure 7.17 Pre- and Post-Program Oral Proficiencies: Arabic OFC



◆ Figure 7.18 Reading Proficiencies: Arabic OFC

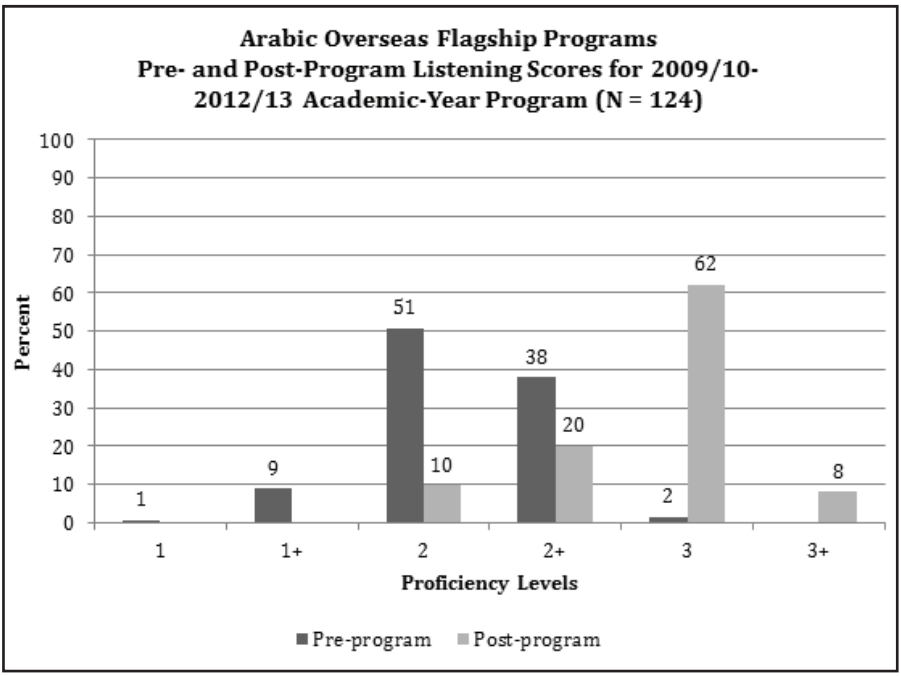


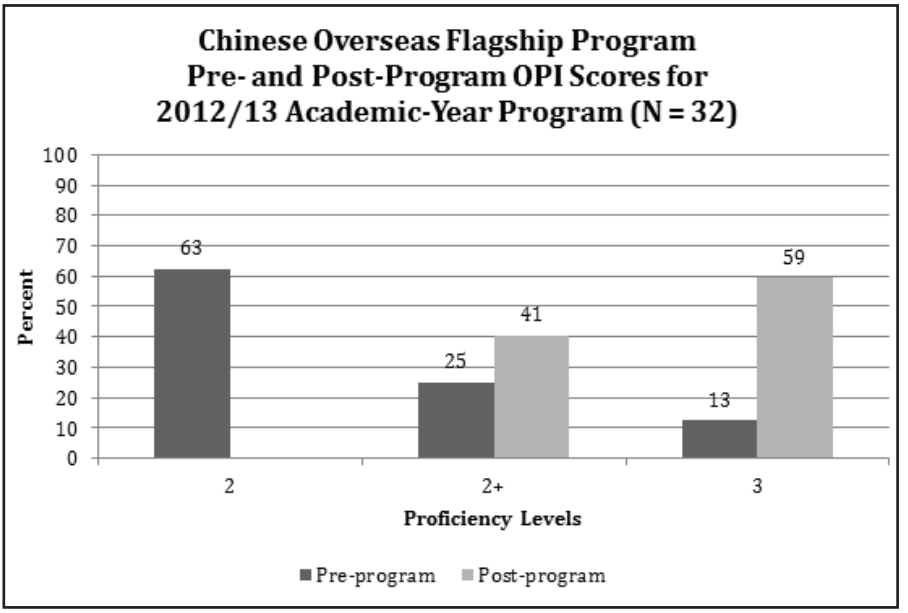
Figure 7.19 Listening Proficiencies: Arabic OFC

71 percent of the group achieved the Professional level in speaking, and 70 percent achieved the Professional level in listening comprehension. In comparison, 51 percent achieved the Professional level in reading, while 38 percent tested at Level 2+. The teaching and testing of Arabic reflects both the diglossic and, in some cases, bidialectal nature of the host country sites where Flagship students are placed. These complexities are reflected in the testing environment for Arabic as well.

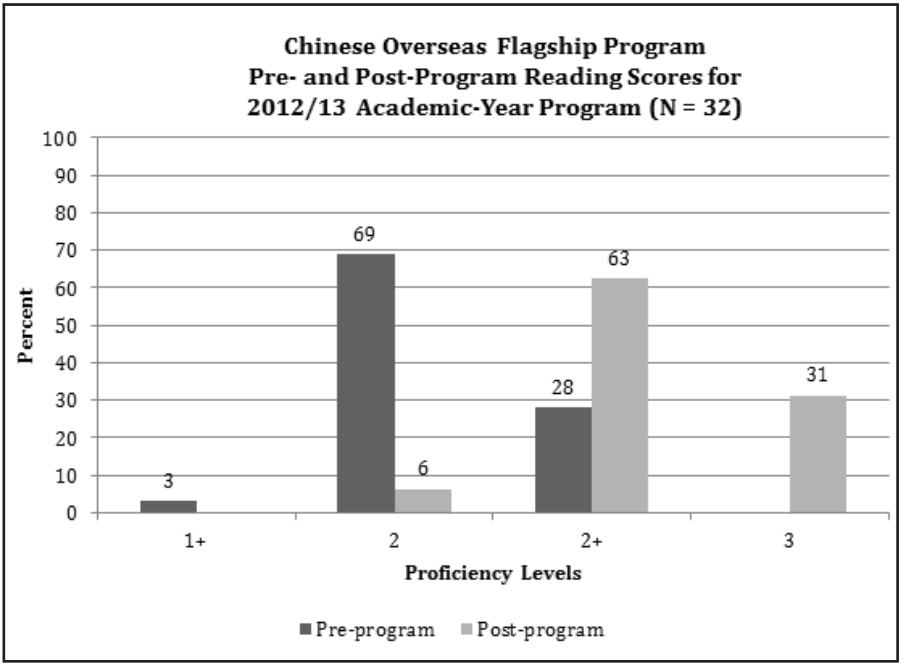
Results of the Chinese OFC are presented next; however, for technical reasons, only one year of program data is available at this time: 2012–2013.

While the proficiency measurements of entering Chinese OFC students were relatively strong in comparison to all other OFCs, the production of Level-3 speakers (and above) was somewhat lower than both the Arabic and the Russian programs. On the other hand, no participant in the Chinese OFC scored lower than 2+ in speaking at the close of the program (see Figure 7.20).

Reading and listening proficiency ratings for Chinese OFC students show that 31 percent of the cohort attained Level-3 proficiency in reading, while 38 percent achieved that level in listening comprehension (see Figures 7.21 and 7.22). Entering levels for both modalities were primarily at 2. A majority of participants in the program were rated at 2+ on the post-program test in reading, while program-final listening tests also included a cohort of 25 percent of the learners who scored at Level 2. Seven out of thirty-two students registered null gain in listening. Speaking scores exceeded reading or listening scores for the group as a whole by more than twenty



◆ Figure 7.20 Pre- and Post-Program Oral Proficiencies: Chinese OFC



◆ Figure 7.21 Pre- and Post-Program Reading Proficiencies: Chinese OFC

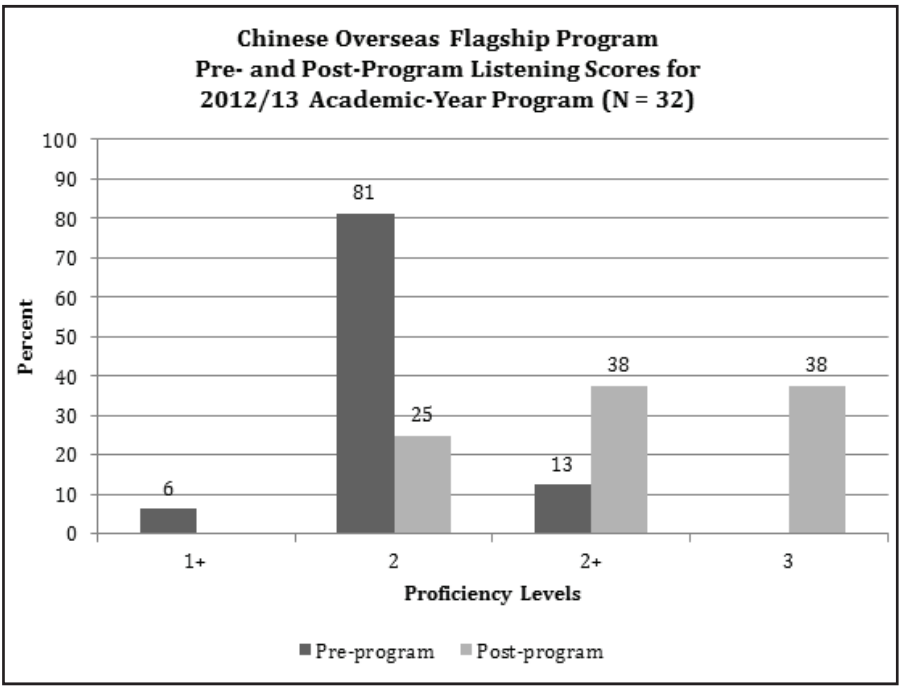


Figure 7.22 Pre- and Post-Program Listening Proficiencies: Chinese OFC

percentage points. It should be noted that the testing of reading at Level 2+ and above requires knowledge of both traditional and simplified characters.

Discussion

The data presented here and elsewhere in this volume make clear that the pathway to professional-level competencies in the major world languages is open and available to Americans with the motivation and support to engage the necessary mechanisms of immersion study and domestic learning. The foregoing report demonstrates that access to high-quality overseas immersion programs, such as those supported by the US government through NSLI-Y, CLS, and the Language Flagship, can affect proficiency growth across skills at every stage of the language learning career, from absolute novices to professionals working at Levels 3, 3+, and 4. Overseas immersion learning is not the only pathway for language acquisition, but where the less commonly taught languages are concerned, it is difficult to construct a comprehensive curricular model within our existing educational system without recourse to one or more immersion models.

Returning then to the research questions with which this study began:

1. We have seen that threshold gains at the Novice-to-Intermediate levels and across sublevels within the Intermediate range are a clear effect of the eight-week

intensive summer immersion model, implemented by the NSLI-Y programs for seven major world languages. We have seen that the NSLI-Y academic-year program is producing L2 speakers who test at the Advanced level, a level of language mastery typical of upperclassmen and graduating seniors at many of our major universities. We have seen the summer CLS program for university students providing a similar opportunity for achieving threshold gains to the Intermediate or Advanced level for students of a wide range of language-study backgrounds across thirteen languages. A 50+ percent success rate in producing Intermediate-level speakers from Novices, and Advanced-level speakers from Intermediates, is a notable achievement in second language acquisition in the context of a summer program, a time frame that has not previously been regarded as sufficient for threshold-level gains. Finally, in response to the need to produce greater numbers of Americans capable of functioning as "global professionals," we have observed that it is possible to reach Level 3 and above on a systematic basis, year after year, with US undergraduate students who have attained Level 2 at the time of their participation in the program in the year-long OFC model.

2. We have also seen that reading and listening gains generally align with the acquisition of speaking skills in all three federal programs. In some cases, interpretive listening and reading appear to lag slightly behind the corresponding interpersonal listening and reading functions. In other cases, such as the Russian OFC, reading actually exceeded speaking-skills production at Level 3 and above. The actual variation of speaking, reading, and listening skills across summer and academic-year programs is presented in detail for both aggregate and single-language groups in the foregoing report, in the course of which both program duration and initial levels of participants are identified and discussed.

The study has identified no differences (positive or negative) in terms of an age or gender effect for language gain in the present data. Nor have we identified sufficient basis in these data to identify any one of the critical languages as objectively more difficult to acquire than other NSLI languages in terms of time-on-task. Further research will be needed to assess the potential effects of the diglossic, and, in some cases, bidialectal environment of the acquisition and assessment of Arabic at Level 3 and above. The data on the Chinese programs are still insufficient to confirm whether issues regarding the dual writing systems in contemporary Chinese affect the acquisition and assessment of reading at Level 3; further research is needed here as well.

Conclusion

A recent international conference, "Languages for All? The Anglophone Challenge," brought together a distinguished group of language education and policy experts from the United States, the United Kingdom, Australia, and Canada to envision the L2 needs of our respective societies in the context of the interconnected and highly competitive world of the twenty-first century. Conference organizer

Richard D. Brecht (2013, 4) summarized the views of the conference, starting with the assumption that all Americans should have access to L2 training and that the country as a whole will require persons of differing levels of proficiency, depending on their roles in society and the economy. Among other recommendations, the conference report envisions a multilingual American population in the not-too-distant future in which

- Level 1: 50 percent have some exposure to international perspectives, culture, and/or language in order to inform lifelong decisions about work and learning and to support language and international efforts broadly in society;
- Level 2: 30 percent have some basic language skills in order to work, for example, in the service industry and to travel freely;
- Level 3: 15 percent have global professional skills in order to practice their work successfully at a high level internationally;
- Level 4: 5 percent have expert skills in order to perform necessary research and design and implement language education and training.

The federally supported overseas immersion programs provide a model, and also a proof of concept, that the United States has the capability to produce the multiple levels of language proficiencies across the major world languages needed by the nation well into the twenty-first century. The responsibility of policymakers and educators is to expand these and other successful models now in place and move forward with the preparation of a new generation of citizens that possess the full range of linguistic, cultural, and regional skills essential for success in the global economy.

Appendix 1

NSLI for Youth Programs
www.nsliforyouth.org

The National Security Language Initiative for Youth (NSLI-Y) programs offer intensive language immersion in a variety of locations around the world. Scholarships are available for students wanting to learn the following languages: Arabic, Chinese (Mandarin), Hindi, Korean, Persian (Tajiki), Russian, and Turkish.

Programs may take place in the following locations: China, India, Jordan, Korea, Morocco, Oman, Russia, Taiwan, Tajikistan, Turkey, and other locations around the world.

Eligibility Requirements:

- US citizen
- Grade point average (GPA) of 2.5 or higher on a 4.0 scale, or the equivalent
- 15–18 years of age at start of program
- Enrolled in high school (including home school)
- Not an immediate family member of an employee of the US Department of State who works in the Youth Programs Division of the Bureau of Educational

and Cultural Affairs (ECA) or an employee at an NSLI-Y administering organization whose duties involve the NSLI-Y program
- Have not previously traveled outside the United States on a long-term (more than eight weeks) program sponsored by the Bureau of Educational and Cultural Affairs, Department of State
- Previous NSLI-Y *summer* program participants or participants of ECA-funded short-term programs are only eligible to apply for a NSLI-Y academic-year program.

Previous language study is not a requirement. Students of all levels of language ability are encouraged to apply.

The NSLI-Y program seeks applicants who represent the diversity of the United States. Students of all racial, ethnic, religious, and socioeconomic backgrounds are encouraged to apply, as are students with disabilities.

Appendix 2

Critical Language Scholarship Program
www.clscholarship.org

Overview and reference to funding and cooperating agencies (from the website)
A program of the US Department of State, Bureau of Educational and Cultural Affairs, the Critical Language Scholarship (CLS) Program offers intensive summer language institutes in thirteen critical foreign languages. The selection process is administered by the American Councils for International Education, with awards approved by the US Department of State, Bureau of Educational and Cultural Affairs. The CLS program is administered by the American Councils and Ohio State University/Ohio University.

Summary of Key Eligibility Requirements:
Applicants to the Critical Language Scholarship program must be US citizens and enrolled in an accredited degree program in the United States at the undergraduate (associate's or bachelor's degree) or graduate level (master's or doctoral degree). Participants must be at least eighteen years of age by the beginning of the program. Language proficiency requirements vary across programs.

Those applying for the Chinese, Japanese, or Russian programs must demonstrate the equivalent of at least two years of college-level study, while those applying for the Arabic or Persian programs must demonstrate the equivalent of at least one year of college-level study by the start of the program. There are no minimum language-proficiency requirements for applicants to Azerbaijani, Bangla, Hindi, Indonesian, Korean, Punjabi, Turkish or Urdu.

Appendix 3

The Language Flagship Capstone Program
www.thelanguageflagship.org

The Language Flagship Capstone full-year immersion program is open to all undergraduate students who are committed to attaining Professional- or Superior-level language proficiency through an intensive language training program tailored to their professional interests and academic specialization. It may occur during the third, fourth, or fifth year of a student's undergraduate program. The model also assumes that, in addition to full-year study, some students will require additional periods of immersion overseas to accelerate their language learning and to accommodate academic schedules. The program also accepts applications from part-time, non-degree-seeking students.

Applicants should have a strong academic record, a demonstrated interest in advancing their Arabic, Russian, Persian, Chinese, Hindi/Urdu, Korean, Portuguese, Russian, Swahili, and Turkish skills and using these languages in their future career, and a desire to share their understanding of this language and culture within the larger community.

Undergraduate Program
All students who are enrolled at one of the Domestic Programs and reach the required proficiency level ILR-2 in their language are accepted to the Overseas Program, upon recommendation of the Overseas Academic Council.

Post-BA Program
This program accepts applicants who did not participate in a Domestic Flagship program and already have a bachelor's degree. The participants are selected on the basis of their language skills, academic merits, work experience, and ability to demonstrate how advanced Russian skills are going to help their career plans. Applicants to the Flagship Post-BA program must either possess a BA degree or expect to receive one before starting the program. Successful applicants who are not heritage speakers must have completed at least three years of the language at a college level.

Notes

1. Overseas language offerings vary among the three federal programs, although all support training in Arabic, Chinese, Hindi, Korean, Persian, Russian, and Turkish. The US Department of State funds the National Security Language Initiative for Youth (NSLI-Y) program for high school students, the Critical Language Scholarship (CLS) program funds the overseas summer institutes for university students and graduate students, and the Defense Language and National Security Education Office (DLNSEO) funds the Language Overseas Flagship Capstone programs.
2. In order to ensure consistency in the presentation and reporting of program designs and program data, only those overseas NSLI-Y, CLS, and Flagship programs administered by the American Councils have been included in the present study.

3. This schema is adapted from a more detailed discussion of overseas language immersion models for traditional and heritage learners (Davidson and Lekic 2012).

References

Bazarova, Saodat, Maria D. Lekic, and Camelot Marshall. 2009. "The Online Proficiency-Based Reading, Listening, and Integrated Writing External Assessment Program for Russian: Report to the Field." *Russian Language Journal* 59: 59–78.

Brecht, Richard, D. Davidson, and Ralph Ginsburg. 1995. "Predictors of Foreign Language Gain during Study Abroad." In *Second Language Acquisition in a Study Abroad Context,* edited by Barbara F. Freed, 37–66. Amsterdam: John Benjamins.

Brecht, Richard D. et al. 2013. "Languages for All? The Anglophone Challenge." University of Maryland. Accessed May 10, 2014. http://www.casl.umd.edu/sites/default/files/LFA_WhitePaper_fnl.pdf.

Byrnes, Heidi. 2013. "Notes from the Editor." *Modern Language Journal* 97 (4): 825–27.

Davidson, Dan E. 1982. "Assessing Language Proficiency Levels of American Participants in Language Programs in the Soviet Union." *Russian Language Journal* 36 (125): 221–32.

———. 2007. "Study Abroad and Outcomes Measurements: The Case of Russian." *Modern Language Journal* 91: 276–280.

———. 2010. "Study Abroad: When, How Long, and with What Results? New Data from the Russian Front." *Foreign Language Annals* 43 (1): 6–26.

Davidson, Dan E., and Maria D. Lekic. 2012. "Comparing Heritage and Non-Heritage Learning Outcomes and Target Language Utilization in the Overseas Immersion Context: A Preliminary Study of the Russian Flagship." *Russian Language Journal* 62: 47–78.

Davidson, Dan E., and Nadra Garas. 2009. "The ACTR Nationwide Survey of Russian Language Instruction in U.S. High Schools in 2009." *Russian Language Journal* 59: 3–20.

Dewey, Dan P. 2004. "A Comparison of Reading Development by Learners of Japanese in Intensive Domestic Immersion and Study Abroad Contexts." *Studies in Second Language Acquisition* 26 (2): 303–327.

Dwyer, Mary M. 2004. "More Is Better: The Impact of Study Abroad Program Duration." *Frontiers: The Interdisciplinary Journal of Study Abroad* 10: 151–63.

Higgs, T. V., and Ray Clifford. 1984. "The Push toward Communication." In *Curriculum, Competence and the Foreign Language Teacher,* edited by T. V. Higgs, 51–79. Lincolnwood, IL: National Textbook.

Kinginger, Celeste. 2008. "Language Learning in Study Abroad: Case Studies of Americans in France." *The Modern Language Journal,* 92 (Monograph).

———. 2013. "Identity and Language Learning in Study Abroad." *Foreign Language Annals* 46 (3): 339–58.

Magnan, Sally Sieloff, and Michele Back. 2007. "Social Interaction and Linguistic Gain during Study Abroad." *Foreign Language Annals* 40 (1): 43–61.

