APPLIED LANGUAGE LEARNING

Applied Language Learning

2003 VOLUME 13 NUMBER 1

VOLUME 13 · NUMBER 1

Applied Language Learning

PB 65-02-01
0306408

The mission of Professional Bulletin 65, *Applied Language Learning (US* ISSN 1041-679X and ISSN 1041-6791 for the online version), is to provide a forum for the exchange of ideas and information on instructional methods and techniques, curriculum and materials development, assessment of needs within the profession, testing and evaluation, and implications and applications of research from related fields such as linguistics, education, communications, psychology, and the social sciences.

Applied Language Learning, published semiannually by the Defense Language Institute Foreign Language Center and Presidio of Monterey, presents professional information. The views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material.

JOEL B. HUDSON
ADMINISTRATIVE ASSISTANT
TO THE SECRETARY OF THE ARMY

ERIC K. SHINSEKI
GENERAL
UNITED STATES ARMY
CHIEF OF STAFF

Availability

To access *Applied Language Learning* on the Internet type:
http://www.monterey.army.mil

Additionally, you may obtain the journal on microfilm from ERIC Clearinghouse on Language and Linguistics, Center for Applied Linguistics, 1118 22nd Street, NW, Washington, DC 20037.

Bulk-rate postage is paid at DLIFLC. The basis of official distribution is one copy per training instructor and one per five military linguists.

Postmaster
Send change-of-address information to:
Applied Language Learning
Defense Language Institute Foreign Language Center
Presidio of Monterey, CA 93944-5006

United Parcel Customers
Location is:
Applied Language Learning
Tin Barn
Bldg. 518, Room 7
Presidio of Monterey, CA 93944-5006

Readers
Contact Editor, Dr. Woytak (ATFL-CD-AJ), *Applied Language Learning*
Defense Language Institute Foreign Language Center
Presidio of Monterey, CA 93944-5006
E-mail: AJ@monterey.army.mil
Telephone: (831) 242-5638 DSN: 768-5638
Fax: (831) 242-5850

Cover Design ***Barney Inada***
Al Macks
Assistance in Graphics ***Elaine Koppany***

From the Editor

Reviewers for *Applied Language Learning*

The individuals listed below served as reviewers of manuscripts submitted to *Applied Language Learning* in 2002 and 2003. We express our gratitude for expert service to:

Stephen Butler
Defense Language Institute
Foreign Language Center
Christine M. Campbell
Defense Language Institute
Foreign Language Center
John B. Carroll
University of North Carolina
Marianne Celce-Murcia
University of California
Los Angeles
Ray T. Clifford
Defense Language Institute
Foreign Language Center
Tracey M. Derwing
University of Alberta
Edmonton
Dan Douglas
Iowa State University
Donald Fischer
University of New Mexico
Robert C. Gardner
University of Western Ontario
Luba Grant
Defense Language Institute
Foreign Language Center
Evelyn Hatch
University of California
Los Angeles
John S. Hedgcock
Monterey Institute of
International Studies
Eli Hinkel
Seattle University
J. Ward Keesling
Defense Language Institute
Foreign Language Center
Gordon Jackson
Defense Language Institute
Foreign Language Center
Renee Jourdenais
Monterey Institute of
International Studies
James F. Lee
University of Indiana
Ronald P. Leow
Georgetown University
Paul Nation
Victoria University of Wellington
Thomas Parry
Defense Language Institute
Foreign Language Center
Wilga Rivers
Harvard University
David J. Shook
Georgia Institute of
Technology
Richard Sparks
College of Mount Saint Joseph
Leo Van Lier
Monterey Institute of
International Studies
Swathi Vanniarajan
San Jose State University
Maryann Weber
Missouri Southern State College

Applied Language Learning

Volume 13 Number 1

Articles

News and Views

Reviews

General Information

Applied Language Learning
2003, Vol. 13, No. 1, pp. 1-17

The Effects of Pronunciation Instruction on the Accuracy, Fluency, and Complexity of L2 Accented Speech

Tracey M. Derwing
University of Alberta
and
Marian J. Rossiter
Simon Fraser University

We identified changes in 48 non-native speakers' (NNSs') pronunciation over a period of 12 weeks as a result of the type of instruction they received – Global (primarily prosodic features), Segmental (focus on consonants and vowels), No Specific Pronunciation instruction. Expert judges assessed speech samples from Time 1 and 12 weeks later at Time 2 for accentedness, comprehensibility and fluency. In addition, the judges classified errors as likely to interfere with comprehensibility, as merely bothersome, or as merely salient errors. They also provided an overall impression of the speech samples. The only group that was judged to have improved significantly had received Global instruction. It is argued that although the Segmental group demonstrated fewer phonological errors at Time 2, they paid so much attention to phonological accuracy that they were unable to demonstrate improvement in other areas. The Global group's productions were perceived to have improved, apparently at little cost to attentional resources required for non-prosodic aspects of speech. Implications for pronunciation instruction are drawn.

Recently a great deal of attention has been placed on the effects of a focus on form in the second language (L2) classroom (see reviews by Long and Robinson, 1998; Spada, 1997). Spada concludes from her extensive review of classroom and laboratory studies that form-focused instruction has a beneficial effect on L2 learning. Although there appear to be differential effects, depending on the nature and extent of L2 learners' exposure to the forms in question after instruction has taken place, where exposure is frequent the forms taught appear to be maintained (e.g., Spada & Lightbown, 1993). While several studies have examined the role of instruction focused on well-defined grammatical structures, (e.g., Pienemann, 1985, 1989; White, 1998; White, Spada, Lightbown & Ranta, 1991), there is, as yet, very little corresponding literature that deals with the *effect of instruction* on specific forms in L2 pronunciation. Materials developers and teachers initially used a contrastive analysis approach

to pronunciation instruction, identifying segmental differences between L1 and L2 (e.g., Nilsen & Nilsen, 1973). Emphasis was placed on both discrimination and production skills, particularly of individual sounds embedded in single syllable words. Many materials still in use today focus primarily on segments, with very little attention given to prosodic factors (e.g., CALI, 1998; English Computerized Learning, 1996; Orion, 1997).

In recent years, more attention has been paid to prosodic aspects of pronunciation; advocates of a global approach to instruction (e.g., Celce-Murcia, Brinton & Goodwin, 1996; Firth, 1992; Gilbert, 1993; Pennington & Richards, 1986) have argued that factors such as rate, intonation, and rhythm are important contributors to comprehensibility. Indeed, several studies suggest that nonnative prosody does affect the comprehension of the native listener. For instance, Lennon (1990), in a study of 4 L2 speakers who attended university in an English environment (but who received no explicit pronunciation instruction), asked NSs to judge picture descriptions recorded at the beginning and end of the NNSs' stay. The judges gauged the L2 learners' later renditions to be more fluent. Subsequent analysis of the NNSs' productions indicated significant improvement in speech rate, in filled pauses per T-Unit, and in the percentage of T-Units followed by a pause. Lennon concluded that "it is fluent delivery in performance that is probably the overriding determiner of perceived oral proficiency" (1990, p. 391).

Anderson-Hsieh, Johnson, and Koehler (1992) have also argued for the primacy of prosodic factors. They conducted a study of 60 nonnative speakers, from 11 language backgrounds, reading a passage from the SPEAK test. When three ESL instructors rated the oral productions for pronunciation comprehensibility and acceptability, the authors found that the ratings correlated very highly ($\underline{r} = 0.9$) with an independently measured overall prosody score (based on impressionistic ratings of two transcribers and the first author). Indeed, it was the prosodic factor that exerted the strongest effect as opposed to phonemic deviancies.

In an investigation of task performance by 12 learners of French before and after residence abroad, again with no pronunciation instruction, Towell, Hawkins, and Bazergui (1996) attributed gains in fluency to "an increase in the length and complexity of the linguistic units that are uttered between pauses. This suggests that what has changed is the rapidity with which syntactic and discourse knowledge can be accessed for on-line speech production" (pp. 112-113). Their study consisted of several before and after measures (e.g., speaking rate, articulation rate, phonation/time ratio, average length of pause, and mean length of run); however, the experiment did not include NS judgments.

Further support for the influence of prosody on NS comprehension is provided by Munro and Derwing (1998), who have shown that some L2 speakers may be perceived to be more comprehensible at rates that are somewhat faster than those they typically produce.

Although the studies reviewed here all point to the importance of prosodic factors for comprehensibility, the effects of instruction were not measured. Wennerstrom (1998), in an innovative study of 18 Mandarin international teaching assistants (ITAs), found that the judgment scores of overall comprehensibility of oral production (using scales closely related to the SPEAK) were correlated with successful use of the paratone, "a phenomenon whereby speakers expand their pitch range at the beginning of a new topic and compress it at the end" (1998, p. 5). Wennerstrom suggests that some of the ITAs likely benefited from global speaking instruction, including intonation patterns in English.

Of course, there are factors other than pronunciation which affect comprehensibility, including grammar (Varonis & Gass, 1982), discourse markers (Tyler, 1992; Williams, 1992), and lexical specificity (Tyler, 1992). Nonetheless, there is a need to determine which aspects of pronunciation are most crucial to intelligibility and which forms of instruction are most effective.

Role of Pronunciation Instruction

Assuming that a focus on form is necessary to alter NNSs' pronunciation patterns, the question arises "On which forms should we focus?" Numerous studies have documented changes in L2 speakers' pronunciation (e.g., Elliott, 1997; MacDonald, Yule & Powers, 1994; Perlmutter, 1989), but without empirical measures of the *effects* of those changes on listeners, there is still no clear indication of where instructors should place their emphasis. Intuitively appealing arguments have been put forward, such as the role of functional load (factors such as degree of sound similarity, frequency and distribution of phonemes). Brown (1991), for example, argues that some phonemes are more important than others, depending on the number of minimal pairs that they distinguish. There is, however, a need to determine which features of pronunciation actually influence intelligibility. Since Abercrombie first argued that most "language learners need no more than a comfortably intelligible pronunciation" (1949, p. 120), it has been generally agreed that intelligibility should be the foremost goal of any pronunciation class (e.g., Celce-Murcia et al., 1996; Firth, 1992; Gilbert, 1993; Morley, 1991, 1994). Munro and Derwing (1994) and Derwing and Munro (1997) have suggested a hierarchy of importance, based on NS judgments of NNS speech: intelligibility (whether or not a particular form interferes with actual understanding) is paramount, then comprehensibility (the perceived degree of difficulty involved in understanding a given form), with accentedness the least important consideration.

Derwing, Munro, and Wiebe (1997) undertook a study in which ESL learners, who had been in an English-speaking environment for an average of 10 years, participated in a speaking improvement course that focused on global speaking strategies (e.g., stress, rhythm, intonation), as advocated by Firth (1992) and Gilbert (1993). Thirty-seven native listeners transcribed speech samples (true/false sentences) taken at the beginning and end of the 12-week course in order to assess the ESL learners' intelligibility. In addition, the samples were rated for ease of understanding (comprehensibility) and degree of accentedness. Overall, there was a significant improvement in

intelligibility in all sentences and better ratings over time of comprehensibility and accentedness for the true sentences. This study indicated that long-term language learners can alter their pronunciation in a reading task, but, because transfer to spontaneous speech was not measured, the effectiveness of the instruction for free production could not be assessed.

One study which did investigate transfer to extemporaneous speech was that of Elliott (1997), who conducted a study in which American learners of Spanish were asked to mimic individual words and sentences, to read isolated words aloud, and to perform a short picture description task at the beginning and end of a semester of instruction. During the instructional period, students received practice on several segments believed to be difficult for nonnative speakers of Spanish. The post-test indicated that these students were able to improve their productions in the mimicking and reading tests, but that there was no significant improvement in their spontaneous productions.

Following an exploratory study by MacDonald, Yule, and Powers (1994), who compared different types of pronunciation correction over a very short period (2 days), Derwing, Munro, and Wiebe (1998) evaluated the effects of three distinct forms of instruction (No Special Pronunciation [NSP] instruction, Segmental instruction, and Global instruction) on the oral production of three groups of learners of English as a second language (ESL). Forty-eight learners provided sentences which they read aloud, and extemporaneous narrative descriptions of a standard picture story (Derwing & Munro, 1997; Munro & Derwing, 1994, 1995) at the beginning (Time 1) and end (Time 2) of a 12-week period of instruction. The sentence data were presented to 48 NS listeners, who rated before and after samples for comprehensibility (very easy to understand – impossible to understand) and accentedness (no accent – very strong accent), on 9-point scales. They found that both the Segmental and the Global groups improved significantly on the comprehensibility ratings, but that, although all three groups were rated as being less accented at Time 2, the Segmental group improved to a significantly greater extent than the others.

In the same study, excerpts of 45 seconds from the picture story narratives were rated by six NS listeners, all of whom had extensive ESL teaching experience, on 9-point scales for accentedness, comprehensibility and fluency (NS-like fluency – extremely dysfluent). The analyses of the three judgment tasks showed that the NSP and Segmental instruction groups made no significant improvement; however, the Global instruction group improved significantly in comprehensibility and fluency over time. Derwing et al. attributed these results to the differential effects of instruction. They hypothesized that the prosodic training transferred to the extemporaneously produced narratives, despite the increased cognitive demands of creating the utterances. The segmental improvement in the sentences which students read aloud did not transfer when they were required to change the focus of their attention from pronunciation to lexical and syntactic choices. However, the authors didnot examine the students' productions to see what changes actually took place over time.

We undertook the present investigation in an attempt to determine how the improvements in comprehensibility and fluency reported by Derwing et al. (1998) for the narrative task were manifested in students' oral productions before and after pronunciation instruction. In this study we report a second phase of the narrative experiment, in which we asked the same six judges to listen to the excerpts again, this time with a transcription in front of them. We asked the listeners to identify errors and code them as to whether they interfered with comprehensibility, and whether they were bothersome or merely salient. In addition, the judges were asked to provide their overall impressions of the speakers' productions. We also examined the phonetically transcribed data ourselves for error types (details below).

-

Method

Participants

ESL Students

The NNSs were 48 adult learners registered in full-time ESL classes in a local college. They were all of intermediate proficiency and ranged in age from 18 to 44, with a mean of 31.7 years. They had spent from seven months to 15 years in English-speaking Canada, with a mean of 3.3 years. All had started learning English as adults. This was a typical program with a typical mix of students in the Alberta immigrant context. Sixteen learners were assigned by the institution to each of three classes, which were balanced as closely as possible for English language proficiency, native language, gender, length of time in Canada, and age upon arrival (see Table 1).

Table 1. *Speaker Group Information*

	NSP Group	Global Group	Segmental Group
L1			
E. European	8	10	9
E. Asian	4	2	1
Spanish	2	3	3
Other	2	1	3
Gender	6 male 10 female	4 male 12 female	6 male 10 female
Time in Canada (years) M = 3.86	M = 3.19	M = 2.71	
Age at Arrival (years)	M = 29.05	M = 26.98	M = 27.45

One of the investigators met with the learners to explain the project and answer any questions they might have had. All of the learners volunteered to participate, although alternate arrangements had been made for that portion of regular class time if they declined. Their commitment to the study was evident in the questions they asked and in the enthusiasm they showed during data collection sessions, both at the beginning and end of the study. The majority also requested a summary of results at the end of the project.

Teachers

The principal investigator met with the ESL program director at the college before the beginning of term to discuss the study. The names of three interested teachers were put forward at that time. The principal investigator met with these teachers and explained the nature of the research project. She indicated that both the Global and Segmental groups were likely to show improvement in pronunciation during the course of the study, and suggested that teachers choose the instructional condition with which they felt most comfortable. Two teachers with academic coursework in linguistics and TESL offered to teach the experimental groups; the third, who had no formal education in linguistics, chose the NSP class. The teachers spent time with the first author before the beginning of the course, negotiating instructional materials and procedures, and again at mid-point in the term to report on their reactions and those of their learners. At this point they confirmed their willingness to continue with the experiment until the end of the course.

Listeners

A small number of expert judges were recruited to participate in the listening part of the experiment. The listeners were six female ESL teachers who had extensive experience in instructing learners from a wide variety of native language backgrounds and proficiency levels. All were native speakers of English, ranging in age from 39 to 52 years, with self-reported normal auditory acuity. The first author participated as one of the six judges; because the speech samples were randomized as to condition and time, it is unlikely that her ratings could have been influenced by her knowledge of the task.

Procedure

Instruction

All three groups of students (NSP instruction, Segmental instruction, Global instruction) were registered in an ESL program for 20 hours per week. All classes followed a skills-based communicative curriculum of listening, speaking, reading, writing, and grammar. The control group followed the regular curriculum (with no pronunciation instruction) for four hours a day. The experimental groups spent an average of 20 minutes per day on pronunciation instruction (leaving three hours and 40 minutes for the standard curriculum). Thus, in total, the experimental groups received 20 hours of pronunciation instruction over the course of the 12 weeks. The students in the Segmental

condition used pronunciation materials in the language laboratory and participated in teacher-fronted pronunciation activities focusing on the production of individual phonemes (Corbett, 1992): identification and discrimination of individual sounds, and repetition exercises using minimal pairs. Emphasis for the students in the Global condition was on suprasegmentals: word and sentence stress, intonation and rhythm, projection, and speech rate. This class used commercial materials such as *Jazz Chants* (Graham, 1978) and *Sounds Great* (Beisbier, 1995) as a basis for both prescribed and innovative instructional techniques. Both instructors employed listening discrimination tasks, production tasks, and a combination of the two; the only difference was the nature of the content: segmental in one case and prosodic in the other. One of the researchers met with the teachers in the experimental groups at the outset and provided them with suggestions and materials for use in their classes. She kept in contact with them over the course of the study, and met with them again at the six-week point and after the second test period. The teachers confirmed that they had adhered to the guidelines they had been given throughout the instruction period.

Collection of listening stimuli

The ESL students provided speech samples for investigators in the college language laboratory on a Sony Console LLC-9000 system. They did this near the beginning of the study (Time 1) and again 11 weeks later (Time 2). We gave the speakers the same amusing cartoon story consisting of eight frames and asked them to describe the events depicted. We considered it unlikely that the learners would recall their productions of the story at Time 2. They were not told in advance that they would see the story again. However, if they did benefit from familiarity with the pictures, the degree of familiarity was held constant across groups; thus any differences could not be attributed to the stimuli. Neither the teachers nor the students had access to the stimuli during the instruction period.

Stimulus preparation

The speech samples collected were digitally recorded and saved as audio files on a Macintosh computer at 22kHz with 16-bit (i.e., CD) resolution. They were re-recorded randomly on tapes for use during two professional judgment sessions. Excerpts of 45 seconds' duration were selected from the beginning of each of the extemporaneous narratives, one from Time 1 and another from Time 2.

The judges listened to the excerpted speech samples on two separate days for a total of 5.5 hours, in two sessions interspersed with breaks to mitigate fatigue. The listeners heard three warm-up items at the beginning of each session.

After each speech sample was played, the judges rated it for *comprehensibility*, on a scale from 1 (very easy to understand) to 9 (impossible to understand), and for *accentedness*, on a scale from 1 (no accent) to 9 (very strong accent). Judges also assigned a *fluency* rating from 1 (very fluent) to 9

(extremely dysfluent); fluency was assessed on the basis of temporal factors, as opposed to proficiency level. (Judges were asked to interpret fluency in terms of rate of speech and hesitation phenomena.)

Immediately after the judges had assigned ratings on a sample for a speaker, they were given a written transcript of the same speech sample, which was played for them again. They were instructed to circle errors and classify them, using a numbering scheme, as comprehensibility errors = 1 (likely to cause problems for understanding), bothersome = 2 (annoying, irritating), and salient errors = 3. (The raters did not appear to have difficulty assigning error types in the time available.) Salient errors were defined as noticeable, but did not interfere with comprehension and were not judged to be annoying. The raters were then asked to write their overall impressions of the quality of the speech sample in a few words.

Results

Inter-rater reliability Pearson coefficients (r) for the six judges on the scalar ratings were as follows: comprehensibility, .72; accentedness, .69; and fluency, .74. In this study, five judges' identification of errors by type were in the same direction, i.e., they identified more salient errors than bothersome ones, and more bothersome ones than comprehensibility errors. One judge noted considerably more bothersome than salient errors, and more salient than comprehensibility errors, but in debriefing she indicated that she had been sensitized to error identification through extensive placement testing of ESL students in a pronunciation program.

Analyses of the data were conducted to determine what factors had influenced the significant improvement in comprehensibility and fluency ratings in the narratives for students in the Global condition (see Table 2).

Table 2. *Professional Judges' Ratings of NNSs' Narratives*

_____	NSP	Segmental	Global
Comprehensibility	No change	No change	Improved*
Accentedness	No change	No change	No change
Fluency	No change	No change	Improved*

*p <.05

Analyses

The total number of words in each entire narrative was calculated. A repeated measures ANOVA indicated no interaction and no significant difference between groups; however, all groups had a significant decrease in words used to describe the pictures between Time 1 and Time 2, F (1, 45) = 4.94, p = .0314. That is to say, overall, students were more efficient the second time around in describing all the pictures.

We also calculated the number of words in the 45-second speech samples. A repeated measures ANOVA, with condition as the between factor and time as the within factor, showed no interaction and no significant differences between groups, but a highly significant increase in number of words in a 45-second period between Time 1 and Time 2, F (1, 45) = 23.52, p <.0000; i.e., all students' productions were faster.

