

Change in Foreign Language Skills Over Time

Amber N. Bloomfield (abloomfi@umd.edu)

Megan C. Masters

Steven J. Ross

Stephen P. O'Connell

Kassandra Gynther

University of Maryland, College Park
7005 52nd Avenue, College Park, MD 20901

Abstract

Foreign language professionals invest considerable time and effort in acquiring foreign language skills. Of key interest is how these skills change over time, and which sustenance, or language training activities, are efficacious in maintaining or improving proficiency. This paper discusses the results of mining more than 800 test/re-test records of foreign language professionals. Analyses investigated the extent to which lag time between test occasions and formal language training impacted changes in listening and reading proficiency ratings. Results indicate that certain factors, such as initial proficiency level, affect both patterns of change and the rate at which foreign language skills manifest evidence of loss.

Keywords: foreign language attrition, foreign language assessment, language training

Introduction

The study of how foreign language proficiency changes over time is a relatively new area of research, often noted as receiving its first major impetus from a conference at the University of Pennsylvania in May 1980 (e.g., Clark & Jorden, 1984; De Bot & Weltens, 1995; Lambert & Freed, 1982; Weltens, 1987). In the few decades during which research on this topic has been active, a variety of papers have been published, though many of these studies have focused on which aspects of the foreign language are lost, and in what order (e.g., syntax vs. lexical knowledge first; Jordens, De Bot, & Trapman, 1989). Although this approach is valuable for describing the language aspects most vulnerable to loss, it does not target the question of what factors (e.g., conversing informally with friends in the L2) increase or decrease the rate of loss or how quickly general language abilities, such as reading comprehension, begin to show loss. Studies which have investigated factors influencing the loss of foreign language skills have examined the duration of the period of reduced input (i.e., time elapsed since peak language ability was attained), achieved proficiency level prior to the period of reduced input, and target language use during the period of reduced input, as well as other factors. The current study expands on previous research by examining the language skills of adults with foreign language proficiency in a variety of languages at multiple points of time to determine the rate at which loss

occurs, and how change in proficiency over time is affected both by formal language training and starting proficiency level.

Period of reduced input and change in language skills

Duration of the period of reduced input has been defined as time since the end of formal training (e.g., Bahrlick, 1984) or time since the end of intensive language exposure, such as an immersion experience (e.g., Snow, Padilla, & Campbell, 1988). More generally, this factor has been conceptualized as the amount of time since learners achieved their peak proficiency (Bardovi-Harlig & Stringer, 2010), though this characterization can be problematic when learners' abilities actually improve during the period (Gardner, Lalonde, Moorcroft, & Evers, 1987; Murtagh & van der Slik, 2004). The amount of time the learner has had to lose language skills is a particularly intuitive factor in explaining degree of foreign language loss, with the common sense prediction being that loss increases with elapsed time. Several studies have found evidence for such a relationship (e.g., Murtagh & van der Slik, 2004; Nagasawa, 1999; Reetz-Kurashige, 1999), though there are some findings suggesting the rate of loss over time may not be linear (Bahrlick, 1984). However, it is important to take into consideration other factors, such as the target language-specific activities the learner has engaged in during the period of reduced input, rather than just the time elapsed since some benchmark of language learning was attained.

One issue with research on the duration of the period of reduced input is that these studies explore duration as a discrete factor, investigating language skills at a few specific points in time (often at only one or two time points after the period of reduced input has begun). In general, there is a tradeoff in the literature between sample size and the number of time points at which the language skills of the participants are measured: those studies with sizable *n* (e.g., Clark & Jorden, 1984; Gardner et al., 1987; Murtagh & van der Slik, 2004) tend to measure language skills only twice, once at the beginning of the period of reduced input (i.e., the baseline) and again a set amount of time later. Cross-sectional studies, like Bahrlick (1984) and Snow et al. (1988), take only one measure of language skills for each participant and compare across groups that have