Martinsen, Rob A. 2010. "Short-Term Study Abroad: Predicting Changes in Oral Skills."*Foreign Language Annals* 43 (3): 504–530.

Open Doors. 2013. "Report on International Educational Exchange Online." Accessed November 11, 2013. http://www.iie.org/Research-and-Publications/Open-Doors.

Watson, Jeffrey R., Peter Siska, and Richard L. Wolfel. 2013. "Assessing Gains in Language Proficiency, Cross-Cultural Competence, and Regional Awareness during Study Abroad: A Preliminary Study." *Foreign Language Annals* 46 (1): 62–79.

Part Three

Future Directions in Assessment, Program Design, and National Policy

This page intentionally left blank

Chapter 8

From Proficiency to Expertise: Using HR Evaluation Methods to Assess Advanced Foreign Language and Culture Ability

PATRICK MCALOON

SHARON WHITMORE STARES AT a short stack of résumés on her desk. The vice president for Business Development at the largest electricity generation company in the United States just sent her another email asking when she will name someone to lead their new US–China Clean Energy research initiative. From a technical standpoint, each of the American and Chinese candidates is clearly qualified and each has relevant project leadership experience, but what does it mean when their résumés all say "fluent in English and Mandarin?" Sharon has accompanied staff to China in the past and learned quickly the limits of her company's "bilingual" employees' language skills. Sure, everyone got along fine at those endless banquets, but an expensive interpreter is always needed for technical discussions, and Sharon knows from human resources (HR) research that simple staff meetings in an international office that is not truly bilingual can result in many misunderstandings. How can she know which of the candidates is truly able to work in a bilingual setting?

Since few hiring managers are fluent in all the languages spoken where their employees work, it would be difficult for them to recognize professional-level foreign language ability, even when they see it. In lieu of being able to evaluate professional foreign language skills herself, Sharon needs a method of evaluating how well each candidate can appropriately apply relevant job expertise, as well as language and cultural knowledge, to specific target tasks. From what she can tell, measurements of "proficiency" in the foreign language education community do not seem to fit the bill.

Foreign language educators commonly use the idea of "proficiency" to define foreign language ability. Reading the proficiency guidelines published by the American Council on the Teaching of Foreign Languages (ACTFL)[1] reveals that proficiency is primarily a measure of an individual's functional familiarity with the structure of

the target language as native speakers use it. Furthermore, the proficiency guidelines define advanced skills in terms of an individual's approximation of communicating like a "highly-educated" native speaker. (For an extended discussion of ACTFL proficiency guidelines, see Martin, this volume.)

As currently written, the proficiency scale may be well-suited for defining and assessing foreign language ability at the Novice and Intermediate levels. At such levels, the ACTFL proficiency scale can tell teachers and school administrators what they need to know about their students. At more advanced levels, the ACTFL proficiency scale and its accompanying methods of assessment may be useful for evaluating the abilities of foreign language educators (i.e., people in the same fields as those who created the proficiency scale), but for other occupations, someone like Sharon would need measures and assessments that describe a candidate's ability to successfully accomplish specific professional tasks in the foreign language environment, such as resolving a labor dispute or producing a profit and loss statement. In short, the ACTFL/Interagency Language Roundtable (ILR) proficiency scale focuses on generalized erudition. Linguistic accuracy may limit the scale's applicability when attempting to describe and measure professional-level foreign language use.

Shakespeare wrote in *As You Like It* "all the world's a stage, and all the men and women merely players" (Act II, Scene VII, lines 139–140), and more recently, Walker and Noda (2000, 14) wrote, "the flow of social life occurs as a series of performances."[2] Individuals are constantly being evaluated by those around them; therefore, building a repertoire of performances that their audiences find valuable becomes the core responsibility of a foreign language learner. In this regard, Walker and Noda (2000, 14) observe that "No one really learns a foreign language. Rather, we learn how to do particular things in a foreign language; and the more things we learn to do, the more expert we are in that language." As such, expertise in professional foreign language use should be gauged by assessing the quality of performances made in the target language and in the target work environment.

Commonly Used Tools for Foreign Language Assessment

The gold standard of foreign language assessment in US education is the ACTFL Oral Proficiency Interview (OPI), a structured conversation that is designed to elicit speech production that indicates the proficiency level of the individual. Though the ACTFL OPI is not designed to measure how well an individual operates in any specific professional domain, the proficiency standards do indicate that use of scholarly language differentiates a Superior performance from an Advanced one. Insofar as the use of scholarly language is expected of those in scholarly professions, the ACTFL proficiency scale and its attendant OPI measurement tool implicitly measure foreign language ability in the academic domain. Individuals working in professions whose performances do not occur in a scholarly register may be limited to an "Advanced" rating on an OPI, regardless of the appropriateness of their performances to their respective professional domains.

Oral Proficiency Interviews appear to be acontextual, meaning that the context defaults to that of a tester chatting with an interviewee. For example, in one sample Advanced-level response, the interviewer asked the interviewee to discuss what people in White Plains, New York, are talking about.[3] The most appropriate conversation on this topic should vary depending on who is talking and why. One would anticipate that the language used by a congressman from White Plains in a meeting with the governor would differ from that of a bartender chatting with a traveling salesman. Nevertheless, their respective abilities to describe current events in White Plains in a register appropriate for their respective audiences could have a significant impact on their compensation. According to the ACTFL proficiency scale, the bartender would be effectively barred from testing above Advanced in an OPI if he does not use scholarly or educated language. The bar owner, on the other hand, may believe it is more useful to measure the bartender's ability to generate revenue for the bar instead of his use of educated language. Unless the bar was located next to a university, the bartender's tips could even *decrease* if use of highly-educated speech created distance between him and his customers.

The ACTFL proficiency guidelines are based on the ILR scale, a proficiency scale developed by US government agencies with foreign language needs. Such agencies rate individuals using their own version of the OPI that is correlated with ILR scale proficiency-level descriptors. The designs of the ILR and ACTFL OPI for measuring proficiency are very similar, but differ primarily in the mechanics of how the interviews work, e.g., the number of raters involved in an OPI and how differences of opinion between the two raters are handled (Kang, Lee, and Kim 2007). ILR OPIs allow for more than one role-play, whereas role-plays have a limited function in ACTFL OPIs.

Another commonly used proficiency test is the Defense Language Proficiency Test (DLPT). The DLPT tests the passive skills of reading and listening. More commonly taught languages are tested with multiple-choice questions; less commonly taught languages, such as Dari, are tested with short-answer questions about the prompts.[4] The DLPT was designed to assess the skills of military language analysts whose jobs included consuming foreign language media and producing English-language reports about them. The default context of the DLPT—a test taker trying to understand decontextualized samples of the target language—is probably close to the context in which the test takers will actually be using their language skills.[5] If DLPT test takers were to produce reports on assigned media samples rather than answer short questions, then this proficiency test could probably be considered a performance-based test, in the sense that it evaluates test takers on their ability to actually perform their target jobs.

One language assessment tool that seeks to describe what learners can *do* in the language is LinguaFolio. Developed by the Center for Applied Second Language Studies at the University of Oregon, LinguaFolio is a portfolio for language learners that is composed of three pieces: self-assessments, biographical information, and a "dossier" that holds samples of a student's work. The self-assessments are a series of checklists in which a student answers whether or not she or he can perform certain tasks. These tasks are explicitly tied to ACTFL proficiency descriptors, e.g.,

"I can give and seek personal views and opinions on a variety of familiar topics," with samples of familiar topics being offered as "favorite celebrities," "favorite sports team," and "the importance of recycling in my community" (National Council of State Supervisors for Languages 2009).

One advantage of portfolio assessment tools such as LinguaFolio is that they can capture data about performances made over a period of time. A portfolio provides a platform on which to rate performances of target abilities whenever they are needed, not just when artificially elicited for an interview. Thus, LinguaFolio can capture many more data points than an OPI can in defining a learner's proficiency level on the ACTFL scale. In addition, the ease with which students and teachers alike can use the system makes accumulating and reviewing this data fairly simple.

LinguaFolio ratings are self-assessments, the validity of which has received mixed reviews in the scholarly world (cf. Oskarsson 1984, Ross 2006, and Saito 2003). Individual researchers' theories on language tend to guide the process of validating self-assessments. If a researcher believes that the purpose of language is self-expression, then the final arbiter of successful communication is the nonnative speaker. If the researcher believes that language represents a part of a culturally situated act performed to elicit from listeners a sympathetic response and generate feelings of "being on the same team," then the native-speaker listener's assessment may be more valuable. For instance, an American who can use Chinese to discuss his or her favorite American sports team and describe the importance of recycling in the community may impress Chinese teachers, but if those conversations are not meaningful to a Chinese listener (e.g., the favorite sport is lacrosse) then there could be a fatal disconnect between the American's self-perceived ability to communicate and the Chinese native listener's willingness to converse on a topic to which she or he cannot relate. To eliminate this disconnect, the can-do statement could be revised to read something like "I can exchange views and opinions about a sport familiar to my native listeners."

If the opinions of native-speaking professional peers are most important in evaluating foreign language expertise, then interviews by educators and self-assessments by nonnatives may fall short in answering the question "Is the nonnative speaker able to function successfully in the target-language workplace?" Because performance serves as the basis for evaluation and assessment in professional organizations, HR management can offer some useful methods for evaluating foreign language expertise.

Professional Performance Evaluation in the Human Resources Field

Organizational psychologists Borman and Motowidlo describe two types of professional performance: 1) task performance, which is the ability to successfully carry out a technical skill needed by the organization, and 2) contextual performance, or the ability to successfully negotiate the social aspects of work in order to meet organizational goals, e.g., conducting a meeting (Sonnentag and Frese 2005, 6). If we define professionally useful levels of foreign language ability as the ability to perform one's

job successfully with native-speaking peers and in a target-culture work environment, then professional-level language assessment should assess contextual performance. Depending on the tools used for assessment, assessing the task performances of a nonnative employee (or prospective employee) also may be possible.

There are several different methodologies used for workplace performance appraisal, including Management by Objective and Critical Incidents. Most are methods of generating "downward feedback" (Lepsinger and Lucia 1997, 7), in which a supervisor records and measures an employee's performance. The inclusion of upward feedback (employee self-assessments) in the evaluation process started gaining popularity in the 1980s (Lepsinger and Lucia 1997, 7). Many companies still use a combination of upward and downward feedback for their performance appraisals, giving employees an opportunity to report accomplishments of which their supervisors may not be aware, as well as to respond to remarks made by their supervisors in their evaluations. Upward feedback represents an important step forward in terms of including additional perspectives on individual performance, but still lacks input from the wide variety of individuals whose lives and livelihoods are impacted by an employee's performance, including suppliers, colleagues in other functional areas, and internal and external clients.

Changes in the corporate world pushed HR professionals to seek an assessment tool capable of capturing the opinions of a wider spectrum of people with whom an individual employee interacts. The result was the creation of 360-degree feedback: a process that incorporates downward feedback, self-assessment, and peer assessment to provide a full picture of an individual's professional performance.

A multi-rater feedback process such as 360-degree feedback is more complex and takes more time to use than supervisor-only or supervisor-plus self-assessments. As a result, this tool is generally used to evaluate and provide feedback to people whose skills are highly valuable in the organization, especially executives and upper-level managers (Lepsinger and Lucia 1997, 17). Managers represent a significant investment in salary and benefits, as well as in time spent accruing the experience and expertise needed to perform complex tasks. Similarly, high-performing speakers of foreign languages represent a significant investment of time and money, which makes them reasonable candidates for a similar assessment tool.

The complexity of 360-degree feedback is twofold: first, it requires careful design in order to obtain feedback on the types of performances that the organization needs and seeks to develop; second, it requires the involvement of more people than traditional supervisor-subordinate performance reviews, thus complicating the collection and analysis of data.

The feedback tool must be carefully designed to include the organization's target performance dimensions. Lepsinger and Lucia (1997, 10) divide these dimensions into three categories: skill, knowledge, and style, as described in Figure 8.1.

Because this tool was developed to evaluate and provide feedback to managers, many of the target dimensions are relatively universal, e.g., good managers will have to negotiate with supervisors, subordinates, peers, and outside stakeholders regardless of the industry in which they work. At the same time, not all management positions are created equally, and not all professional roles require the same volume of

Skill	Proficiency at performing a task; degree of mastery (for example, ability to think strategically, communicate in writing, delegate work, influence, negotiate, operate a machine)
Knowledge	Familiarity with a subject or discipline (for example, knowledge of an industry or business)
Style	A pattern of characteristics or ways of responding to the external environment (for example, self-confidence, energy level, self-sufficiency, emotional stability)

◆ Figure 8.1 Working Definitions for the Types of Data Collected by 360-Degree Feedback (from Lepsinger and Lucia 1997, 10)

common behaviors. Many vice presidents in banks have few or no people who report directly to them, whereas vice presidents in manufacturing companies have many. Their relative needs for negotiating with subordinates versus peers may be quite different, and these differences should be taken into account when an organization designs a 360-degree feedback tool. The skill category, in particular, may be highly domain-specific and require differentiation between industries, function areas, and even locations within a company, e.g., branch offices versus a home office.

For 360-degree feedback to work well, items included in rating questionnaires must be observable (Tornow et al. 1998, 161). Asking raters to rate performances that they have not witnessed can become frustrating and leave raters no alternative other than to infer the quality of unobserved performances based on other behaviors or, worse, on vague impressions of the subject.

Figure 8.2 is an example set of rating items (Tornow et al.1998, 79, Exhibit 4.1).

The second challenge to implementing 360-degree feedback has to do with the number of people required to participate: the subject's supervisor(s), the subject him- or herself, and at least three peers (a minimum number of peers is required in order to maintain anonymity; Scott and London 2003, 178). Because at least five people rate each subject, obtaining and collating performance ratings consumes a considerable amount of time. The raters must be identified and invited to participate; they must be trained in how to understand and complete the feedback process, and sometimes, they need to be reminded to complete the rating process, as well. Survey takers' data must then be analyzed, collated, and reported in such a way that anonymity of peers is maintained, while the ratings from different categories of raters remain transparent.

Advantages of 360-Degree Feedback

Despite the time and expense involved in implementing 360-degree feedback, there are a number of advantages to using such a process for assessing professional performance. These advantages include adjusting the attitudes of participants toward their professional roles and increasing the ease with which the assessments can be converted into constructive behavioral change.

Perhaps the most important advantage of 360-degree feedback systems, from the standpoint of foreign language educators, is that such systems help participants understand that everything they do professionally—and in a target language/culture—has end-users or "customers." Simply satisfying the boss ignores the fact that effective professional conduct requires satisfying customers inside and outside of one's place of employment. Language learners who become accustomed to the

Consider this manager's effectiveness in the following items. How satisfied are you with the way this manager…

	Highly Dissatisfied	Dissatisfied	Neither Satisfied or Dissatisfied	Satisfied	Highly Satisfied	No Information
Serving Customers						
1. Creatively responds to customers' needs	HD	D	N	S	HS	NI
2. Responds to customers' needs in a timely manner	HD	D	N	S	HS	NI
3. Works to meet commitments that have been made to customers	HD	D	N	S	HS	NI
Processes Are the Problem, Not People						
1. Involves employees in decisions that affect their work	HD	D	N	S	HS	NI
2. Asks people what they need to do their work better	HD	D	N	S	HS	NI
3. Strives for continuous improvement	HD	D	N	S	HS	NI

Reprinted from Maximizing the Value of 360-Degree Feedback: A Process for Successful Individual and Organizational Development, p. 79, Exhibit 4.1, with permission from Jossey-Bass Publishers/A Wiley Imprint. (c) 1998.

◆ Figure 8.2 Example 360-Degree Feedback Rating Items

assessments of educators can fall into the same trap as employees assessed exclusively by their bosses: they may forget that professional success depends on pleasing both the boss *and* professional peers.

The idea that numerous people already evaluate our daily performances is significant to foreign language educators because it informs the theory and design of assessment tools. If meaning-making and professional-level communication occur in a sociocultural context, then that context should be taken into account when professional-level foreign language ability is asssessed. As Rao and Rao (2005, 89) point out, 360-degree feedback tools remind people that they make an impression on others. In a cross-cultural context, the distance between the impression the nonnative is trying to leave and the impression that she or he actually leaves on native colleagues can be highlighted by the effective use of 360-degree feedback tools designed for such contexts.

The 360-degree feedback tools have an additional advantage over standard proficiency-based tests in that they collect opinions about a professional's performance from multiple types of consumers, all of whom are familiar with certain professional requirements of the test subject's job. A traditional OPI that is conducted by foreign language educators may be able to assess a nonnative speaker's *general* language and cultural knowledge, but it is limited in its ability to assess a nonnative speaker's *professional* language and cultural knowledge, unless the profession in question relates to foreign language education. The 360-degree feedback assessment format allows for the inclusion of foreign language educators' assessments, but also requires the input of individuals who are familiar with the test subject's profession. This is a crucial advantage because, by definition, "professionally useful foreign language ability" is an ability for which other professionals are willing to pay. Industries, organizations, and function areas within organizations have specific professional and cultural expectations to which their members must adhere in order to succeed; the judges of that success are other members of those groups. For example, chemists are aware when other chemists know what they are talking about; drug dealers can tell who is and is not behaving according to the standards of their community—a distinction with which undercover police must be familiar.

The assessment of professionally useful foreign language ability has been held back by the tension inherent in employing industry outsiders (foreign language educators) to evaluate the performance of foreign language speakers in a variety of professions. In recent years, proficiency-level descriptors have adopted the idea that higher levels of ability should include the ability to communicate effectively in the subject's field of expertise. However, assessing one's ability to perform in professional contexts will continue to pose a challenge as long as foreign language educators continue to conduct proficiency-based foreign language assessment interviews out of context. Take, for example, a fireman: though an ILR 4 or 5 rating requires that a fireman tailor his language to his audience, most foreign language educator OPI raters would be unable to assess the appropriateness of that language tailoring, even if the OPI were redesigned to elicit such performances. A 360-degree feedback tool allows foreign language educators to avoid pegging performance to the highly-educated native speaker ideal and increases their focus on how others in the

test subject's field feel about his or her professional performance (what we can call expertise).

Information regarding how an organization defines expertise can be communicated via assessment tools themselves. Analogously, "expert students" focus their efforts on mastering material that will be covered on the test, and the same can be said of aspiring professionals. Three hundred sixty–degree feedback can guide performance expectations in an organization (Tornow et al. 1998, 26; Lepsinger and Lucia 1997, 44) by showing employees what is on the test, so to speak. Accordingly, organizations that utilize 360-degree feedback tools often find that their employees become more sensitive to the impressions that they are leaving on all the people that they interact with, and they come to understand that their organization values these interactions.

The selection of an assessment tool tells an organization's members what the organization wants from them, as do the assessment tools adopted by end users of professional-level foreign language ability. Foreign language speakers in the American armed services are given pay increases for climbing the ILR proficiency scale, but is there a direct correlation between movement up the ILR scale and the ability to successfully accomplish military tasks? Is a tank commander who speaks Iraqi Arabic at ILR 4 necessarily worthy of pay higher than another tank commander who scores only an ILR 3 but is able to establish such a rapport with the local community that he is told ahead of time where someone has hidden an Improvised Explosive Device? Considering the fact that many, if not most, members of the Iraqi local community do not speak their own language at a level above ILR 3, expertise in accomplishing target tasks in the target culture environment may be a better rubric against which to establish bonus pay.

One can argue that proficiency scales are not designed to measure this kind of ability, and, therefore, criticism for failing in this regard is misdirected. As long as such scales remain the standards against which professional performances by foreign language speakers are held, foreign language learners will likely focus on moving up the proficiency scale rather than moving up in the hearts and minds of their peers.

Disadvantages of 360-Degree Feedback Tools

There are two main disadvantages to using 360-degree feedback tools: the cost and the validity of the text-based format on which most such tools are based. These drawbacks can be overcome in ways that make this method a robust choice for evaluating the skills of the relatively few Americans who achieve professionally useful foreign language ability.

Implementing a 360-degree feedback type of assessment tool requires a considerable amount of time and, therefore, cost. Before the evaluation process begins, an organization must identify the performances it wants to evaluate. Questions intended to reflect the test subject's ability to perform those tasks must then be designed. Also, before evaluation begins, those participating in the assessment process must receive training regarding the tool's intended purpose and benefits. After the tool is designed, all participants in the process (including supervisors, peers, and subordinates) must be trained on how to use the system to provide feedback.

Once the assessments are completed, collation and analysis begin, and when the data finally appear in a digestible format, they should be shared with the subjects in a constructive way. Because test subjects tend to selectively remember results that already fit with their self-conceptions and forget those that conflict, coaching sessions offer the most effective format for sharing evaluation results (Atwater et al. 2008, 216). Retesting should be carried out on a regular basis in order to establish which parts of the coaching produce results and to identify areas in which the professional still can improve.

Second, traditional 360-degree feedback tools tend to be text-based, which limits respondents to evaluating their recollections of the test subject's behaviors. Asking questions about behaviors previously observed means that such tools can be used only by those who have already worked with the individuals in question. As a result, questionnaire-based evaluations are of limited use to a hiring organization, as none of its own trusted employees will have had an opportunity to observe and therefore have a recollection of an applicant's previous performance.