Rate of speech in syllables per second was computed for each speech sample. A repeated measures ANOVA showed a significant improvement over time for all conditions, F (1, 47) = 20.775, p=.0000, but no interaction and no difference between groups.

All the judge-identified comprehensibility, bothersome, and salient errors were counted and categorized as errors of phonology, morphology, syntax, semantics, pausing, or repetition (see Table 3). An "other" category was created for ambiguous or uninterpretable errors, of which there were very few. So few specific prosodic errors were identified that we did not create a category for them. We attribute the judges' reluctance to circle what they perceived to be prosodic errors to the fact that prosody is not represented orthographically in English. Prosodic factors were overwhelmingly cited in the overall impressions, however, as described below.

Table 3. *Classification of Identified Errors*

Error type	Examples
Phonological	substitution ("zen" for "then"); omission ("fen" for "friend") ; insertion ("esnow" for "snow")
Morphological	word formation ("foots" for "feet"); verb tense or form ("He get" for "He got/gets")
Syntactic	errors in sentence formation at the word level ("reached to the forest" for "reached the forest"); omission of auxiliary verbs ("He kissing" for "He is kissing")
Semantic	incorrect lexical choice ("They matched the fire" for "They lit the fire"); word coinage ("snackwich" for "sandwich"); production errors leading to difficulty in interpretation of meaning ("sleeves" for "tracks")
Filled pauses	"um", "uh"
Repetition	words, phrases, or clauses repeated with or without modification in syntax, morphology, or word order
Prosodic	stress, intonation, vowel length

A hierarchy of error gravity was established to facilitate classification; if an error had been assigned multiple codes, an error involving comprehensibility was considered to be the most serious, followed by bothersome, followed by salient. Thus, if an error was coded as both affecting comprehensibility and bothersome, it was assigned to the comprehensibility category. Of all the identified errors, 43.7% were classified as salient, 32.0% as bothersome, and only 24.1% as comprehensibility errors. Error counts were then normalized for the samples by dividing by the number of words for each 45-second segment of speech. When we looked at the normalized errors, there were some sizable differences across the three error categories. Over conditions, times, and judges, comprehensibility errors tended to be mostly phonological, bothersome errors were mostly due to filled pauses, and salient errors were predominantly morphological.

We decided to restrict further analyses of the data to phonological errors and filled pauses, as these were the types most clearly related to pronunciation instruction. The data were normalized by dividing each speaker's phonological error counts by the number of words in the 45-second excerpts produced in either the pre or post conditions. We calculated the degree of improvement between normalized errors at Time 1 and Time 2 for each of the three groups using Wilcoxon signed rank tests (Hatch and Lazaraton, 1991, p. 297). Although there was no significant improvement in phonological differences from Time 1 to Time 2 for the NSP and Global groups, the Segment group made significantly fewer errors at Time 2 (see Table 4). The Wilcoxon test on filled pauses for each group from Time 1 to Time 2 did not show a significant difference in any of the three groups, although the Global group's score most closely approached significance.

Table 4. *Wilcoxon Signed Rank Tests of Improvements in Phonology and Pausing in Normalized Errors From Time 1 to Time 2*

Error Type	NSP	Global	Segmental
Phonological	$Z = .625$, $p = .532$	$Z = .233$, $p = .816$	$Z = 1.196$, $p = .049$
Filled Pauses	$Z = 1.099$, $p = .272$	$Z = 1.726$, $p = .084$	$Z = .245$, $p = .807$

An investigation of actual and perceived errors was also carried out. All ninety-six 45-second speech samples were phonetically transcribed. Using a stratified proportional random sampling technique, we selected six different participants, two from each condition, and analysed both their Time 1 and Time 2 performances, for a total of 12 separate transcripts. These speech samples were coded for actual errors, using the phonetic transcriptions and audiotapes. This error count was then compared to the number of errors identified by the judges. The numbers of actual and perceived errors were proportional; the listeners identified errors within all categories of actual errors, but in every case the judges identified fewer errors than actually occurred (they identified

67% of all 235 errors).

The speech samples were also analyzed for number of propositions at Times 1 and 2 as a measure of complexity. Using the criteria proposed by Tomlin (1984), a proposition was defined as:

> a semantic unit consisting of a predicate plus its arguments about which a truth value can be obtained…[and] an utterance in the text was counted as realizing a proposition if and only if it is realized by a full clause or by a partial clause for which missing arguments are readily recoverable. (p. 122)

A repeated measures ANOVA with instruction type as the between factor and time as the within factor revealed a significant interaction of instruction type and time, $F(2, 46) = 3.436$, $p = .04$. Tests of simple effects indicated that there were no group differences in the number of propositions at Time 1, but both the Global and NSP groups' narratives had a significant increase in number of propositions at Time 2, $F(3, 45) = 11.132$, $p = .002$ and $F(3, 45) = 26.936$, $p = .000$ respectively.

We categorized the judges' overall impressions as prosodic (pausing, stress, intonation, repetition, vocal fry[creaky voice]), phonological (segmental), and morphological (grammatical) categories. The prosodic comments far outweighed those related to phonological and morphological errors: of 190 total comments at Time 1, 160 (84%) were negative prosodic impressions; at Time 2, 156 of the 197 (79%) comments overall were also negative about the prosody of the samples. All judges demonstrated a similar pattern in their impressions as exemplified in one sample, which elicited the following comments from the six judges: "intonation"; "long pauses, lack of cohesion, pause problem"; "rate of speech – slow"; "redundancy"; "halting speech interrupts flow"; "long pauses mid-sentence". Because the comments regarding phonological and morphological impressions were so few, they could not be analysed. We conducted a repeated measures ANOVA on the negative prosodic comments from Time 1 and Time 2 for each of the groups. There was a significant interaction of time and group ($F[2, 190] = 7.417$, $p < .001$), such that the negative prosodic comments for the Global group decreased by 8% at Time 2, while they increased by 8% for the Segmental group. There was no appreciable change in the judges' impressions of the Control groups' prosody over time (see Figure 1).

Discussion

Research in the area of second language acquisition and cognition (Foster & Skehan, 1996; Lennon, 1990; Skehan & Foster, 1997; VanPatten, 1990) suggests that human information processing resources are limited and that, where second language learners have difficulty attending to all aspects of language, trade-offs must be made. Lennon (1990), for instance, noted that fluency and syntactic complexity show an inverse relationship in language learners' pro-

ductions. In a study in which they measured oral fluency, accuracy and syntactic complexity, Skehan and Foster concluded that learners cannot attend equally to all three performance aspects. In their words, "[a]chieving more highly in one seems mostly to be at the expense of doing well on the others" (1997, p. 207). In a similar vein, VanPatten (1990) has argued that beginning and intermediate level learners cannot attend to both form and and meaning at the same time.

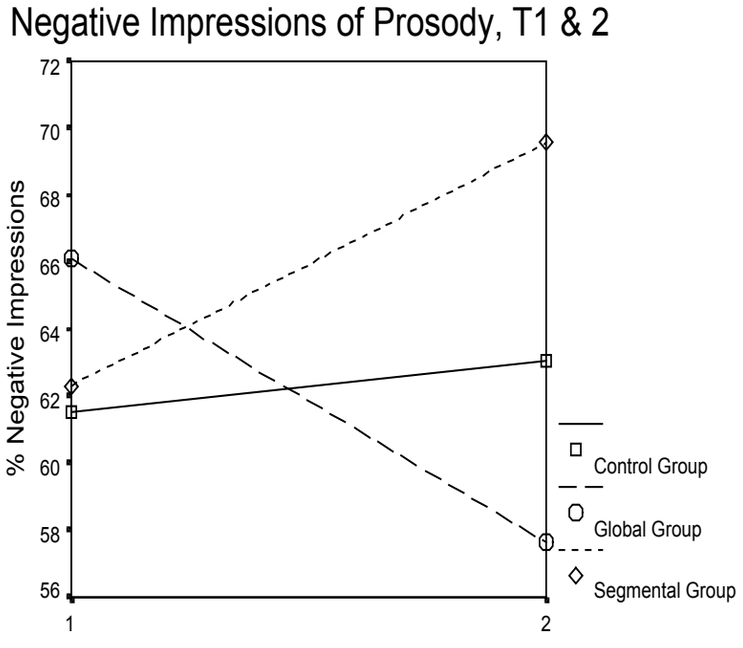

Figure 1.

In the present study, we examined factors of accuracy, fluency, and complexity in the narratives. Results suggest that learners in the Segmental group, whose pronunciation instruction had focused specifically on the identification, discrimination, and production of discrete sounds and minimal pairs, devoted more attention during task processing at Time 2 to accuracy of pronunciation, as evidenced by the reduction of phonological errors over time. This focus evidently left fewer attentional resources available to allocate to complexity and fluency. It is interesting to note that the speakers in the Segmental group showed no significant improvement in the total number of propositions produced in a 45-second speech sample at Time 2. Furthermore, the prosodic aspects of their speech were considered by the judges to be worse at Time 2 than at Time 1. In other words, the students appeared to have been focusing on phonological form at the expense of processing resources.

Learners in the Global condition, on the other hand, seem to have focused more on fluency skills developed during the course, as measured by the reduction of pausing and the number of negative impressions reported by the judges. As context is important to prosodic instruction, they may have attended more to content complexity in the narrative task, as indicated by the significant increase in number of propositions. The results suggest that the activities in the Global pronunciation instruction condition appeared to have promoted an automaticity that resulted in improved fluency and the release of attentional resources for other purposes. As has been pointed out (McLaughlin, Rossman & McLeod, 1983), once a skill becomes automatic, it requires little or no attention. The difference between the two experimental groups here may be attributed to the fact that segmental production has a strong motor component that may necessitate greater attention to output than prosodic features demand.

The NSP group also showed a significant increase in the number of propositions produced at Time 2. This improvement is not altogether surprising, given that the ESL instruction that all three groups received in their regular curriculum for the duration of the experiment was communicative in nature. Although the Segmental group received a great deal of communicative instruction, the focus on phonology in their pronunciation classes may have caused a trade-off in which concern for the accurate production of consonants and vowels interfered with syntactic processing.

Conclusion

In this study we identified changes in NNSs' pronunciation over a period of 12 weeks as a result of the type of instruction received –- Global, Segmental, No Specific Pronunciation instruction. We examined these factors in relation to the Time 1 and Time 2 judgments of comprehensibility and fluency that were assigned to the speech samples by six expert judges. Moreover, the judges classified errors as ones that they perceived would interfere with intelligibility, bothersome errors, or merely salient errors; they also provided an overall impression of the speech samples. An error analysis indicated that although the Segmental group made notable gains in phonological accuracy, this did not translate into improved ratings: the only group that showed overall improvement over time was the Global instruction group, whose ratings of comprehensibility and fluency were significantly higher at Time 2. Furthermore, they had many fewer filled pauses, they produced significantly more propositions at Time 2 than at Time 1, and the judges' impressions of the Global group's prosody improved over time.

We do not advocate eliminating segment-based instruction altogether, but, if the goal of pronunciation teaching is to help students become more understandable, then this study suggests that it should include a stronger emphasis on prosody. In order to ensure that teachers' and students' time in the pronunciation classroom is well spent, the factors most crucial to improved production must be empirically identified. We think that this is especially

important, given that the majority of pronunciation errors identified in this study were judged to be neither annoying nor to interfere with comprehensibility. Several research directions are suggested by this study, including the description of developmental patterns in pronunciation, the effectiveness of specific activities in pronunciation instruction, and the ongoing investigation of factors that affect comprehensibility.

Finally, we recognize the social nature of interaction; reactions to accented speech are affected by many factors other than comprehensibility. Although we have not addressed these issues here, we certainly acknowledge their importance. In the final analysis, our students must communicate in real-life contexts with native speakers and other nonnative speakers. Our ultimate goal as teachers and researchers must be to reconcile the many factors that influence successful communication, taking into account the people with whom our students come into contact.

References

Abercrombie, D. (1949). Teaching pronunciation. *English Language Teaching, 3*, 113-122.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of non-native pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*, 529-555.

Beisbier, B. (1995). *Sounds great: Intermediate pronunciation and speakin*g (Book 2). Boston: Heinle and Heinle.

Brown, A. (1991). Functional load and the teaching of pronunciation. In A. Brown (Ed.), *Teaching English pronunciation: A book of readings* (pp. 211-224). London: Routledge. (Reprinted from *TESOL Quarterly, 22*, 593-606).

CALI, Inc. (1998). ELLIS master pronunciation. Provo, UT: Author.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.

Corbett, K. (1992). *Better English pronunciation for immigrants: Consonant and vowel corrections*. Calgary, AB: Coralex.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 1*, 1-16.

Derwing, M. J., Munro, M. J., & Wiebe, G. E. (1997). Pronunciation instruction for fossilized learners: Can it help? *Applied Language Learning, 8*, 185-203.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*, 393-410.

Elliott, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania, 80*, 95-108.

English Computerized Learning, Inc. (1996). *Pronunciation power*. Edmonton, AB: English Computerized Learning, Inc.

Firth, S. (1992). Pronunciation syllabus design: A question of focus. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 173-183). Oxford: Oxford University Press.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*, 299-323.

Gilbert, J. B. (1993). *Clear speech: Pronunciation and listening comprehension in North American English. Student's Book* (2nd ed.). New York: Cambridge University Press.

Graham, C. (1978). *Jazz chants: Rhythms of American English for students of English as a second language*. New York: Oxford University Press.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*, 387-417.

Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15-41). Cambridge: Cambridge University Press.

MacDonald, D., Yule, G., & Powers, M. (1994). Attempts to improve English L2 pronunciation: The variable effects of different types of instruction. *Language Learning, 44,* 75-100.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning, 33*, 135-158.

Morley, J. (1991). The pronunciation component of teaching English to speakers of other languages. *TESOL Quarterly, 25*, 481-520.

Morley, J. (Ed.). (1994). *Pronunciation pedagogy and theory: New views, new dimensions*. Alexandria, VA: TESOL.

Munro, M. J., & Derwing, T. M. (1994). Evaluations of foreign accent in extemporaneous and read material. *Language Testing, 11*, 254-66.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning, 45*, 73-97.

Munro, M. J., & Derwing, T. M. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning, 48*, 159-182.

Nilsen, D. L. F., & Nilsen, A. P. (1973). *Pronunciation contrasts in English.* New York: Regents.

Orion, G. F. (1997). *Pronouncing American English: Sounds, stress, and intonation*. Pacific Grove, CA: Heinle & Heinle.

Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly, 20*, 207-25.

Perlmutter, M. (1989). Intelligibility rating of L2 speech pre- and postintervention. *Perceptual and Motor Skills, 68*, 515-521.

Pienemann, M. (1985). Learnability and syllabus construction. In K. Hyltenstam & M. Pienemann (Eds.), *Modeling and assessing second language acquisition* (pp. 23-75). Clevedon, Avon: Multilingual Matters.

Pienemann, M. (1989). Is language teachable? *Applied Linguistics, 10*, 52-79.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, *1*, 185-211.

Spada, N. (1997). Form-focussed instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching, 29*, 1-15.

Spada, N., & Lightbown P. M. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition, 15,* 205-224.

Tomlin, R. S. (1984). The treatment of foreground-background information in the on-line descriptive discourse of second language learners. *Studies in Second Language Acquisition, 6*, 115-42.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of English. *Applied Linguistics, 17*, 84-119.

Tyler, A. (1992). Discourse structure and the perception of incoherence in international teaching assistants' spoken discourse. *TESOL Quarterly, 26*, 713-729.

VanPatten, W. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition, 12*, 287-301.

Varonis, E. M., & Gass, S. (1982). The comprehensibility of nonnative speech. *Studies in Second Language Acquisition, 4*, 114-136.

Wennerstrom, A. (1998). A study of Chinese speakers of English. *Studies in Second Language Acquisition, 20*, 1-25.

White, J. (1998). Getting the learners' attention: A typographical input enhancement study. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 85-113). Cambridge: Cambridge University Press.

White, L., Spada, N., Lightbown, P. M., & Ranta, L. (1991). Input enhancements and second language question formation. *Applied Linguistics, 12*, 416-432.

Williams, J. (1992). Planning, discourse marking, and the comprehensibility of international teaching assistants. *TESOL Quarterly, 26,* 693-711.

Acknowledgements

We are very grateful to the staff and students in the ESL department at Grant MacEwan College for their participation in this study. We also thank the judges for their contribution. Murray Munro's advice was very welcome, as always. The assistance of Grace Wiebe, Kama Jamieson, and Martha Gibson was much appreciated. Thanks to Bruce Derwing and Leila Ranta for their comments. Finally, we would like to thank SSHRC and Alberta Advanced Education and Career Development for providing funding for this research.

Authors

TRACEY M. DERWING, Professor, Educational Psychology, University of Alberta, 6-102 Education North, University of Alberta, Edmonton, Alberta T6G 2G5, Canada. Specializations: NS-NNS intelligibility, L2 pronunciation, ESL pedagogy.

MARIAN J. ROSSITER, Post-Doctoral Fellow, Department of Linguistics, Simon Fraser University, 8888 University Drive, Burnaby V5A 1S6, Canada. Specializations: Second language acquisition, classroom-based research, language pedagogy, motivation.

Applied Language Learning
2003, Vol. 13, No. 1, pp.19-32

Utilizing Non-Cognitive Predictors of Foreign Language Achievement

Daniel L. McCollum
Pennsylvania State University

In the present study, a scale to measure self-efficacy and learning-goal orientation in the foreign language-learning domain was constructed. In addition, the relationships between self-efficacy, learning-goal orientation, and foreign language achievement were investigated. Previous research on self-efficacy within a foreign language-learning context is limited and research on learning-goal orientation in the foreign language achievement domain does not exist. One-hundred-twenty- eight German language students at Pennsylvania State University completed the new instrument – the Measure of Foreign Language Achievement Potential (MOFLAP). The MOFLAP was found to be highly reliable and evidence of the instrument's construct, discriminant and predictive validity was found. Individually, self-efficacy is highly predictive of course grades and learning orientation is moderately predictive of course grades. The MOFLAP can be used to explore these constructs in the foreign language achievement context and to assess students' foreign language achievement potential.

Across the past decade, researchers of foreign language-learning have become further engaged in the task of identifying cognitive and non-cognitive variables that predict success in foreign language learning (i.e., Dornyei, 1994; Tremblay & Gardner, 1995; Lett & O'Mara, 1990). In a recent study, Sparks, Javorsky, Patton, and Ganschow (1998) identified three cognitive variables that, when combined, were moderately predictive of foreign language achievement -- verbal memory, phonological coding/recoding, and cognitive speed. The authors did not include non-cognitive variables in their study; rather, they discounted the predictive worth of non-cognitive variables, positing that in past research non-cognitive variables have been poor predictors of foreign language achievement.

In light of the claim of Sparks et al. (1998) and a resurgence of interest in the role of non-cognitive variables in foreign language learning (see Dornyei, 1994), the purpose of the present research is to investigate the influence non-cognitive factors in foreign language achievement. This was achieved through a brief review of previously studied, non-cognitive variables and an investigation of the relationships between additional non-cognitive variables

and foreign language achievement. The new variables are studied via the development of a new scale.

Previous Research on Non-Cognitive Factors

For over four decades, a social psychological theory of foreign language achievement has dominated foreign language learning research. Gardner and Lambert (1959) posited the foundation of the theory by introducing the primary components of what developed into the socio-educational model of foreign language achievement. Building on earlier research, Gardner, Clement, Smythe, and Smythe (1979) developed the Attitude/Motivation Test Battery (AMTB), which is an instrument designed to assess five non-cognitive variables theorized to influence foreign language achievement. The variables measured by the AMTB are: effort towards a goal, desire to learn the language, satisfaction with studying the language, attitude toward the language-speaking group, and attitude towards the learning environment (Gardner, 1985). Upon a search of the relevant literature, to the best of this author's knowledge, predictive validity evidence of the AMTB is deficient – there is a lack of empirical evidence supporting relationships between foreign language achievement and the constructs measured by the AMTB. Yet, for nearly three decades, the socio-educational model – as measured via the AMTB – has been a dominant approach to assessing non-cognitive variables in foreign language achievement. The strict focus on the socio-educational model may have prevented researchers from identifying useful non-cognitive predictors of foreign language achievement.

Foreign language learning researchers have recently suggested that additional constructs may underlie foreign language achievement, such as the constructs identified in the educational psychology literature on self-concept and motivation (e.g., Dornyei, 1994). With the agenda of expanding upon and bringing up to date the non-cognitive variables associated with foreign language achievement, the present researcher looked to the educational psychology literature for constructs predictive of academic achievement that may be predictive of foreign language achievement. Two constructs were included in the present study; each is described below.

Self-Efficacy

Self-efficacy refers to "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances" (Bandura, 1986, p. 391). One's self-efficacy is an indicator of one's abilities, one's willingness to exert effort, one's likelihood to persist towards an achievement goal, and one's perceptions of task difficulty (Bandura, 1982). Factors such as past performance, experiences with others' performances, persuasion from others and physiological changes can influence one's self-efficacy judgments in a given context (Schunk, 1984). Essentially, self-efficacy is one's beliefs about one's capability to master a given task.

The self-efficacy construct should not be confused with the more global self-confidence construct, which reflects a belief that one can cope with almost any task.