experienced different durations of reduced input (note that this method does not take into account potential differences in initial proficiency levels). By contrast, some studies compare language skills for the same individual at a number of time points, but tend to involve only very small groups of participants, generally children (but see Russell, 1999, for an exception), and to focus on the specific aspects of the language that are lost rather than loss of general language ability (e.g., Hansen-Strain, 1990; Reetz-Kurashige, 1999; Yoshitomi, 1999). The reason for the small number of studies exploring multiple time points with adult foreign language learners is likely attributable to practical constraints: it is difficult to longitudinally track and maintain contact with the same group of language learners over a long period of time. Yet, to explore how foreign language skills change over time, it is desirable to examine the skills of the same set of individuals repeatedly during the period of reduced input.

Language use during the period of reduced input

The extent to which foreign language is used during the period of reduced input is likely to be an important determiner of the amount of knowledge lost at the end of this period because it determines just how “reduced” the input during this period is for the learner. Clark and Jorden (1984) found that the learners of Japanese who did not show loss months after formal language training ended reported using the language more regularly than those who did show loss. In a similar study, Murtagh and van der Slik (2004) demonstrated that use of the target language after leaving formal language training predicted strength of language skills for learners of Irish eighteen months after leaving school. In a study with employees of the Canadian government, French-dominant bilinguals reported more opportunities to use their less dominant language (English) and also showed less loss of skills in their weaker language than did English-dominant bilinguals (Edwards, 1977, discussed in Oxford, 1982). Snow (1982) found that, at the group level, the lowest amount of loss was exhibited by Spanish immersion student groups that had the highest proportion of learners who continued to study the target language after the immersion ended (reported in Snow et al., 1988).

Achieved proficiency and change in language skills

Achieved proficiency in the L2 is important for assessing change in foreign language skills because it provides the baseline against which to compare current ability. This factor has also been considered in its own right as a potential predictor for foreign language loss. Having higher target language proficiency may lead to decreased loss over time (for reviews, see Bardovi-Harlig & Stringer, 2010; Weltens, 1987). One potential reason for this is that having greater foreign language proficiency may provide a learner with more available strategies to compensate for loss of specific foreign language knowledge. For example, a learner could use target language morphological knowledge to

uncover the meaning of a forgotten lexical item in the same way children use this type of information to comprehend unfamiliar words (Carr & Johnston, 2001). In addition, some theories of foreign language acquisition suggest that language knowledge, once it reaches a critical threshold, simply becomes resistant to loss (Bardovi-Harlig & Stringer, 2010). However, there are somewhat mixed results for the relationship between achieved proficiency and change in language skills over time, with several studies suggesting that higher proficiency learners do indeed experience less loss over time than do lower proficiency learners (Clark & Jorden, 1984; Gardner et al., 1987; Kaufman, 1995; Nagasawa, 1999) and others suggesting there is no difference in rate of language loss between higher and lower proficiency learners (Bahrick, 1984; Weltens & Grendel, 1993).

The Current Study

The current study explored change in general foreign language skills (reading and listening comprehension) for a large number of foreign language professionals who were tested multiple times (2-7 test occasions, with the majority having 3 or more) over a period as long as 6 years. The dataset included test histories for nearly 50 different languages; because of the small number of people testing in any one language, all analyses were completed in aggregate across the tested languages. Most people in the dataset were tested annually, but there was considerable variation in the frequency of testing. In addition to test records, information on participation in official language training was available for individuals in the dataset.

Method

The dataset included 1084 test histories for listening and 1085 for reading.¹ Each test event was associated with a rating of proficiency based on the raw score from the test; the raw score was not included in the dataset. Ratings are based on the Interagency Language Roundtable (ILR) scale for that particular skill: possible scores can range from 0 to 5, with 0 = *No Proficiency* and 5 = *Functionally Native Proficiency*; between each pair of adjacent levels is a “plus level” (e.g., 0+, 2+), *assigned when proficiency substantially exceeds one skill level and does not fully meet the criteria for the next level* (Interagency Language Roundtable, 2011).² For the purpose of analyses, all ILR ratings were recoded into numeric values.³ All scores were associated with a test date, so it was possible to calculate the amount of time, in days, from the first test administration to each subsequent test. Whether or not a person received language training, dummy coded as 0 = *no training* and 1 = *training*, was included in all analyses. Only a small subset

¹ A minority of individuals tested in more than one language; analyses treat each test history as a separate case.