Yet neither of these shortcomings need prevent one from using 360-degree feedback principles for evaluating professional-level foreign language speakers. Furthermore, the expense incurred in using this kind of tool is minor compared to the investment already made in developing the kinds of high-level skills such tools assess.

360-Degree Assessment for Foreign Language Speakers

The need to assess foreign language speakers who are applying for a job requires an assessment tool that allows evaluators to observe performances without meeting the test subject. The Advanced Language Performance Portfolio System (ALPPS) created at Ohio State University (OSU) provides evaluators with samples of actual performances rather than questionnaires about recalled performances. This section will examine how this tool applies the principles of 360-degree feedback to the needs of foreign language assessment.

Designed in the mid-2000s, ALPPS was initially tested with Chinese Flagship MA students at OSU. This online system stores foreign language performance samples, evaluators' assessments of those performance samples, and students' scores on standardized tests all in a portfolio created for each individual. Along with serving as a repository of such raw data, the system also produces reports that show an individual's strengths and weaknesses, charts a learner's improvement or decline over time, aggregates data, and produces reports for selected populations within the database of portfolios.

The system is designed to store video and text samples in each portfolio, thus allowing assessors to watch the videos, read the texts, and submit online evaluations for each category. Uploaded videos demonstrate speaking, listening, and behavioral skills, whereas text samples demonstrate writing and reading skills. As documents can be uploaded in Word and PDF formats, even scans of handwritten materials produced for a given work task (e.g., a memo, notes for a speech) can be included in an individual's portfolio.

In ALPPS, raters evaluate performance samples based on a mastery scale. For example, after a portfolio evaluator watches a video of an individual giving a speech in the target language, the evaluator then rates various aspects of the performance using a scale of 0 to 5 (no relationship to the ILR scale) as seen in Figure 8.3.[6]

The differences in rating levels are clearly subjective, but performance evaluation is also inherently subjective: one person's "strong" is another person's "passable." The subjectivity of interpersonal evaluation largely explains why a tool designed to elicit the opinions of native-speaking professionals needs the input of multiple raters. Expecting busy professionals who are not foreign language educators to commit to participating in training sessions that attempt to establish objective criteria for each rating level simply is not practical. Moreover, it is unlikely that such training would lead to objective evaluations.

As designed for use with Chinese Flagship students, ALPPS portfolio samples are stored under the following categories: Presentation, Conversation, Composition, Reading, Occasional/Spontaneous Events, Oral Interpretation, and Resolving Conflicts. Any organization commissioning the creation of portfolios can establish its own categorization system based on the organization's own target skills. For instance, a hospital could create categories such as Bedside Manner, Communication during Surgery, Presentations at Professional Conferences, Q&A at Professional Conferences, and Journal Articles. Although these specific categories already fall under general categories designed for Flagship students, by giving them specific professional descriptors, an organization shows participants which skills it values most.

As a web-based tool, ALPPS allows participants to evaluate portfolios from remote locations. Organizations whose employee portfolio samples may hold proprietary data can commission versions that exist only on an internal network. Built by foreign language educators for use by academic and nonacademic professional communities, ALPPS initially was designed to differentiate ratings made by the portfolio subject (nonnative language speaker), professionals in the subject's field, and foreign language educators who are native or near-native speakers of the target language. Because ALPPS can categorize raters according to the commissioning

Rubric in ALPPS English evaluation interface		**Rubric in ALPPS Chinese evaluation interface (English translations in italics added for this book)**	
5	Superior Competence/Superior Performance	5	表现非常好 *Excellent performance*
4	Strong Competence/Strong Performance	4	表现挺好 *Very good performance*
3	Competence/Passable Performance	3	表现不错 *Pretty good performance*
2	Some Competence/Some Performance	2	表现一般 *So-so performance*
1	Minimum Competence/Minimum Performance	1	表现有进步 *Performance shows improvement*
0	No Competence/No Performance	0	表现差 *Poor performance*

◆ Figure 8.3 Sample ALPPS Evaluation Rubric (McAloon 2008, 229)

organization's requirements, "native-speaking colleague" could be further subdivided into rater categories such as boss, peer, or subordinate.

Tools like ALPPS are particularly useful for *prospective* employers of foreign language speakers. The first piece of information that an employer receives about an applicant's language ability is a self-assessment on the applicant's résumé, where self-promotion may take precedence over accuracy. Some résumés include third-party ratings such as a Test of English as a Foreign Language (TOEFL) score for English or a Hanyu Shuiping Kaoshi (HSK) score for Chinese. These ratings can establish whether a job candidate meets a minimum level of ability, but they may not describe what the candidate can actually *do* in the target language. In the professional interview process, role-plays and case analyses elicit performances similar to those needed for specific jobs. When performed in a foreign language, recordings of these performances can be included in a tool such as ALPPS and reviewed by the potential employer.

Employers of foreign language speakers also can use tools such as ALPPS to provide feedback on the performance of current employees. In the 1990s, when 360-degree feedback was becoming popular, textual questionnaires about recalled performances were the only medium through which organizations could assess those performances. Today, employers can use mobile technology to conveniently record and upload performance samples to online portfolios that raters can access from anywhere in the world. As one group of HR professionals said about 360-degree feedback, "[o]ne thing that has been shown to particularly impact reliability and validity is the degree to which the measure is observable. Wherever possible, it is important that measures are based on observation rather than assumption, judgment and intuition" (Edwards, Scott, and Raju 2003, 55). Tools like ALPPS make accomplishing such a task much easier for employers.

Language instructors can likewise benefit from using multi-rater, multimedia, portfolio-based assessment tools. Currently, foreign language instructors access their students' scores indirectly after certified testers have evaluated them in a clinical environment. Assessment tools like ALPPS offer language instructors direct access both to assessed performances and to the evaluations made by native speakers who either are, or could be, the learner's professional colleagues. Insofar as they are commonly used by many language educators, portfolios in platforms such as ALPPS should still include standardized scores as reference points for understanding the learner's general language abilities.

ALPPS-like platforms also appeal to language program administrators because they offer unmediated access to performance evaluations made by native speakers in the learners' target professional domains. Such platforms can turn evaluation data, originally written in the target language, into reports in the administrator's native language. For example, if a language program designed to produce German-speaking marketers is housed in a department of European languages, a department chair can use the system to produce individual reports in English based on the evaluations of native-speaking observers. Even if the department chair cannot understand the language contained in the performance samples, rubrics written in German for German-speaking evaluators can be displayed in English when reports are produced.

Finally, foreign language learners benefit from the use of feedback tools like ALPPS. Multi-rater feedback tools effectively communicate an organization's target behaviors and provide ongoing feedback regarding the evaluation subject's improvements in those areas. These tools can motivate language learners to improve their performance in specific areas. Once foreign language learners graduate from their final formal learning institution, native speakers with whom they work will assess their language proficiency.

Foreign language learners also can benefit by reviewing samples of others' performances that are rated highly by native professionals. In addition, if the portfolio tool is used to capture samples of native-speaker professional performances (e.g., if a hospital uses an ALPPS-like system to collect samples of nursing work performances, regardless of nurses' native languages), then nonnative members of the community can use the samples to learn more native-like means of communicating with superiors, peers, and customers.

Challenges and Further Work

As designers of 360-degree feedback tools point out, creating and using an effective multi-rater assessment platform presents a number of challenges. What skills should be included for evaluation? What behaviors are most important to successful completion of the target job? Is the target job even clearly defined? Are there desired behaviors that can be generalized across numerous professions?

One challenge encountered during the initial use of ALPPS was that sometimes the most culturally appropriate reaction to a situation entails *not doing or saying anything at all*. What kind of video or text sample can reflect an appropriate *absence* of observable behavior? A similarly difficult skill to observe is the successful negotiation of a situation that occurs over the course of several interactions or during one interaction too long to reasonably ask any evaluator to observe in its entirety. Such skills still require the use of interviews to elicit impressions of the target subject because two-minute video clips and short documents cannot possibly convey the entire picture.

Another challenge involves determining how many and what type of samples sufficiently represent an individual's abilities for the kind of work she or he will be expected to do. Members of the target professional community, in conjunction with designers of the assessment system, can best answer both of these questions because establishing a workable solution requires coordination.

After it is determined which samples to include, capturing video performance samples introduces additional challenges. Individuals need to receive training in how to capture quality video samples, how to minimize the elicitation of unnatural behaviors caused by the act of video recording, how to identify which performances are portfolio-worthy and which are not, and how to edit videos down to a reasonable and watchable length.

Identifying and successfully inviting qualified assessors presents another challenge of implementing multi-rater portfolio assessment. Lepsinger and Lucia (1997, 127) believe that a professional colleague should know the assessment subject for at least four months before being allowed to participate in a 360-degree feedback

program. However, at the time Lepsinger and Lucia expressed this opinion, recording and replaying videos of professional performances was inconvenient. Now that videos are more easily shot, stored, and replayed in digital format, establishing a minimum period of familiarity for portfolio raters may no longer be needed. In fact, *not* knowing the subjects may be preferable for raters, as personal experience with them could make already subjective ratings more subjective.

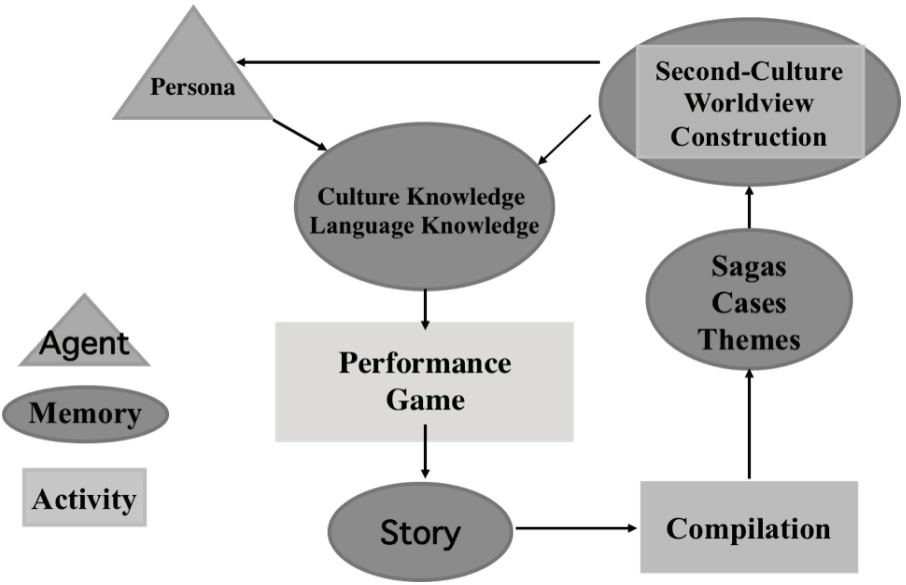
In an ideal world, at least ten evaluators would assess a portfolio evaluation subject and additional evaluations would follow for the duration of the subject's working life. In the real world, however, evaluators can experience evaluation fatigue if asked to evaluate more than three people at a time (Lepsinger and Lucia 1997, 121), and native-speaking colleagues of nonnative speakers may be in such diverse locations that contacting and training ten people to complete the process can require an enormous investment of time and money. Yet the organizations that commission 360-degree evaluations consider the process a valuable investment.

Convincing institutions to see the value of a multi-rater performance evaluation poses a further challenge. If users of the system do not see the value of the tool, then their participation will be half-hearted and will limit the utility of the portfolios. Regardless of how valuable the data from such tools might be, if decision makers and check writers find their current choice of evaluations is satisfactory, then persuading organizations to implement tools like ALPPS will prove difficult.

Conclusion

Multi-rater portfolio assessment provides a practical reminder to those involved in high-level foreign language education that the end users of foreign language ability are actually the target language native speakers whom the learners must impress with their expertise. If foreign language learners hope to achieve their and their organization's goals in the target culture, they must adapt their messages to the native audience. 360-degree assessment tools inform foreign language learners how their audience perceives their professional-level performances. By providing industry professionals direct access to video and textual samples of learners' performances of domain-specific skills, portfolio-based 360-degree assessment tools effectively determine how well foreign language learners can complete job-specific tasks in the target language that may or may not correlate with the descriptors on the ACTFL/ILR proficiency scales.

Three hundred sixty–degree multi-rater portfolio assessment tools also allow organizations to identify and reinforce specific sets of behaviors that they believe are necessary for success in the target industry or organization. For example, a portfolio system commissioned by an advertising agency to assess account managers could be set up to collect samples of pitches for advertising campaigns. For foreign language speakers operating at high levels, recognizing and internalizing target culture behaviors and worldviews play a key role in their success. Walker and Noda (2000, 32) describe a cycle of compilation in which foreign language learners turn personal observations and experiences in the target culture into behaviors that are adapted for that target culture, as shown in Figure 8.4.[7]



By Mari Noda and Galal Walker. Reprinted by permission of the authors.

◆ Figure 8.4 Walker and Noda's Cycle of Compilation

Arguably, the most effective way to evaluate a learner's ability to construct a second-culture worldview and translate it into a culturally appropriate action requires observing the action itself. With portfolios that hold samples of performances recorded at different times, systems like ALPPS can generate reports that show how portfolio subjects have progressed in producing performances that align with native audiences' expectations.

A further advantage of the transparency of multi-rater portfolio systems is that ratings can be isolated by rater identity. Portfolio evaluation reports can compare ratings made by industry insiders, such as peers and superiors, against those of industry outsiders, such as language instructors. Each of these comparisons can provide valuable data to program administrators, instructors, employing organizations, and the portfolio subjects themselves, who generally want to leave a good impression. If a portfolio system's reports show that instructors are consistently giving learners higher ratings than potential or actual professional colleagues, the instructors can adjust their program of instruction accordingly, e.g., the addition of domain-specific field experience versus additional class work under the direction of language instructors.

Looking Ahead

Particularly among the less commonly taught languages, the American foreign language education community has largely left the achievement of professionally

useful levels of foreign language ability up to individual learners after they graduate. Likewise, responsibility for assessing the uppermost levels of foreign language ability and their specific professional applications have been left to organizations possessing a critical mass of foreign language evaluators, such as the Department of Defense. Despite the concrete need for a means of assessing professional-level foreign language ability, the proficiency movement has had difficulty ameliorating this need with its existing focus on evaluating general language ability at all levels of application and on using an educated native speaker ideal. Even the question of how to evaluate Distinguished or Superior ability has fixed the discussion squarely in terms of "proficiency," which may have prevented educators from looking outside the foreign language community for inspiration.

Organizations that have invested in producing high-value foreign language speakers, such as the Language Flagship and multinational corporations, can build on the experience of the 360-degree feedback practitioner community to perfect similar tools for assessing performance in foreign language environments. The fact that the behaviors of highly trained people have ripple effects throughout the organization makes the high cost of multimedia portfolio assessment a reasonable investment.

Foreign language educators and end users of foreign language ability need not settle for proxies when seeking to establish individuals' ability to do job-related tasks in a target language environment. The manager at the beginning of this chapter should someday be able to log on to her ALPPS account and pull up the portfolios of everyone applying for the position. She will read a report in English telling her what native speakers thought of the candidates' job performance in the target language, and she will be able to see for herself how the candidates appear when working in the language, even if she does not understand them. The person she hires may be responsible for a billion-dollar project, and she will not want it to fail because of cross-cultural miscommunication issues. Thankfully, her company will have invested in creating and evaluating portfolios for all the job candidates. Now, she will know who demonstrates both language proficiency and *expertise*.

Notes

1. "ACTFL Proficiency Guidelines 2012," American Council on the Teaching of Foreign Languages, accessed June 23, 2013, http://actflproficiencyguidelines2012.org/.
2. An elegant summary of how the term "performance" is used in a number of fields relevant to foreign language educators can be found in Zeng (2011, 5).
3. "ACTFL Proficiency Guidelines, Speaking," American Council on the Teaching of Foreign Languages, accessed June 23, 2013, http://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012/english/speaking
4. "FAQs about the DLPT V," *United States Coast Guard*, accessed June 23 2013, http://www.uscg.mil/hr/cgi/eso/courses_and_tests/dlpt/FAQ_DLPT.pdf.
5. According to one Army linguist, much of the oral interpretation in the Middle East is done by private contractors, whereas those in the Army who use their language skills do so from bases in the United States and translate target language material. http://www.wired.com/dangerroom/2011/08/lost-in-translation-how-the-army-wastes-linguists-like-me/.

6. This rubric was taken from a 2006 iteration of ALPPS. Notice that the Chinese and English rubrics do not exactly match, which would be a cause for concern when one is comparing evaluations submitted by Chinese speakers and English speakers, i.e., nonnative speakers of Chinese.
7. Since the original publication of the chart, Walker and Noda have since added "themes" to cases and sagas.

References

Atwater, Leanne E., Joan F. Brett, and James W. Smither. 2008. "What Do Leaders Recall about Their Multisource Feedback?" *Journal of Leadership & Organizational Studies* 14 (3): 202–18.

Edwards, Jack, John Scott, and Nambury Raju. 2003. *The Human Resources Program Evaluation Handbook*. Thousand Oaks, CA: SAGE Publications.

Kang, Sahie, Si Yen Lee, and Hee-Sun Kim. 2007. "OPI Familiarization, Facts or Myths?" Presentation at the Annual Workshop and Conference of the American Association of Teachers of Korean, Chicago, IL, June 2007.

Lepsinger, Richard, and Antoinette Lucia. 1997. *The Art and Science of 360° Feedback*. San Francisco: Pfeiffer.

McAloon, Patrick. 2008. "Chinese at Work: Evaluating Advanced Language Use in China-Related Careers." PhD diss., Ohio State University, 2008. Retrieved from Ohio LINK ETD Center. Document number: osu1218548897.

National Council of State Supervisors for Languages. 2009. "Interpersonal Communication Self Checklist," last modified October 2009. Accessed February 6, 2014. http://www.ncssfl.org/docs/InterpersonalPerson2PersonNCSSFLTemplate_October2009FINALPROTECTED.doc.

Oskarsson, Mats. 1984. *Self-Assessment of Foreign Language Skills: A Survey of Research and Development Work*. Strasbourg, France: Council for Cultural Cooperation.

Rao, T.V., and Raju Rao. 2005. *The Power of 360-Degree Feedback*. Thousand Oaks, CA: Response Books.

Ross, John A. 2006. "The Reliability, Validity, and Utility of Self-Assessment." *Practical Assessment, Research & Evaluation* 11 (10): 1–13.

Saito, Yoko. 2003. "The Use of Self-Assessment in Second Language Assessment." *Working Papers in TESOL & Applied Linguistics* 3 (1). http://www.tc.columbia.edu/academic/tesol/WJFiles/pdf/Saito_Forum.pdf.

Scott, John C., and Manuel London. 2003. "The Evaluation of 360-Degree Evaluation Programs." In *The Human Resources Program Evaluation Handbook*, edited by Jack Edwards, John Scott, and Nambury Raju, 177–199. Thousand Oaks, CA: Sage Publications.

Shakespeare, William. 1904. *As You Like It*, edited by Cyrus Lauron Hooper. Chicago: Ainsworth & Company. http://books.google.com/books?id=P8Q5AQAAMAAJ&pg=PA62&dq=as+you+like+it&hl=en&sa=X&ei=RAP0UqPOK4HkyAGy40H4Cw&ved=0CFAQ6wEwBA#v=onepage&q=as%20you%20like%20it&f=false.

Sonnentag, Sabine, and Michael Frese. 2005. "Performance Concepts and Performance Theory." In *Psychological Management of Individual Performance*, edited by Sabine Sonnentag, 1–25. Chichester, UK: John Wiley & Sons, Ltd.

Tornow, Walter, Manuel London, and CCL Associates. 1998. *Maximizing the Value of 360-Degree Feedback: A Process for Successful Individual and Organizational Development*. San Francisco: Jossy-Bass Publishers.

Walker, Galal, and Mari Noda. 2000. "Remembering the Future: Compiling Knowledge of Another Culture." In *Reflecting on the Past to Shape the Future*, edited by Diane W. Birckbichler and Robert M. Terry, 187–212. Lincolnwood, IL: National Textbook Company.

Zeng, Zhini. "Second-Culture Worldview Construction: Integrating In-Class and Out-of-Class Activity in in-China Study." MA thesis, Ohio State University, 2011. Retrieved from Ohio LINK ETD Center. Document number: osu1308270341.