Historically, self-efficacy has been a reliable predictor of academic success and success in various other domains (e.g., Bandura, 1986). For example, Schunk (1984), found a high-moderate correlation between self-efficacy and a measure of academic achievement. In addition, Zimmerman, Bandura, and Martinez-Pons (1992) offer a similar, significant correlation within the context of social studies learning. Neither of these studies was done in a foreign language-learning context. Tremblay and Gardner (1995) included self-efficacy in a study on foreign language achievement. Using a causal model, they found that self-efficacy influences motivated behavior, which in turn influences achievement, however the model was only a moderate fit and they did not offer a direct relationship between self-efficacy and foreign language achievement. Contrary to Tremblay's and Gardner's (1995) use of the self-efficacy construct, it can be used as a direct predictor of academic achievement.

In the predictive function of self-efficacy, the nature and difficulty of the achievement task and the goals one has play a role in determining success (Bandura, 1986). When considering self-efficacy, a specific task must be identified -- theoretically, there is no general sense of self-efficacy, only self-efficacy to complete a specific task. Therefore, in the present research the self-efficacy construct is confined to the task of learning a foreign language. Given the specific task, one's beliefs about the task interact with one's self-efficacy, to produce performance. In a study by Bandura (1982), superior performance was found to occur when there was a combination of high self-efficacy and the belief that the task to be completed was a challenging one. In addition, when one has set goals and one is self-efficacious, one tends to exert effort and achieve (Bandura & Cervone, 1983). Furthermore, self-efficacious people tend to set more challenging goals (Zimmerman, Bandura, & Pons, 1992). Combining these findings, self-efficacy can be maximized as a predictor of achievement when one is seeking a challenging goal. Therefore, a construct capturing one's desire for challenging goals should be included with a measure of self-efficacy. This leads to the second construct in the present study.

Learning-goal Orientation

Learning-goal orientation (sometimes called mastery goal orientation) encompasses the choice of challenging tasks, developing greater competence of the material being learned, and the belief that one has control over one's achievements (Dweck & Leggett, 1988). A student with a learning-goal orientation is seeking challenging tasks in order to develop competence and an appreciation of the material being learned. Additionally, learning-oriented individuals tend to feel self-efficacious for the task and persist through difficult tasks. These qualities are indicators of the prevailing characteristic of a learning orientation – an adaptive behavioral process that results in achievement (Dweck & Leggett, 1988; Elliot & Dweck, 1988; Elliot, McGregor &

Gable, 1999).

While self-efficacy has been found to be correlated with learning orientation, learning orientation has also been indicative of one's choosing tasks that will increase competence, regardless of one's level of self-efficacy (Elliot & Dweck, 1988). A reasonable deduction from this finding is that a learning orientation may have its own predictive power, separate from self-efficacy. In as much, Elliot et al. (1999) found a significant, moderate correlation between learning orientation and overall grade point average in college courses. However, little is known about the interaction between self-efficacy and learning orientation, and there has been no prior research on this in the foreign language-learning domain.

Within a theoretical perspective similar to learning-goal orientation, Noels et al. (2000) investigated intrinsic motivation in foreign language learning. (Intrinsic motivation refers to the belief that the performance of a task is valuable in itself.) The authors stated that intrinsic motivation is not predictive of success in foreign language learning. However, the authors did not address the relationship between intrinsic motivation and the more comprehensive learning-goal orientation. In research outside of foreign language learning, it has been revealed that the development of learning-goals is related to increased intrinsic motivation (Elliot & Church, 1997; Heyman & Dweck, 1992), but intrinsic motivation does not encompass the broad set of adaptive processes that learning-goal orientation incorporates. Intrinsic motivation is only one indicator of a learning orientation. Therefore, intrinsic motivation is deemed to be subsumed by learning-goal orientation. The present research introduces learning-goal orientation as an element of achievement in the foreign language-learning domain.

Summary

The purpose of the present research is to investigate self-efficacy and learning-goal orientation as they relate to foreign language achievement. This investigation includes the construction and evaluation of the psychometric properties of an instrument intended to measure the aforementioned constructs in the context of foreign language learning. The name of the new instrument is Measure of Foreign Language Achievement Potential (MOFLAP). The self-efficacy and learning-goal orientation constructs are expected to be individually predictive of foreign language achievement and maximally predictive when combined.

Method

Participants

The 128 participants in this study were undergraduate students at Pennsylvania State University enrolled in a foreign language course at the time they completed the instrument. Participants were selected from eight German

language classes. The course levels ranged from German I through German III. There were 42 females and 86 males. Ages ranged from 18 to 44 with a mean age of 20.84 *(SD* = 3.28).

Materials

The author of this paper created the 18 items in the two intended subscales of the MOFLAP (shown in the appendix). Educational measurement and foreign language experts reviewed the items. Items were worded to assess the level of presence of each construct. Responses to items indicate level of agreement on a Likert scale of 1 to 5, with 1 being strongly disagree and 5 being strongly agree. Half of the items were reverse worded to control for acquiescence.

Procedure

Participants volunteered for eight group administrations. After giving informed consent, all respondents were instructed to carefully read each item and respond according to the scale, which was listed at the top of the instrument. Final course grades - the criterion measure in this study - were obtained from the university's registrar. There were no missing data.

Results

SPSS 10.0 was used to score and analyze the MOFLAP. Scores were calculated using average subscale scores (dividing the summed item scores for each subscale by the number of items on the subscale). Higher scores on a component indicated more presence of that component, and lower scores indicated less presence. After scoring, the items and intended subscales were analyzed.

Table 1 contains the item statistics for the MOFLAP. None of the items has restricted range, extreme skewness, or extreme kurtosis. Item 2, a learning item, has the lowest mean ($\underline{M}$ = 1.91, $\underline{SD}$ = .91), and Item 7, a self-efficacy item, has the greatest mean ($\underline{M}$ = 4.19, $\underline{SD}$ = .80). Item 3, from self-efficacy, has the highest correlation with class grades, while item 12, from learning, has the lowest correlation with class grades.

Exploratory Factor Analysis

An exploratory factor analysis (EFA), using principal axis factoring, was used to investigate the factor structure (construct validity) and discriminant validity of the new instrument. The factor selection was based on parallel analysis, because there is empirical evidence that parallel analysis is the most accurate factor selection method (Zwick & Velicer, 1986). The factors were expected to be correlated, so the oblique rotation method of promax rotation was used. Promax rotation is a recommended procedure, in part because of

the simple structure it provides (Gorsuch, 1983).

Table 1
Descriptive Statistics for Items

Subscale	Item	Correlation w/ course grade[a]	Mean	SD	Skew	Kurt
Self-Efficacy	1	.41	3.84	1.01	-0.69	0.36
	3	.57	3.88	1.12	-1.10	0.70
	5	.40	3.95	1.05	-1.08	0.78
	7	.37	4.19	0.80	-0.91	0.63
	9	.36	3.98	0.78	-0.77	0.72
	11	.39	3.79	0.95	-0.78	0.24
	13	.41	3.83	1.02	-0.82	0.14
	15	.46	3.54	1.23	-0.51	-0.84
	17	.47	4.10	0.88	-1.13	1.28
Learning	2	.27	1.91	0.91	0.55	-0.80
	4	.35	3.04	1.29	-0.14	1.13
	6	.31	2.75	1.34	0.21	-1.14
	8	.33	3.03	1.13	-0.26	-0.76
	10	.27	2.52	1.20	0.41	-0.79
	12	.26	2.45	1.41	0.53	-1.07
	14	.33	2.69	1.23	0.18	-0.98
	16	.27	2.82	1.27	0.04	-1.03
	18	.40	2.94	1.24	-0.15	-1.01

Note. Skew = skewness; Kurt = kurtosis.
[a] = Pearson's r was used; all correlations were significant at the .01 level.

Based on parallel analysis two factors were selected for rotation. (For more on parallel analysis in SPSS, see Thompson & Daniel, 1996). The two factors accounted for 58.8% of the total variance. Factor 1 accounted for 48.5% of the variance. All of the self-efficacy items loaded together on factor 1, with a small secondary loading of learning item 4, so it was clearly interpretable as the self-efficacy factor. Factor 2 -- the learning factor -- accounted for 10.3% of the variance.

In consideration of the adequacy of the sample size for factor analysis, empirically derived criteria presented by MacCallum, Widaman, Zhang and Hong (1999) were followed. With communalities generally above .6 and an item to factor (p:r) ratio of approximately 7:1 (with more items per factor being better), a sample size of 60 is good for stable recovery of population factors. In the present study, communalities are generally above .6, the p:r ratio is 9:1 and the sample size is 128, therefore the sample size is deemed very good for a factor analysis. Table 2 lists the factor loadings and communalities. Absolute values less than .30 have been suppressed because these values are meaningless to interpretation (Gorsuch, 1983).

Table 2
Exploratory Factor Analysis – Principal Axis Factoring with Promax Rotation

Subscale	Item	Factor 1	Factor 2	Communality
Self-Efficacy	17	.87		.74
Self-Efficacy	5	.82		.68
Self-Efficacy	11	.81		.60
Self-Efficacy	3	.79		.74
Self-Efficacy	9	.79		.66
Self-Efficacy	7	.78		.64
Self-Efficacy	15	.74		.64
Self-Efficacy	1	.70		.62
Self-Efficacy	13	.70		.53
Learning	18		.87	.75
Learning	6		.85	.66
Learning	14		.84	.72
Learning	10		.79	.64
Learning	12		.77	.62
Learning	8		.65	.61
Learning	16		.59	.52
Learning	2		.56	.35
Learning	4	.32	.39	.50

After the EFA, the subscales' reliabilities were determined using coefficient alpha – both reliability coefficients were very good. Table 3 shows the reliability analysis.

Following reliability analysis, descriptive statistics, intercorrelations, and correlations with class grades were obtained – these are shown in Table 4. The learning subscale had a lower mean (M = 2.68, SD = .94) than the self-efficacy subscale (M = 3.90, SD = .80). The correlation between the subscales is .62. The self-efficacy subscale had a larger correlation with class grades (r = .53, $p < .01$) than the learning subscale (r = .40, $p < .01$), but both were significant.

Table 3
Reliability of Subscales

Subscale	Reliability	Item	Scoring[a]	a - item[b]	
Self-Efficacy	.93	1		0.93	
		3	-	0.92	
		5		0.92	
		7	-	0.92	
		9		0.92	
		11	-	0.92	
		13		0.93	
		15	-	0.93	
		17		0.92	
Learning	.91	2			0.91
		4	-	0.91	
		6		0.90	
		8		0.90	
		10		0.90	
		12	-	0.90	
		14	-	0.89	
		16	-	0.91	
		18		0.90	

[a] = Items marked with – are reversed scored. [b] = Cronbach's Alpha with the item removed.

Table 4
Descriptive Statistics and Correlations

Measure	Mean	SD	Learning		Class Grades
Self-Efficacy	3.90	0.80	.62**	.53**	
Learning	2.68	0.94	-		.40**
Class Grades	4.20	3.12	-	-	

Note. ** = $p < .01$.

Linear Regression

Simple regressions and a multiple regression of self-efficacy and learning orientation revealed that each variable is predictive of course grades and that learning orientation makes a small contribution to the combined prediction. The R^2 with self-efficacy as a predictor was .282, with an adjusted R^2 of .276, F (1, 126) = 49.51, p = .00. The R^2 using learning orientation as a predictor was .162, with an adjusted R^2 of .155, F (1, 126) = 24.30, p = 00. The multiple regression resulted in an R^2 of .291 and an adjusted R^2 of .279, F (2, 125) = 25.63, p = .00. Since the correlation between the variables is high, a variance inflation factor (VIF) was used to check for multicollinearity. Multicollinearity did not exist among the variables (VIF = 1.622, for both variables), based on a standard of no multicollinearity if the VIF is less than 10 (Neter, Kutner, Nachtsheim & Wasserman, 1996). Table 5 shows the multiple regression model.

Table 5

Subscale	SE	B	t	p	VIF	
Learning			.08	.12	1.24	.22
1.622						
Self-efficacy		.09	.46	4.77	.00	1.622

Note. VIF = variance inflation factor.

Discussion

The expected two-subscale structure of the MOFLAP was supported by the EFA. The clear emergence of the two factors in the EFA provides evidence of the construct and discriminant validity of the subscales. Though the correlation between the subscales was high, it is not high enough to weaken support for the subscale's discriminant validity. In addition, the items' and subscales' significant correlations with class grades are evidence of predictive validity. As in past research outside of the foreign language learning domain, there was a strong correlation between self-efficacy and class grades. The correlation in the present research is even more powerful than those presented by Schunk (1984) and Zimmerman, Bandura, and Martinez-Pons (1992) who investigated self-efficacy for learning social studies. Additionally, learning orientation was highly correlated with class grades as it has been in past studies outside of the context of foreign language learning. Again, the correlation between learning orientation and class grades in this study is more powerful than in past research using overall college GPA as the criterion (i.e., Elliot et al., 1999). Moreover, the reliabilities of the self-efficacy and learning-goal orientation scales are very good and superior to those of the scales used by previous researchers. This may account for the more powerful correlations in the present study. Overall, there is evidence of construct, discriminant, and predictive

validity for both of the subscales and the subscales are highly reliable.

The psychometric evidence provided for the MOFLAP supports its use for the other purpose of this research -- to determine whether self-efficacy and learning-goal orientation are significant predictors of class grades. As previously stated, the variables are correlated with class grades, therefore they are predictive. The linear regression leads to the conclusion that of the two variables, self-efficacy is the superior predictor. Learning orientation adds to the prediction of course grades, when combined with self-efficacy, but it is only a very slight improvement in the prediction using self-efficacy alone. The multiple regression of the two, non-cognitive variables, accounts for slightly less of the variability in foreign language course grades from the present study, than the combination of three cognitive variables measured by Sparks, Javorsky, Patton and Ganschow (1998) in a previous study. Therefore, there is potential that non-cognitive variables can account for much of the variance in foreign language achievement. A model that includes the constructs in the MOFLAP, as well as cognitive variables, may offer a superior method of predicting foreign language achievement than using cognitive or non-cognitive constructs alone.

Theoretical Implications

The more powerful correlation between self-efficacy and class grades in the present research is not a nullification of the potential importance of learning orientation in foreign language learning. While the additional variance accounted for by the learning orientation variable, when paired with self-efficacy, is very small, it should be considered that past research in which self-efficacy and learning orientation were investigated, a learning orientation has been indicative of one's choosing tasks that will increase competence, regardless of one's level of self-efficacy (Elliot & Dweck, 1988). That is, a learning orientation begets mastery learning even when one perceives one's self to have low ability. As previously noted, learning orientation is indicative of a set of adaptive learning behaviors that lead to success (e.g., Dweck & Leggett, 1988; Elliot & Dweck, 1988), whereas self-efficacy alone is predictive of success (e.g., Bandura, 1986). However, academic success may be obtained without self-efficacy when one is learning-oriented. Therefore, learning orientation may be a broader reflection of the learning situation than self-efficacy, though self-efficacy better predicts grades.

While self-efficacy is a more powerful predictor of class grades than learning orientation, learning orientation reflects a more dynamic approach to understanding students' foreign language achievement. One important component of learning orientation is the development of networks of long-lasting friends (e.g., Dweck & Leggett, 1988), which clearly reflects beliefs about prosocial behavior. Wentzel (1992) found that students often give social goals higher priority than academic goals in an academic setting. In addition, Schneider, Ackerman and Kanfer (1996) state that prosocial goals help students to achieve more fully, that there is more in a learning situation than

academic performance.

Future Directions

Given the results of the present study, the limited sample of learners and languages, it is suggested that the relationships between self-efficacy and learning-goal orientation be further explored in multiple foreign language learning situations. The present research reveals the potential of these variables as substantial non-cognitive predictors of foreign language achievement. Future research including these variables could be conducted using causal modeling to investigate if these variables have direct, causal relationships with foreign language achievement. As alluded to above, a prosocial behavior variable may impact the relationship between learning orientation and achievement, therefore this variable may be included in future projects. Moreover, variables such as effort and persistence--both of which are known to correlate with the variables in the present study--may be included in models predicting foreign language achievement.

Before moving forward by including more non-cognitive variables, a study combining the foreign language achievement-related, non-cognitive variables in the AMTB (see Gardner et al., 1979), the cognitive variables identified by Sparks, Javorsky, Patton and Ganschow (1998), and the variables investigated in the present study, may be useful. A multiple regression using all of the previously identified non-cognitive and cognitive factors may lead to a better model than those previously noted--a model that reflects a fuller range of variables that affect foreign language learners.

Conclusions

The psychometric properties of the MOFLAP warrant use by educators and researchers to measure self-efficacy and learning-goal orientation, as they directly relate to the success of foreign language learners. Researchers should consider using the MOFLAP to explore further the nature of these relationships, with the inclusion of other cognitive and non-cognitive variables. There is more to a learner and learning situation than cognition, and as shown in the present study, non-cognitive factors are contributors to foreign language learner's successes. A well-developed understanding of students' characteristics, both cognitive and non-cognitive, may contribute to the development of a better learning environment. The MOFLAP can be used to help develop a more in depth understanding of the variables that affect success in foreign language learning.

Appendix

1. I am confident that I can learn the material in my foreign language class.
2. I am learning a foreign language to become more cultured.
3. Regardless of my efforts, it is unlikely that I will do well in my foreign language class. (-)
4. I dislike the challenge involved in learning a foreign language. (-)
5. Learning a foreign language is too difficult for me to do. (-)
6. My desire to understand a foreign language is the reason I am taking this class.
7. I am capable of learning a foreign language.
8. Getting a good grade without developing an appreciation for the foreign language will be disappointing to me.
9. By making an effort I will learn the material in my foreign language class.
10. I am learning a foreign language mostly for the sake of learning something new.
11. It is highly unlikely that I will be able to overcome difficult events on my way to learning a foreign language. (-)
12. The only reason I am taking a foreign language class is to fulfill a requirement towards my degree. (-)
13. Whatever difficulties students in my foreign language class may have, the difficulties will be even worse for me. (-)
14. Enrolling in a foreign language class had very little to do with my appreciation for foreign languages. (-)
15. I have some serious doubts about my ability to learn a foreign language. (-)
16. Getting a good grade in my foreign language class is far more important than truly learning the language.
17. I believe I have the ability to learn a foreign language.
18. I am learning a foreign language because it is interesting.

Note: Odd items are self-efficacy and even items are learning. (-) indicates reverse scoring.

References

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122-147.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Dornyei, Z. (1994). Motivation and motivating in the foreign language classroom. *The Modern Language Journal, 78*¸ 273-284.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256-273.

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality & Social Psychology, 72*(1), 218-232.

Elliot, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology, 54*(1), 5-12.

Elliot, A. J., McGregor, H. A., & Gable, S. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. *Journal of Educational Psychology, 91*(3), 549-563.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gardner, R. C., Clement, R., Smythe, P. C. & Smythe, C. L. (1979). Attitudes and motivation test battery-revised manual (Research Bulletin # 15). London, Ontario: University of Western Ontario.

Gardner, R. C., & Lambert, W. E. (1959). Motivational variables in second language acquisition. *Canadian Journal of Psychology, 13*¸ 266-272.

Gorsuch, R. L. (1983). Factor Analysis. Hillsdale, NJ: Lawrence Erlbaum Associates.

Heyman, G. D., & Dweck, C. S. (1992). Achievement goals and intrinsic motivation: Their relation and their role in adaptive motivation. *Motivation and Emotion, 16*(3), 231-247).

Lett, J., & O'Mara, F. (1990). Predictors of success in an intensive foreign language learning context: Correlates of language learning at the Defense Language Institute Foreign Language Center. In T. Perry & C. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 222-260). Englewood Cliffs, NJ: Prentice Hall.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4(*1), 84-99.

Neter, J., Kutner, M. H., Nachttsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models*. 4th edition. New York, NY: McGraw Hill.

Noels, K. A., Pelletier, L. G., Clement, R., & Vallerand, R. J. (2000). Why are you learning a second language?: Motivational orientations and self-determination theory. *Language Learning, 50*(1), 57-85.

Schneider, R. J., Ackerman, P. L., & Kanfer, R. (1996). To "act wisely in human relations": exploring the dimensions of social competence. *Personality and Individual Differences, 21*¸ 469-481.

Schunk, D. H. (1984). Self-efficacy perspective on achievement behavior. *Educational Psychologist, 19*(1), 48-58.

Sparks, R. L., Javorsky, J., Patton, J., & Ganschow, L. (1998) Factors in the prediction of achievement and proficiency in a foreign language. *Applied Language Learning, 9*(1), 71-105.

Strahan, R., & Gerbasi, K. C. (1972). Short, homogenous versions of the Marlowe-Crowne social desirability scale. *Journal of Clinical Psychology, 28*¸ 191-193.

Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement, 56*(2), 197-208.

Tremblay, P. F., & Gardner, R. C. (1995). Expanding the motivation construct in language learning. *Modern Language Journal, 79*(4), 505-520.

Wentzel, K. R. (1992). Motivation and achievement in adolescence: A multiple goals perspective. In D. H. Schunk & J. L. Meece (Eds.), *Student perceptions in the classroom* (287-306). Hillsdale, NJ: Lawrence Erlbaum Associates.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*(3), 663-676676.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of fiverulesfodetermining the number of components to retain. *Psychological Bulletin, 99*(3), 432-442.

Author

DANIEL L. MCCOLLUM, M.S., Ph.D. Candidate, Department of Educational and School Psychology and Special Education, Pennsylvania State University. Specializations: test development and psychometrics.