² For more information on the ILR proficiency scales, see <http://www.govtirl.org/Skills/ILRscale1.htm>.

³ For example, 1 = 10, 1+ = 16, 2 = 20, 2+ = 26, etc.

of the individuals in the dataset had formal language training on record (~14%).

Initial target language proficiency was defined as the first ILR rating on record for any given language in the relevant skill (i.e., reading or listening). Due to the sparseness of individuals at some levels of proficiency, each person's first rating was coded as either *high*/2 = a rating of 2+ or greater or *low*/1 = a rating of 2 or lower. This recoded variable was used in analyses of skill loss. In addition, information about whether or not the test version changed during the testing history was available for the dataset. This factor is likely to be important when examining change in ratings, as the introduction of a new, unfamiliar test may decrease ratings even in the absence of loss of language skills. This factor was coded as 1 = *version change* and 0 = *no version change*.

Results

Overall patterns of change To explore the pattern of change across test administrations, a latent growth analysis (Bollen & Curran, 2006) was used to examine the direction and trajectory of change over time. In these latent growth analyses, the first score or intercept (*ICEPT*) and the growth over time (*SLOPE*) were dependent variables, while participation in formal language training (*Training*) and the occurrence of a test version change (*verdif*) were included as independent variables (see Figure 1). Because few test histories contained more than four tests, only the first four tests in the dataset were included in the latent growth analyses.

For listening test histories, 398 cases in the dataset contained at least 4 test records and 729 contained at least 3. The model fit the data adequately ($RMSEA = 0.053$, $CFI = 0.984$). The average slope was $M = 1.35$ ($SE = 0.15$; $t = 8.94$, $p < .001$), indicating a significant positive pattern of change (improvement) in listening scores over test administrations. The correlation between the intercept and

the slope was also significant for listening scores ($r = -0.37$, $p < .001$), indicating that people who started with lower initial ILR ratings had a faster rate of improvement over time than did those who started with higher ILR ratings. The *training* variable had a significant negative relation (-0.15) to the intercept for listening scores ($t = -4.27$, $p < .001$), indicating the people who participated in formal training had lower initial ILR ratings, perhaps indicating that these individuals self-selected for training. In addition, training had a significant positive relationship (0.20) with slope ($t = 3.09$, $p < .01$), indicating that the trajectories of improvement for professionals who had some kind of language training were higher than for those who had not participated in a training activity. The version change variable also had a significant relationship with slope, but this relationship was negative (-0.19 ; $t = -2.88$, $p < .01$). The direction of this relationship is intuitive: the introduction of a new test version decreases the rate of improvement. What is less intuitive, however, is the significant positive relationship (0.09) between the version change variable and the intercept ($t = 2.76$, $p < .01$). This relationship indicates that people who experienced a version change at some time in their testing history also tended to have higher initial ILR ratings. A new test version was introduced for only a subset of languages during the time covered by the test history data, so whether a version change occurred was partially dependent on the language tested. It may be that initial proficiency in this dataset was higher in those languages that experienced a version change, leading to the relationship between version change and the intercept. In fact, chi-square tests revealed that significantly more people with starting proficiency of 2+ or above experienced a version change for both listening and reading ($\chi^2(1) = 8.08$ and $\chi^2(1) = 8.32$, respectively; both $ps < .01$).