This page intentionally left blank

Chapter 9

Professional Language Skills: Unprecedented Demand and Supply

RICHARD D. BRECHT, WILLIAM P. RIVERS, JOHN P. ROBINSON, AND DAN E. DAVIDSON

THIS CHAPTER PRESENTS AN overview of the current state of language education at the advanced levels in the United States, starting with a brief assessment of the need for language skills, driven by the decade-long increase in demand in the public and private sectors for language professionals—individuals with the ability to perform as professionals in a second language across a wide range of disciplines. As perspective to meeting this demand, we provide evidence for the traditional supply system—the language proficiency output of higher education over the past few decades—which has been a product of two traditional pathways to achieving advanced-level skills: intensive instruction, with some in-country immersion, and very modest extended sequences built upon a weak K–12 base. However, we then provide evidence for current and projected extraordinary enhancements of these pathways, which derive from current best practices in classroom and immersion programs, as well as major advances in science and technology. While still some distance from fulfilling the promise of universally available, efficient, and effective programming for professional language skills, this evidence lays out common parameters for program design that can be used in settings across industry, government, heritage communities, and education to dramatically improve proficiency outcomes. Important for policymakers, the second pathway of long sequence is supported by broad popular opinion in favor of language study in primary and secondary schools.[1]

The Enhanced Need for High-Level Linguistic and Cultural Skills

In the United States, the traditional rationale for federal support of language education and training has focused on national security, beginning with the Second World

War and the subsequent National Defense Education Act of 1958 (Brecht and Rivers 2012). To this end, the US government has built an extensive recruitment and language training system for defense, diplomacy, and intelligence needs, with announced goals of producing graduates with professional levels of language proficiency (at or above the 2+ level on the Interagency Language Roundtable [ILR] Proficiency Guidelines). While a series of Government Accountability Office (GAO) reports and congressional hearings cite linguistic deficiencies across the government (see, for example, United States Government Accountability Office 2010), there is little reason to expect this focus on high-level skills to change for the foreseeable future. The US government has well-documented needs across more than eighty agencies and for an equally large number of languages. The general requirements are at professional levels of proficiency in two or more languages (English and one other), as well as additional mission- or job-specific skills.

Industry Needs

What has changed dramatically in the last decade or so is the fact that the language industry, and the major sectors of the US economy that it serves, faces a major talent gap. The language industry continues to grow at 8–10 percent per year, some three-to-four-times faster than the overall US economy, and is worth at least $15 billion per year (Kelly and Stewart 2011). Industry experts and observers expect this growth to continue, if not accelerate, due to the explosion in content, particularly from social media and the use thereof by major industry clients. Accordingly, there is an intense need for skilled professionals to meet this burgeoning demand for multilingual, multimodal, multidirectional communications. Among the language professionals in demand are translators and interpreters, who must possess professional skill levels in at least two languages and must also be adept at using the modern technology that these professions now require. Skilled professionals are equally needed and are scarce in the supporting business disciplines of the industries, including information technology, project and program management, customer relations, and business operations. These personnel all require high degrees of linguistic and cultural skills as well as other critical technical and business skills.

Foreign Languages (FL) and STEM (Science, Technology, Engineering, and Math)

Of late, FL advocates have noted that the intersection between FL and STEM has attracted significant attention across the industry and government, in particular in the White House Office of Science and Technology Policy. According to the Joint National Committee for Languages and the Globalization and Localization Association (2013), language is increasingly intertwined with the information economy and the research and development that support it. National educational policy has focused on increasing the STEM capacity of the educational system and increasing and retaining STEM graduates in the United States. The argument advanced by advocates is essentially that language has, for some time, been part of STEM, or at

least intersects with STEM disciplines in novel ways. Among these intersections are the increasing use of blended learning integrated into FL programs; the high degree of technology use in translation, which as an industry relies on terminological databases, mixed stochastic and rule-governed machine translation, and crowd-sourced translations as enablers and adjuncts to human translators; and web-based conference interpreting.

Finally, because of the global nature of the language industry, individuals with these skill sets are in high demand worldwide and are not tied to any one locale. As a result, the US language industry and its customers face fierce global competition for the available talent. As the language industry continues to grow at an 8–10 percent annual rate (well ahead of the growth of the US economy and the US workforce), so does the concern that the US education system is unable to develop and sustain capacity in foreign languages from kindergarten through college. To address this, the Globalization and Localization Association established the industry wide Task Force on Global Talent, with representation from major language-service providers and customers such as Google, eBay, Manpower, and others.

Current Baselines of Supply

The last comprehensive, nationally representative study of language proficiency outcomes was conducted in the mid-sixties by John Carroll (1967), using the Modern Language Association Cooperative tests. No national study at any level of the educational system has been conducted since. The United States stands apart from industrialized nations in this respect, as the European Union (2012), among others, has far more recently assessed the outcomes of FL education. These assessments receive governmental support, whereas even the sole postwar study of FL outcomes in the United States did not merit such support, as it was funded by the Ford Foundation.

Nevertheless, we do find indicators of outcomes of FL education programs. While a small number of institutions test their graduates in a limited range of languages, such as George Washington University and Brigham Young University, there is no collation of such data sources that would allow inferences on outputs of the system. One reliable source of information (in statistical terms) is the data collected by the American Council of Teachers of Russian (ACTR), which has administered pre- and post-program proficiency tests to more than five thousand sojourners since 1976. Accepting that these individuals represent a self-selected sample of FL learners in that they chose to apply for study abroad and were presumably more motivated than the average Russian student (and, perhaps, more proficient, although that hypothesis remains untested), we can view these results at least as an indicator of the general level of proficiency obtainable from FL education at a US institution of higher education prior to study abroad. Moreover, as the data collected by ACTR include the number of years of study, we can gauge an expected proficiency outcome from four years of Russian study. The results of nearly forty years of data collection indicate a remarkably stable, and disappointingly low, median proficiency of Intermediate High, roughly equivalent to an ILR rating of 1+, across listening,

reading, and speaking (Davidson 2010). Here are the results of data collected for the years 1994–2009:

- Speaking score after a minimum of six semesters of study or equivalent: Mean is 4.58 (standard deviation [SD 1.55]) for semester students; 4.78 (SD 1.70) for academic year students, where *5.0 = Intermediate Low*.
- Reading score after a minimum of six semesters of study or equivalent: Mean is 7.03 (SD 2.27) for semester students; 7.43 (SD 2.46) for academic year, where *7.0 = Intermediate High*.
- Listening score after a minimum of six semesters of study or equivalent: Mean is 5.72 (SD 1.57) for semester students; 5.80 (SD 1.47) for academic year, where *6.0 = Intermediate Mid*.

Without study abroad, US Russian programs have consistently produced Intermediate-level speakers, readers, and listeners in the target language over the fifteen-year period of 1994–2009 (Davidson 2010).

Watson, Siska, and Wolfel (2013) confirmed this indicator in a study of outcomes from the United States Military Academy (West Point). Using the Defense Language Proficiency Test for Reading and Listening and the American Council on the Teaching of Foreign Languages Oral Proficiency Interview to measure proficiency before and after study abroad for third- and fourth-year students at West Point in seven languages (Arabic, Chinese, French, German, Portuguese, Russian, and Spanish), the baselines proved consistent with the ACTR studies. Language programs at West Point yield a median proficiency of Intermediate High in Watson et al.'s study. From the limited number of available studies, we can tentatively conclude that a FL curriculum spanning four years yields a median proficiency of Intermediate High.

The New Generation of Professional Language Skills

We turn now to promising programmatic innovations that offer significant affordances for student outcomes near or at the professional level of proficiency (American Council on the Teaching of Foreign Languages [ACTFL] Superior, ILR 3 and above). These innovations have resulted in vast improvements of the two pathways to such high-level language proficiency: more efficient extended learning sequences beginning at an early age and high-quality intensive instruction, with more effective immersion, whether begun in early school years or later.

Pathway: Extended Sequences

We start with a pyramidal sketch of components of extended sequences in language education in the United States, shown in Figure 9.1:

Our contention is that the only way to ensure an adequate and constant supply of professional and expert language capabilities is to have policies that essentially broaden the base and sharpen the points of this pyramid. To do this, policies must ensure that all students in elementary school are exposed to global perspectives coupled with an understanding of the roles of language abilities and cultural skills, while

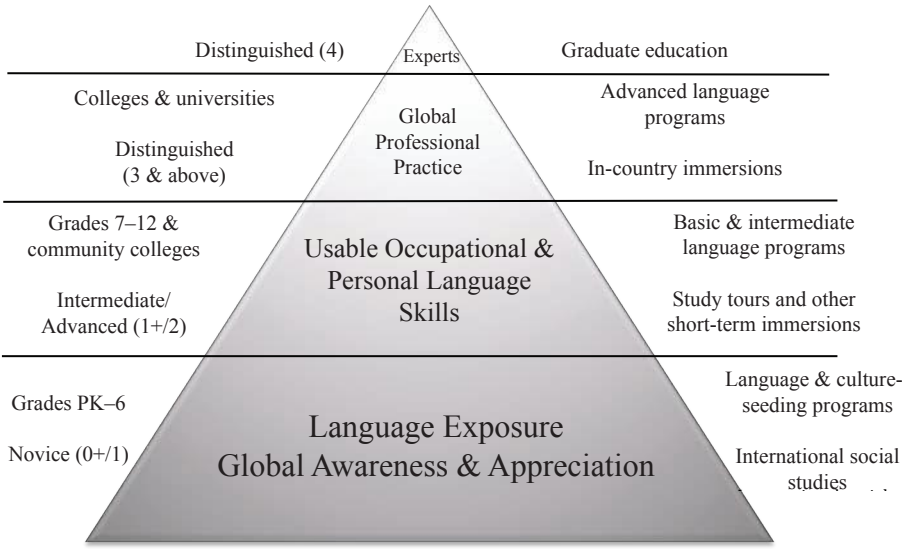


Figure 9.1 Language Capacity for Plurilingual Citizens: The Pyramidal Base

a significant number of students receive basic communication skills in some foreign language through Foreign Language Elementary School (FLES) or immersion programming. Following this initial exposure, secondary schools would be responsible for providing a majority of their learners with usable language skills. Colleges and universities would then be in a position to graduate global professionals who are able to practice across the international environment with adequate language skills. Finally, a combination of undergraduate and graduate programs in language and literature would produce the expertise that would enable practice at all the aforementioned levels. That this language-education system is possible depends on many factors, not the least of which is broad-based support from parents and other constituents of the voting public, as discussed previously.

Pathway: High-Quality Intensive Instruction

The characteristics of high-quality intensive instruction are a staple of higher education and include a six-credit course that meets five days a week, with practice sessions and extramural cultural programs. It may include summer intensive programs, but inevitably the program assumes one or more in-country programs for a summer, semester, or even an entire academic year. In some cases, the in-country program may entail internships or direct enrollment in an institution of higher learning in the host country. Such programs may build upon abilities acquired in secondary school or experience in the target culture, or they may, with less probability of engendering professional skills, attempt to take learners to high levels starting with a beginning language course in freshman year.

Both pathways are now producing more highly proficient graduates than they have in the past, as recent developments have made these pathways much more effective

and efficient. These developments, which we will call "enablers," include (1) broad public support for language education across the United States, (2) research in cognitive science and second language acquisition, (3) language learning technology, (4) best practices in primary and secondary FL education, and (5) best practices in higher education.

Enabler 1: Broad Public Support

Let us start with the game-changer in extended sequence: public support for language learning at the K–12 level. Without question, the extended sequence depends on a strong demand for language education for children in this country. We see this growing demand manifest in the number of states providing secondary school credit for language competence, regardless of how that language may have been acquired, as well as in the spread of programs across the states for a Seal of Biliteracy on high school diplomas. Most importantly, we see this in the attitudes of parents across the country supporting FL education for their children before they graduate from high school.

This grassroots support for FL education seems to have sparked a generational shift toward broad support for FL education. Public opinion with respect to language and international education may be ahead of the general position of educational leadership, and of the language field itself. This very tentative conclusion is based on a series of surveys conducted in 2000, 2008, and 2013. The 2013 survey was commissioned by the "Languages for All" effort and conducted by the Joint National Committee for Languages. Overall, the data indicate continued support for FL education, with a sustained increase in support among younger respondents in the last five years.

The seven FL policy questions shown in Table 9.1 were first developed in the 2000 General Social Survey (GSS). The 2000 GSS is an in-home, 90-minute personal survey that has been conducted at one- to two-year intervals since 1972 by the National Opinion Research Center (NORC) at the University of Chicago. These items were included in a 2008 national election survey conducted by the Public Opinion Research Laboratory at the University of North Florida (UNF) in October–November of 2008 and were repeated again in the summer of 2013 through a commercial survey firm, Knowledge Networks.

Robinson, Rivers, and Brecht (2006) used the 2000 GSS data to derive some new perspectives and insights about the state of FL policy opinions in the public. These include the significant age as well as education differences in FL opinions, their interconnection with a respondent's ability to speak a FL, and the two different and distinct dimensions that underlie them—namely, support for "English-only" policies (items 1, 2b, 2c, 2e, and 2f in Table 9.1) and support for secondary-school students taking FL courses (items 2a and 2d).

Table 9.1 shows the trends over the past thirteen years for each of the seven language policy attitude questions.

As can be seen in Table 9.1, there was a wide variation in the support for these policy positions across the thirteen years covering them. Figure 9.2 provides this information graphically.

◆ Table 9.1 2000–2013 Changes in Foreign Language Policy Questions

1. Do you favor a law making English the official language of the United States, or do you oppose such a law?

2000	Favor: 78%	Oppose: 22%
2008	Favor: 72%	Oppose: 28%
2013	Favor: 81%	Oppose: 19%

2. Please tell us whether you strongly agree, agree, disagree, or strongly disagree with each of these statements:

a. Children in the United States should learn a second language fluently before they finish high school.

2000	Strongly Agree: 27%	Agree: 49%	Disagree: 22%	Strongly Disagree: 3%
2008	Strongly Agree: 40%	Agree: 40%	Disagree: 15%	Strongly Disagree: 6%
2013	Strongly Agree: 17%	Agree: 49%	Disagree: 30%	Strongly Disagree: 4%

b. Bilingual education programs should be eliminated in American public schools.

2000	Strongly Agree: 6%	Agree: 16%	Disagree: 50%	Strongly Disagree: 28%
2008	Strongly Agree: 10%	Agree: 13%	Disagree: 41%	Strongly Disagree: 36%
2013	Strongly Agree: 10%	Agree: 20%	Disagree: 49%	Strongly Disagree: 21%

c. Speaking English as the common national language is what unites all Americans.

2000	Strongly Agree: 26%	Agree: 50%	Disagree: 21%	Strongly Disagree: 3%
2008	Strongly Agree: 39%	Agree: 38%	Disagree: 17%	Strongly Disagree: 6%
2013	Strongly Agree: 36%	Agree: 44%	Disagree: 16%	Strongly Disagree: 46%

d. Learning a foreign language is as valuable as learning math and science in school.

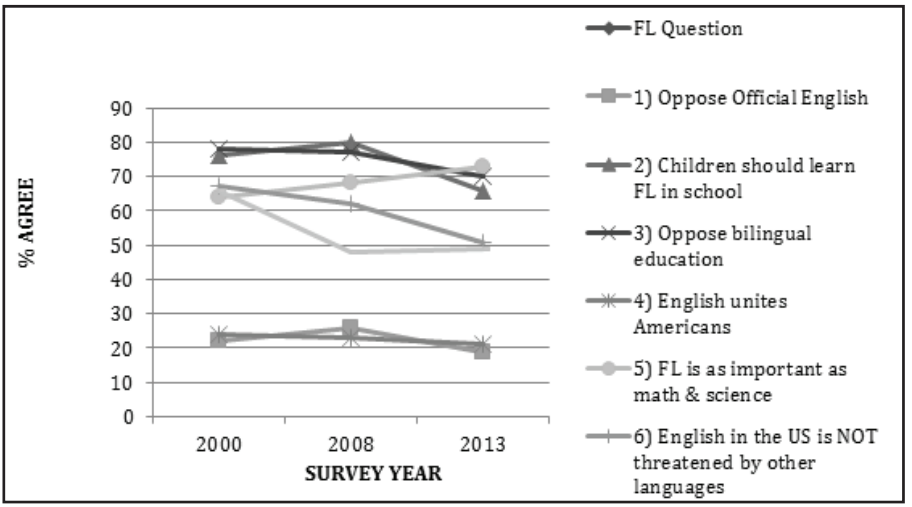
2000	Strongly Agree: 21%	Agree: 43%	Disagree: 31%	Strongly Disagree: 5%
2008	Strongly Agree: 32%	Agree: 36%	Disagree: 24%	Strongly Disagree: 8%
2013	Strongly Agree: 24%	Agree: 47%	Disagree: 25%	Strongly Disagree: 4%

e. English will be threatened if other languages are frequently used in large immigrant communities in the United States.

2000	Strongly Agree: 9%	Agree: 24%	Disagree: 51%	Strongly Disagree: 16%
2008	Strongly Agree: 16%	Agree: 22%	Disagree: 40%	Strongly Disagree: 22%
2013	Strongly Agree: 17%	Agree: 31%	Disagree: 39%	Strongly Disagree: 13%

f. Election ballots should be printed in other languages in areas where lots of people don't speak English.

2000	Strongly Agree: 17%	Agree: 49%	Disagree: 22%	Strongly Disagree: 12%
2008	Strongly Agree: 18%	Agree: 30%	Disagree: 27%	Strongly Disagree: 25%
2013	Strongly Agree: 11%	Agree: 34%	Disagree: 31%	Strongly Disagree: 24%



◆ Figure 9.2 Public Opinion Toward Foreign Language, 2000–2013

Combining the "strongly" and "less strongly" positions from the responses in Table 9.1, close to or more than 70 percent agreed in all three years that English should be the official language of the United States (Q1), that high school students should become fluent in a foreign language (Q2b), and that FL has the same educational value as math or science (Q2d); some 75 percent also agreed that English unites Americans (Q2c). In contrast, less than 25 percent agreed that bilingual education should be eliminated (Q2a), and 32–36 percent agreed that immigrant use of foreign languages posed a threat to English (Q2e). The overall import of these data is that public opinion supports language instruction, and that support has in general held steady for the past decade. With respect to public attitudes toward official language policy, support for foreign language use and instruction does not contradict a general recognition among the public that English is the *de facto* official language of the United States. Moreover, our results consistently show that, while public opinion favors making English the *de jure* official language of the United States, the use of other languages is not seen as a threat to English.

In the summer of 2013, a coalition of language organizations coordinated by the Center for Advanced Study of Language at the University of Maryland produced a metastudy entitled *Languages for All? The Anglophone Challenge* (Brecht et al. 2013), which addressed Enablers 2–5, listed here. The basic trends and conclusions cited in this study are summarized here.

Enabler 2: Research in Cognitive Science and Second Language Acquisition

An abundant amount of research demonstrates the cognitive advantages of *the use* of two or more languages, regardless of how, where, or when acquired, for instance, during early childhood or later in life (see, for example, Council on Foreign Relations 2012 and National Academies 2007). The underlying hypothesis, demonstrated by

functional magnetic resonance imaging experiments (fMRI), event-related potential experiments (ERP), and other advanced cognitive psychology and neurophysiology work, is that the use of more than one language creates stronger and more robust functional connectivity in networks of brain regions important for language learning and its cognitive control. Among the demonstrated benefits are striking findings, such as the delay of the onset of dementia symptoms, better financial choices, better outcomes in subsequent language learning, and better educational outcomes,[2] among other recent discoveries. While the cognitive and educational benefits of language use are better understood than before, this research has yet to translate into programmatic or methodological changes in language learning.

A major advance in the field of second language acquisition has been the recognition that, like other complex cognitive abilities such as mathematics, language is best learned experientially ("learning by doing") rather than didactically ("language as object"). Just as with math and science, the shift has moved away from memorization of bodies of content toward intensive cognitive engagement (e.g., Task-Based Language Learning, Language for Specific Purposes, Language Immersion).

Enabler 3: Language Learning Technology

Through the Internet, the student has nearly unlimited opportunities for time-on-task in a growing variety of activities and modalities, from traditional (and somewhat deprecated, as seen above) drills and memorization, through exposure to authentic materials, to live interactions with native speakers via web-enabled video technology. Assuming connectivity, the costs of delivery are low to nonexistent. Moreover, the US government has produced a significant number of resources that are available free of charge. The Defense Language Institute Foreign Language Center's (DLIFLC) website at www.dliflc.edu has numerous resources for the beginning and advanced language learner. The DLIFLC Headstart program covers twenty-two languages, providing listening, reading, speaking, and writing experiences to the learner. Additionally, more than six thousand learning objects to practice listening and reading are available at no cost in the Global Language Online Support System (GLOSS). Telephone conversations among native speakers are available in Arabic and Spanish. Language Survival Kits and Cultural Orientations are available at no cost in nearly all major world languages. SCOLA (Satellite Communications for Learning, www.scola.org) provides broadcasts of more than 170 languages at low cost via the web, along with thousands of learning objects based on these authentic materials. The industry also makes available at no or low cost a growing number of language-learning tools and programs that are currently used by thousands of learners, and this access is only at the early stages of development. The opportunity to be on-task is ubiquitous; the task of instruction is to make the opportunities fruitful, and the task of program and curricular design is to integrate blended learning in the curriculum in meaningful and pedagogically appropriate ways while not overburdening the classroom instructor. Language-learning technology provides affordances in terms of enhancing current practice, as well as extending classrooms virtually, thus enabling far more Americans to learn another language.

Enabler 4: Best Practices in Primary and Secondary FL Education

Needless to say, the base of the pyramid described above represents kindergarten, elementary, and secondary schools across the country. The Pre-K–12 language teaching and learning sector is experiencing a period of notable increase in effective program models and classroom practices, encouraging research results, and broadening innovation. One of the most effective modes of providing usable language skills at the Pre-K–12 level is language-immersion programming. Successful exemplars of dual-language immersion programs beginning at the elementary level with articulated sequencing through grade 12 can be found throughout the United States.