Applied Language Learning
2003, Vol. 13, No. 1, pp. 33-56

Does a Quiz Facilitate or Spoil Language Learning?
Instructional Effects of Lesson Review Quizzes

Yoshinori Sasaki
Ochanomizu University
and
Harumi Hayakawa

The research project to be reported intends to clarify the instructional effects of review quizzes as a part of assessment. In Study One, a survey was administered with approximately 100 second-year Japanese language students, inquiring whether they would like to see quizzes as a part of assessment. Several months later, a follow-up survey of the same nature was conducted. The two surveys repeatedly demonstrated students' consistent and overwhelming preference for quizzes, with an expectation that quizzes would help them to study regularly.

In Study Two, a different set of students in the following year's second-year Japanese language course were allowed to choose whether they would take quizzes as a part of assessment (final marks of students who chose not to take quizzes more heavily depended on other assessment components, such as mid-term and other kinds of tests along with assignments). Subsequently performance scores of quiz takers and non-quiz takers were compared. Students who chose to take quizzes outperformed the rest by a significant margin in class performance after the quiz sessions started, whereas a difference was not detected in terms of the scores beforehand.

In sum, many students expected that quizzes would help them to study more regularly (Study 1). Students who preferred taking quizzes benefited from them, and outperformed those who avoided them (Study 2).

Quantitative language testing research has long centered around such notions as reliability, concurrent validity, and predictive validity of standard proficiency tests. Recently, some researchers have paid more attention to another type of validity, namely, *consequential validity* (McNamara, 1996), which addresses the influence of a test in the real educational world.

*Presented by Dr. Sasaki at the 23rd Applied Linguisics Association of Australia (ALAA) Annual Conference, Griffith University, Queensland, Australia, July 1998.

> Aspects of the consequential validity of language tests have been recognized under the heading "washback" (or "backwash"). The term is used to refer to the impact of tests on the teaching program that leads up to them. This is particularly relevant for public tests which are established not primarily as part of a curriculum but which nevertheless become the focus of teaching program's which may be organized around them (TOEFL preparation courses are an important example). (McNamara, 1996, p. 23)

While McNamara cites TOEFL (proficiency test) as an example, the notion of consequential validity is equally useful, and crucial, in so-called achievement tests, namely, in-class tests designed to measure the mastery of a particular curriculum. Most importantly, it is of primary concern for teachers to learn whether such tests indeed contribute to achieving the curricular goals.

It has long been a rather common practice for teachers to give short quizzes to students. In many instances, the primary objective of such practices is to motivate learners to review/preview regularly, so that students will eventually achieve a higher level of mastery. The efficacy of quizzes for this instructional objective seems obvious: Indeed, it is so obvious that it is not easy to identify empirical studies which seek to validate this claim.

On the other hand, skeptics may invoke the psychological theory of intrinsic learning motivation (e.g., Lepper, Greene & Nisbett, 1973) to question the appropriateness of this scheme. A possible objection against frequent quizzes could be that: Giving external incentives to learners to do a certain task diminishes their intrinsic motivation, and thus once the incentive is discontinued, they will feel less inclined to do it. Also, they will learn to pay less efforts on something that does not bring about much academic incentives. In short, quizzes/tests will do more harm than good in the long run.

If this is an accurate prediction, teachers should give as few tests/quizzes as possible, in order not to spoil students' motivation. The feasibility of this recommendation in real instructional situations is another problem, though.

It is unfortunate that such controversies on curricular decision making at many institutions are settled "politically" (in the word's negative sense) often in favor of seniority, combative argumentation style, tenacious personality, and bargaining skills, rather than through scientific reasoning and disciplined empirical methodology. We feel that it is important to set a model to reach a mutually agreeable instructional decision in a scientifically sound manner, as debates over instructional matters may potentially provide an opportunity to make important discoveries in applied linguistics, educational psychology, and other related areas.

Indeed there is an accumulation of psychology literatures in the area of incentive and motivation. In this area, there has been a fierce debate between two camps each emphasizing (Lepper, Keavney, & Drake, 1996) and qualifying (Cameron & Pierce, 1994; Cameron & Pierce 1996) external incentive's possible negative effects on a learner's intrinsic motivation.

Its pedagogical implication is obvious: If all forms of external incentives (including review quizzes) would diminish learners' subsequent motivation to study when the incentive is no longer provided, its negative effects in the long run are likely to outweigh immediate facilitating effects. Thus it would be prudent for teachers to study pertinent literatures to see whether quizzes are indeed harmful.

Despite the apparent differences between the above-mentioned two competing camps, however, they agree on several important issues. Even Cameron & Pierce (1994; 1996), strong skeptics of incentive's supposedly sweeping negative effects on intrinsic motivation, openly admit that in some instances incentives can be harmful, namely: "… when subjects are offered a tangible reward (expected) that is delivered regardless of level of performance, they spend less time on a task than control subjects once the reward is removed." (Cameron & Pierce, 1994, p. 395)

Thus the practice of giving students a credit for solely submitting an assignment (e.g., essay) regardless of its quality should be conducted, if ever, only sparingly.

On the other hand, Lepper et al. (1973), three of the first discoverers of incentive's potentially negative effects, nevertheless recommended effective uses of external incentives under some circumstances:

> ... extrinsic incentives may often be effectively used to increase interest in certain broad classes of activities. On the present line of reasoning, this proposition should be particularly true when (a) the level of initial interest in the activity is very low and some extrinsic device is essential for producing involvement with the activity or (b) the activity is one whose attractiveness becomes apparent only through engaging in it for a long time or only after some minimal level of mastery has been attained. In fact, such conditions characterize the prototypical token-economy program, if the tangible extrinsic rewards are necessary to elicit to the desired behavior. Hence, it would be a mistaken overgeneralization from the present study to proscribe broadly the use of token-economy programs to modify children's behavior (p. 136)

Thus it is important for practitioners to find the optimal middle ground, where the positive effects of incentives are maximized and their negative side-effects, if any, should be minimized. The present research project intended to seek an answer to this question empirically. Specifically, it addressed whether it is worthwhile to give quizzes (i.e., quizzes provide positive effects on students' subsequent academic performance) or not (i.e., quizzes provide non-existent or negative effects). The two studies to be reported here attempted firstly to investigate how language learners perceive the effects of quizzes on their learning habits, and secondly to objectively measure the impact of quizzes on

learners' subsequent academic score.

Study 1: Survey Research

The following two hypotheses assume that learners subjectively believe in the quiz-as-facilitator position.

Hypotheses

Hypothesis 1: Majority of university learners of Japanese prefer a quiz as a part of the assessment system.
Hypothesis 2: Students' attitude toward a quiz is longitudinally stable across academic terms.

Method

Participants

Participants of Phase 1 were 3rd-semester (second year) students attending a Japanese language course at the University of New South Wales. Most of the students had received approximately 130 hours of formal instruction in Japanese before they joined the class. Participants of Phase 2 were students in a 4th-semester Japanese course in 1996, the majority of whom were the respondents of the Phase 1 survey.

The students experienced frequent in-class review quizzes in the first term of the previous academic year (1995). This practice was discontinued in the second term, because some of the staff members raised a strong objection against it on the basis of the speculation outlined above. Thus, the majority of students had experienced Japanese language courses both with and without review quizzes before they responded to the survey.

Procedure

Timing: The study consisted of two phases. Phase One was conducted in the first term of the 1996 academic year, and Phase Two (follow-up) was conducted in the second term of the same year.

Materials: An anonymous opinion survey about quizzes was conducted in Phase 1 and Phase 2 repeatedly. The questions on a questionnaire sheet were as follows:

1. Which of the following three assessment plans do you think is the most appropriate for this course? (Circle one)

Plan T: a small number of long tests only
Plan Q: frequent short quizzes only (e.g., vocabulary, kanji, grammar)
Plan Q&T: a combination of both quizzes and long tests

2. Why?

Results

Results are reported in Table 1 below. More than 75% of students constantly supported incorporating a quiz in evaluation in one way or another.

Table 1. *Students' Choice of Grading Plan (1996 Academic Year)*

Students' Choice		Number of Respondents (%)			
		Phase One (Term 1)		Phase Two (Term 2)	
Would rather not have quizzes	Plan T (test only)	19 (21.8%)		11 (21.2%)	
Would rather have quizzes	Plan Q (quiz only)	68 (78.2%)	21 (24.1%)	41 (78.8%)	9 (17.3%)

Many students who preferred quizzes stated that they expected quizzes to motivate them to study regularly. (See the Appendix for examples of the statements.)

Discussion

These two surveys clearly demonstrate that students consistently and overwhelmingly supported short quizzes as a part of assessment, with an expectation that quizzes would motivate them to study regularly.

It is noteworthy that the ratios of choices are remarkably similar

despite the several months' interval: Support for quizzes reached almost 80% in both instances. These results support the constancy of students' belief in the positive effects of quizzes on learning, although the anonymous nature of the survey prevents us from demonstrating the stability at the individual's level.

Study 2: Action Research

Purpose

Study 1 revealed students' subjective expectation that quizzes would enhance learning by shaping a steady learning habit. In Study 2, this expectation was tested against students' actual academic performance. The following hypothesis represents the "quiz-as-facilitator" position.

Hypothesis: Students who regularly take assessable lesson review quizzes throughout an academic term will outperform those who do not in terms of academic test scores at the end of the term.

Method

Participants:
Participants are attendants of a fourth-semester (second year) Japanese language course at the University of New South Wales, ranging over 14 weeks. Most of the students had received approximately 200 hours of formal instruction in Japanese before they joined the class.
Instructional Background:
The class met five hours a week, consisting of one 1-hour lecture and two 2-hour tutorials. The textbook assigned by the School on the instructional team was: Neustupny, J. Okabe, M. & Muraoka, H. (Eds.) (1997). *Interacting with the Japanese: Book 4*. Monash University Japanese Studies Centre. The instructional team made substantial efforts to adapt this textbook, which was originally designed for Monash University students in the Melbourne vicinity, to meet the needs of University of New South Wales students in Sydney by localizing activities, etc.

Procedure

Timing: The study was conducted between July and November of 1997 (Table 2).

Quizzes: In Week 5, each student was instructed to choose one of the following three quiz options as their assessment scheme.

Option 1 (No quiz): The student would not take quizzes, and the weights on other assessment components (e.g., mid-term hourly written tests; final examination) are proportionally higher.

Table 2. *The Timeline of the Study (1999 Academic Year)*

Week	Major in the Class assessment Other than Quiz	Quiz
5	Oral Test 1 (W-F)	Quiz Option Choice Due (Friday)
6	Hourly Test 1 (Monday)	Semi-Weekly assessable Quiz (due each Friday)
7-12		

Option 2 (Non-assessable quiz): The student would take quizzes, but the results do not affect her/his final grade. Weights on assessment components are same as Option 1.

Option 3 (Assessable quiz): The student would take quizzes, which would account for the 20% of her/his final grade. Weights on other components are proportionally lower by 20%.

Students' choices were collected on the following Monday. Some students failed to submit the option choice sheet and their data are not included in the following analyses.

Students took the quizzes during the self-access time of the Language Resource Center (LRC) at their leisure, any time before the pre-specified due date. The first set of assessable review quizzes were due on the Friday of Week 6. A quiz battery of each unit consisted of the following four mutually independent quizzes, and students were allowed to take these on different days:

Kanji (Chinese Characters) reading (Fill-in-the-blank);
Kanji usage (multiple-choice);
Passage reading (multiple choice);
Structure and grammar (multiple choice).

Thus, half of the quizzes addressed knowledge about Chinese characters and their compounds.

Technological Background of the Quiz Administration

Quiz files were developed using Question Mark Company's Question Mark 2 for Mac (authoring software). Students accessed the quiz files on a local server from a client workstation via the local-area network (LAN), using "Question Mark Presenter" (browser software). Each quiz file contained several to ten of discrete question items, and the system randomly chose three or five items among them each time. Students had five minutes to complete each quiz (the quiz automatically terminated after five minutes), and students learned their score immediately. They also had an opportunity to see the correct answers after the quiz. The quiz scores stored on the server were subsequently transferred to a student record database, developed with FileMaker Pro 3 (Sasaki, 1997) (Figure 1)

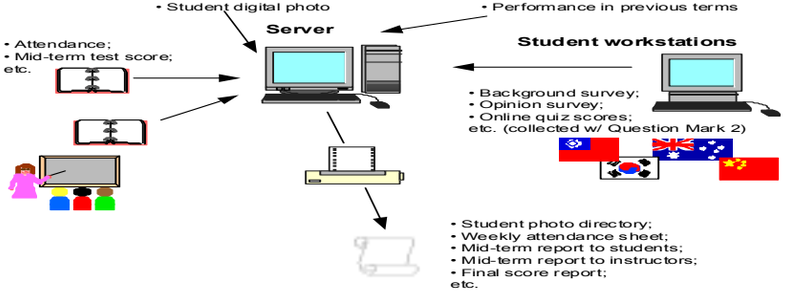

Figure 1. Conceptual Diagam of the Student Record Database

Table 3. *Weights on Assessment Components*

Name	*Feature*	*Options 1 (No quiz) & 2 (non-assessable quiz)*	*Option 3 (Assessable quiz)*
Oral Test 1	Pre-specified role play	5%	4%
Oral Test 2	Same as above	10%	8%
Hourly written test 1	Typical achievement test (e.g., strictly controlled vocabulary and sentence patterns, familiar text, mostly closed-end questions)	10%	8%
Hourly written test 2	Same as above	10%	8%
Final written examination	Achievement test with partial proficiency-test ingredients (purposely loosened control on vocabulary and sentence patterns; incorporation of texts unfamiliar to the testees, some open-ended questions, including an essay writing)	20%	16%
Interview Project	Interviewing volunteer students recruited from the 4th-year Japanese class	10%	8%
Writing 1	Self-introduction message to a campus-based email list	5%	4%
Writing 2 (Project)	Writing a story based on a set of pre-specified pictures	10%	8%
Attendance & Participation		10%	8%
Quiz	Multiple-choice or fil lin-the-blank computer tests (semi-weekly) Kanji reading Kanji usage Passage reading Structure & grammar	0%	20%
Total		100%	

Whereas the first and second hourly written tests typified an achievement test (e.g., strictly controlled vocabulary and sentence patterns, familiar text, mostly close-ended questions), the final examination partly resembled a proficiency test in some respects (purposely loosened control on vocabulary and sentence patterns, incorporation of texts unfamiliar to the testees, some open-ended questions, including a short essay writing task).

Results

Demographic Compositions

Table 4. *Gender of Students by Quiz Options*
Total Chi-square = 2.31 (n.s.)

Students' Performance	*Option 1 (No quiz)*	*Option 2 (Non-assessable quiz)*	*Option 3 (Assessable quiz)*	*Total*
Female	18 (66.7%)	16 (84.2%)	27 (65.9%)	61 (70.1%)
Male	9 (33.3%)	3 (15.8%)	14 (34.1%)	26 (29.9%)
Total	27 (100.0%)	19 (100.0%)	41 (100.0%)	87 (100.0%)

Table 5 presents the first language breakdown of students by their Options (language background data are based on student background survey conducted in the first week; backgrounds of students who did not attend the first week class are classified "unknown"). Clearly, students from East-Asian language backgrounds (those familiar with Chinese characters) tended to choose Option 3 (assessable quiz), whereas those without such backgrounds (i.e., Indonesian, Vietnamese, English, Bulgarian) are more likely to choose Option 1 or 2. Among the East-Asian language speakers, Chinese (Cantonese and Mandarin) speakers showed a particular inclination to choose Option 3, compared to Korean speakers.

Table 5. *First-language Backgrounds of Students by Quiz Option*

First Language		Option 1 (No quiz)	Option 2 (Non-assess-able quiz)	Option 3 (As-sess-able quiz)	TO-TAL
East-Asian language	Cantonese	3 (11.1%)	2 (10.5%)	9	14
	Mandarin	1 (3.7%)	2 (10.5%)	5	8
	Korean	3 (11.1%)	1 (5.3%)	3 (7.3%)	7
Others	Indonesian	1 (3.7%)	3 (15.8%)	1 (2.4%)	5
	Vietnam-ese			1 (2.4%)	1
	English	7 (25.9%)	5 (26.3%)	5	17
	Bulgarian	1 (3.7%)			1
Unknown		11 (40.7%)	6 (31.6%)	17	34
TOTAL		27 (100.0%)	19	41	87

Table 6. *Ratio of Students Who Took Quizzes*

	Option 1 (No quiz)	Option 2 (Non-assessable quiz)	Option 3 (Assessable quiz)
By person	0/27 (0.0%)	2/19 (10.5%)	41/41 (100.0%)

Table 6 represents how many students in each option group took a quiz at least once before the end of the academic term. Quite naturally, all of Option 3 (assessable quiz) students took quizzes, and none of Option 1 (no quiz) students did so. Among the 19 students who chose Option 2 (non-assessable quiz), only two took a quiz and others did not do so.

Table 7. *Student Academic Performance by Option*

		Average Score (SD) Number of participants			Options 1 vs. 3				
						t-			
		Option 1 (No quiz)	Option 2	Option 3 (Assessable		t	p		
Final		67.9	62.5	68.7					
		66.3	67.1 (5.5)	67.5 (8.8)					
		73.1 (9.8)	70.7 (9.1)	74.0 (9.5)					
		25.2 (3.2)	25.7 (3.2)	24.9 (4.0)					
		64.1	58.0	62.3					
		25.5 (4.6)	26.0 (3.4)	25.3 (3.3)					
		69.1	67.1	74.9					
		4.97	6.71 (9.1)	11.79					
	Final								
		71.0	66.5	70.5					

Table 7 presents the average scores of the three groups in performance areas before and after the quiz session started. Among these three groups, the comparison between Option 1 (No quiz) vs. Option 3 (Assessable quiz) students should provide the most straightforward answer to the initial research question as to the instructional effects of quizzes, whereas results from Option 2 (Non-assessable quiz) students tend to trigger multiple possible interpretations because of their equivocal position. For this reason, the rest of this research report centers around these two groups, namely Option 1 (No quiz) vs. Option 3 (Assessable quiz) students. Given the fact that this study ultimately sought to decide whether it was worthwhile to give quizzes (i.e., quizzes provide positive effects on subsequent learning) or not (i.e., quizzes provide non-existent or negative effects on subsequent learning), one-tailed statistical tests were used for testing the hypothetical positive quiz effect on performance scores.

Importantly, Option 1(no quiz) and Option 3 (assessable quiz) students

marked similar scores before the quiz session started (Figure 2): No significant differences were detected in their final marks for the three earlier academic terms (’96 1A; ’96 1B; ’97 2A), nor for the hourly written test and oral test administered by Week 6, making their subsequent performance following treatment easier to compare. On the other hand, Option 3 (assessable quiz) students outperformed Option 1 (No quiz) students in their second written test score ($t(68) = 1.76$, $p < 0.05$, one-tailed).

Figure 2. Test scores of Option 1 and Option 3 students

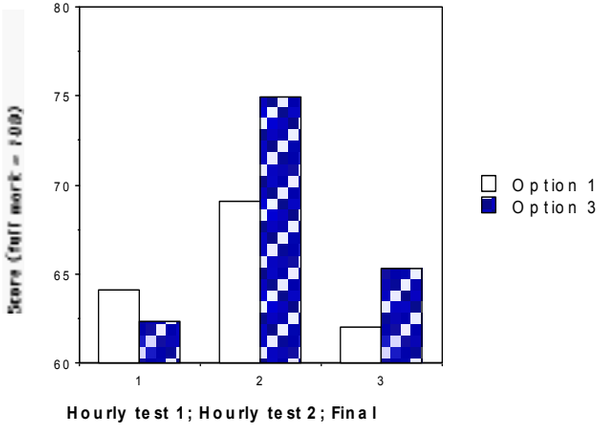

Table 8.

	Average Score			Options 1 vs. 3				
				Direction	One-way tests			
					t-test		Mann-Whitney (Non-parametric)	
	Option 1 (No quiz)	Option 2 (Non-assessable quiz)	Option 3 (Assessable quiz)		t	p	Z (corrected)	p
Kanji writing	13.0	15.8	12.9	O1>O3	0.064	ns	0.018	ns
Listening	4.5	3.9	5.1	O1<O3	1.598	ns	1.269	ns
Expressions	14.3	13.3	14.2	O1>O3	0.029	ns	0.219	ns
Reading	25.8	22.1	25.7	O1>O3	0.013	ns	0.23	ns
Comic	7.9	6.5	7.1	O1>O3	0.731	ns	0.707	ns
Total %	64.1	58.0	62.3	O1>O3	0.3	ns	0.646	ns

Table 9.

	Average Score			Options 1 vs. 3				
				Direction	One-way tests			
					t-test		Mann-Whitney (Non-paramet-ric)	
	Option 1 (No quiz)	Option 2 (Non-as-sessable quiz)	Option 3 (As-sessable quiz)		t	p	Z (cor-rected)	p
Kanji writing	12.2	12.1	13.7	O1<O3	2.026	<.05	2.489	<.05
Listening	14.7	15.1	14.8	O1<O3	0.401	ns	0.9	ns
Expres-sions	10.4	10.7	11.3	O1<O3	0.943	ns	0.886	ns
Reading	31.9	29.2	34.8	O1<O3	1.316	ns	1.08	ns
Total %	69.1	67.1	74.9	O1<O3	1.76	ns	1.352	ns

Table 10. *Final Examination Breakdown of Scores by Section*

	Option 1 > Option 3	Option 1 < Option 3	
Hourly written test 1	4	1	
Hourly written test 2	0	4	
Final examination	1	4	

Tables 8, 9 and 10 provide breakdowns of test/examination scores by sub-sections. In the first hourly test (Table 8), average scores of Option 1 (no quiz) students were higher than Option 3 (assessable quiz) students in four of the five sections, although none of those differences were statistically significant. In hourly test 2 (Table 9), Option 3 students outperformed Option 1 students in all four sub-sections, and in the final examination (Table 10), Option 3 students outperformed Option 1 students in four of the five sub-sections. However, the margins provided significant differences in the Kanji sections only. In total, the distribution before and after the treatment, as is summarized in Table 11, provides a highly significant difference (Chi-square (2) = 15.25, $p<0.0005$).