A number of states, including California, Florida, Illinois, Massachusetts, New York, Texas, and Florida, have recently launched a new initiative that includes the "Seal of Biliteracy," which is affixed to high school diplomas. As has been mentioned, many states now award secondary school credit for language competence, no matter how the competence is acquired. This kind of program indicates market forces at work involving, to this point, unrecognized language assets in the country. It also points to the fact that, in recognition of the difficulty of making room for language in K–8 programs, parents are finding other ways to move their children along the path of language learning.

Enabler 5: Best Practices in Higher Education

The Language Flagship Program, funded by the Defense Language and National Security Education Office, is the most visible embodiment of the new trend of sustained and intensive language programming. On twenty-five campuses across the United States, primarily public institutions, this relatively inexpensive federal initiative has demonstrated how the combination of innovative curricular offerings, smart technologies, overseas immersion programming, external accountability for results, and raised expectations for students, faculty, and administrators can produce professional-level speakers (ILR/Speaking, Reading, and Listening level 3 and above) across ten critical languages.

Over the past three academic years, 107 undergraduate Flagship students have completed the overseas capstone programs in Arabic (75 students) and Russian (32 students).[3] Sixty-six percent of these students (71 students) achieved an oral proficiency score of ILR 3 or higher upon completion of the overseas program. The percentage of Arabic and Russian Flagship students over the last three academic years who achieved ILR 3/3/3 (ILR 3 in speaking, reading, and listening) or higher is 45 percent (48 students), and the percentage who achieved 2+/2+/2+ or higher in each modality is 76 percent (81 students). For the most recent 2012–2013 capstone year, the Arabic program results were: 89 percent (25 out of 28 students) achieved ILR 3 in oral proficiency, and 43 percent (12 students) achieved ILR 3/3/3 or above. For the most recent 2012–2013 Russian capstone year, 94 percent (15 out of 16 students) achieved ILR 3/3/3 or above.[4] The Russian data are presented in Figure 9.3.

The Flagship and similar strong FL programs have laid to rest the long-standing myth that FL study in US higher education cannot produce Professional-level literacy in a foreign language. The data on Flagship show that the opposite is true, that

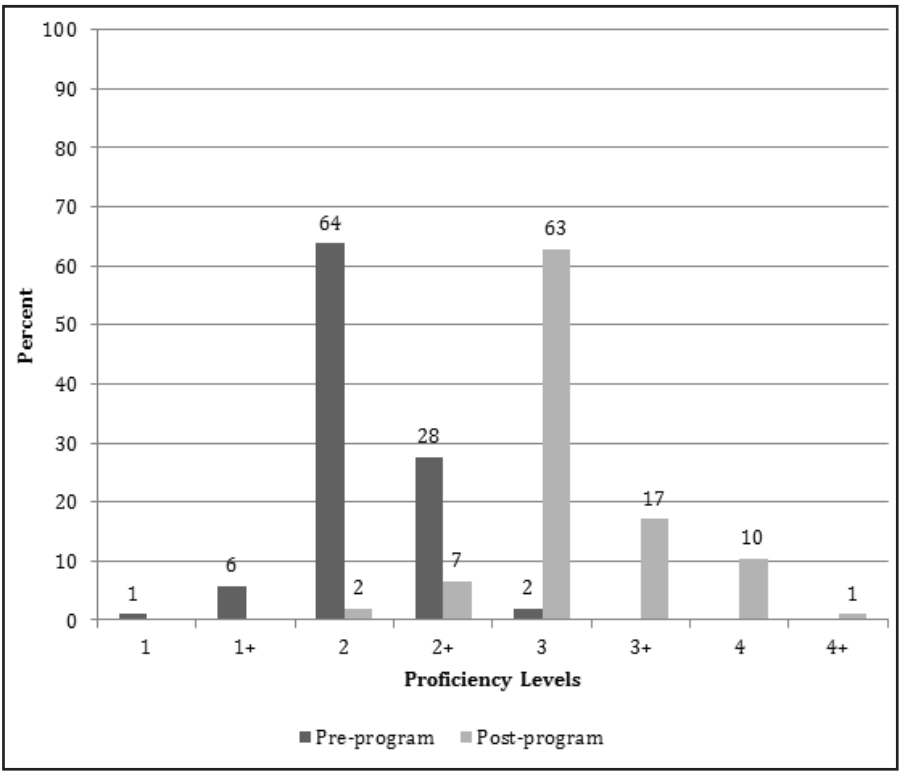


Figure 9.3 Russian Overseas Flagship Program Pre- and Post-Program OPI Scores for 2004–05 through 2012–13 Academic Year Program (N = 105)

the curricular resources are available, and that high-level outcomes are possible not just for a narrow group of linguists but for students in any field of study.

Finally, there is a budding movement to diffuse cutting-edge innovation in higher-education language learning, which is known as the Partnership for Language in the United States (PLUS). More than fifty university programs met in October 2013 to discuss the formation of a coalition whose members will be distinguished by high-level and rigorous standards ensured by peer evaluation. This may be a sign of movement outside of the federally funded Flagship programs for dramatic change in language learning at the higher-education level.

Outside the Formal Education System: The Government

Evidence for the success of language programming outside of the formal education system is sparse, as such efforts are usually not measured by standard metrics. The exception to this is the government language education system, which encompasses the defense, diplomacy, and intelligence spheres. For example, the DLIFLC,

the primary source of language training for the US military, consistently and rigorously documents student outcomes across almost two dozen languages, including the hardest for English speakers to learn, such as Arabic, Korean, Chinese, and Japanese. From 2007 to 2012, the DLIFLC has contributed to the field 9,000+ 2/2/1+ qualified linguists, 2,884 2+/2+/2 qualified linguists, and 950 3/3/2 qualified linguists from the Basic Course Program. The measures for success in the last two years include not only 2/2/1+ proficiency. The data below show the percentage of students from the entering classes who graduate with 2/2/1+ proficiency on the Defense Language Proficiency Test in listening, reading, and speaking, regardless of attrition:

Results FY12:	Proficiency: 77.3%	Production: 61.0%
Results FY13:	Proficiency: 86.4%	Production: 66.1%

These data demonstrate that a disciplined and accountable program can reliably take high school graduates to significant proficiency levels in specified periods of instruction. To be sure, this success is due to a range of factors, including low student-to-faculty ratios, superior technology (see Enabler 3 above), most students being selected by language aptitude, native-speaking teachers, constant formative assessment and mentoring, and intensive course work. On the other hand, the institute works mostly with high school graduates and includes military-dictated scheduling, rare languages, and a stressful environment for learners who must succeed or be reassigned or released from service. In sum, we take these data as evidence that language training has grown in effectiveness, utilizing all emerging advantages, in spite of significant challenges.

Conclusion

Language education in the United States has undergone significant improvement over the past two decades, making the attainment of high-level Professional language skills more possible and certainly more accessible. With increasing public support and advances in science, technology, and practice, the future looks bright—that is, if policymakers, leaders, and managers take full advantage of these enabling factors. Finding common ground among the five sectors of the "Language Enterprise" (education, government, industry, heritage communities, and overseas partners) in vision, message, and strategy for moving forward represents an important step in realizing such language skills. This joint effort on behalf of language has a new force behind it: the growing demand on the part of industry for language-enabled global professionals.

Notes

1. The authors acknowledge contributions by the participants in the "Languages for All? The Anglophone Challenge" effort, including the White Paper and the International Forum held at the University of Maryland on September 30, 2013.

2. On delaying dementia symptoms, see Alladi et al. 2013; on financial decisions, see Boaz, Haykawa, and An 2012; on third language acquisition, see Rivers and Golonka 2010 and the references cited there; on closing the achievement gap, see Collier and Smith 2012.
3. We cite preliminary results from the Flagship program when only Russian and Chinese had sufficient sample sizes for meaningful reporting; additional data are represented in Davidson, this volume.
4. For more detailed data on the immersion component of the Flagship program, see Davidson 2010 and this volume.

References

Alladi, Suvarna, Thomas H. Bak, Vasanta Duggirala, Bapiraju Surampudi, Mekala Shailaja, Anuj Kumar Shukla, Jaydip Ray Chaudhuri, and Subhash Kaul. 2013. "Bilingualism Delays Age at Onset of Dementia, Independent of Education and Immigration Status." *Neurology* 81 (22): 1938–44.

Boaz, Keysar, Sauyri L. Haykawa, and Sun Gyu An. 2012. "The Foreign-Language Effect: Thinking in a Foreign Tongue Reduces Decision Biases." *Psychological Science* 23 (6): 661–68.

Brecht, Richard, and William P. Rivers. 2012. "Language Policy in Defense and Attack." In *Handbook of Language Policy*, edited by Bernard Spolsky, 262–277. Cambridge: Cambridge University Press.

Brecht, Richard, with Martha G. Abbott, Dan Davidson, William P. Rivers, John Robinson, Robert O. Slater, Amy Weinberg, and Anandini Yoganathan. 2013. *Languages for All? The Anglophone Challenge.* College Park, MD: University of Maryland Center for Advanced Study of Languages.

Carroll, John B. 1967. "Foreign Language Proficiency Levels Attained by Language Majors near Graduation from College." *Foreign Language Annals* 1 (2): 131–51.

Collier, Virginia, and Wayne Smith. 2012. *English Learners in North Carolina*. Raleigh, NC: North Carolina Department of Education.

Council on Foreign Relations. 2012. *U.S. Education Reform and National Security.* Independent Task Force Report No. 68. Chaired by Joel I. Klein and Condoleeza Rice. Washington, DC: Council on Foreign Relations.

Davidson, Dan E. 2010. "Study Abroad: When, How Long, and with What Results?" *Foreign Language Annals* 43 (1): 6–26.

European Union. 2012. *First European Survey on Language Competencies: Final Report.* Brussels: European Union.

Joint National Committee for Languages, the National Council for Languages and International Studies, and the Globalization and Localization Association. 2013. *Languages Are Vital to US STEM Capabilities.* http://www.languagepolicy.org/wp-content/uploads/2013/11/LSTEM.pdf.

Kelly, Nataly, and Robert G. Stewart. 2011. *The Language Services Market: 2010*. Lowell, MA: Common Sense Advisory.

National Academy of Sciences, National Academy of Engineering, and the Institute of Medicine of the National Academies: Committee on Prospering in the Global Economy of the 21st Century: An Agenda for American Science and Technology; Committee on Science, Engineering, and Public Policy. 2007. *Rising above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. http://www.nap.edu/catalog.php?record_id=11463.

Rivers, William, and Ewa Golonka. 2010. "Third Language Acquisition—A Qualitative Investigation of Factors Influencing Program Development." In *Mnemosynon: Studies on Language and Culture in the Russophone World. Presented to Dan E. Davidson by his Students and Colleagues/Мнемозинон: язык и культура в мире руссофонии. Сборник статей к юбилею Дэна Дэвидсона от его учеников и коллег*, edited by Richard D. Brecht, Liudmilla A. Verbitskaja, Maria D. Lekić, and William P. Rivers, 415–31. Moscow: Azbukovnik.

Robinson, John P., William Rivers, and Richard Brecht. 2006. "Demographic and Sociopolitical Predictors of American Attitudes towards Foreign Language Policy." *Language Policy* 5 (4): 421–42.

US Government Accountability Office, Reports and Testimonies. 2010. *Departments of Homeland Security, Defense, and State Could Better Assess Their Foreign Language Needs and Capabilities and Address Shortfalls*. GAO-10-715T, July 29, 2010. http://www.gao.gov/products/GAO-10-715T.

Watson, Jeffrey R., Peter Siska, and Richard L. Wolfel. 2013. "Assessing Gains in Language Proficiency, Cross-Cultural Competence, and Regional Awareness during Study Abroad." *Foreign Language Annals* 46 (1): 62–79.

Chapter 10

Expanded Understandings and Programmatic Approaches for Achieving Advanced Language Ability[1]

FREDERICK H. JACKSON

ALTHOUGH ENGLISH HAS COME to dominate such fields as commerce, science, technology, law enforcement, and diplomacy in the international arena, scholars and decision-makers have documented a national need for individuals who possess advanced abilities in languages other than English, as well as a deep understanding of the cultures in which the languages are spoken (see, e.g., citations in Jackson and Malone 2009).

"Advanced" language capability, as discussed in such documents, most typically refers to language proficiency that is equal to or higher than the abilities described for Level-3 ("General Professional Proficiency") in the Interagency Language Roundtable (ILR) Language Skill-Level Descriptions (n.d.a., see also Martin, this volume.) Over the last two decades, however, some scholars (e.g., Byrnes 2002b, 2006a; Norris and Pfeiffer 2003) have characterized the ILR-type descriptions of language ability as too limiting in scope.

The expansion in the understanding of linguistic communication components in recent years is extremely relevant to program design and instruction for learners who want to achieve advanced levels of language ability. On the most basic levels, the American Council on the Teaching of Foreign Languages (ACTFL) and associated organizations have revised and expanded upon the influential volume *National Standards for Foreign Language Learning* (ACTFL 1996). While it still focuses on K–16 language learning, this influential volume explicitly incorporates development of knowledge of cultures, communities, and other disciplines within the language curriculum. James Bernhardt (Dubinsky, Bernhardt, and Morley 2005, 25) pointed out that the understanding of ILR-3 language proficiency by senior staff at the Foreign Service Institute's (FSI) School of Language Studies has been deepened and "refined" in recent years, resulting in improved training. As a consequence, FSI's

"three-level speakers [today] are better speakers of other languages than they were ten years ago." At the same time, Heidi Byrnes (2002b, 2006a, 2006b) and several colleagues have also rigorously examined the construct of advanced language ability (AL2), arguing that the models of language proficiency set forth in the ILR descriptions and the related ACTFL guidelines (2012) are overly simplistic and privilege oral interactions and transactions (Byrnes 2002b, 38–9). In contrast, they have proposed a model based on "multiple literacies," which, in turn, is anchored in different communicative contexts, where *public contexts* require the language user to take part in broad societal and professional arenas and to make language and cultural choices based on the requirements of those contexts. Another recent development that almost certainly will affect our understanding of advanced language communication ability as well as instruction and assessment is the ILR's 2012 publication of Skill Level Descriptions for *Competence in Intercultural Communication*. The latter incorporates both linguistic and extralinguistic elements at each level and refers "to the content and form of communication, both verbal and nonverbal" (Federal Interagency Language Rountrree 2012 n.d.b.). A complement and supplement to the longstanding ILR Language Proficiency descriptions, these new guidelines incorporate both transactional and public communications and in several ways reflect an extension of ACTFL's *National Standards*, for example, by explicitly referring to successful interpretation of language and behaviors in cultural context, presentation of self and content in said context, and interaction with members of the culture. In the first section of this paper, we shall examine aspects of the evolving understanding of the construct of Advanced Language capability (AL2).

A number of language programs have incorporated insights such as these into effective curriculum design and learning activities. In the second section of the paper, three such programs that have achieved strong results are described. Byrnes (2006a) and her colleagues in the German program at Georgetown University have dedicated most of the last two decades to developing an undergraduate curriculum based on multiple genres and literacies to prepare students to fully take part in German-speaking social and professional contexts. FSI's "Beyond-3" language-enhancement programs have been offered for several years for motivated individuals who have returned from overseas service where they used the language; two examples of such programs are described here, for German and Thai. Another FSI program that prepares Hebrew language students for intensive advanced study in Israel before they take up their assigned positions at the US Embassy is also described. Finally, the paper looks at the very impressive achievements of the Language Flagship programs that combine study in the United States with long-term intensive residence, work, and study in a foreign country (see, e.g., Davidson and Frank 2012).

In concluding, the paper looks briefly at some developments that show promise for the future and also to some areas where investment is needed.

What Does "Advanced Language Ability" Mean, and What Does It Take to Achieve It?

In the introduction to this volume, Martin summarizes the proficiency descriptors for Advanced- and Superior-level speakers according to the ACTFL scale and notes that they were based on the ILR scale (see respective websites for full descriptors). The ACTFL definitions are similar to but not isomorphic with the ILR guidelines. Although both speaking scales emphasize the language of spoken interpersonal interaction, especially in conversation, in comparison to the ILR guidelines, the ACTFL Speaking descriptions make little reference to comprehension of spoken language. Somewhat surprisingly, the ACTFL descriptions make no explicit reference to the spoken language of *presentation* (viz. the *National Standards for Foreign Language Learning*, ACTFL 1996), other than to state that a speaker's language at the Distinguished level "typically resembles written discourse," thereby appearing to imply (wrongly) that the levels are descriptions of language samples and not of the abilities of language users.[2] Only the ILR descriptions at the highest level refer to "functionally equivalent [proficiency] to that of a highly articulate well educated native speaker." In contrast, the ACTFL description makes the point that the proficiency of a Distinguished speaker may be characterized by a "non-native accent," a "lack of native-like economy of expression," or "an occasional isolated language error."

Byrnes has frequently pointed out (2002a, 2002b, and elsewhere) that the construct of an idealized "highly articulate well educated native speaker" does a disservice to language educators and learners by appearing to set up an ideal as the anchoring goal of language learning. She notes the similarity between the ILR's concept and the "ideal speaker-hearer" that Noam Chomsky and his followers posited in setting up the distinction between linguistic *competence* and actual language *performance*, which is often characterized by infelicities of expression and linguistic mistakes or slips of the tongue (Chomsky 1965). As described later in this paper, Byrnes's point is that language educators need to help learners develop the ability to "perform" appropriately and successfully in a variety of real-world contexts. For her, the learner's goal is to become an effective "advanced [second] language user," who is able to function successfully in a range of societal contexts.

In 2012, the ILR developed and published Descriptions of Competence guidelines in Intercultural Communication (ICC), which significantly expands the scope of effective communication and thus of AL2. The preface to those guidelines contains the following statements:

> Intercultural communication is a complex activity that combines several abilities and incorporates both cross-cultural expertise and language skills. For the purposes of this document, the term refers to the content and form of communication, both verbal and nonverbal, among people of different cultures. Competence in intercultural communication is the ability to take part effectively in a given social context by understanding what is being communicated and by employing appropriate language and behavior to convey an intended message . . . [S]ince communicative effectiveness entails more than language, these Descriptions of

> Competence in Intercultural Communication incorporate both linguistic and extralinguistic elements at each skill level. Values, beliefs, traditions, customs, norms, rituals, symbols, taboos, deportment, etiquette, attire, and time concepts are some of the extralinguistic elements that typically shape the form and content of interactions (http://www.govtilr.org/Skills/Competence.htm).

The ICC guidelines refer explicitly to effective communication in a wide range of interpersonal and public contexts and to functions of interaction, presentation, and, critically, of awareness, interpretation, and understanding. They include all four language-skill modalities. Thus, effective ICC entails both polished communication skills and knowledge and understanding of the self and of the "other."

The facilitation of AL2 learning is also a complex enterprise. In a review of the research literature on language teaching, Rod Ellis (2008) described ten core "Principles of Instructed Second Language Acquisition," which can plausibly inform teaching at advanced levels.[3] The ten core principles are as follows:

1. Ensure that learners develop both formulaic expressions and rule-based competence.
2. Have learners focus predominantly on the meaning.
3. [And] also have learners focus on linguistic form, and help them to notice and attend to the "gap" between their own language production and that of other more advanced users.
4. Focus on helping learners to develop implicit knowledge, while not neglecting explicit knowledge.
5. Take into account and respond to the learners' built-in developmental "syllabus," i.e., not to attempt to "teach" aspects of the language that a learner is not yet ready to acquire.
6. Provide extensive input in the second language at levels that the learner is able to comprehend.
7. [And] also provide opportunities for the learners to produce meaningful language and use it to accomplish tasks.
8. [Encourage] the opportunity for learners to interact with each other and with others, which is critical to enable them to communicate and negotiate meaning.
9. Take account of individual differences among the students and endeavor to provide students with the kinds of learning experiences that will maximize their learning.
10. [Conduct] assessment frequently and examine both free and controlled production.

As implied in point (4) of Ellis's list, one of the important individual variables research has demonstrated to be essential and that has affected the level of advanced mastery a learner is able to achieve is openness to the implicit learning (Doughty 2011, N. Ellis 2005). The control of language that is represented by the ILR Level-4 designation is massive in scope and must be understood and produced with automaticity, allowing the brain to focus on the meaningful content and context of the communication. Its many constituent elements cannot all be taught or learned piecemeal; they must be absorbed from experience in rich authentic contexts. However,

it is possible for feedback and instruction to help some learners notice and attend to aspects of communication that might otherwise not be observed (cf. Schmidt 2001; N. Ellis 2005). As Schmidt (1994) pointed out, instruction can help learners "notice the gap" between their own communicative performance and those of native speakers. For learners who are motivated to take advantage of such facilitated noticing of language features, it appears to open doors to deeper and accelerated learning.

Cathy Doughty's research (2011) has added to the understanding of successful advanced language users. She reports three predictors of success for students who wish to develop AL2: cognitive factors, or language learning aptitude; experiential factors, particularly an early start at learning the language or other prior success at learning a language to a high level; and such noncognitive factors as motivation, health, focus, and freedom from distractions. Much of Doughty's current research (2011) focuses on what can be done to maximize the probability of success in developing AL2. For example, steps can be taken to improve learners' aptitude by "training the brain" in a way that increases short-term memory and sharpens perceptual acuity. Once training has begun, she reports that it is necessary to tailor instruction to the individual learning strengths and preferences of students but to do so within theoretically sound pedagogical and methodological principles. Learning progress needs to be monitored and feedback needs to be provided to learners and educators so as to sustain and enhance learning.

Three Language Programs

Overview—Shared Characteristics of the Programs

The three language learning programs described here differ in a number of respects, but they also share the following commonalities:

- Attention is paid from the beginning to "content," with learners focusing on how meaning is conveyed effectively in cultural contexts.
- Attention is also paid to the personal and professional interests and needs of the students. All three programs are characterized by students with diverse interests and different purposes for studying the language; the programs provide scope and opportunity for the learners to focus on developing the language competence necessary to participate in those fields.
- There are processes of frequent assessment to evaluate student progress and to provide formative information for the instructional program.
- There is a need for constructive meaning-based individualized feedback that is emphasized to the learner in all three programs, with the aim of helping students understand and select from a full repertoire of linguistic expressions in order to achieve their intended communicative effect.
- There is at least one extended immersion experience provided in a community where the students will use the language for social and professional purposes in a wide range of authentic contexts.
- There is intensive formal instruction both in classes in the United States and, particularly, during the overseas immersion experiences.

- There are many opportunities for pair- and small-group interactions among students that are typically rooted in meaningful tasks.
- Language is used meaningfully from the beginning of instruction at least 90 percent of the time by teachers and students to carry out realistic and authentic tasks.

The Georgetown University German Program[4]

Beginning in 1997, the Georgetown University German Department (GUGD), under the leadership of Professor Heidi Byrnes, initiated an extensive curricular reform to develop an integrated 4-year genre-oriented and task-based curriculum under the title, "Developing Multiple Literacies" (Pfeiffer and Byrnes n.d.). The new reconfigured curriculum consists of three levels of thematically based sequenced courses: *Contemporary Germany, Experiencing the German-Speaking World*, and *Stories and Histories ("Advanced German")*, followed by a course titled *Text in Context: Reading Germany* and one or two other Level IV courses on such topics as economics, management and marketing, or literature. These fourth-level courses share an emphasis on discourse phenomena and textuality that is manifested differently in each course depending on the genres that are likely to occur in the specified content areas. Beyond that commonality, they set their own focus (e.g., speaking in a professional context or in public settings, or writing that reflects the approaches to and conventions of literary/cultural interpretation) (http://german.georgetown.edu/page/1242716518600.html). There are also a number of Level V courses that, again, are conducted entirely in German. After the *Text in Context* course, the other course offerings in Levels IV and V are listed as nonsequenced courses, and, as the description suggests, they may or may not be taken by students in any order.[5] The program eliminates the distinction between learning goals for nonmajors and majors of German, and nonmajors make up as much as 50 percent of total enrollment in the upper division classes.

As will be described shortly, the primary focus of study in the GUGD is *not* on Global Oral Proficiency. It is instead to produce students who are "competent and culturally literate users of German" (Byrnes 2003). Nonetheless, a Simulated Oral Proficiency Interview (SOPI) was administered to the students. The results showed that after twelve credits of study (180 classroom hours), only 2.5 percent of the students scored lower than Intermediate Low on the ACTFL scale (ILR-1), and 7.7 percent scored at the Advanced level (ILR-2). After completing Level III (270 hours), none of the twelve students who were tested scored lower than Intermediate High (ILR-1+), and two of them (17 percent) achieved oral proficiency ratings of Advanced High (ILR-2+) (Norris and Pfeiffer 2003, 577). All five students in the study who had completed one semester at Level IV (315 hours) received proficiency ratings of Advanced High (ILR-2+) or Superior (ILR-3). All these ratings are significantly higher than those typically achieved elsewhere in the same amount of class time (see, e.g., Jackson and Kaplan 1999 for expected results at FSI).

The GUGD specifies the "advanced L2 user" as the program goal, in contrast to a goal of functioning as a "near-native" or native speaker. It rejects "global proficiency" as an instructional context, but instead it is grounded in capacities that enable certain "environments of use" in societal contexts. Byrnes writes that the driving constructs for preparing students for such environments of use are "what,

for what purpose, how well" the language will be used (Byrnes 2003). In this respect, it has something in common with efforts elsewhere to teach Language for Specific Purposes. Every stage of instruction is mindful of the program goals, not just those of the particular course.

Every course at every level of the curriculum explicitly incorporates attention to genre characteristics, cultural and social contexts, and linguistic form both on the sentence level and in spoken and written discourse. The program instructs students in every skill modality, with a particular emphasis on writing, where close attention must be paid to the nature and background of the anticipated reader in choosing how to express content. Level III marks a clear shift in focus from the spoken and written language of personal interaction in the first two levels to a focus on the language of "public discourse." Byrnes cites James Gee (1998a, 56) in distinguishing between "'primary discourses' of familiarity that all of us acquire largely unconsciously . . . and 'secondary discourses' that are developed . . . 'by having access to and practice with' the secondary institutions of schools, workplaces, offices, businesses, and churches" (Byrnes 2002b, 49). As Byrnes has pointed out (personal communication [p.c.]), the advanced language user must know how to establish credibility in a professional or vocational forum; it is a matter of language, content, genre, tone, and, at least in person, cultural behavior.

Assessment in the GUGD program is formative, not summative, and the information obtained is used to provide feedback to help the student, teacher, and program as a whole improve. Language functions that are used to carry out a task are evaluated at the lexico-grammatical level, and are also evaluated at the sentence and discourse levels, in terms of accuracy, fluency, and (appropriate) complexity. The metaphor of choice making is employed in matching the meaning with the content. Choice making, of course, depends both on having a repertoire of choices to use and on knowing the contextual pragmatic meanings of each item in the repertoire so one can choose appropriately.

In anchoring the syllabus in preparation of an Advanced L2 user, Byrnes (2002b) states that such a person brings "multiple literacies" to language tasks, including those of the first language and any other languages the individual may know. Byrnes sees this first language background knowledge not as problematic and symptomatic of a "deficit model" of language acquisition, but rather as an asset for the advanced language learner, in that it expands perspectives of understanding. Byrnes (2002b, 41) has referred to the advanced L2 user as a "multicompetent speaker."

FSI's School of Language Studies

The State Department's FSI has been engaged for more than sixty-five years in preparing employees to function overseas in now more than seventy languages at the "Professional" level (ILR-3). Despite frequent requests, FSI has never offered programs with the explicit training goal of getting individuals to the ILR-4 level of "Advanced Professional Proficiency." There are too many individual variables that affect the likelihood of a learner achieving such a level through instruction. FSI does, however, offer tailored "Beyond-3" training to individuals who return from the target language country, and during the last decade, in-country intensive "top-off" training

has been arranged for students in various languages who have achieved ILR-3 proficiency in training at FSI and can arrange time for further training before taking up their assignment at the US Embassy or consulate. Proficiency levels of ILR-3+ and ILR-4 have been achieved by motivated students in both of these kinds of programs, which are described here.

Although FSI training results are expressed in terms of ILR Speaking and Reading scores for global language proficiency, the training itself focuses primarily on preparing the students to perform the representational duties that their assignment and position will require. Thus, for example, students in training would not normally be asked to play a role that they would be unlikely to perform in their real lives.

General language learning at FSI.

The FSI was established in 1947 as the training arm of the State Department. At present, the School of Language Studies teaches more than seventy languages—from world languages like Spanish, Arabic, and Russian to national languages such as Turkish, Urdu, and Thai, to regional languages like Pashto, Kurdish, and Malayalam. The School of Language Studies trains a variety of different students, from officers new to the Foreign Service to ambassadors, clerical staff to security personnel. Most students come to FSI with a specific language requirement that they are expected to meet. Students are selected for skills pertinent to the identified needs of the foreign affairs community—and not necessarily for strong language aptitude. FSI students are typically highly motivated; they know that proficiency in the language they are studying is crucial to their success in their jobs—and therefore to their competitiveness within the Foreign Service.

Learning a language to the professional level that the Foreign Service demands requires a great deal of hard work. Classes are small and intensive, with six or fewer students meeting together with a teacher for a minimum of five classroom hours every day. Students are expected to commit to at least three hours daily of independent study outside of class. To get to the threshold level that most overseas jobs require, a good learner starting from scratch in Spanish or Dutch devotes about twenty-four weeks (600 hours) of class-time. To get to that same level in such languages as Thai, Hungarian, or Russian requires about forty-four weeks (1,100 hours). Japanese, Chinese, Korean, and Arabic require more than eighty weeks (2,200 hours) in class; the second year typically takes place at an overseas institution.

FSI instructors were all educated to adulthood in a country where the language they teach was spoken and the culture was fully reflected. One of their tasks is to serve as an explicit representative of the language and culture, and they are also trained to guide their students to understand the perspectives that are represented. To do this requires them to understand American students' attitudes and beliefs, both generally and for the specific students in their classrooms. To help faculty understand their individual students, FSI provides a Learning Consultation Service that diagnoses each student's learning style preferences, discusses them with the student, and, if the student permits, shares the findings with the language program to enable it better to meet the student's learning needs (Ehrman 2001).

Because of the small class size and the abilities of teachers to differentiate instruction to meet individual students' learning needs, almost every class becomes a small-group activity, where students learn how to communicate what they will need in their overseas assignments. Although there is usually a grammatically-based or functional syllabus that forms the spine of an FSI course, most of the learning is typically implicit or incidental, occurring naturally as a student's need to understand, recognize, or express content arises.

Again, because of the intense classes and small class sizes, individual students often request and receive specific learning experiences. Students have asked to spend class time in traditional audio-lingual drills, while others have primarily wanted to listen or read, and still others want to speak the language from day one. Earl Stevick's book *Success with Foreign Languages* (1989) describes several individual cases of successful FSI students talking about what worked for them as second language learners.

Three brief examples of "Beyond-3" training programs at FSI

Students in initial training courses at FSI normally have a training goal of General Professional Proficiency (ILR-3).[6] While some highly skilled students may achieve higher than ILR-3 in one skill modality, it is believed that the achievement of language proficiency significantly above ILR-3 normally only occurs in learners who have been physically immersed in the culture and who have experienced language use and references in an authentic cultural context for an extended period of time and, thus, have anchored the language in that context. "Beyond-3" training is provided in several languages for those who have already served in a country where the 3-level language is spoken and have returned to FSI for additional training prior to another assignment to a country where the language is indigenous. Such training is often individualized for the student, with a mix of guided independent study, including extensive reading, watching authentic media broadcasts, and one-on-one tutoring, all designed to assist the learner to address diagnosed needs.

In one example of a student of Thai, the focus that the student wanted was to greatly expand her useful lexicon in areas of her professional specialization. With an instructor's guidance, this student read both broadly and deeply in her topics of interest, then compiled a number of topical glossaries of frequent lexical items and common expressions used in writing and in formal speech. She then met with the instructor to practice using the forms appropriately in such professional contexts as group meetings, briefings, negotiations, media interviews, press conferences, and other examples of public speaking. She later reported that her glossaries proved invaluable to her in Thailand in preparing for professional events on her schedule. Another "Beyond-3" Thai student spent some of his study time reading political cartoons and jokes to help him understand and employ Thai humor. The same student, whose position would be as a senior officer in the US General Consulate, learned from a teacher about a newly published Thai paperback that presented tongue-in-cheek advice to Thais who wanted to obtain visas to travel and work in the United States. He read the book and discussed its contents with his teacher and with other Thais whom he met in the course of his study. In a third case, teachers showed an advanced Thai student how many of his extended discourse structures in formal

speech were heavily influenced by his native English, and, while intelligible, sounded coarser to the Thai ear than what would be used by a senior Thai. In this example, the student was helped to become more aware of features of formal spoken discourse in Thai and also of ways in which his own speech lacked many of those features. As part of the work with this student, he was given several opportunities to practice reading aloud, as if giving public remarks, from short speeches that he had composed with the assistance of a teacher. This helped him to automatize not only the expressions and style of formal Thai but also the appropriate rhetorical style and rhythm of delivery. In all three of these cases, the intense individualized immersion in Thai, together with their rich previous experience in Thailand, resulted in significantly better scores in the speaking proficiency assessment, and, more importantly, in successful tours as Foreign Service Officers in Thailand.

Another example involves a student of German, a language where Professional-level proficiency requires consistent accurate control of morpho-syntactic forms, such as gender and case marking on nouns and pronouns, verb forms, agreement, and word order, especially in complex sentences. One German student at FSI had learned the language as a child but had left Germany before the age of ten. This student was fluent in the language, with native-like control of the sounds of the language, sentence rhythm, and intonation. However, she lacked consistent control of gender agreement, dative and genitive noun case markings, and word order in sentences with complex verbs or in dependent clauses. She constructed nearly every sentence in the Subject-Verb-Object order, obligatory in English but often altered by speakers of German to front other important information, such as the direct object or expression of time, place, or manner. She also demonstrated a rather limited vocabulary, particularly for work-related topics. She had received an FSI proficiency rating in German well below the ILR-3 that her assignment in Germany would require, and she was assigned to training to upgrade her skills. The teachers feared that when she was in Germany, because she was so fluent, her German counterparts would conclude that her errors were the result of lack of education.

In this extreme case, and in others similar to it, the student needed to unlearn much of what she had unconsciously acquired as a child. She was required to learn the patterns of educated German speech and writing that German children master in school, and at the same time she had to monitor her own production—much of it automatic and unconscious—and notice where it differed from the models that she needed to internalize. It may well be imagined that this was an extremely frustrating process for both the student and teachers. This student had developed a process that had previously worked fluently and without much thought, and with some success according to her friends; it then had to be taken apart and put slowly together again. As a result, much of what had before been fluent and automatic became slow, halting, and clumsy. She was required to listen to models of educated speech and to listen to and critique recordings of her own speech. Teachers designed means for signaling to her every error that she made of a certain type until she no longer made them, and then they began signaling errors of another type. This proved a tedious and painful process for all involved, but it was ultimately successful, and by the end of her tour in

Germany she had regained much of her fluency along with a much more acceptable control of the grammatical structures of German.

In the last decade, a few language programs at FSI have made it possible for students, who developed strong language ability toward the end of their assigned training at FSI, to continue their learning for several weeks at an institution in the country.[7] At the request of the US ambassador to Israel, FSI initiated a Hebrew program for students who had developed a proficiency in reading and speaking the language of at least ILR-2+ to study at an Israeli *ulpan* for four to six weeks before taking up their assigned position at the embassy. The *ulpanim* are schools established by the Israeli government primarily to teach the Hebrew language and Israeli social norms to recent immigrants in the country. Classes are intensive, typically for four or five hours, five days a week, and are offered at three levels of language ability.

The embassy requested this program in order to accustom new officers to Hebrew as it is spoken in the very diverse communities of Israel, where about 40 percent of citizens are native speakers of other languages. The two most urgent concerns of the embassy were to achieve improved listening comprehension in a variety of common contexts and enhanced fluency in cross-cultural interaction to get information efficiently, take part successfully in many-sided group conversations with overlapping turn-taking, and communicating important content in culturally effective ways. Among the contexts where skilled listening comprehension was viewed as especially important were cell-phone usage, including comprehension of voice-mail messages; media broadcasts; public speeches, as in the Israeli *Knesset*; overheard conversations; and in group settings in which Israelis, each with a different speech variety or accent, all talk at once with frequent unexpected changes of topics.

FSI developed the program in consultation with the embassy and with an *ulpan* in Tel Aviv. The minimum Hebrew-speaking proficiency for admission to the program was set at ILR-2+ (ACTFL Advanced High), following two marginally successful experiences when students with lower proficiency levels had been sent. Students were housed by themselves in Hebrew-speaking areas of Tel Aviv and were provided with Israeli newspapers, televisions, and radios to receive Hebrew language broadcasts. They were expected to attend and report on cultural events in their communities. The programs were scheduled for a minimum of six weeks, during which time the students were instructed not to go to the embassy where they would be working. In addition to five hours daily of *ulpan* instruction in a class of fifteen to twenty students, they were also required to arrange for sixty to ninety minutes each day of individual mentoring in the language, where they could push their growing abilities and receive feedback. In addition, they were asked to meet with the embassy's Post Language Officer at least once every two weeks, preferably more often if their program needed improvement in any way. FSI staff prepared the students to take maximum value from their individual tutoring and their immersion. Also, FSI was in contact with the Curriculum Director at the *ulpan* to help monitor the students' progress. Following the conclusion of the program, students were given the FSI language proficiency test (Jackson and Kuhnreich 2005).

Logistics for the Hebrew program were time-consuming, and the program was reportedly discontinued after approximately four years. The results, however, were

quite positive. Out of the seven officers who participated in it during the first two years, all of whom had begun their study of Hebrew at FSI without prior knowledge of the language, all but one had post-program scores on the FSI test that were better by at least one-half level on the ILR scale than their pre-program results and achieved at least ILR-3 in speaking. The most advanced learner also took the Israeli National Hebrew Examination, on which she scored in the 94th percentile. An Inspector General's evaluation of the embassy in Tel Aviv reported this program as a "Best Practice," stating that "officers emerg[e] from the four-week [sic] immersion program with significantly higher test scores as well as more fluency, confidence and familiarity with the multiple Israeli accents. The ambassador and OIG believe [FSI] should formalize the in-country Hebrew immersion program, and use it as a model worthy of emulation in other posts" (Jackson and Kuhnreich 2005).

The Language Flagship Programs

The Language Flagship program is a little over ten-years old at this writing. A component of the federally funded National Security Education Program (NSEP), the Flagship's goals are to create a pool of college graduates from all majors with Professional-level proficiency (ILR-3, ACTFL Superior) in critical languages to develop the next generation of global professionals and to "change the expectations for language learning." Its motto is "Grounded Theory and Innovative Practice for Superior-level Language Learning" (Eisen 2013).

The Flagship has twenty-six US Program Centers and ten overseas programs. It also includes three K–12 Flagship programs. The programs provide instruction for ten languages that have been identified as critical for more Americans to have proficiency in: Arabic, Mandarin Chinese, Hindi, Korean, Persian, Portuguese, Russian, Swahili, Turkish, and Urdu.

Once primarily a graduate student program that required entering students to have demonstrated at least ILR-2-level proficiency in their language, the Flagship has evolved over time. The US Centers are now undergraduate programs that require no prior experience in the language for entering students. At the end of the four years as an undergraduate in intensive language study (in addition to the required study in their majors), students who have achieved proficiency levels of at least ILR-2 (ACTFL Advanced) in speaking and at least one other skill modality continue their study in a "capstone year" at one of the overseas program centers. Results have more than met expectations, with, for example, 89 percent of forty-seven Russian Flagship students achieving oral proficiency of ILR-3 or better, including 23 percent who achieved ILR-3+ and 15 percent who received official proficiency ratings of ILR-4 or better (Davidson 2010, 14; also Davidson, this volume). Students in other languages are seeing similar results (see, e.g., Davidson and Frank 2012; Falsgraf and Spring 2007; Spring 2012a; Spring 2012b; and Van Olphen 2008).

Although there is variation among the different overseas programs, the structure of the capstone year typically includes housing in a dormitory, an apartment, or with a family where other residents are non-English speakers; formal targeted language training approximately three hours a day where writing and public speaking are stressed; direct enrollment in courses at a local partner university; and internships

with a business, organization, or agency that works in the student's field. Many programs also include regular meetings with a tutor/mentor, where each student can be helped to address specific concerns.

Some Key Principles of the Language Flagship

- Content learning in the target language and integrated learning across disciplines

  Flagship students are not typically "language majors;" among the fields of concentration reported for recent Flagship students' majors in addition to language and literature were, in alphabetical order, accounting, biology, business, chemistry, community development, comparative literature, economics, engineering, environment, German, history, international studies, journalism, law, marketing, medical microbiology, nuclear engineering, political science, public health, and public relations.

  Several of the US programs have been able to institute a Language Across the Curriculum model (Freels 2012), where students take courses in their field that are taught in their Flagship language. Falsgraf (2012) points out that the use of technology has been critical for the infusion of content in Chinese for the different fields of concentration. Bourgerie et al. (2012), and Christensen and Bourgerie (this volume), liken this aspect of the program to "Language for Special Purposes at the professional level" (see also Spring 2012a). In the capstone year overseas, all students enroll in courses with native speakers of the language.

- Coordinated domestic coursework and interventions beyond the classroom

  The University of Wisconsin features "intentional integration of in- and out-class learning" that makes use of the resources of physical and virtual communities (Murphy 2012). In the capstone year, in addition to taking language training and courses in their fields, students do work in their fields as interns (Lekic 2012).

- Group tutoring, individual tutoring, and guided cocurricular activities

  Regular small-group tutoring and individual tutoring in the undergraduate program and overseas program provide abundant opportunities for meaningful language use and culture learning. They also provide valuable opportunities for students to ask questions about something they have noticed and to seek and receive constructive feedback. In several instances, students have had separate tutors for their language and content areas.

- Rigorous assessment tied to anticipated learner outcomes

  Students enrolled in the Flagship know in advance that they will receive frequent formal and informal assessments of their progress in relation to established benchmarks and milestones. There are tests of proficiency and of achievement. Much of Flagship learning activity overseas is task-based, typically taking place outside the classroom and providing real-time assessment of the student's ability to function successfully in the language and culture. Feedback from task-based assessments is both global and very specific.