As opposed to these written test scores, the Oral test scores did not provide significant differences between the two groups (Oral test 1: $t(67) = 0.327$, ns; Oral test 2: $t(67) = 0.332$, ns).

Table 8. *Hourly Written Test 1 Breakdown of Scores by Section*Table 9. Table 9. *Hourly Written Test 2 Breakdown of Scores by Section*

Discussion

In agreement with the initial hypothesis, the analyses revealed that students substantially improved their written test score with a relatively narrow focus on specific lessons (hourly written test) after they started taking review quizzes. On the other hand, the effect was not visible enough on the final examination score which had a wider scope with a more open-ended nature; the hourly tests were typical achievement tests, whereas the final examination partly resembled a proficiency test. Also, the between-group difference was not significant in the oral interaction test score.

As expected, Option 1 (no quiz) students never took quizzes and Option 3 (assessable quiz) students took quizzes most of the time. It is slightly more surprising that only a few Option 2 students (non-assessable quiz takers) ever took quizzes. One possible interpretation is that learners at this stage do not regularly accumulate efforts in the absence of external incentives.

It is noteworthy that Option 2 (non-assessable quiz) students were already outperformed by the other two groups before the academic term started. With lower performances and potentially less motivation, students in this group therefore may have been less inclined to take the lab quizzes.

We have yet to come across a principled explanation of why Option 2

(non-assessable quiz) attracted a large number of low-performance students. A subsequent study is needed to test whether this pattern is replicable, or a one-time coincidence.

General Discussion

Major Findings

The present study has yielded two major findings: First, students overwhelmingly supported short quizzes as a part of assessment, with an expectation that quizzes would motivate them to study regularly (Study 1).

Second, it has been confirmed in Study 2 that participants who chose to take quizzes as a part of their assessment subsequently outperformed other

Table 11. *Number of Sub-sections in which Option 1 (no Quiz) Students or Option 3 (Assessable Quiz) Students Outperformed the Other Students.*

Total Chi-square (2) = 15.25 (p<0.0005

learners in their academic score (as indicated by the second hourly test), whereassuch superiority was not evident before the treatment (indicated by the first hourly test and the first oral test). In other words, students who preferred assessable quizzes benefited from them, and outperformed those who preferred not to take quizzes subsequently. These results are amenable to the aforementioned quiz-as-a-facilitator position. In addition, Option 3 (assessable quiz) students' average final exam score was higher than Option 1 (no quiz) students', although the difference did not reach the statistically significant level.

Causal Relations

Since students were allowed to choose a grading option (allocation was not random), it could be possible to argue that Option 3 (assessable quiz) students outperformed Option 1 (no quiz) students because of their higher motivation rather than the quizzes *per se*. In particular, there are at least two possible causal accounts:

Causal Model 1: Taking quizzes regularly -> Forming a steady studying habit -> High academic score

Causal Model 2: High motivation -> High academic score AND Preference for quizzes

However, those "quiz avoiders" (Option 1 students) initially performed as well as others before the quiz treatment started. Even after many quizzes, performance of Option 3 (assessable quiz) students in an Oral Test 2 was not significantly better than others. In other words, motivation alone does not always guarantee a higher academic performance in a language class. A possible explanation for this phenomenon is that motivation requires certain learning practices to facilitate its effect on learning. The administration of

regular quizzes triggered such practices in certain sub-skill areas (literacy skills in this instance) so that Option 3 (assessable quiz) students showed an advantage in those areas. This hypothesis can be summarized as follows:

Causal Model 3: High motivation -> Choosing to take quizzes -> Forming a steady studying habit -> High academic score

In this scenario, a "token economy" technique is effective not only for children and people with psychological difficulties. Even mature adults can take advantage of it, by purposely incorporating it in their life to achieve their long-term goals. This perspective may provide a possible clue to untangling the perplexing relation between motivation and language acquisition (Crookes and Schmidt, 1991).

It should be noted that the Option 3 (assessable-quiz) students outperformed the rest in terms of hourly test and examination scores, despite the fact that the weights on those particular components were *lower* for them. A simple law of economy would predict the opposite: Option 3 (assessable quiz) students should have invested *less* time and effort on an hourly written test because it would yield less return. Lesser investments would in turn have yielded lower scores. The exact time each student spent for test/examination preparation was not measured, but it is clear that the assessable-quiz takers were better prepared than others when they showed up in the test room.

On the other hand, Option 2 (non-assessable quiz) students performed even worse. Moreover, they seldom took quizzes, as Table 6 indicates. This may have something to do with the fact that student had to pay a visit to the computer lab outside of the class hour to take quizzes.

Ethics and Education Research

With all the statistical evidence in congruence with the position emphasizing the positive effects of quizzes, as cited above, existing doubts might still be hard to eliminate until a strictly controlled experiment (where conditions are randomly allocated) replicates our results. This limitation of the research design (lack of random allocation to conditions) reflects a dilemma of experimental educational research, where students are paying tuition/tax for the best education they can reasonably expect. It is difficult to treat students differently for an extended period without their consent and/or a legitimate educational reason, particularly when grading is concerned. Some students may feel that they are unfairly treated, when their final grades are calculated differently compared to their classmates. Once it is ethically questionable to allocate different assessment methods to students irrespective of his/her preference, it is difficult to run a rigorously controlled study in this regard.

Number of Assessable-Quiz Takers

It is also noteworthy that the ratio of the students in Study 2 who chose to take assessable quizzes (41 out of 87 students; 47.1%) was not as large as the ratio of students who expressed their preference for short quizzes in Study 1 survey (78%). Although the sum of Options 2 (non-assessable quiz) and 3 (assessable quiz) students in Study 2 approached 69%, few Option 2 students took them subsequently. One obvious reason for this discrepancy between the survey result and real-life behavior is the difference of quiz administration methods: Whereas the survey respondents in Study 1 would have had in mind in-class quizzes which they had experienced in the first academic term, Study 2 quizzes were administered outside of the class, which took students extra time to take. It is likely that this extra workload discouraged some (including part-time students) to take the quiz option. Giving quizzes in class may have yielded somewhat different results (but this may provide a serious class management difficulty if some students were allowed not to take quizzes).

In this respect, some may want to argue that the high performance of assessable-quiz takers (Option 3) in Study 2 partly stemmed from the very fact that they had *voluntarily chosen* to take quizzes, which would have raised their intrinsic motivation.

Intrinsic Motivation and Quizzes

Some of self-learning practices teachers recommend to language learners (e.g., memorizing the reading and stroke order of Chinese characters; reading grammar notes in the textbook) on their own may not appear intrinsically attractive, despite the fact that they provide useful preview/review of classes. Also, the importance of such preparations in communication often become most visible after a learner reaches a certain threshold level of proficiency so that s/he can daily use those skills in authentic communication: It is particularly so with regard to the literacy skills with non-alphabetical languages of university students, whose high intellectual maturity requires accordingly sophisticated linguistic and orthographical devices to be properly expressed. Thus it is reasonable, as Lepper *et al.* (1973) suggest, to use some form of external incentives until learners reach that level of mastery.

It is unfortunate that some popular psychology books, which teachers in pre-service and in-service training stages may browse, oversimplify the nuance of these literatures, to present a false impression that external incentives *always* spoil intrinsic motivation. Psychological research into intrinsic motivation is a very fertile field, which is not only of theoretical significance but also yields useful pedagogical implications. It is desirable for curriculum developers to become familiar with original research reports in the area (rather than their misleading second-hand paraphrases), so that they can make disciplined effective uses of incentives in the classroom.

Remaining Issues

1. Long-term effects

Whereas the present report focuses on effects of quizzes within a single academic term, its logical extension is investigating their long-term effects. It is particularly important to examine whether the positive effects will remain after the treatment is removed. In the field of clinical psychology, it is a common practice for behavior therapists to discontinue external incentives after a client establishes a favorable target behavior. Often the target behavior persists in the absence of further reinforcements. Presumably this is more likely to happen if the client recognizes the intrinsic merit of taking the behavior in question.

In the context of the current project, hopefully learners, even those who were seemingly once motivated to study for the sheer sake of raising quiz scores, will eventually recognize that such regular learning habits will contribute to mastering the target language. To test these hypotheses, a future study would require a multiple-term design (as opposed to the single-term design of the present study) with a cognitive methodology including introspective protocols (in addition to the primarily psychometric approach of Study 2).

2. Source of option choices

It is intriguing to see in Table 5 that students with East-Asian language background (Cantonese, Mandarin, Korean) account for a greater proportion in Option 3 (assessable quiz takers) than in Option 1 (no-quiz takers). One might want to attribute this to the supposedly "discipline-oriented" East-Asian education system which many of those students would have gone through.

Another possible explanation is more pragmatic: Those without a previous exposure to Chinese characters (e.g., English and Indonesian speakers) avoided Option 3 (assessable quiz) because they did not like the heavy weight on review quizzes, half of which addressed knowledge about Chinese characters and their compounds.

This explanation of Quiz option preference differentiated by language background in terms of their familiarity with Chinese characters would also explain why Chinese (Cantonese and Mandarin) speakers showed a more evident preference for Option 3 than Korean speakers. In comparison with Chinese speakers, Korean students were less intensively exposed to Chinese characters, in the presence of Korean Alphabet characters (Hangul).

If this is a valid interpretation, it is ironic that students with weaker backgrounds in Kanji, who most keenly needed regular review practices, tended to avoid taking assessable quizzes. This is a possible instructional drawback of the assessment option system.

Kanji writing question scores in Tables 8, 9 and 10 provide data to test this hypothesis. In Hourly Test 1, Option 1 (no quiz) students' average score (12.96) and Option 3 (assessable quiz) students' (12.89) did not differ significantly; whereas in the second Hourly Test, Option 3 students (13.67) clearly outperformed Option 1 students (12.16) by a statistically significant

margin. In addition, Option 3 students (16.23) maintained their superiority over Option 1 students (13.08) in the Final Examination as well. In short, the hypothetical group difference of Kanji knowledge between Option 1 and Option 3 students before the quiz session started was not confirmed.

3. Transfer of Kanji skills

Incidentally, this analysis also suggests an interesting transfer of Kanji skills from reception to production: Namely, despite the fact that review quiz questions on Kanji were restricted to their reading (fill-in-the-blank format) and usage (multiple-choice format), Option 3 students who undertook such quizzes eventually attained a higher mark in Kanji writing questions (expressive skill) as well. Further research is needed to confirm this tendency, and to clarify the psycholinguistic mechanism underlying this possible transfer.

On the other hand, the limitation of skill transfer is also obvious from the fact that Option 1 and Option 3 groups did not differ significantly with respect to their second oral test score.

4. Real momentum — incentive or feedback?

Moreover, the present study fails to confirm whether the observed effects of quizzes stems from the incentive (i.e., contribution to the final grade) or feedback to a learner's performance (i.e., score report). A comparison of scores between Option 2 (non-assessable quiz) and Option 3 (assessable quiz) students should have clarified this, but the unexpectedly low baseline score of Option 2 students and their subsequent failure to take quizzes discouraged further analyses.

At any rate, it would be desirable to conduct a replication study which controls the above-mentioned factors (learners motivation; first language background) more rigorously, by allocating the conditions randomly to participants. Such control studies would be more feasible where learners are paid for their participation (e.g., Defense Language Institute), given the ethical consideration mentioned before.

5. Academic Performance and Language Proficiency — How to Assess?

Review quiz's differential effects on different types of subsequent test (i.e., a significant improvement on a typically achievement-focused hourly written test, as opposed to weaker effects on a partly proficiency-focused final written exam) may trigger the suspicion that quizzes simply inflate the achievement-type test scores, but have little to do with eventual enhancement of proficiency (i.e., acquisition).

On the other hand, the same phenomenon can be seen from a very different angle, namely: A proficiency test score, by definition, is designed to represent accumulation of efforts for a long time period, and is thus inevitably (and rightfully) less sensitive to the efficacy of this or that particular learning activities within a single academic term. As a matter of fact, some proficiency tests (e.g., ACTFL's Oral Proficiency Interview) accompany a manual which

discourages specific preparations before taking it (Buck, 1989).

Day-to-day progress taking place in a learner's mind is often subtle and hardly visible to naive eyes; it is one of a teacher's important missions to make the progress visible to learners, to encourage them to actively participate in activities which will eventually lead to an overall enhancement of proficiency. This is precisely the instructional *raison d'etre* (external value) of achievement tests, and language curriculum development is ultimately an attempt to select, sequence and coordinate activities instrumental to enhancing eventual proficiency.

To this view, the outcome of the current study reflects the difference of objectives each pursued by an achievement test and a proficiency test. Although the difference of the overall final exam scores between no-quiz takers (120.4) and assessable quiz takers (130.5) did not reach the statistically significant level, an analysis from a different angle (Table 11) showed a difference beyond a chance level. Thus as an extrapolation of the present study's results, it is reasonable to hypothesize that teachers' continued and consistent efforts to encourage learner participation will eventually reveal itself in the learner's proficiency test performance as well. A further study is needed to confirm this hypothesis.

Speaking of test sensitivity, this project also hints at the difficulty of applying a conventional experimental design to confirm the efficacy of a certain instructional approach in a classroom situation: Among the several sub-sections of written tests/examination, statistically significant difference at the 5% level was detected only in Kanji writing questions (Tables 9 and 10). Given the many variables contributing to students' academic performance in a certain area, it is a formidable task to develop a test which is sensitive enough to clearly detect effects of a single instructional intervention.

Fortunately, an additional analysis in terms of the number of sections in which one or the other group outperformed the other indicated a clear edge of the Option 3 (assessable quiz) group (Table 11). Seemingly primitive analyses like this may provide a useful tool for pedagogy researchers working within the field of a real-life classroom, where data are inevitably subject to a lot of "noise".

Concluding Remarks

With all the qualifications outlined above, it is fair to claim that the results of the present project highlight positive effects of short quizzes, whereas little, if any, evidence has been obtained to demonstrate, or even hint at, their hypothetical negative side effects. Given the balance of accumulated evidence which is overwhelmingly in favor of quiz proponents, it is now very much on the opponents' shoulder to substantiate their concerns about such common practices.

Finally, it should be noted that the quantitative data in Study 2 were all downloaded from the School's student record database (Figure 1). The integrated database provided convenience not only for administrative and instructional purposes, but it also yielded intriguing research data. Indeed the combination of the research traditions of instructional applied linguistics and educational psychology, Computer-Aided Language Learning (CALL), and

desktop database management technology provides great research potential for classroom-based applied linguists. Further exploration of such resources for research purposes is encouraged.

Appendix

Examples of students' comments, who favored quizzes in the Study 1 Survey:

Because for the small quizzes, we can revise the thing we learned within that week. For the long tests, just give use pressure to study harder.
Variation to suit all students is best accommodated by both long tests and short quizzes.

Quiz facilitate continuous learning and revision.
Long tests simulate exam situations.
Can have a clear concept of every topic, keep on reviewing each topic.
Quizzes given regularly before final tests.

Acknowledgements

Part of this study was supported by Australian Research Council (ARC) research grant to the first author as the principal researcher.

We acknowledge our colleagues at the University of New South Wales Japanese Language Program directed by Dr. Chihiro Kinoshita Thomson for sharing their thought-provoking speculations about quiz's influences on subsequent learning, which have motivated us to settle the debate through a disciplined empirical methodology, rather than politically. We also thank Eiko Ushida, Julia Read and Kazumi Hatasa, and anonymous reviewers of *Applied Language Learning f*or their comments on earlier versions of this paper, and editorial assistance by Angelo Alonzo. All remaining errors are ours.

References

Buck, K. (Ed.) (1989). *The ACTFL Oral Proficiency Interview Tester Training Manual*. The American Council on the Teaching of Foreign Languages.

Cameron, J. & Pierce, D. (1994); Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research, 64,* 1: 363-423.

Cameron, J. & Pierce, D. (1996). The debate about rewards and intrinsic motivation: Protests and accusations do not alter the results. *Review of Educational Research, 66*, 1: 39-51.

Crookes, G., & Schmidt, R. (1991). "Motivation: Reopening the research agenda." *Language Learning, 41*, 469-512.

Lepper, M. Greene, D. & Nisbett, R. (1973). Undermining children's intrinsic motivation with extrinsic rewards: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology, 38*: 129-137.

Lepper, M. Keavney, M. & Drake, M. (1996); Intrinsic motivation and extrinsic rewards: A commentary on Cameron and Pierce's meta-analysis. *Review of Educational Research, 66,* 1: 5-32.

McNamara, T. (1996). Measuring second language performance. London: Longman.

Sasaki, Y. (1997). Developing a course information database for research and instruction. Paper presented at the 22nd Annual Applied Linguistics Association of Australia (ALAA) Congress. University of Southern Queensland, Toowoomba, QLD Australia, 4th October.

Discussion Forum

Transnationalism and Language-in-Education Planning in Mexico
Response to Robert B. Kaplan's "Language Teaching and Language Policy"

Patrick H. Smith and Natalia Martínez León
Universidad de las Américas-Puebla

We read with interest Robert Kaplan's discussion of language teaching and language policy (*ALL, 12*/1), and were especially struck by the observation that attempts to modify the use of Spanish in the U.S. should pay greater attention to conditions in Latin America. The article reminded us that while much has been written about the language education of Mexican immigrants living in the U.S., less is known about the language and education futures of those who return to Mexico. As teacher trainers in Mexico, we are also interested in the other side of the coin, namely, how language-in-education policy in the U.S. effects language policy regarding immigrants who return to live in Mexico. If the U.S. has been, in Kaplan's words, "profligate with respect to the richness of its language resources" (p. 83), what of the bilingual competence many returning migrants bring back to Mexico? How is their bilingualism regarded upon return, and how do schools respond to these transnational language resources?

The Atlixco Valley region in central Mexico provides an illustrative example. Migration from the region to the New York City/New Jersey area began in the 1970s, and has accelerated considerably since the early 1990s (Binford, 1999). Today, approximately 50 percent of the Mexicans living in New York City are from Atlixco and other communities in the State of Puebla (Consulado General de Mexico en Nueva York, 2000). Employed primarily in low-paying service jobs, many dream of saving enough money to open a small business that would permit them to leave the impoverished agricultural sector (Gendreau & Giménez, 2000). Although the migrant population continues to be dominated by unmarried males, a growing number of *poblano* women and families with children live in the greater New York area. Return rates, typically higher for families (Cortés, 2001), have understandably increased since the events of September 11, 2001 and subsequent economic downturn in the U.S. As a result, Mexican officials are preparing for the return of as many as 400,000 transnational students who have spent at least some of their growing up and school years living in the U.S. (Reyes, C., 2001, October 5).

Known locally as "*retornados*" [returnees], some children have never lived outside the U.S. Indeed, those have lived a significant portion of their daily lives in English return to Mexico with differing degrees of competence and schooling in English. Because only a minority of Latino children in NYC schools receive instruction in Spanish (MacSwan, 2000), *retornados* are also likely to be less prepared for academic instruction in Spanish than their non-migrant peers, a finding attested in the case of *poblano* parents sending

children home to Mexico in order to help them maintain proficiency in Spanish (Malkin, 2000).

Our research with Poblano transnationals suggests that schools in Mexico face considerable challenges in planning for the language needs of this growing population.

In terms of human resources, *retornados* are likely to be taught by teachers whose training does not contemplate the linguistic diversity of such students. Despite changes in the national curriculum mandating English language instruction in primary school, many teachers lack the proficiency in that language to effectively implement the communicative approach (Domínguez Betancourt, 1995). In contexts where *Normalista* [university-level teacher education] programs are not allowed to set proficiency levels as a condition of admission and students reportedly select teaching careers as a means of improving their own English (Vega, 2002), we have observed few teachers with the language proficiency to support those *retornados* with highly developed English. Similarly, with the exception of those trained to work with indigenous students, few teachers have the necessary training or experience in classrooms where Spanish is the students' second language.

How then can educators who see themselves as language planners accommodate the particular language needs of *los retornados*? One intriguing possibility lies in tapping the funds of linguistic knowledge held by English-fluent students and their families, to incorporate them into instruction for all students. We have seen that, with proper support, such efforts can be highly effective in the context of two-way immersion programs in the U.S. (Smith, in press). Linguistically balanced populations, necessary for two-way programs, are rare in Mexico; Puebla communities with high concentrations of transnational students may prove to be exceptions.

This dilemma brings us to a final point about language policy as expression of language ideology, a notion implicit, but not directly addressed, in Kaplan's article. Although ideologies are inherently in flux rather than fixed, *retornados* and their teachers face a set of competing and particularly powerful language ideologies with respect to bilingualism and the acquisition of English. True, English fluency is highly prized in Mexico, functioning as a type of social passport for elite groups and as an instrument of (or impediment to) advancement for others. However, despite the obvious economic advances achieved by individual families, we see little evidence that returned transnationals are consequently perceived as models of or resources for language development.