The major proficiency assessments of Flagship students include pre- and post-administrations of the ACTFL Oral Proficiency Interview (OPI), scored on the ILR scale; online Reading and Listening tests developed by the American Councils for International Education; and, at the end of the study, the FSI test of Speaking and the Defense Language Proficiency Tests (DLPT) for Reading and Listening. Russian students take the European Union-developed Test of Russian as a Foreign Language (Davidson 2010, 2012). Chinese summer students are given pre- and post-program criterion-referenced proficiency assessments of reading and listening that were developed at Brigham Young University (Bourgerie p.c.; also http://chineseflagship.byu.edu/node/10).

Articulation from domestic programs to overseas programs

It is obviously critical for the overseas programs to mesh their instruction and learning activities appropriately with the abilities and experience of the students from the domestic programs. The established benchmarks help in achieving that meshing, and senior staff from the American Councils coordinates between the domestic and overseas program and the central Flagship office. Joint US and local professional academic and student support teams are comprised of US and local senior language specialists and US and local students to provide support and oversight. One of these teams will pay particularly close attention to any program where test results or reports from students or instructional staff indicate that something may be amiss (Eisen p.c.).

Some Programmatic Issues in the Flagship

The Language Flagship does not publicly mandate or endorse any specific language teaching approach or pedagogical principles, except for emphasizing a focus on the common end goal of Professional-level proficiency. If the program produces the desired results, and if the students are pleased with their learning progress, Flagship rarely interferes. Eisen (p.c.) has reported that a few domestic Flagship programs at some US universities tolerate classes that are larger than they should be for effective learner-centered instruction and that the instruction in some classes is quite teacher-centered, with little opportunity for students to interact with each other in group work that is task-based and performed in pairs or small groups. Also, some teachers in those classes often tend to use English. The Flagship would like to change these behaviors.

Flagship has not tried to directly confront issues like these, concerning professional development and training for teachers, but the Flagship may need to do so in the future. If the most successful research-based Flagship approaches to instruction can be disseminated more widely, both inside the Flagship and among teachers of other languages, it would be a great benefit to the field.[8]

Effective overseas programs encourage a self-reflective stance in their students: self-assessments, regular goal setting, regular feedback from facilitators/advisors, analysis of events from both the local and the student's own perspective, debriefing on situations, discussions of attitudes and appreciation, as well as holistic and analytic assessments of L-2 linguistic competencies and related abilities.

The task requires expert balancing of challenge and support to the learner, who will invariably experience both cultural shock and cognitive overload (Davidson 2012).

Conclusion: Some New Resources for Advanced Language Learning

The highly motivated language learner now has access online to a rich panoply of resource tools that can be used to support independent and/or instructed language study. Among these tools are the guided self-assessment instruments that have been developed by the ILR (e.g., http://www.govtilr.org/Publications/speakingsa.html), *LinguaFolio* (https://linguafolio.uoregon.edu/documents/LFGrid.pdf), and ACTFL Progress Indicators for Language Learners (http://www.actfl.org/publications/guidelines-and-manuals/ncssfl-actfl-can-do-statements). These give learners and their teachers clear information about what constitutes functional proficiency.

Malone and her colleagues (2004) have reported that distance learning can combine text, video, CD-ROM, and synchronous and asynchronous use of the Internet in effective ways, as exemplified by the advanced online courses for Chinese, Japanese, and Korean developed at the University of Hawaii (http://nflrc.hawaii.edu/certificates/).
In a recent issue of the online journal *Language Learning & Technology*, Godwin-Jones (2013) describes promising emerging technologies for the teaching and learning of Less Commonly Taught Languages. Thorne and Reinhardt (2007) have written a paper on the topic "Technologies for Advanced Foreign Language Proficiency."

The federally funded Language Resource Center for Open Educational Resources and Language Learning at the University of Texas has the mission of identifying, developing, and disseminating open-source materials for language learning and teaching (http://www.coerll.utexas.edu/coerll/). The ILR also hosts a web page "Webliography" of language and area studies resources at http://www.govtilr.org/Web_LCTL/ILR%20WLCTL.htm, and the Defense Language Institute hosts a Global Language Online Support System (GLOSS) for forty languages at http://gloss.dliflc.edu.

One of the most promising developments for motivated independent learners of language is tools from the field of corpus linguistics that can be used to analyze and deconstruct written and spoken texts in almost any language that appears on the Internet. One example is the SEALang Lab site for Southeast Asian languages, which was developed with Department of Education funding by Doug Cooper (http://www.sealang.net/lab/index.htm.). The site is linked electronically to core dictionaries for such languages as Thai and Lao that a user can access instantly online to help understand any online text in the language. Concordance software at the site also permits a learner to identify high-frequency vocabulary in topic areas of interest and to analyze sentence and discourse structures in different genres. Unfortunately, however, the funding for this excellent project was discontinued after 2009, and the site has not been developed further or updated since then.

Some Crying Needs for (Re-)Investment

The SEALang Lab site is not alone in losing funding for the improvement of language learning and teaching. With the exception of federal funding for STARTALK and the Language Flagship from the National Security Language Initiative of 2006, national investment in language education has been very scarce.[9] Such investment will be critical in order to provide the kinds of advanced language instruction described in this chapter and this book to more learners. Indeed, universities across the country are continuing to cut back on offering third- and fourth-year so-called "advanced" courses in many languages in order to reduce costs, which further reduces opportunities for students to continue developing their language abilities toward truly advanced levels.

The areas of greatest need include investment in

- Beginning use-oriented language instruction in elementary schools and continuing it through high school.
- Developing closely articulated learning materials for upper-level students, such as the early computer-mediated tutorials to teach advanced skills in Russian that were developed at the University of Wisconsin–Madison (http://imp.lss.wisc.edu/rails) (Malone et al. 2004). Flagship programs are taking the initiative in this for the ten Flagship languages.
- Developing digital language resource materials and analytical tools for the less commonly taught languages, such as those that were started for the SEALang Lab project and the excellent courseware for recognizing Arabic dialects that was developed at the National Foreign Language Center by Gerald Lampe and his colleagues (http://www.nflc.umd.edu/node/55#.Ur-gef3UMjE).
- Investing in teacher education and professional development of teachers who are asked to instruct advanced language learners. A part of this development must be to strengthen the language proficiency of non-native-speaking teachers and the intercultural understanding of *all* teachers. The current minimum proficiency to teach language to Americans has been recommended by ACTFL as Advanced Low (ILR-2) in languages cognate with English and Intermediate High (ILR-1+) for languages such as Chinese, Japanese, Korean, or Arabic. Such a low level of language ability is clearly inadequate to instruct students who are striving to achieve professional proficiency and better. Byrnes (2002a) has reported on one very successful intensive immersion-based seminar for nonnative teachers of German to strengthen their control of professional language, but those participants had attained at least Superior-level proficiency in German.[10]

The next area of need for professional development for all language teachers is in understanding the research findings of advanced-level second language acquisition and its implications for instructional and delivery design. As we have seen, effective advanced language instruction is about facilitating the students' learning by creating an excellent learning environment with activities appropriate to the learners' current abilities, differentiating instruction for the individual students and providing constructive feedback that enables them to continue to progress.

Notes

1. Thanks and appreciation are due to anonymous reviewers for their comments on an earlier draft of this article, to Sam Eisen of the Defense Language National Security Education Office for taking time to deepen my understanding of the Language Flagship programs, and to Heidi Byrnes, Chair of the GUGD, for sharing her insights into AL2 over several years. It is needless to say that any remaining errors of interpretation or understanding are solely my responsibility.
2. As the description also states, one of the features of a speaker at the Distinguished level is the ability to "tailor language to a variety of audiences" and, presumably, contexts. It would seem unlikely that a Distinguished speaker's speech with friends at a picnic or around a bridge table would "resemble written discourse."
3. The federal STARTALK program, which provides brief exposure in summer courses to proficiency-based training in ten critical languages for students in grades five through sixteen, has identified six core principles for language instruction. Although these principles are not at all targeted at advanced language learning, they are nonetheless applicable in that context:

    - Implementing a Standards-based curriculum organized around thematic content
    - Facilitating a learner-centered classroom
    - Using the target language at least 90 percent of the time and providing comprehensible input for instruction
    - Integrating culture, content, and language in the instruction
    - Adapting and using age-appropriate authentic materials from many sources
    - Conducting frequent performance-based assessment and using it to inform the program and provide feedback to teachers and students

4. Full details about the Georgetown German program can be found online at http://german.georgetown.edu/page/1242716510124.html, as well as in many of the writings of Heidi Byrnes and her colleagues.
5. The courses in Levels I–III may be taken intensively, for six credits per course. As a result, the sequence through Level IV can be completed in as few as two academic years at Georgetown. Students choosing intensive study achieve somewhat higher mean oral proficiency ratings after twelve semester hours of instruction than do those in nonintensive courses (Norris and Pfeiffer 2003, 578).
6. Some students—particularly first-tour or second-tour Junior Officers—have a training goal of Limited Working Proficiency (ACTFL "Advanced;" ILR-2) and are correspondingly given less training time.
7. Among the FSI language programs that have implemented this kind of program at least once are Spanish, Russian, Thai, and Hebrew.
8. The experience of the STARTALK program, which is also federally funded, has shown that a small investment in professional development of language teachers can have a profoundly positive affect on the teaching of those languages and, through the sharing of the insights and techniques gained from STARTALK with teachers of other languages in the teacher's home institution, on the teaching of those languages as well. These programs have been instrumental in helping language teachers and teacher educators of different languages recognize that they are members of a single professional community of practice.
9. While federal support for world language education has suffered during the last five years, some states and many local communities have stepped up strongly to implement programs in public schools. Especially noteworthy is the statewide Utah Dual Language Immersion programs for Chinese, French, Portuguese, and Spanish (see http://utahimmersion.org/?page_id=2). The Chinese program is an initiative of the Language Flagship, Brigham Young University, and the state to implement "K-16 Chinese education" (http://utahchinesedli.org/). It is hoped that the program may serve as a national model.

10. Teachers who are native speakers of the language and who grew to adulthood in a country where it is the language of education normally do not need further knowledge of the language and understanding of the culture. However, what they do need is an understanding of American cultural and professional expectations as represented in their students, colleagues, and the school or college department where they are employed.

References

American Council on the Teaching of Foreign Languages. 1996. N*ational Standards for Foreign Language Learning*. ACTFL. http://www.actfl.org/sites/default/files/pdfs/public/StandardsforFLLexecsumm_rev.pdf.

———. 2012. *ACTFL Proficiency Guidelines 2012: Speaking, Writing, Listening, and Reading*. http://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012.

Bourgerie, Dana Scott, Matthew B. Christensen, and Rita Cortex. 2012. "Customizing and Individualizing Domestic Training at the Advanced Level." Presented at preconvention workshop on Increasing Language Proficiency at the Postsecondary Level Using Flagship Principles at the ACTFL Annual Convention, Philadelphia, PA, November 15, 2012.

Byrnes, Heidi, ed. 1998a. *Learning Foreign and Second Languages: Perspectives in Research and Scholarship*. New York: The Modern Language Association of America.

———. 1998b. "Constructing Curricula in Collegiate Foreign Language Departments." In *Learning Foreign and Second Languages: Perspectives in Research and Scholarship*, edited by Heidi Byrnes. New York: The Modern Language Association of America. 262–95.

———. 2002a. "Contexts for Advanced Foreign Language Learning: A Report on an Immersion Institute." In *Developing Professional-Level Language Proficiency*, edited by Betty Lou Leaver and Boris Shekhtman, 61–76. London: Cambridge University Press.

———. 2002b. "Toward Academic-Level Foreign Language Abilities: Reconsidering Foundational Assumptions, Expanding Pedagogical Options." In *Developing Professional-Level Language Proficiency*, edited by Betty Lou Leaver and Boris Shekhtman, 34–60. London: Cambridge University Press.

———. 2003. "Teaching Toward Advanced (Professional-level) L2 Performance." Unpublished Professional Development Institute workbook. Offered by the National Capital Language Resource Center, May 27–28, at Georgetown University, Washington, DC.

———, ed. 2006a. *Advanced Language Learning*. London: Continuum Press.

———. 2006b. "Locating the Advanced Learner in Theory, Research, and Educational Practice: An Introduction." In *Educating for Advanced Foreign Language Capacities*, edited by Heidi Byrnes, Heather D. Weger-Guntharp, and Katherine Sprang, 1–16. Washington, DC: Georgetown University Press.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Davidson, Dan E. 2010. "Study Abroad: When, How Long, and with What Results? New Data from the Russian Front." *Foreign Language Annals Special Focus Issue: Language Learning and Study Abroad*, edited by Dan E. Davidson, 43(1) 6–26.

———. 2012. "Reinventing Study Abroad: What the Overseas Flagships Have Taught Us about Integrating Assessment into Overseas Immersion Learning." Presented at preconvention workshop on Increasing Language Proficiency at the Postsecondary Level Using Flagship Principles at the ACTFL Annual Convention, Philadelphia, PA, November 15, 2012.

Davidson, Dan E. and Victor Frank. 2012. "The Overseas Component of the Language Flagship: Addressing Learner Needs within an Acquisition-Rich Environment." *Journal of Chinese Teaching and Research in the U.S., Special Issue for the Language Flagship*, vol. 4: 8–21.

Doughty, Cathy. 2011. "Redesigning Language Learning." Plenary Presentation at monthly meeting of the Interagency Language Roundtable, Arlington, Virginia, October 11, 2011.

Dubinsky, Inna, James Bernhardt, and Joan Morley. 2005. "High-Level Proficiency Acquisition: Oral Skills." In *Teaching and Learning to Near-Native Levels of Language Proficiency II: Proceedings of the Fall 2004 Annual Conference of the Coalition of Distinguished Language Centers*, edited by Inna Dubinsky and Richard Robin, 23–32. Salinas, CA: MSI Press.

Ehrman, Madeline. 2001. "Bringing Learning Strategies to the Learner: The FSI Language Learning Consultation Service." In *Language in Our Time: Bilingual Education and Official English, Ebonics and Standard English, Immigration and the Unz Initiative*, edited by James E. Alatis and An-Huei Tan, 41–58. Washington, DC: Georgetown University.

Eisen, Samuel. 2013. Personal communication with Frederick Jackson regarding the Language Flagship. Arlington, VA, August 8. Followed by email correspondence August 11–26.

Ellis, Nick C. 2005. "At the Interface: Dynamic Interactions of Explicit and Implicit Language Knowledge." *Studies in Second Language Acquisition*, 27(2), 305–52.

Ellis, Rod. 2008. "Principles of Instructed Second Language Acquisition." *CAL Digest* December 2008.

Falsgraf, Carl. 2012. "Content-based Language Learning (Second Language Content Learning)." Presented at preconvention workshop on Increasing Language Proficiency at the Postsecondary Level Using Flagship Principles at the ACTFL Annual Convention, Philadelphia, PA, November 15, 2012.

Falsgraf, Carl, and Madeline K. Spring. 2007. "Innovations in Language Learning: The Oregon Chinese Flagship Model." *Journal of the National Council of Less Commonly Taught Languages*, vol. 4: Spring, 1–16.

Freels, Sandra. 2012. "Language Across the Curriculum: Adding Breadth to the Undergraduate Language Program." Presented at pre-convention workshop on *Increasing Language Proficiency at the Postsecondary Level Using Flagship Principles* at the ACTFL Annual Convention, Philadelphia, PA, November 15, 2012.

Godwin-Jones, Robert. 2013. "Emerging Technologies: The Technological Imperative in Teaching and Learning Less Commonly Taught Languages." *Language Learning & Technology*, 17 (1). http://llt.msu.edu/issues/february2013/emerging.pdf.

Jackson, Frederick H., and Sarah Kuhnreich. 2005. "Hebrew In-Country Transition Immersion Training Program." Presented at the ILR 2005 Year of Languages Showcase: Preparing for the Global Language Challenge. Howard Community College, Columbia, MD.

Jackson, Frederick H., and Marsha A. Kaplan. 2001. "Lessons Learned from Fifty Years of Theory and Practice in Government Language Teaching." In *Georgetown University Round Table on Languages and Linguistics: Language in Our Time*, edited by James E. Alatis and An-Huei Tan, 71–87. Washington, DC: Georgetown University Press.

Jackson, Frederick H., and Margaret E. Malone. 2009. "Building the Foreign Language Capacity We Need: Toward a National Framework for Language Education in the United States." Posted on http://www.nflc.umd.edu/publications/Building_The_Foreign_Language_Capacity_We_Need

Lekic, Maria D. 2012. "Maximizing Results of the Overseas Study Environment." Presented at pre-convention workshop on Increasing Language Proficiency at the Postsecondary Level Using Flagship Principles at the ACTFL Annual Convention, Philadelphia, PA, November 15, 2012.

Malone, Margaret E., Benjamin Rifkin, Donna E. Christian, and Dora E. Johnson. 2004. "Attaining High Levels of Proficiency: Challenges for Language Education in the United States." *Journal for Distinguished Language Studies* 2: 67–88.

Murphy, Dianna. 2012. "Designing Learning Environments to Promote L2 Use and Culture Learning Outside of the Classroom." Presented at pre-convention workshop on Increasing Language Proficiency at the Postsecondary Level Using Flagship Principles at the ACTFL Annual Convention, Philadelphia, PA, November 15, 2012.

Norris, J. M., and Peter C. Pfeiffer. 2003. "Exploring the Use and Usefulness of ACTFL Oral Proficiency Ratings and Standards in College Foreign Language Departments." *Foreign Language Annals* 36: 572–581.

Pfeiffer, Peter C., and Heidi Byrnes. n.d. "Curricular Reform, Learning Outcomes, and Identity of Majors: A Case Study of German Program Outcomes Assessment." Georgetown University. http:// german.georgetown.edu/scholarship/programevaluation/.

Schmidt, Richard. 1994. "Implicit Learning and the Cognitive Unconscious: Of Artificial Grammars and SLA." In *Implicit and Explicit Learning of Languages*, edited by Nick C. Ellis. London: Academic Press, 165–209.

———. 2001. "Attention." In *Cognition and Second Language Instruction*, edited by Peter Robinson. Cambridge: Cambridge University Press. 3–32.

Spring, Madeline K. 2012a. "Languages for Special Purposes Curriculum in the Context of Chinese-Language Flagship Programs." In *Modern Language Journal Special Issue: Focus Issue: Languages for Specific Purposes: Update on Über Grosse and Voght (1991) in a Global Context*, vol. 96 s1, edited by Barbara A. Lafford, 140–57.

———. 2012b. "Linking Curriculum, Assessment, and Professional Development: Articulation in Chinese Flagship Programs." In *Teaching Chinese as a Foreign Language: Theories and Applications* (2nd edition.), edited by Michael Everson and Yun Xiao, : 195–236. Boston, MA Cheng & Tsui.

Stevick, Earl W. 1989. *Success with Foreign Languages*. New York: Prentice Hall.

Thorne, Steven L., and Jonathan Reinhardt. 2007. "Technologies for Advanced Foreign Language Proficiency." Position Paper. Center for Advanced Language Proficiency Education and Research, Pennsylvania State University. http://calper.la.psu.edu/docs/pdfs/alppositionpapers/CALPER_ALP_Technologies.pdf.

Van Olphen, Herman. 2008. "A New Model for Teaching South Asian Languages: The University of Texas Hindi-Urdu Flagship." http://hindiurduflagship.org/assets/pdf/2008_01_01_SALPAT.pdf.

Conclusion

To Advanced Proficiency and Beyond: Charting a New Course in the Twenty-First Century

TONY BROWN AND JENNIFER BOWN

THE BREADTH OF TOPICS in this volume demonstrates the complexity of developing professional levels of foreign language proficiency. Foreign language study in the United States became unprecedentedly urgent in the decade following 9/11. Brecht et al. (this volume) point out an increasing need for learners with Advanced-level proficiency (those at Level 3 and beyond on the Interagency Language Roundtable [ILR] scale), yet show that the US educational system has generally fallen short in its efforts to produce such learners. In response to findings from a nationwide Russian language survey, Davidson and Garas soberly concluded: "The pre-occupation with preparing a generation for the globalized economy of the 21st century in the foreign affairs community of the U.S., and in virtually every world center today from Beijing and Brussels to Moscow, Shanghai, Seoul, and Tokyo appears to have left most of the American heartland untouched" (2009, 17). As early as 2000, Richard Robin identified the "proficiency paradox," bemoaning the lack of real progress in proficiency since the publication of the American Council on the Teaching of Foreign Languages (ACTFL) guidelines and the subsequent "proficiency movement" in language teaching. The data put forward by Brecht et al. in this volume suggest that the traditional four-year college language curriculum is not capable of yielding Advanced second language (L2) speakers without a serious redress of current teaching practices.

Fortunately, as Brecht et al. demonstrate, the push to develop Advanced levels of language proficiency has gained momentum, and technology and curricular design are catching up to the desired standard (this volume). Studies presented in this volume indicate that learners *can* reach ILR Level 3, and even ILR Level 4, with instructors' focused attention on task design, assessment, and individualized instruction. In addition, those studies consider a number of areas for future research

and curriculum development, including (1) challenges of assessment, (2) immersion experiences, (3) individualized instruction, (4) curricular materials, and (5) selective versus open enrollment.