References

Binford, L. (1999). Mexican migrants in New York and Mexico: New analytical and practical perspectives on transnationalization and incorporation. *La Vitrina, http://www.lavitrina.com/html/current/curren9/binford.html.*

Consulado General de México en Nueva York. (2000). *Informe sobre migración: Características de la comunidad mexicana en Nueva York* [Migration report: Characteristics of the Mexican community in New York City].

Cortés, S. (2001, October). *Flujo migratorio laboral hacia Estados Unidos: Puebla en el decenio de los noventa* [Labor and migration flow to the United States: Puebla in the 1990s]. Paper at the V Congreso de las Américas. Puebla: Universidad de las Américas-Puebla.

Domínguez Betancourt, E. (1995). *La efectividad del enfoque comunicativo para la enseñanza del inglés en las escuelas públicas secundarias de México* [The effectiveness of the communicative approach in the teaching of English in Mexico's public secondary schools]. Unpublished Masters thesis. Universidad de las Américas-Puebla.

Gendreau, M., & Giménez, G. (2000). Impacto de la migración y de los media en las culturas regionales tradicionales [The impact of migration and mass media on traditional regional cultures]. In M. A. Castillo, A. Lattes, & J. Santibáñez (Eds.), *Migración y fronteras* (pp. 173-196). Tijuana, Mexico: El Colegio de la Frontera Norte.

Kaplan, R. B. (2001). Language teaching and language policy. *Applied Language Learning, 12(*1), 81-86.

MacSwan. J. (2000). *The New York City Schools Research Report: Implications for Arizona's Proposition 203, the English-Only initiative.* Unpublished manuscript: Arizona State University.

Malkin, V. (2000). Divergent visions of education and anxieties over 'los jovenes en Estados Unidos': Mexican parents in New Rochelle, New York. *La Vitrina (http://www.lavitrina.com/html/curren9/program.htm).*

Ortega, L. (1999). Language and equality: Ideological and structural constraints in foreign language education in the U.S. In T. Huebner, & K. A. Davis (Eds.), *Sociopolitical perspectives on language policy and planning in the USA*, 16. Sociopolitical perspectives on language policy and planning in the USA (pp. 243-266). Amsterdam: Johns Benjamins.

Reyes, C. (2001, October 5). Diseña SEP apoyo a hijos de migrantes [SEP programs designed to support migrant children]. *Reforma*, pp. A-6.

Smith, P. H. (2001). Community language resources in dual language schooling. *Bilingual Research Journal, 25(*3), 375-404. [special issue devoted to recently completed dissertations].

Vega, A. (2002). *Práctica reflexiva y actitudes de profesores EFL en formación inicial hacia la superación profesional en México* [Reflective practice and the attitudes of Mexican EFL teachers in training towards professional development]. Unpublished manuscript, Departamento de Lenguas, Universidad de las Américas-Puebla.

Applied Language Learning
2003, Vol. 13, No. 1, pp. 85-90

Say, *Yes*! to The National Museum of Language

Lidia Woytak
Defense Language Institute Foreign Language Center

The National Museum of Language was dreamed up several years ago by Amelia Murdoch, Glenn Nordin and several other language professionals. Using their own savings and volunteering their time, they put their dreams into action. Because I truly support their efforts, I welcomed the opportunity to describe my vision of the museum..

I see the museum as a place dedicated to the preservation of our language heritage. I also see it as a place where every American can find information and inspiration on all language matters. Finally, I see it as a place bursting with such activities as language Olympics, poetry readings, and language discussion groups.

American English

At the museum, visitors will have an opportunity to acquaint themselves with a variety of documents illustrating the development of American English from its infancy, through the colonial times, the Civil War era to the present. Visitors can examine colonial documents written in a variety of languages. In 17th century Manhattan, for example, people used 18 different languages to communicate with one another. Visitors will have a chance to view samples of American English written in Danish, Dutch, Russian, French, German, English, and Spanish. They will note the use of Dutch and French in communications between Governor Bradford and the Dutch plantation in Manhattan. In those documents, they will also note certain Dutch words that are still in use today, including *bowery, cookie, scow, sleigh,* and *span*.

In many documents, visitors will note frequent use of English-based pidgins as a medium of communication between the indiginenous peoples and the explorers. Visitors will encounter Native American words such as *moccasin, parsimony, opossum, tomahawk,* and *hickory,* which are used to this day in mainstream American English. 18th century documents such as Captain Cook's Diaries include such Native American words as *tattoo, kangaroo,* and *taboo.*

Visitors to the National Language Museum will also have a chance to get acquainted with letters and diaries written by American pioneers. These letters describe their daily worries and successes. Visitors may wonder how these letters were delivered: by foot, by train, by stagecoach? Are they going to be lucky to see an authentic stagecoach from the period?

At the National Museum of Language , visitors will have an opportunity to review documents exemplifying the founding fathers' keen interest in the language of the new nation. In particular, they can skim through a proposal for an American Language Academy by John Adams, James Madison's call for freedom of the press, Benjamin Franklin's writings on national publishing, Thaddeus Kosciuszko's bequest for providing language education to African-Americans, and Noah Webster's *Dissertation of the American English*. Visitors will note how succinctly Webster, the famous "schoolmaster of America," summarized the sentiment of the nation in this statement: "As an independent nation, our honor requires us to have a system of our own, in language as well as in government."

Visitors can likewise glance through early editions of Webster's dictionaries, as well as the *Blue-backed Speller*, a *Compendious Dictionary of the English Language,* and the famous *American Dictionary of the English Language*. In the *American Dictionary,* Webster was the first to document distinctive American vocabulary such as *skunk, hickory, caucus,* and *chowder*. Moreover, in the introduction to this first American dictionary, Webster explained that speakers who used these forms determined their meanings. Therefore, such entries as *congress, senate, assembly,* and *court* reflect one meaning for U.S. users and another for British users. Thus, two centuries ahead of contemporary lexicologists and semanticists, Webster wrote that meanings were not to be found in words, but in the language as spoken by the people. Visitors will learn that Webster's dictionary of 70,000 entries surpassed Samuel Johnson's British masterpiece not only in scope but in authoritativeness as well.

Museum visitors may enjoy skimming through a variety of monolingual, bilingual, and multilingual dictionaries, noting the impact of immigrant and indigenous populations on American English. In the dictionary, *Americanisms*, they will note a number of words borrowed from French, German, Finnish, Spanish, and Polish.

Visitors also will have an opportunity to glance at the *Dictionary of American Regional English*, the product of 50 years of study begun in the 1930s by Kurath and completed by Carver in 1987. By viewing maps in the dictionary delineating American dialects, visitors can identify what dialects are spoken in their area. Additionally, they can examine the *Phonological Atlas* by Labov, which provides the first national survey of American phonology.

Literature

At the National Museum of Language visitors can view the exhibits of such American literary masters as Henry David Thoreau, Herman Melville, Mark Twain, Jack London, Henry James, Theodore Dreiser, Ernest Hemingway, William Faulkner, John Steinbeck, and many others.

Visitors can renew their acquaintance with Mark Twain, the author of *Adventures of Tom Sawyer* and *The Adventures of Huckleberry Finn*. They will have an opportunity to view a multitude of domestic and foreign editions

of these books. Just by looking at them, they can identify with these colorful youngsters, just like readers all over the world. In the *Adventures of Tom Sawyer*, Mark Twain created a real rather than an ideal boy. He went even further *in The Adventures of Huckleberry Finn.* By giving Finn, a backwoods boy, the power of a narrator, Twain preserved the regional speech of the period.

Visitors can also reflect on the exhibits devoted to masters of American poetry, including Walt Whitman, Vincent Benet, and Emily Dickinson. Approaching the portrait of Dickinson wearing a white dress, they can add details to her image, evoking her own words, "small, like the wren; and my hair is bold, like the chestnut burr; and my eyes, like the sherry in a glass that the guest leaves." Looking through the glass, they can decipher words of her poems scribbled on discarded envelopes and old bills. Visitors will encounter startling figures of speech, familiar words in unfamiliar uses, sudden shifts of tone, metrical irregularities, and deliberately imperfect rhymes. They can look at the first edition of her book published several years after her death.

At the museum, visitors can enhance their appreciation of African-American literature by glancing at poetry by Phillis Wheatley, and works of various genre by W. E. B. Du Bois and Frederick Douglas.

Language Awareness

Visitors will have an opportunity to examine charts, graphs, and maps denoting the multitude of 5,000 to 6,000 languages spoken on our planet. They will have the opportunity to view charts of families of contemporary languages, namely Proto-Indo-European, Semitic, and Niger Congo. They will note American English and 5,000 other languages of the Proto-Indo European family. Visitors will have an opportunity to skim through seminal research articles of these language families and view artifacts documenting them. Through lectures and exhibits, visitors will be made aware of the consequences of language decline and death, including the loss of identity and cultural heritage.

At the museum, visitors will come to view language as a living entity. It is born with its people, is used by them as a communication tool, and dies with them. Languages become endangered through the cultural assimilation of their users, frequently ethnic minorities with powerful neighbors who are at times subject to invasion. The list of endangered languages currently includes the tribes of Papua New Guinea (900 languages), the native people of the Americas (also 900 languages), as well as national and tribal minorities of Africa, Asia, and Oceania. European peoples such as the Irish, the Frisians, the Provençal, and the Basques are also included in this list.

Visitors to the Museum will have an opportunity to view a spread of several hundred Native American languages from the first European contact to the present day. Visitors will also have the opportunity to view samples of extant native writings such as the Cherokee, Cree, and Chippewa syllabaries. Moreover, visitors can listen to the staged recordings of American trade jargons such as Chinook and Mobilian. Visitors can also participate in simulations of

Native American sign language and the Planius developed by Kiowas.

Through exhibits and lectures, the visitors will learn that many Native American languages are facing extinction. Of the 187 languages still spoken in the United States and Canada, 149 are no longer learned by children as the primary language of the household. Of the 100 languages spoken in California in 1800, only 50 still have speakers, but today there is not a single California native language that is being learned as a primary language. Visitors will also learn that, if they do nothing about it, 80% of the remaining North American languages and all of the California native languages will become extinct with the passing of the present generation. The loss of any of these languages represents the loss of a human intellectual heritage, of all that could have been learned through that language about local language, culture, and social values.

Visitors will also be introduced to natural and artificial languages. An example of an artificial language, Esperanto, created for international communication, will be exhibited. Other artificial languages such as Cobol, Fortran, and Pascal — used in computer programming — will be also displayed.

Through exhibits and recordings, the museum will highlight the history of linguistics by focusing on such fields as semantics, structural linguistics, applied linguistics, sociolinguistics, and computational linguistics.

In the section on semantics, the study of meaning, visitors will learn that this field was introduced by Alfred Korzybski in the 1930s and popularized by S. I. Hayakawa, Stuart Chase, and Wendell Johnson in the 1940s and 1950s. Korzybski's observations of indigenous peoples led him to the conclusion that language with its structure and rhythm reflects the needs of its users. Consequently, he proclaimed that meaning is in people and that their environment dynamically influences shared meanings.

At the displays on descriptive linguistics of the 1950s, visitors will see photographs of field linguists talking to Native Americans. Linguists attempted to describe the grammars of languages that had no writing systems, including Chippewa, Ojibwa, Apache, or Mohawk. The display presents descriptive grammars and other artifacts related to these languages.

Nearby visitors can review a display on structural linguistics, where they can view constituent structure diagrams of sentences mapped by structural linguists. Focus on structure was inspired by the 1957 edition of Noam Chomsky's *Syntactic Structures*. Chomsky believed that the study of syntax should take precedence over the study of meaning.

Museum visitors will gain a better understanding of sociolinguistics, which focuses on the role of language in society. In this section, visitors can review seminal books and articles on American dialects, accents, and speech patterns. Visitors can listen to recorded samples of African American Vernacular English, Chicano Spanish, and an Appalachian variety. Visitors can also participate in debates on official language use, view exhibits on gender differences in speech, and explore ways of avoiding sexist language.

Future displays will introduce related fields such pragmatics and discourse analysis, which are based on the principle that one cannot interpret language in isolation from social and cultural contexts.

Visitors will also learn about connections between language and biology. Exhibits provide information about speech impairments such as dyslexia, autism, stuttering, and language loss. They will also learn about current research aimed at overcoming these impairments.

Foreign Language Teaching

The Museum also will house exhibits on foreign language learning and teaching. Visitors will see charts surveying the languages taught and learned throughout the world. Visitors can view exhibits exemplifying a variety of approaches to teaching foreign languages. These include "scientific," structural, audiolingual, notional and functional, communicative, and — prevalent in the last decade — proficiency-oriented approaches. Visitors can learn about the merits and pitfalls of these approaches by reading student and teacher journals kept during the instructional process.

Visitors will have an opportunity to view TV capsules of authentic language use in play or in action. Just listening to a TV broadcast in its native setting will make visitors aware of the inseparability of language and culture. Visitors can also participate in activities introducing them to a foreign language through a simple conversation, a song, or a game.

Visitors may enjoy a display of media used in language teaching. They will note a progression from heavy to light pocketsize tape recorders. They can review the progression of typewriters from small manual models, through electric keyboards, to word processors.

The Museum also features the achievements of outstanding American foreign language teachers and linguists. Visitors will learn more about career opportunities in foreign languages, acquainting themselves with the great need for specialized translators and interpreters in fields such as commerce, defense, medicine, and law.

Conclusion

I see the museum as a liaison for a variety of international and national language organizations, such as the American Association for Applied Linguistics, the American Language Society, the Modern Language Association, and American Council on the Teaching of Foreign Languages. Visitors also will have a chance to review language journals such as *Applied Language Learning, Foreign Language Annals, Language, Psycholinguistics, Sociolinguistics, Language and Society, Language Planning, Communication,* the *Modern Language Journal,* and many others.

I see the National Language Museum as a place where all language treasures are preserved, cherished, and shared. I also see it as a place where every American can find ancestral information and inspiration. Moreover, I see it as a place of reflection, and site for using language creatively. Finally, I see it as a place to which all Americans will be willing to contribute their own linguistic artifacts such as original manuscripts, records, letters, diaries or recordings of regional stories and interviews.

I hope my vision of the National Museum of Language will inspire you to share your thoughts about the museum with the readers of *Applied Language Learning*. Please submit your comments about the museum to:

Applied Language Learning
Defense Language Institute Foreign Language Center
Presidio of Monterey, CA 93944-5006
Email: AJ@pom-emh1. army.mil

For information on joining the museum, please contact,

The National Museum of Language
7100 Baltimore Avenue, Suite 202
College Park, Maryland 20740

Reviews

Contemporary Chinese Place Names: Names of Administrative Divisions at County and City Level. By Irena Katuzynska. Peter Lang (2002), ISBN 3-906762-67-X.

Reviewed by JIM JIELU ZHAO
Defense Language Institute Foreign Language Center

A common problem for many Chinese learners is to correctly pronounce and remember names of places. Much of the problem we have had in teaching names of places is that many of us may have lost sight of the fact that every place name has its own origin, its own legend, its own semantic features, and its own particular phonetic and morphological form. If we could incorporate such information into the teaching of place names, then learning Chinese place names would be definitely an interesting and productive process. *Contemporary Chinese Place Names* is just such an informative book that could really help us with the toponymical learning and teaching in Chinese. By analyzing research material which covers 1973 names of administrative units at county and city level of twenty provinces of the People's Republic of China, the author provides the reader a fairly detailed description of the various aspects of contemporary Chinese place names in terms of structure and meaning.

This book consists of 7 chapters, including an introduction, summary, bibliography and appendices. The author begins with a historical overview of Chinese toponymical research. Even though that many scholars claimed that toponomastics as a field of study originated in Europe in the first half of the 19th century, the history of Chinese toponymical studies can be traced back to the centuries Before Christ. "The first collections containing some information on place names are supposed to come down from Zhou and pre-Han times, although large portions of them evidently originated during the Han Dynasty. The list of those regarded by Chinese toponymy scholars as of the highest worth begins with *Shanhaijing* (Classic of the Mountains and Rivers), c. 4th cent. B.C., a treasury of early myths and legends, registering about 450 names of mountains and more than 200 names of bodies of water"(p.11).

The second and third chapters are the two most substantial and important ones in this book. In the second chapter, the author presents a thorough discussion of structural features of Chinese toponym. The author starts with the discussion of the constituent elements of place-name words: a specific part and a generic part. She then discussed different types of generic-specific structures and analyzed the generics and specifics from a morphological point of view. A great merit of this structural study on Chinese toponyms lies in the fact that the author examines place names from different perspectives and divides them into different categories. Having fully realized the limitations of structural-grammatical classification, the author studied the names of Chinese counties and cities as forms diachronically derived from appellatives and proper names. According to the author, "Chinese toponyms can be divided into primary place names (arisen from appellatives and proper names other than

geographical ones) and secondary names as the second and third-order place names (arising respectively from primary and the second-order toponyms" (pp. 87-88). From the point of view of syntactic constructions, Chinese place names can also be classified as "simple place names (irrespective of their morphological structure, on the basis that their specific parts are derived from appellative and proper names without any formal devices) and compositional formations (as their specific parts consist of elements derived from appellatives or proper names, and modified by other elements with the use of the formal device–the compounding method)" (p. 88). The findings of this study show that the only formal linguistic device used in Chinese place-naming is the compounding method, the method of producing compound words. That is why all typical contemporary Chinese place-name words comprise at least two syllables/morphemes/words.

In the third chapter, the author investigated the semantic contents of Chinese place names. By analyzing the meanings of the names of counties and cities, the author divided place names into nine major categories according to their semantic classifications. They are (1) descriptive names, (2) associative names, (3) incident-commemorative names, (4) possessive names, (5) commendatory names, (6) combined names, (7) names resulting from substitution by homophones, (8) mistake-names, (9) shift-names. Of all these nine major types of place names, the names shifted from other geographical entities form the largest group. Such names make up 32 percent (630 cases) of all the names included in the research material, because shifting-naming has proved to be the easiest and the most effective process to create an administrative division. The second and third largest groups of toponyms are made up of commendatory names (372 cases-19%) and descriptive names (296 cases-15%). This chapter includes many legends and historical facts that vividly explained how and why place names were formed in certain ways. This inclusion of such interesting materials definitely adds to the cultural and historical flavor of the text.

Chapter 7 with a wealth of bibliography is of great value for anyone who is interested in further studies on Chinese toponyms. This bibliography consolidates almost all available sources of research on Chinese place names. It consists of 371 items of Chinese references and 135 references in English and other languages.

All in all, *Contemporary Chinese Place Names* is a highly informative and enjoyable book that can serve as a useful reference for Chinese learners and teachers. However, one thing we have to keep in mind is that this book is not intended as a tool book or a dictionary on place names and the reader should not expect to use it as a handbook that is consulted only when a question arises. This book is a scholarly study on various aspects of contemporary Chinese place names. The analysis of the structural and semantic features of these place names provides us not just a good understanding of the linguistic forms of these names, but also a full comprehension of the origin and meaning of these names. In addition, by probing into the structures, origins, and changes of Chinese place names, we can also get a glimpse of social evolution throughout different dynasties in Chinese history.

***The Language Bridge to the Future: Army Language Master Plan*.** (2000). By Office of the Deputy Chief of Staff for Intelligence, Department of the Army. Washington, DC: DA, ODCSINT.

Kurt E. Müller
COL, USA, Ret.

The *Army Language Master Plan* (ALMP) is an important document for both researchers and language educators. Anyone researching language use in the military, language-education policy, or operations research and systems analysis for military personnel qualifications will find useful data as well as discussion of concepts that drive linguist management. Faculty in both the resident basic courses and the diversity of refresher programs that exist under various Command Language Programs will find in the ALMP indications of user skills to inform the development language curricula. But the major value of the document is as a departure point for exploring the challenges of linguist management rather than as a blueprint for resolving the issues it raises. Salient points of the study include:

- linguist requirements are routinely understated in both absolute numbers and proficiency levels required;
- translator-interpreter skills are growing in importance;
- documentation and management of language needs suffer in consequence of revised strategy, reduced end-strength, deactivation of units, and change in perception of threat;
- language skills are interdependent with common-soldier skills, job (MOS) skills, and team integration;
- linguist requirements grow during small-scale contingencies (SSC);
- language requirements are significantly over-weighted toward Russian and Spanish;
- in large-scale and protracted deployments language needs will not be fulfilled solely by military personnel.

Organized into five chapters, the body of the study is a series of briefing slides accompanied by interspersed text rather than a discursive presentation of issues. This option may be a strength that facilitates note-taking and review of the various issues. The chapters proceed to lay out a background for the study, a description of the current linguist force, a proposed force, an "implementation roadmap," and a short conclusion. The presentation acknowledges the interdependence of various communities that recruit linguists, define language requirements, teach language, manage personnel, and report readiness.

Among the strengths of the report are its attention to the development and management of a linguist force. Although the discussion of numerous points remains undeveloped, the enumeration of issues is highly useful. As an example, the study suggests a series of questions to determine whether a specific language requirement should be filled by a (US) military linguist or another source. The likely alternatives are contract civilians and allied or coalition military personnel. The US military linguist pool is foreseen as a mix of active and reserve component soldiers, and the authors devote attention to the determination of mobilization languages.

There are excellent reasons for reliance on the RC for languages required in response to a major theater war (MTW), but, with one exception that is based on an erroneous assumption, they go unstated. As a point of departure, force-development analysts should look at issues of constraints on availability of contract personnel, legal status of contract personnel (which may differ as a consequence of nationality), and a balance of language proficiency and dedication to mission. The study notes the provision of linguists by allied and coalition forces but fails to project circumstances that would constrain this source. For example, during the Korean War, augmentees to US forces were threatened with a draft into their own national forces. Consequently, SSCs are likely to exert fewer constraints than MTW, particularly in some languages.

The authors offer a few instances of staffing problems with multi-apportioned forces. An excellent example is drawn from the 101st Air Assault Division, which is projected to respond to two disparate theaters and fills its linguist requirement half for one theater, half for the other. A thoughtful mobilization plan might fill the shortfall from RC organizations, but there is no discussion of AC-RC integration, either tested in exercises or validated during operations. Moreover, rather than concentrating on perceived differences in response to MTW and SSC, the authors might have distinguished between unit call-up with planned support relationships and individual RC fillers in AC organizations.

The study asks whether translator-interpreter is an intelligence discipline and answers in the affirmative with a willingness to further develop this job skill under the supervision of the intelligence community. Embracing this skill as belonging in MI responds to a protracted problem in the National Guard linguist community, for which deployments to provide language support have sometimes been questioned as not contributing to MI disciplinary skills. The need for translators and interpreters (defined as one specialty, not separate ones, as the language profession defines these) recurs in several discussions, one of which illuminates the contract option by offering categories of local nationals and US personnel with varying security clearances.

We who study various aspects of linguist management have institutional biases that color our perception of mission challenges. In this instance, the intelligence bias seriously undermines the utility of the ALMP as an Army-wide document. This bias affects both the appreciation of the range of circumstances requiring language facility and the projected response. Had the study been undertaken by the Army Staff rather than conducted under contract, one would hope that pre-release coordination between Intelligence and Operations would have corrected the erroneous assumption that Presidential Selected Reserve Call-up (PSRC) is only applied to MTW, not SSC. Most of the authors' examples of contingencies, e.g., Haiti, Bosnia, Kosovo, were all undertaken with the assistance of PSRC, though perhaps not in MI.

The G-2 perspective is also evident in the authors' conclusion that the linguist force should concentrate on adversaries' languages. A G-3/J-3 perspective would conclude differently, particularly in the light of studies such as those of multilateral operations in Korea. The Operations Research Office of Johns Hopkins University investigated "language problems in Korea" and Major William Fox produced an enlightening work we would today recognize as a study of the lead-nation model of multilateral operations. Both studies offer significant evidence of the contribution that proficiency in Allies' or coalition partners' languages makes to command and control and unity of effort. To add recent evidence, DoD's final report to Congress on the Gulf War credits the language proficiency of Special Operators in the coordination center (C^3IC) with the achievement of unity of effort between multinational Arab forces under one commander and Western multinational forces under another. Since such coordination is no 0+ task, research is needed to reconcile the SF community's requirement for 0+ proficiency with DoD's post-conflict review. Moreover, the 0+ level applies to special forces, a branch and one of five SOF disciplines; it certainly does not apply to Civil Affairs or Psyop, two other SOF disciplines.

The study's greatest disappointment is in its treatment of special operations, and that community is largely responsible for this gap. SOCOM has long been reviewing its language requirements, and current data hamper consideration of the community's needs. But at 45% of the linguist requirement—the same as MI—a SOF perspective is essential.

In its summary numbers, the study appears to have counted the billets in Civil Affairs and Psyop, but in its consideration of the CA community (notably, in table 3.49), it appears to exclude an MOS and branch that account for 5,000 personnel of whom approximately 1,000 require language. Subsequent to the appearance of this study, the Civil Affairs community engaged in a Language Needs Analysis and specified proficiency levels for (team-based) mission-essential tasks. In that analysis, tactical battalions required L, R, and S level 2, and brigades and commands (which work at operational and theater levels) required L 3, R 3, and S 2+ or 3. Note the speaking requirements reflect the operations focus of these personnel.

Bulk reporting of SOF language requirements under SOCOM as compared to reporting languages by theater for the MI community does a disservice to planners who must project by language the need for language-maintenance materials. Since SOCOM supports regional theaters, the study would be much more useful if it compared SOF language requirements by region as it does for MI requirements.

The MI community undoubtedly has the most mature picture of its language needs, and its numbers comprise the prime influence on resident training at DLI, funding for maintenance training, and allocation of materials. That dominance militates toward some conclusions that are questionable for the rest of the military linguist community. Notably, the authors have taken issue with coding language billets without specifying a language and conclude that these codes should be replaced by language codes specific to one of two MTWs. This issue requires more discussion. As an example, consider a Foreign Area Officer with a specialty in South Asia and repetitive assignments in Bangladesh and India. If the personnel system accepted the ALMP recommendation and coded this position for an MTW language appropriate to his theater, he would be restricted to Chinese or Korean. The emphasis on MTW languages for psyop is reasonable, but for CA may be misplaced: theater-level units need a variety of languages, but to designate one position in public health for French and another in public transportation for German is about as effective as filling only half the 101st Division's MI requirement. CA employment has been so extensive that its force structure is scheduled to expand. Deployment experience demonstrates that its language needs are much closer to those of the AC agencies that comprise only 3% of the linguist requirement than to those of mobilization forces foreseen for MI.

The shortcomings of the ALMP are attributable to attempts to extend across the army observations appropriate to one branch. But a continuing dialogue among the diverse constituencies that work with language issues can use this study fruitfully to improve the Defense Foreign Language Program.

***Issues in English Language Education.* (2002).** By Marianne Nikolov. Bern: Peter Lang. Pp. 238.

Reviewed by JOHN HEDGCOCK
Monterey Institute of International Studies

Issues in English Language Education inquires into "second language acquisition (SLA) and foreign language learning (FLL) in general, and English language education in Hungary in particular" (p. 12). This book presents theoretical and empirical accounts of child and adolescent SLA/FLL, with a primary emphasis on the critical period hypothesis (CPH) (Lenneberg, 1967; Singleton, 1989; Singleton & Lengyel, 1995), attitudinal and motivational influences on language acquisition, and the impact of strategic skill development in classroom FLL. Nikolov's contribution is unique in that she draws extensively on her experience as a classroom EFL teacher in Hungary, where she has conducted numerous action research projects. She reports and synthesizes the results of these studies in support of her conclusions. Although the volume addresses important theoretical and practical issues of special concern to English as a foreign language (EFL) professionals, the book's title suggests a broader thematic scope than the chapters actually cover. This book's weak points unfortunately tend to compromise the insights that researchers and practitioners might seek from a work with such an ambitious title.

The first chapter essentially provides a critical examination of arguments for and against the CPH, as well as its impact on foreign language education programs. The author carefully appraises neurological, cognitive, social, and linguistic accounts of the critical period literature, referring explicitly to strong and weak versions of the CPH as supported by seminal studies of the last three decades (with the striking exception of UG-based accounts, which Nikolov apparently elected to forgo entirely). She observes that psycho-sociological variables, instructional methods, pedagogical materials, and teaching styles all represent a "smorgasbord of factors" that play a role in studies supporting or rejecting the CPH (p. 42). The author sensibly expresses reservations about the strong version of the hypothesis (i.e., that complete SLA/FLL is possible beyond an as yet undetermined biological age). In line with recent studies that refute the strong version of the CPH (e.g., Ioup, Boustagui, Tigi, & Moselle, 1994), Nikolov maintains that "native proficiency in SLA" is available to both adults and children, further arguing that "both SLA and FLL at an early age can be beneficial in the long run, but [that] there are certain pedagogical consideration to be borne in mind" (p. 43). Though accurate, the author's synthetic assessment of the critical period literature is less than comprehensive. Notably absent from her discussion, for example, are references to Birdsong (1999) and Herschensohn (2000), two recent and influential books on this topic.

Chapter two sets out to provide further empirical evidence to refute the strong version of the CPH by reviewing two of her own studies of successful adult learners of EFL and Hungarian. In line with Ioup et al. (1994), the author discovered that several learners in her research who had begun FLL

after the age of 14 were judged by native speakers to have native fluency in their respective target languages. Describing these participants as "exceptionally successful learners" (p. 84), Nikolov hypothesizes that their impressive SLA/FLL largely reflects integrative motivation and a strong desire to "sound like natives" (p. 84).

Chapter Three takes up the motivational thread by surveying research on the role of motivation in SLA in general and in child FLL in particular. Adopting an observably uncritical stance with regard to the validity of current motivational constructs, the author introduces a series of longitudinal investigations of Hungarian primary pupils' attitudes and motivational profiles. Her primary research questions focus on why children believe they learn English, which motivational factors influence their learning, and the extent to which these factors shape learning outcomes. Not unexpectedly, findings suggest that, as learners mature, integrative orientations give way to instrumental incentives. Nikolov consequently argues that the FL curriculum should reflect this natural evolution.

The topic of Chapter Four, children's strategy use, relates thematically to the motivational and attitudinal research summarized in the preceding chapter. Regrettably, the author frequently refers to the chapter as a "paper," and, in fact, it reads like a conference presentation. It thus suffers from rhetorical and theoretical weaknesses that undermine its value. Rather than surveying relevant strategy-based research and presenting empirical findings, the chapter unconventionally blends a discussion of cognitive and metacognitive strategy research with disparate (and therefore uninterpretable) bits of data from the author's numerous prior works.

Although Chapter Five does not exhibit the serious coherence problems of Chapter Four, it nonetheless presents anecdotal findings on negotiated classroom interaction that are difficult to evaluate and generalize. In this chapter, Nikolov endeavors to characterize how she implemented a process syllabus in EFL classes with adolescent and preadolescent pupils. Although the author refers to the work reported here as a "study," design features and methodology are almost entirely omitted, thereby limiting the chapter's usefulness to readers seeking replicable outcomes and meaningful implications. In the chapter summary, the author writes that "children acquired a lot of language, developed a favorable attitude towards English, the teacher, and language learning in general. They became self confident and responsible for their own learning" (p. 156). Although Nikolov is to be commended for reporting several "negative outcomes" (p. 156), she presents tenuous and impressionistic conclusions, which are supported by scant empirical data.

In contrast to Chapters Four and Five, Chapter Six presents comparatively more substantial data. The author reports on Hungarian EFL learners' pair and group interactions in 111 class sessions, with the explicit purpose of analyzing linguistic behaviors (primarily code-switching) and variables influencing pupils' willingness to interact in English. Following a survey of research on learner interaction in second language, immersion, and FL settings,

Nikolov describes her data collection procedures. In her analysis, she presents illustrative extracts from transcribed peer interactions, leading her to conclude that "on the whole, interaction and the amount of input in English were on a surprisingly low level" (p. 187). She insightfully attributes excessive L1 use to linguistic, social, affective, and situational factors that inhibit target language use in Hungarian EFL classrooms.

Chapter Seven presents an account of observations conducted in 118 EFL classes at 55 school sites throughout Hungary. Based on 12 research questions, the study was designed to "find out as much as possible about what goes on in classrooms" (p. 196) and to discern the Hungarian school system's readiness to implement an Examination Reform program. Though backed by descriptive statistical data, the insights summarized in this chapter are rather general, with the majority of the data tables lacking sufficient explanatory information to discern scales or units of measurement. The author's synthesis of observational and survey findings paints a discouraging picture of EFL education in Hungary, where teacher training, learner motivation, and educational resources are sorely lacking.

The flaws apparent in *Issues in English Language Education* cannot be overlooked. One of its chief defects is that it constitutes a patchwork: The author has apparently pieced together a number of papers and projects, a few of which have undergone little or no editing. Completely lacking a summary or concluding chapter, the volume does not draw its contents together in a coherent manner and leaves open a number of rather serious conceptual, theoretical, and empirical gaps. Compounding these weaknesses and the numerous flaws enumerated above are frequent stylistic errors and non-idiomatic features in the prose, which compromise the text's credibility, comprehensibility, and fluency. Finally, despite the author's judicious use of headings throughout the text, the volume's unconventional typography and format damage its presentational quality. For instance, rather than providing a single reference list or bibliography at the end of the volume, each chapter is followed by its own reference list. These individual lists frequently duplicate the contents of bibliographies elsewhere in the book and contribute to the impression that the chapters are self-contained and disconnected, rather than unified. Several chapter bibliographies also contain errors. Finally, numerous chapters contain pages with inexplicable blank spaces, leading the reader to wonder about the extent and professional quality of the editing that preceded the book's production.

Notwithstanding its disappointing faults, *Issues in English Language Education* reveals a number of interesting (though distressing) realities concerning the current state of English language instruction in Hungary, a setting that presents daunting challenges for classroom teachers. Readers specifically interested in exploring the nitty-gritty details of EFL teaching and FL learning in underrepresented contexts such as Hungary should find this volume to be a helpful resource.

References

Birdsong, D. (Ed.). (1999). *Second language acquisition and the Critical Period Hypothesis.* Mahwah, NJ: Erlbaum.

Herschensohn, J. (2000). *The second time around: Minimalism and L2 acquisition.* Amsterdam: John Benjamins.

Ioup, G., Boustagui, E., Tigi, M. E., & Moselle, M. (1994). Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition, 16,* 73-98.

Lenneberg, (1967). *Biological foundations of language.* New York: Wiley & Sons.

Singleton, D. (1989). *Language acquisition: The age factor.* Clevedon, UK: Multilingual Matters.

Singleton, D., & Lengyel, Zs. (Eds.). (1995). *The age factor in second language acquisition.* Clevedon, UK: Multilingual Matters.

General Information

ALL Index

Authors and Articles

Davis, Lynne. (1998). *Essay Scores as Instruments for Placement and Advancement in an Intensive English Program. 9*(1 & 2), p. 107.
Derwing, Tracey M. (1997). *Pronunciation Instruction for "Fossilized" Learners: Can It Help? 8*(2), p. 217.
Douglas, Dan. (1989). *Testing Listening Comprehension. 1*(1), p. 53.
Dunkel, Patricia A. (1992). *The Utility of Objective (Computer) Measures of the Fluency of Speakers of English as a Second Language. 3*(1 & 2), p. 65.
Dupuy, Beatrice. (1993). *Incidental Vocabulary Acquisition in French as a Foreign Language. 4*(1 & 2), p. *55.*
Dupuy, Beatrice. (1997). *Voices from the Classroom: Intermediate-Level French Students Favor Extensive Reading over Grammar and Give Their Reasons*. *8*(2), p. 285.
Dutertre, Ayça. (2000). *A Teacher's Investigation of Her Own Teaching. 11*(1), p. 99.
Ehrman, Madeline. (1998). *The Modern Language Aptitude Test for Predicting Learning Success and Advising Students. 9*(1 & 2), p. 31.
Eisenstein Ebsworth, Miriam. (1997). *What Researchers Say and Practitioners Do: Perspectives on Conscious Grammar Instruction in the ESL Classroom. 8*(2), p. 237.
Ellis, Rod. (1994). *Factors in the Incidental Acquisition of Second Language Vocabulary from Oral Input: A Review Essay. 5*(1), p. 1.
Feyten, Carine M. (1999). *Consciousness Raising and Strategy Use. 10*(1 & 2), p. 15.
Flaitz, Jeffra J. (1999). *Consciousness Raising and Strategy Use. 10*(1 & 2), p.15.
Ganschow, Leonore. (1992). *Factors Relating to Learning a Foreign Language among High- and Low-Risk High School Students with Learning Disabilities. 3*(1 & 2), p. 37.
Gardner, Robert C. (1991). *Second-Language Learning in Adults: Correlates of Proficiency. 2*(1), p. 1.
Garrett, Nina. (1989). *The Role of Grammar in the Development of Communicative Ability. 1*(1), p. 15.
Garrett, Nina. (1991). *Language Pedagogy and Effective Technology Use. 2*(2), p. 1.
Glass, William R. (1997). *The Effects of Lexical and Grammatical Cues on Processing Past Temporal References in Second Language Input. 8*(1), p. 1.
Gonzallez-Bueno, Manuela. (2001). *Pronunciation Teaching Component in SL/FL Education Programs: Training Teachers to Teach Pronun ciation. 12*(2), p. 133.
Granschow, Leonore. (1998). *Factors in the Prediction of Achievement and Proficiency in a Foreign Language. 9*(1 & 2), p. 71.
Han, Youngju. (2000). *Grammaticality Judgment Tests: How Reliable and Valid Are They? 11*(1), p.177.

Hedgcock, John. (2000). *Overt and Covert Prestige in the French Language Classroom: When Is It Good to Sound Bad. 11*(1), p. 75.
Hinkel, Eli. (2001). *Matters of Cohesion in L2 Academic Texts. 12* (2), p. 111.
Hinkel, Eli. (1994). *Pragmatics of Interaction: Expressing Thanks in a Second Language. 5*(1), p. 73.
Hodges, Rosemary. (1995). *Examining the Value of Conversation Partner Programs. 6*(1 & 2), p. 1.
Hokanson, Sonja. (2000). *Foreign Language Immersion Homestays: Maximizing the Accommodation of Cognitive Styles. 11*(2), p. 239.
Holznagel, Donald C. (1991). *Managing Innovation and Change for Instructional Technology. 2*(2), p. 45.
Hughes Wilhelm, Kim @ Rivers, Marilyn. (2001). An Audience Approach to EAP Writing Assessment: Learners, Teachers, Outsiders. *12*(2), p. 67.
Hussein, Anwar S. (1995). *Sociolinguistic Patterns of Arabic Native Speak ers: Implications for Teaching Arabic as a Foreign Language. 6* (1 & 2), p. 65.
Izumi, Shinichi. (2000). *Implicit Negative Feedback in Adults NS-NNS Conversation: Its Availability, Utility, and the Discourse Structure of the Information-Gap Task. 11*(2), p. 289.
Javorsky, James. (1992). *Factors Relating to Learning a Foreign Language among High- and Low-Risk High School Students with Learning Disabilities. 3*(1 & 2), p. 37.
Javorsky, James. (1998). *Factors in the Prediction of Achievement and Proficiency in a Foreign Language. 9*(1 & 2), p. 71.
Johnson, Adm. Jay L. (2000). *Language Training and Naval Operations from the Sea. 11*(1), p. 29.
Johnson, Ruth. (1997). *A Link Between Reading Proficiency and Native-Like Use of Pausing in Speaking. 8*(1), p. 25.
Johnson, Ruth. (1998). *Essay Scores as Instruments for Placement and Advancement in an Intensive English Program. 9*(1 & 2), p. 107.
Johnson, Yuki. (1997). *Proficiency Guidelines and Language Curriculum: Making ACTFL Proficiency Guidelines Effective in Furthering Japanese Language Proficiency. 8*(2), p. 261.
Kaplan, Robert B. (2001). *Language Training and Language Policy 12*(1) p. 81
Kennedy, Lt. Gen. Claudia J. (2000). *Meeting the Army's Language Needs. 11*(1), p. 9.
Kimbrough, Jeania. (1995). *Examining the Value of Conversation Partner Programs. 6*(1 & 2), p. 1.
Kitajima, Ruy. (2001). *Japanese Benefactive Auxiliary Verbs: The Relationship Between Noticing and Use. 12*(1), p. 55
Krashen, Stephen D. (1993). *Incidental Vocabulary Acquisition in French as a Foreign Language. 4*(1 & 2), p. 55.

Krashen, Stephen D. (1997). *Free Voluntary Reading as a Predictor of TOEFL Scores. 8*(1), p. 111.
Kumaravadivelu, B. (1994). *Intake Factors and Intake Processes in Adult Language Learning. 5*(1), p. 33.
LaRocca, Michela A. (1999). *Consciousness Raising and Strategy Use. 10*(1 & 2), p. 15.
Leaver, Betty Lou. (2000). *The World From the Perspective of a Peripatetic Pedagogue. 11*(1), p. 205.
Lee, James F. (1997). *The Effects of Lexical and Grammatical Cues on Processing Past Temporal References in Second Language Input.* 8(1), p. 1.
Lee, Sy-Ying. (1997). *Free Voluntary Reading as a Predictor of TOEFL Scores. 8*(1), p. 111.
Lefkowitz, Natalie. (2000). *Overt and Covert Prestige in the French Language Classroom: When Is It Good to Sound Bad. 11*(1), p. 75.
Leow, Ronald P. (1997). *The Effects of Input Enhancement and Text Length on Adult L2 Readers' Comprehension and Intake in Second Language Acquisition. 8*(2), p. 151.
LoCastro, Virginia. (1997). *Pedagogical Intervention and Pragmatic Competence Development. 8*(1), p. 75.
Lowe, Jr., Pardee. (1993). *Proficiency and Performance in Language Testing. 4*(1 & 2), p. 19.
Lowe, Jr., Pardee. (1998). *Zero-Based Language Aptitude Test Design: Where's the Focus for the Test? 9*(1 & 2), p. 11.
Lunberry, Clark. (1994). *Deviant English and the Para-Poetic. 5*(1), p. 93.
Markee, Numa. (1994). *Curricular Innovation: Issues and Problems. 5*(2), p. 1.
Matsuo, Naoko. (2000). *Varieties of Conversational Experience: Looking for Learning Opportunities. 11*(2), p.265.
McQuillen,Jeff. (1994). *Reading versus Grammar: What Students Think is Pleasurable and Beneficial for Language Acquisition*. *5*(2), p. 95.
Mecarty, Frances H. (2000). *Lexical and Grammatical Knowledge in Reading and Listening Comprehension by Foreign Language Learners of Spanish. 11*(2), p. 323.
Meunier, Lydie E. (1994). *Computer-Assisted Language Learning in Cooperative Learning. 5*(2), p. 31.
Mollering, Martina. (1995). *Pragmatics in Interlanguage: German Modal Particles. 6*(1 & 2), p. 41.
Moore, Rita. (1997). *A Link Between Reading Proficiency and Native-Like Use of Pausing in Speaking. 8*(1), p. 25.
Mora, Raimundo. (1995). *Silence, Interruptions, and Discourse Domains: The Opportunities to Speak. 6*(1 & 2), p. 27.
Munro, Murray J. (1997). *Pronunciation Instruction for "Fossilized" Learners: Can It Help? 8*(2), p. 217.
Nunan, David. (1993). *From Learning-Centeredness to Learning Centeredness. 4* (1 & 2), p. 1.