Challenges of Assessment

Findings from this volume underscore a number of challenges associated with adequately measuring culturally inflected language. For example, Christensen points out that the Oral Proficiency Guidelines, inasmuch as they emphasize quantity of speech, reflect Western European norms of engagement; however, in Asian cultures speakers are often expected to say less than more. Such an incongruity is not limited to the testing of Asian languages; rather it reflects a deeper concern about the Oral Proficiency Guidelines themselves. The guidelines require greater quantities of language as a means of differentiating Superior-level from Advanced-level speech, a violation of the Gricean norm[1] for many western and particularly nonwestern languages. In response to this perceived need to align test design and sociolinguistic norms, the American Council on the Teaching of Foreign Languages/Language Testing International (ACTFL/LTI) recently recalibrated the Oral Proficiency Interview (OPI) for Arabic, which is focused on assessing the balance of colloquial and Modern Standard Arabic in the speech of interviewees at the Superior level and above.[2]

Many of the chapters in this volume point to the use of other assessment measures above and beyond the ACTFL OPI and Writing Proficiency Test (WPT). McAloon, for example, notes that the OPI inadequately addresses speakers' professional abilities in their L2. He proposes an alternative method of language assessment using the HR model of 360-degree feedback. A variety of assessments can also be a more accurate evaluation of abilities. Christensen and Bourgerie report on employing a battery of tests, as well as evaluating student papers and presentations. Similarly, Leaver and Campbell describe the use of multiple assessment instruments, though more for instructional purposes than for the purpose of assigning grades or scores. Though the ACTFL OPI and the similar ILR oral exam may remain the gold standard for assessing language proficiency in the United States, additional instruments can provide an enhanced understanding of learners' language skills.

Another potential problem with current proficiency scales is that the test taker's performance is compared and evaluated using the native speaker as the norm. As Jackson (this volume) notes in citing Byrnes (2002), the very notion of the native speaker suggests the existence of a single, idealized register of the target language. In reality, speech communities, and even speakers, are heterogeneous and use language flexibly, depending upon the particular context and task (Hymes 1972). Accordingly, Jackson emphasizes the importance of establishing credibility in professional domains by learning the norms of particular discourse communities, both linguistic and nonlinguistic. For this reason, Jackson particularly welcomes the ILR's recent publication of the skill level descriptions for competence in intercultural communication. As Jackson notes, the skill level descriptions refer "to the content and form of communication, both verbal and nonverbal." These new guidelines emphasize

the multiplicity of contexts that demand culturally appropriate conduct of a truly Advanced-level speaker (this volume).

Similarly, in this volume, Jackson and McAloon question the criterion of the ACTFL Distinguished level and of ILR Level 4+ to speak like a highly articulate, well-educated native, noting that erudite academic speech is not appropriate in all contexts. Indeed, Lantolf and Frawley (1985) question the "privileging" of academic speech in the ACTFL and ILR scales, citing the following admission from the *Foreign Language Proficiency Assessment* manual: "In universities there are Spanish professors who can speak eloquently about Calderon's dramas, but who could never read a Spanish contract or haggle over the terms of a lease" (ETS, n.d.). In keeping with this assertion, assessments for those who wish to use their L2 professionally should evaluate how appropriately the learners use their L2 in their professional fields.

Immersion Experiences

The studies described by Kennedy and Hansen and by Brown et al. are the only ones in this volume that do not involve some kind of in-country immersion experience, either during the program or at its culmination (as is the case with Flagship programs described herein). Rifkin (2005) points to a ceiling effect in traditional university foreign language programs, arguing that achieving even Advanced on the ACTFL scale requires some kind of intensive immersion experience, either overseas or in a program like Middlebury's. Immersion provides learners with extensive input and opportunities for authentic interaction in the target language. Nevertheless, even immersion experiences must be carefully designed to promote proficiency. The programs described herein carefully structure the immersion experience to provide the best opportunities for gain: they often involve direct enrollment (Christensen and Bourgerie, this volume; Davidson, this volume) or extensive content courses (Leaver and Campbell, this volume; Belnap and Abuamsha, this volume). Moreover, the programs provide learners with additional support; Leaver and Campbell describe the role of the accompanying faculty member who assists learners in their preparations for class, and Belnap and Abuamsha describe several significant interventions designed to help learners find conversation partners, develop coping strategies, and learn reading strategies.

Individualized Instruction

Another hallmark of the programs described in this volume is individualized instruction (i.e., one-on-one sessions with native speakers). Christensen and Bourgerie describe the numerous benefits associated with individualized instruction, in particular, the chance to practice speaking in the target language with a native speaker. Traditional classroom environments do not afford opportunities for extended discourse-level speech that defines ILR Level 3 and above. Rifkin (2003) points out that instructors often attempt to circumvent this problem by having students speak in pairs or small groups; however, as he notes, pair or group work does not allow the instructor to listen attentively to each student and to attend to his or her mistakes.

Jackson notes that tailoring learning to students' specific interests, needs, and strengths is a necessary component of developing advanced language proficiency. In a one-on-one setting, a native speaker can carefully note a learner's errors—grammatical, pragmatic, and sociolinguistic—and direct the learner's attention to those errors through explicit instruction. Leaver and Campbell (this volume) point out the importance of explicit attention to language forms for learners trying to get to ILR Level 3 and beyond: "At these levels, coming close is not good enough—not in expression, morphology, syntax, phonetics, intonation, register, cultural behavior, social reaction, internal emotions, or many other areas that give a person away as a non-native speaker." As such, individualized instruction provides an efficient means of overcoming gaps in learners' knowledge.

Curricular Materials

Most textbooks of language-teaching methodology do not adequately address the teaching of foreign languages at professional levels of proficiency. Omaggio Hadley's *Teaching Language in Context* (2000), which for many years has been one of the standard methods textbooks, only provides examples of listening, reading, writing, and speaking activities up to the Advanced level of proficiency and primarily focuses on Novice- and Intermediate-level tasks. This focus on the lower levels of proficiency is not surprising given that enrollment is highest in the beginning levels of language instruction (Modern Language Association Ad Hoc Committee on Foreign Languages 2007) and that the average proficiency level obtained after four years of college instruction is Intermediate High (Brecht et al., this volume).

In the past decade, a small number of volumes have been dedicated to questions of proficiency at the Advanced levels. (Leaver and Campbell cite these seminal works in this volume.) Nevertheless, the availability of materials for teaching at these levels, as well as manuals discussing *how* to teach such learners, is still inadequate to meet the growing demands on the language industry. As Leaver and Campbell point out, professional development for teachers of Professional-level language courses involves a great deal of "bootstrapping."

Becoming familiar with the criteria for Professional-level language represents an important starting point for teachers. The criteria are, in particular, the ACTFL and ILR guidelines that articulate the tasks that learners must be able to accomplish, the text types with which learners must be familiar, and the contexts in which language is used. Instructors hoping to push their learners to Level 3 and beyond must not, therefore, be satisfied with mere narration and description, nor with paragraph-level speech. However, as discussed above, general proficiency descriptors may not adequately represent the full range of language that learners at Levels 3 and 4 must master. It is essential to understand the professional demands that may be made of learners and, accordingly, to tie instruction and assessment to the norms of the particular discourse communities that learners hope to join. Thus, additional detailed ethnographic research is necessary to flesh out the general proficiency guidelines.

Development of materials at the advanced level and beyond should also become a priority. The publication of *Mastering Russian through Global Debate* and *Mastering English*

through Global Debate represents an important first step in furthering the Professional proficiency movement.[3] The increasing availability of language materials on the Internet (see Brecht et al., this volume, for a list) will also facilitate language learning and teaching and provide instructors with readily available authentic materials on which to base instruction. Even more important than availability of authentic materials, however, is the availability of well-designed learning tasks to accompany said material. As new instructors come on the scene, availability of materials will facilitate teaching advanced-level language courses.

A centralized website for cataloguing resources, including materials, scholarly research, and so forth would dramatically improve the practice in this domain. Ultimately, we hope to see this movement culminate in the publication of widely available textbooks devoted to teaching methodology and in the creation of university-level courses offering training in this area.

Much can be learned from examining methodology in the teaching of *second*, as opposed to *foreign*, languages. In a second language situation, individuals are learning the official language of the country in which they reside. Therefore, they are generally exposed to a greater variety of input and have more opportunities for practice. The field of English as a Second Language has long outstripped foreign language education in the United States in terms of the development of Advanced-level materials. This difference in development largely stems from the need to prepare students to participate in university courses. Similarly, instructors of Russian in the Russian Federation, or of German in Germany, often have more experience developing Level 3+ materials than their second language counterparts at universities in the United States. Thus, partnerships should extend beyond national borders to foreign institutions.

Selective versus Open Enrollment

This volume likewise addresses the issue of selective versus open enrollment, that is, whether to admit students to an advanced foreign language course or program based on fulfillment of prerequisites and successful registration or on course- or program-specific criteria. Indeed, the sheer cognitive capacity required to think critically in one's first language (L1) raises important questions about the ability to transfer those skills into an L2 and whether learning a foreign language can serve as a medium for developing cognitive skills possibly lacking in the L1 (Dewey et al., this volume).

Past research suggests that students with a pre-OPI rating of Advanced High benefit considerably more from a course pitched at the Superior level than those with a pre-OPI rating just one sublevel lower (Brown 2009, 546). To a certain degree, prerequisites fulfill a selection function, and students frequently self-select based on what they know about a course or program. Alternatively, admission to some specialized programs, such as the National Flagship Language Program (Davidson, this volume; Christensen and Bourgerie, this volume; Jackson, this volume) or the English for Heritage Language Speakers (EHLS) program (Kennedy and Hansen,

this volume), entails a thorough and rigorous selection process involving proficiency testing in each of the skills.

Likewise, one can turn to scaffolding theory[4] and Lev Vygotsky's zone of proximal development, or the "distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers," to find support for selective enrollment in advanced foreign language courses or programs (Vygotsky 1978, 86). Learner-initiated activities, such as debate and problem solving, require active engagement in negotiation and compromise—both of which represent Superior-level skills (Brown et al., this volume). ACTFL is currently developing reading and listening tests for a number of languages, including Arabic, Chinese, English, French, German, Russian, Spanish, and Turkish. These tests are tied to ACTFL proficiency guidelines using a criterion-referenced approach (Clifford and Cox 2013). Such tests will contribute valuable and affordable information to a learner's overall language profile.

Closing Thoughts

Because the push for developing professional levels of language ability is relatively new, many of the programs and techniques described in this volume represent various stages of a trial- and error-process. Therefore, it is critical that language professionals share ideas. Partnerships between universities and government agencies certainly play a critical role in facilitating interchange of ideas and development of materials. Additionally, professional conferences devoted entirely to teaching languages at Level 3 and beyond expand the existing literature and contribute innovative theories and practices to the field of foreign language education.

Notes

1. Grice's maxim about quantity of speech states that a speaker should be as informative as he or she can be and provide as much information as needed, and no more (see http://www.sas.upenn.edu/~haroldfs/dravling/grice.html).
2. ACTFL's introduction to the new Arabic annotations states: "ACTFL recognizes that the situation for testing Arabic is a special situation and therefore must be dealt with in a special manner" (see http://actflproficiencyguidelines2012.org/arabic/index.php?p=speaking).
3. Georgetown University Press is scheduled to release both volumes in Fall 2014.
4. Although often attributed to Lev Vygotsky, the concept of scaffolding was created by Jerome Bruner (see Bruner and Sherwood 1975).

References

Brown, N. Anthony. 2009. "Argumentation and Debate in Foreign Language Instruction: A Case for the Traditional Classroom Facilitating Advanced Level Language Uptake." *Modern Language Journal* 93: 534–549.

Bruner, Jerome S., and Victoria Sherwood. 1975. "Peekaboo and the Learning of Rule Structures." In *Play: Its Role in Development and Evolution*, edited by Jerome S. Bruner, Alison Jolly, and Katherine Sylva, 277–285. Harmondsworth, UK: Penguin Books.

Byrnes, Heidi. 2002. "The Role of Task and Task-Based Assessment in a Content-Oriented Collegiate Foreign Language Curriculum." *Language Testing* 19: 425–443.

Clifford, Ray T., and Troy L. Cox. 2013. "Empirical Validation of Reading Proficiency Guidelines." *Foreign Language Annals* 46: 45–61.

Davidson, Dan E., and Nadra Garas. 2009. "ACTR Census of Russian Programs in the U.S." *Russian Language Journal* 59: 3–20.

ETS (Educational Testing Service). n.d. F*oreign Language Proficiency Assessment*. Princeton: Education Testing Service.

Hymes, Dell H. 1972. "On Communicative Competence." In *Sociolinguistics*, edited by J. B. Pride and Janet Holmes, 269–293. Harmondsworth, UK: Penguin Books.

Lantolf, James P., and William Frawley. 1985. "Oral Proficiency Testing: A Critical Analysis." *Modern Language Journal* 69: 337–345.

Modern Language Association Ad Hoc Committee on Foreign Languages. (2007). *Foreign Languages and Higher Education: New Structures for a Changed World*. New York: Modern Language Association of America.

Omaggio Hadley, Alice. 2000. *Teaching Language in Context*, 3rd ed. Boston: Heinle and Heinle Publishers, Inc.

Rifkin, Benjamin. 2003. "Oral Proficiency Learning Outcomes and Curricular Design." *Foreign Language Annals* 36: 582–588.

———. 2005. "A Ceiling Effect in Traditional Classroom Foreign Language Instruction: Data from Russian." *Modern Language Journal* 89: 3–18.

Robin, Richard. 2000. "Foreign Language across the Curriculum and the Proficiency Paradox." In *The Learning and Teaching of Slavic Languages and Cultures*, edited by Olga Kagan and Benjamin Rifkin, 29–43. Bloomington, IN: Slavica.

Vygotsky, Lev S. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

This page intentionally left blank

Contributors

Khaled Abuamsha is a Senior Lecturer of Arabic and Academic Director at the Qasid Institute in Amman, Jordan. His research interests include proficiency measurement, methods of foreign language teaching, teacher training, foreign language program evaluation, and curriculum design. He has written more than twenty books and papers and participated in several international conferences pertaining to Arabic teaching and linguistics.

R. Kirk Belnap is a Professor of Arabic in the Department of Asian and Near Eastern Languages at Brigham Young University, where he has been since 1988. Since 2002, he has served as director of the National Middle East Language Resource Center (NMELRC). His research interests include language policy and planning, second language acquisition, and the history of Arabic.

Dana S. Bourgerie is a Professor of Chinese in the Department of Asian and Near Eastern Languages at Brigham Young University. His research interests are in language variation, sociolinguistics, dialect studies, and language acquisition. Dr. Bourgerie is a past president of the Chinese Language Teachers Association and served as the director of the Chinese Flagship Center at BYU since its inception in 2002 until Fall 2013.

Jennifer Bown is an Associate Professor of Russian in the Department of German and Russian at Brigham Young University. Dr. Bown has published several articles on language learning in nontraditional environments, including individualized instruction, foreign language housing, and study abroad. She recently coauthored two textbooks titled *Mastering Russian through Global Debate* and *Mastering English through Global Debate* (Georgetown University Press).

Richard D. Brecht is the founding executive director of the University of Maryland Center for Advanced Study of Language (CASL) and is professor emeritus of Slavic Languages and Literatures at the University of Maryland at College Park. Dr. Brecht is the author of numerous books and articles on language policy, second language acquisition, and Slavic and Russian linguistics.

Tony Brown is an Associate Professor of Russian at Brigham Young University, where he has taught since 2004. Dr. Brown has published articles in the fields of second language acquisition, language policy, and cultural history and, most recently, coauthored two advanced-level language textbooks titled *Mastering Russian through Global Debate* and *Mastering English through Global Debate* (Georgetown University Press).

Christine Campbell is currently Associate Provost, Directorate of Continuing Education, at the Defense Language Institute Foreign Language Center (DLIFLC). Active in the profession for thirty-five years, she has worked in administration, in education, and in the editorial arena. She has published widely in her areas of interest—language assessment and learner variables.

Matthew B. Christensen is Professor of Chinese in the Department of Asian and Near Eastern Languages at Brigham Young University. Dr. Christensen is active in teacher training and is the coauthor of two books on the topic: *Performed Culture: An Approach to East Asian Language Pedagogy* (with Paul Warnick) and *A Performance-Based Pedagogy for Communicating in Cultures: Training Teachers of East Asian Languages* (with Mari Noda). Since Fall 2013, he has served as the director of the Chinese Flagship Center at BYU.

Ray Clifford is an Associate Dean in the College of Humanities at Brigham Young University and director of the Center for Language Studies. Throughout his career, Dr. Clifford has specialized in the science of creating valid, reliable, and fair tests and has published widely on the subject. He has served as president of many national and international language organizations and received numerous honors.

Troy Cox is an Assistant Professor (Professional) of Linguistics at Brigham Young University. Dr. Cox is a certified ACTFL oral proficiency tester, has used his testing expertise as a forensic linguist and in test development projects for various organizations, teaches assessment courses, and has published articles on automatic speech recognition in assessment. He is currently serving as the Associate Director of Research and Assessment in the Center for Language Studies at BYU.

Dan E. Davidson is President and cofounder of American Councils for International Education and Professor of Russian and Second Language Acquisition at Bryn Mawr College, where he has held the rank of full professor (now part-time) since 1983. Dr. Davidson is the author or editor of forty-four books (including the widely used video-based *Russian in Stages Series*) and more than sixty articles in the fields of Russian language, culture, and educational development, including a major twenty-year longitudinal, empirically based study of adult second language acquisition during study abroad.

Dan Dewey, Associate Professor of Linguistics at Brigham Young University, received a PhD in Second Language Acquisition from Carnegie Mellon University. Dr. Dewey teaches courses in linguistics, second language acquisition, language pedagogy, and language testing. He has published in numerous venues, including *Studies in Second Language Acquisition, Foreign Language Annals, Innovation in Language Learning and Teaching,* and *Frontiers: The Interdisciplinary Journal of Study Abroad.*

Dennis L. Eggett is an Associate Research Professor of Statistics and the Director of the Center for Statistical Consultation and Collaborative Research at Brigham Young University. Dr. Eggett has coauthored over one hundred forty journal articles and collaborated with researchers from many different fields ranging from nursing to engineering.

Christa Hansen, as Director of the English for Heritage Language Speakers Program at Georgetown University, is responsible for the design of its academic structure, the development and delivery of a curriculum to meet National Security Education Program (NSEP)–defined programmatic goals, and management of the instructional program. Her publications include articles on testing issues in the Cambridge Applied Linguistics Series, Language Testing, and NAFSA publications.

Frederick H. Jackson is Senior Research Associate at the National Foreign Language Center at the University of Maryland, where he serves on the STARTALK Advisory Board and was Principal Investigator for the federally funded Read Arabic and Read Chinese online materials development projects. Dr. Jackson has authored or coauthored two books and several articles, and he has made numerous presentations at such conferences as ACTFL, TESOL, AAAL, LSA, the Georgetown Roundtable, and the National Conference on the Less Commonly Taught Languages.

Deborah Kennedy, as Associate Vice President, Adult English Language Education, leads the Center for Applied Linguistics' (CAL) adult language acquisition and adult learning programs and activities. A specialist in English for specific purposes, Ms. Kennedy has extensive experience in designing, developing, conducting, and evaluating English language programs for adults in the workplace. She currently directs CAL's English for Heritage Language Speakers (EHLS) program.

Betty Lou Leaver is Associate Provost for Continuing Education at the Defense Language Institute Foreign Language Center (DLIFLC). Dr. Leaver earned her PhD at the Pushkin Institute in Moscow, Russia, in Second Language Acquisition (SLA). She has published a dozen books on SLA, nearly 200 articles, and, with colleagues, won the 1997 best quality research award from the International Performance Improvement Society for investigations into Russian and Arabic cultures.

Cynthia L. Martin is an Associate Professor of Russian at the University of Maryland, where she teaches courses in language, literature, and culture as well as courses that are part of the general education program. Dr. Martin was the Director of the Domestic Russian Flagship Program at UM (2006–2010). Her research interests include second language acquisition and assessment, methodology and pedagogy, as well as contemporary Russian culture and Russian art from the 1950s to the present.

Patrick McAloon holds a PhD in Chinese language pedagogy from the Ohio State University, focusing on development and assessment of professionally useful levels of foreign language ability. Dr. McAloon currently divides his time between teaching Chinese at Ohio State and providing educational experiences in the United States for Chinese organizations and individuals through his firm, SinoConnect.

William P. Rivers is the Executive Director of the Joint National Committee for Languages and the National Council for Languages and International Studies. Dr. Rivers has twenty years of experience in culture and language for economic development and national security, with expertise in research, assessment, program evaluation, and policy development and advocacy. During his career, Dr. Rivers has also taught Russian at the University of Maryland, worked as a freelance interpreter and translator, and conducted field work in Kazakhstan, where he regularly returns to teach.

John P. Robinson specializes in social science measurement and methodology and has studied the evolution of time consumption since the 1960s. Currently, Dr. Robinson coordinates several studies to test competing theories and hypotheses about the Internet's impact on society. Additionally, Dr. Robinson edited two major attitude measures handbooks—*Measures of Political Attitudes* and *Measures of Personality and Social Psychological Attitudes*.

Index