Nunan, David. (1995). *Pragmatics in Interlanguage: German Modal Particles. 6*(1 & 2), p. 41.
Olive, Floyd. (1998). *Essay Scores as Instruments for Placement and Advancement in an Intensive English Program. 9*(1 & 2), p. 107.
O'Mara, Francis. (1991). *Measurement and Research Implications of Spolsky's Conditions for Second Language Learning. 2*(1), p. 71.
Orr, Joseph. (2000). *Language Training Opportunities: Today and Tomorrow 2000 Command Language Program Manager Seminar. 11*(2), p. 367.
Otto, Sue K. (1991). *Training in Instructional Technologies: Skills and Methods. 2* (2), p. 15.
Overstreet, Maryann. (1999). *Fostering Pragmatic Awareness. 10*(1 & 2), p. 1.
Oxford, Rebecca L. (1992). *Language Learning Strategies: Crucial Issues of Concept and Definition. 3*(1 & 2), p. 1.
Oxford, Rebecca L. (1993). *Instructional Implications of Gender Differences in Second/Foreign Language Learning Styles and Strategies. 4*(1 & 2), p. 65.
Oxford, Rebecca L. (1996). *Employing a Questionnaire to Assess the Use of Language Learning Strategies. 7*(1 & 2), p. 25.
Oxford, Rebecca L. (1997). *A Gender-Related Analysis of Strategies Used to Process Written Input in the Native Language and a Foreign Language. 8*(1), p. 43.
Patton, Jon. (1992). *Factors Relating to Learning a Foreign Language among High- and Low-Risk High School Students with Learning Disabilities. 3*(1 & 2), p. 37.
Patton, Jon. (1998). *Factors in the Prediction of Achievement and Proficiency in a Foreign Language. 9*(1 & 2), p. 71.
Porto, Melina. (2001). *Second Language Acquisition Research: Implications for the Teachers. 12*(1), p. 45
Pusack, James C. (1991). *Software for Language Training: Directions and Opportunities. 2*(2), p. 61.
Rekart, Deborah. (1992). *The Utility of Objective (Computer) Measures of the Fluency of Speakers of English as a Second Language. 3*(1 & 2), p. 65.
Rivers, Marilyn, & Hughes Wilhelm, Kim. (2001). *An Audience Approach to EAP Writing Assessment: Learners, Teachers, Outsiders. 12*(2), pp. 177.
Rossiter, Marian J. (2001) *The Challenges of Classroom-Based SLA Research. 12*(1), p. 31.
Ryan, Gen. Michael E. (2000). *Language Skills in Expeditionary Aerospace Force. 11*(1), p. 13.
Schweers, C. William. (1997). *What Researchers Say and Practitioners Do: Perspectives on Conscious Grammar Instruction in the ESL Classroom. 8*(2), p. 237.

Reviews

van Lier, Leo. (1998). *Review: Cots: Teaching by Chattinb. 9*(1 & 2), p. 147.

Vanniarajan, Swathi. (2000). *Review: Pinker: Words and Rules: The Ingredients of Language. 11*(1), p. 213.

Vanniarajan, Swathi.(2001). *Review: Searle: Mind, language, and Society. 12*(2), p. 191.

White, Philip A. (1998). *Review: Lee and Van Patten: Making Communicative Language Teaching Happen: Directions for Language Learning and Teaching. 9*(1 & 2), p. 151.

Woytak, Lidia. (1995). *Review: Van Lier: Introducing Language Awareness. 6*(1 & 2), p. 91.

Editorials

Devlin, Col. Daniel D. (2000). *Military Linguists for the New Millennium. 11*(1), p. 1.

Money, Arthur L. (2000). *Language Skills and Joint Vision 2020. 11*(2), p. 235.

Mueller, Col. Gunther A. (2000). *Beyond the "Linguist": Global Engagement Skills. 11*(1), p. 15.

Reimer, Gen. Dennis J. (1997). *Army Language Needs for the New Century. 8*(2), p. 147.

Ryan, Michael E. (2000). *Language Skills in Expeditionary Aerospace Force. 11*(1), pp. 13-14.

Shelton, Gen. Henry H. (2000). *Preparing for the Future: Joint Vision 2010 and Language Training. 11*(1), p. 5.

Interviews

Woytak, Lidia. (1997). *Linguists in Action: Interview with Colonel Daniel D. Devlin. 8*(2), p. 295.

Woytak, Lidia. (1998). *Interpreter in Action: Interview with Lieutenant Colonel Richard Francona (Retired). 9*(1 & 2), p. 121.

Woytak, Lidia. (2000). *Leading the U.S. Army into the New Millennium: Interview with General Dennis J. Reimer. 11*(1), p. 33.

Calendar of Events

2003

Practical Applications in Language Corpora, 4–6 April, international conference, Lodz, Poland. Contact: Email: corpora@krysia.uni.lodz.pl

Northeast Conference on the Teaching of Foreign Languages (NECTFL), 10–13 April, Washington, DC. Contact: Northeast Conference, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; (717) 245-1977, Fax (717) 245-1976, Email: nectfl@dickinson.edu Web: www.dickinson.edu/nectfl

American Educational Research Association (AERA), 21–25 April, Chicago. Contact: AERA, 1230 17th St., NW, Washington, DC 20036-3078; (202) 223-9485, Fax: (202) 775-1824. Web: www.aera.net

International Society for Language Studies (ISLS), 30 April–2 May, first conference, St. Thomas, U.S. Virgin Islands. Contact: Email: isls@uconn.edu Web: home.earthlink.net/~isls/conf.htm

Fourth International Symposium on Bilingualism, 30 April–3 May, Tempe, AZ. Contact: ISB4, Arizona State University, PO Box 870211, Tempe, AZ 85287-0211; (480) 727-6877, Fax (480) 727-6875, Email: isb4@asu.edu Web: isb4.asu.edu

WorldCALL 2003, 7–10 May, Banff, Alberta, Canada. Contact: www.worldcall.org

National Association of Professors of Hebrew (NAPH), 18–21 May, Conference on Hebrew Language and Literature, University of South Florida, Tampa, FL. Contact: Gilead Morahg, Department of Hebrew Studies, University of Wisconsin, 1346 Van Hise Hall, 1220 Linden Dr., Madison, WI, 53706; (608) 262-3204, Email: gmorahg@wisc.edu Web: polyglot.lss.wisc.edu/naph/

CALICO 2003, 20–24 May, University of Ottawa, Canada. Contact: info@calico.org Web: calico.org.h/CALICO03/index.html

NAFSA: Association of International Educators, 25–30, May, Annual Conference, Salt Lake City, UT. Contact: Conference@nafsa.org Web: www.nafsa.org

ADFL Summer Seminar East, 12–14 June, New Haven, CT. Contact: Elizabeth Welles, Director, or David Goldberg, Associate Director, ADFL, 26 Broadway, Third Floor, New York, NY 10004-1789; (646) 576-5133, Email: adfl@mla.org Web: www.adfl.org

International Association for Language Learning Technology (IALLT), 19–21 June, Ann Arbor, MI. Contact: www.lsa.umich.edu/lrc/iallt/

ADFL Summer Seminar West, 26–28 June, Snowbird, UT. Contact: Elizabeth Welles, Director, or David Goldberg, Associate Director, ADFL, 26 Broadway, Third Floor, New York, NY 10004-1789; (646) 576-5133, Email: adfl@mla.org Web: www.adfl.org

Fédération des Professeurs de Langues Vivantes / World Federation of Modern Language Associations (FIPLV), 2–6 July, 21st World Congress, Rand Afrikaans University, Johannesburg, South Africa.

Contact: Anna Coetzee, Department of Afrikaans, Rand Afrikaans University, PO Box 524, Auckland Park 2006, Republic of South Africa; (+27) (11) 489-2698, Email: aec@lw.rau.ac.za Web: www.fiplv.org

American Association of Teachers of French (AATF), 4–7 July, annual convention, La Pointe du Bout, Martinique. Contact: Jane Abrate, AATF, Mailcode 4510, Southern Illinois University, Carbondale, IL 62901-4510; (618) 453-5731, Fax (618) 453-5733, Email: abrate@siu.edu Web: www.frenchteachers.org

Australian Federation of Modern Language Teachers Associations (AFMLTA), 10–12 July, annual conference, Brisbane, Australia. Contact: Angela Scarino, President AFMLTA, 9 Stanley Street, North Adelaide SA 5006, Australia; (+61) (08) 8302-4775, Fax (+61) (08) 8302-4774, Email: conference2003@afmlta.asn.au Web: www.afmlta.asn.au/conf2003.htm

Eighth International Pragmatics Conference, 13–18 July, Toronto, Canada. Contact: Jef Verschueren, IPrA Research Center, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk, Belgium; (+32) (3) 820 27 73, Fax (+32) (3) 230 55 74, Email: jef.verschueren@ua.ac.be Web: ipra-www.uia.ac.be/ipra/8th_conference.html

Fourteenth European Symposium on Language for Special Purposes, 18–22 August, Guildford, UK. Contact: LSP 2003, Department of Computing, School of Electronics, Computing and Mathematics, University of Surrey, Guildford, Surrey, UK; Email: lsp2003@surrey.ac.uk Web: www.computing.surrey.ac.uk/lsp2003

EUROCALL 2003, 3–6 September, University of Limerick, Ireland. Contact: June Thompson, EUROCALL Office, The Language Institute, University of Hull, Hull HU6 7RX, UK; Fax (+44) (0) 1482 466180, Email: eurocall@hull.ac.uk Web: www.eurocall.org/confs/cfp/euro2003cfp.htm

Third International Conference on Third Language Acquisition and Trilingualism, 4–6 September, Tralee, Ireland. Contact: Muiris O'Laoire, Dept. of Languages and Communication, School of Business and Social Studies, Institute of Technology, Tralee, Ireland; Email: molaoire@tinet.ie Web: www.spz.tu-darmstadt.de/projekt_L3

European Second Language Association (EUROSLA), 19–21 September, 13th annual conference, Edinburgh, UK. Contact: www.hw.ac.uk/langWWW/eurosla/eurosla03.htm

American Translators Association (ATA), 5–8 November, Phoenix, AZ. Contact: ATA, (703) 683-6100, Fax (703) 683-6122, Email: conference@atanet.org Web: www.atanet.org

American Council on the Teaching of Foreign Languages (ACTFL), 20–23 November, Philadelphia. Contact: ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, Email: headquarters@actfl.org Web: www.actfl.org

American Association of Teachers of Arabic (AATA), 20–23 November,

Philadelphia. Contact: John Eisele, Executive Director, AATA, Department of Modern Languages and Literatures, College of William and Mary, PO Box 8795, Williamsburg, VA 23187; (757) 221-7412, Fax (757) 221-3637, Email: aata@wm.edu Web: www.wm.edu/aata/

American Association of Teachers of German (AATG), 20–23 November, Philadelphia. Contact: AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (856) 795-5553, Fax (856) 795-9398, Email: headquarters@aatg.org Web: www.aatg.org

Chinese Language Teachers Association (CLTA), 20–23 November, Philadelphia. Contact: CLTA Headquarters, Cynthia Ning, Center for Chinese Studies, Moore Hall #416, University of Hawai'i, Honolulu, HI 96822; (808) 956-2692, Fax (808) 956-2682, Email: cyndy@hawaii.edu Web: clta.deall.ohio-state.edu

Modern Language Association of America (MLA), 27–30 December, location to be announced. Contact: MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863, Email: convention@mla.org Web: www.mla.org

American Association of Teachers of Slavic and Eastern European Languages (AATSEEL) and American Council of Teachers of Russian, 27–30 December, location to be announced. Contact: AATSEEL, Kathleen E. Dillon, Executive Director, PO Box 7039, Berkeley CA 94707-2306, Email: aatseel@earthlink.net Web: clover.slavic.pitt.edu/~aatseel

International Association of Teachers of Czech (IATC–NAATC) (formerly: North American Association of Teachers of Czech), 27–30 December, location to be announced. Contact: Neil Bermel, Department of Russian and Slavonic Studies, University of Sheffield, Sheffield S10 2TN, UK; (+44) (0) 114 222 7405, Fax (+44) (0) 114 222 7416, Email: n.bermel@sheffield.ac.uk Web: www.language.brown.edu/NAATC/index.html

2004

Southern Conference on Language Teaching (SCOLT), 18–20 March, Mobile, AL. Contact: Lynne McClendon, SCOLT, 165 Lazy Laurel Chase, Roswell, GA 30076; (770) 992-1256, Fax (770) 992-3464, Email: lynnemcc@mindspring.com Web: www.valdosta.edu/scolt

Southwest Conference on Language Teaching (SWCOLT), 25–27 March, Albuquerque, NM. Contact: Audrey Cournia, SWOLT, (775) 358-6943, Fax (775) 358-1605, Email: CourniaAudrey@cs.com Web: www.learnalanguage.org/swcolt

Teachers of English to Speakers of Other Languages (TESOL), 29 March–3 April, Long Beach, CA. Contact: TESOL, 700 South Washington Street, Suite 200, Alexandria, VA 22314; (703) 836-0774, Fax (703) 836-7864, Email: conventions@tesol.org Web: www.tesol.org

Central States Conference on the Teaching of Foreign Languages, 1–3 April, Dearborn, MI. Contact: Patrick T. Raven, Executive Director, PO Box 251, Milwaukee, WI 53201-0251; (414) 405-4645, Fax (414) 276-4650, Email: CSCTFL@aol.com Web: www.centralstates.cc/

American Educational Research Association (AERA), 5–9 April, San Francisco. Contact: AERA, 1230 17th St., NW, Washington, DC 20036-3078; (202) 223-9485, Fax: (202) 775-1824. Web: www.aera.net

Northeast Conference on the Teaching of Foreign Languages (NECTFL), 15–18 April, New York. Contact: Northeast Conference, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; (717) 245-1977, Fax (717) 245-1976, Email: nectfl@dickinson.edu Web: www.dickinson.edu/nectfl

International Reading Association (IRA), 9–14 May, annual convention, Toronto, Canada: Contact: International Reading Association, Headquarters Office, 800 Barksdale Rd., PO Box 8139, Newark, DE 19714-8139, (302) 731-1600, Fax: (302) 731-1057, Web: www.ira.org

American Association of Teachers of French (AATF), 18–23 July, Atlanta, GA. Contact: Jane Abrate, AATF, Mailcode 4510, Southern Illinois University, Carbondale, IL 62901-4510; (618) 453-5731, Fax (618) 453-5733, Email: abrate@siu.edu Web: www.frenchteachers.org

International Conference on Immersion and CLIC Education, September (dates to be announced), Kokkola, Finland. Contact: Jaana Laitinen, Email: janna.laitinen@kokkola.fi Web: www.kokkola.fi/sivistystoimi/virasto/index.htm

American Translators Association (ATA), 13–16 October, Toronto, Canada. Contact: ATA, (703) 683-6100, Fax (703) 683-6122; Email: conference@atanet.org Web: www.atanet.org

American Council on the Teaching of Foreign Languages (ACTFL), 18–21November, Chicago. Contact: ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, Email: headquarters@actfl.org Web: www.actfl.org

American Association of Teachers of German (AATG), 18–21 November, Chicago. Contact: AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (856) 795-5553, Fax (856) 795-9398, Email: headquarters@aatg.org Web: www.aatg.org

Chinese Language Teachers Association (CLTA), 18–21 November, ChicagoContact: CLTA Headquarters, Cynthia Ning, Center for Chinese Studies, Moore Hall #416, University of Hawai‘i, Honolulu, HI 96822; (808) 956-2692, Fax (808) 956-2682, Email: cyndy@hawaii.edu Web: clta.deall.ohio-state.edu

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning* (*ALL*) is to increase and promote professional communication within the Defense Language Program and academic communities on adult language learning for functional purposes.

Submission of Manuscripts

The Editor encourages the submission of research and review manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Research Article

Divide your manuscript into the following sections:

- Abstract
- Introduction
- Method
- Results
- Discussion
- Conclusion
- Appendices
- Notes
- References
- Acknowledgments
- Author

Abstract

Identify the purpose of the article, provide an overview of the content, and suggest findings in an abstract of not more than 200 words.

Introduction

In a few paragraphs, state the purpose of the study and relate it to the hypothesis and the experimental design. Point out the theoretical implications of the study and relate them to previous work in the area.

Next, under the subsection *Literature Review,* discuss work that had a direct impact on your study. Cite only research pertinent to a specific issue and avoid references with only tangential or general significance. Emphasize pertinent findings and relevant methodological issues. Provide the logical continuity between previous and present work. Whenever appropriate, treat controversial issues fairly. You may state that certain studies support one conclusion and others challenge or contradict it.

Method

Describe how you conducted the study. Give a brief synopsis of the method. Next develop the subsections pertaining to the *participants,* the *materials,* and the *procedure.*

Participants. Identify the number and type of participants. Specify how they were selected and how many participated in each experiment. Provide major demographic characteristics such as age, sex, geographic location, and institutional affiliation. Identify the number of experiment dropouts and the reasons they did not continue.

Materials. Describe briefly the materials used and their function in the experiment.

Procedure. Describe each step in the conduct of the research. Include the instructions to the participants, the formation of the groups, and the specific experimental manipulations.

Results

First state the results. Next describe them in sufficient detail to justify the findings. Mention all relevant results, including those that run counter to the hypothesis.

Tables and figures. Prepare tables to present exact values. Use tables sparingly. Sometimes you can present data more efficiently in a few sentences than in a table. Avoid developing tables for information already presented in other places. Prepare figures to illustrate key interactions, major interdependencies, and general comparisons. Indicate to the reader what to look for in tables and figures.

Discussion

Express your support or nonsupport for the original hypothesis. Next examine, interpret, and qualify the results and draw inferences from them. Do not repeat old statements: Create new statements that further contribute to your position and to readers understanding of it.

Conclusion

Succinctly describe the contribution of the study to the field. State how it has helped to resolve the original problem. Identify conclusions and theoretical implications that can be drawn from your study.

Appendices

Place detailed information (for example, a table, lists of words, or a sample of a questionnaire) that would be distracting to read in the main body of the article in the appendices.

Notes

Use them for substantive information only, and number them serially throughout the manuscript. They all should be listed on a separate page entitled *Notes*.

References

Submit on a separate page of the manuscript a list of references with the centered heading: *References*. Arrange the entries alphabetically by surname of authors. Review the format for bibliographic entries of references in the following sample:

Dulay, H., & Burt, M. (1974). Errors and strategies in child second language acquisition. *TESOL Quarterly*, *16* (1), 93-95.

Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw-Hill.

List all works cited in the manuscripts in *References*, and conversely, cite all works included in *References* in the manuscript. Include in reference citations in the text of the manuscript the name of the author of the work cited, the date of the work, and when quoting, the page numbers on which the materials that you are quoting originally appeared, e.g., (Jones, 1982, pp. 235-238).

Acknowledgments

Identify colleagues who contributed to the study and assisted you in the writing process.

Author

Type the title of the article and the author's name on a separate page to ensure anonymity in the review process. Prepare an autobiographical note indicating: full name, position, department, institution, mailing address, and specialization(s). Example follows:

JANE C. DOE, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St., Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Review Article

It should describe, discuss, and evaluate several publications that fall into a topical category in foreign language education. The relative significance of the publications in the context of teaching realms should be pointed out. A review article should be 15 to 20 double-spaced pages.

Review

Submit reviews of textbooks, scholarly works on foreign language education, dictionaries, tests, computer software, video tapes, and other non-print materials. Point out both positive and negative aspects of the work(s) being considered. In the three to five double-spaced pages of the manuscript, give a clear but brief statement of the work's content and a critical assessment of its contribution to the profession. Keep quotations short. Do not send reviews that are merely descriptive.

Manuscripts are accepted for consideration with the understanding that they are original material and are not being considered for publication elsewhere.

Specifications for Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to:

Applied Language Learning
ATFL-AP-AJ
ATTN: Editor (Dr. L. Woytak)
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006

Manuscripts should be typed on one side only on 8-1/2 x 11 inch paper, double-spaced, with ample margins. Subheads should be used at reasonable intervals. Typescripts should typically run from 10 to 30 pages.

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (4th Ed., 1994) available from the American Psychological Association, P. O. Box 2710, Hyattsville, MD 20784.

Review Process

Manuscripts will be acknowledged by the editor upon receipt and subsequently sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

Specifications for Floppy Disks

Preferably use Windows-based software. Format manuscripts produced on one of the DOS-based or Macintosh systems, as an ASQII file at double density, if possible. Please name the software used. MS Word or text documents preferred.

Copyright

Applied Language Learning
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006

Periodical Postage
Paid at Monterey, CA
And Additional
Mailing Offices

PB-65-02-1
United States Army
PIN: 080668-1/N 000
Approved for public release.
Distribution is unlimited.