For reading test histories, 400 cases in the dataset contained at least 4 test records and 729 contained at least 3. The model fit of these data was successful ($RMSEA = 0.027$, $CFI = 0.997$). The average slope was $M = 1.03$ ($SE = 0.15$; $t = 7.08$, $p < .001$), indicating a modest but significant pattern of positive change over time. Unlike for listening ratings, the significant correlation between the intercept and the slope was positive ($r = 0.24$, $p < .05$); this indicates that people with higher initial ILR ratings tended to show greater improvement over time than people with lower initial ILR ratings. The reason for this difference between listening and reading ratings is not immediately clear.

As for listening, the *training* variable has a significant negative relationship (-0.20) with the intercept for reading ($t = -5.71$, $p < .001$), indicating that people who participated in training tended to have lower initial reading ILR ratings, and has a significant positive relationship (0.16) to the slope ($t = 2.67$, $p < .01$), suggesting that improvement was faster for those individuals who have received some type of language training. The version difference variable has a significant negative relationship (-0.19) to slope ($t = -2.88$, $p < .01$); as for listening, the introduction of a new test version leads to a slower rate of growth in reading ratings

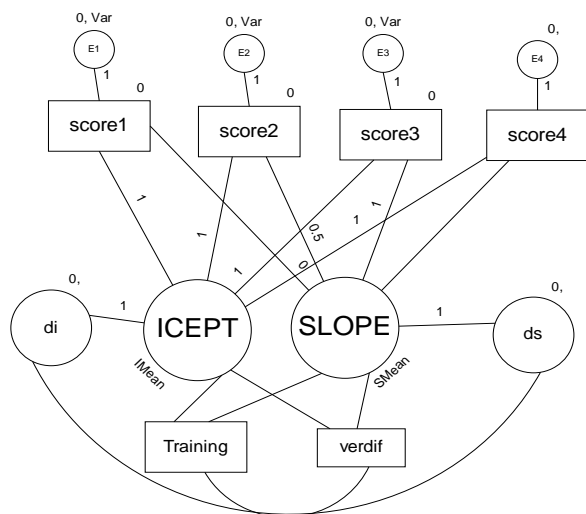


Figure 1. The latent growth model fit to listening and reading ILR ratings

over time. The same non-intuitive positive relationship between version change and the intercept was also present for reading scores (0.12; $t = 3.42$, $p < .001$). Again, the most likely explanation is that those languages for which a new test version has been introduced also tend to be languages where the individuals have a higher initial rating.

Ratings improved over time for both reading and listening, as indicated by the significant positive slopes in the latent growth analyses. Version change negatively impacted this trend, leading to a slower rate of improvement over time for listening scores but did not affect change over time for reading scores. Training, on the other hand, positively affected this trend, indicating that the formal language training resulted in faster improvement in ILR ratings for both listening and reading. While people with lower initial ratings showed a faster rate of improvement in listening, this pattern was reversed for reading. It may be that attaining higher proficiency levels in reading becomes easier as proficiency increases, while attaining higher levels for listening becomes more difficult as proficiency increases. This is a topic for future research.

Loss in ILR ratings To examine how the amount of time between test occasions affected the incidence of loss in ratings, event history analyses were conducted separately for reading and listening. A test history was coded as showing a loss if any subsequent test occasion produced a lower ILR rating than the first test occasion. In an event history analysis, the time between the first test and the test that yielded a lower score (i.e., lag time) is modeled to capture the amount of lag time for cases coded as losses compared to the lag time for cases where scores are sustained or increased from the first to the most recent test (i.e., non-loss cases).⁴ The goal of the event history analysis is an estimation of the average time lag associated with increasing incidences (events) of proficiency loss and the effect of any covariates on the rate of loss.

The version change factor was revised somewhat for the event history analyses. The factor captured whether a person who experienced a loss experienced a version change *prior to that loss*, rather than at another point in their test history. For individuals who did not show loss, this factor continued to indicate whether they had experienced a version change at any time in their test history.

Across all sets of listening test records, 17.7% of the cases were coded as showing a loss (i.e., at least one rating was lower than the first ILR rating on record). The event history analysis estimates the time lag associated with survival (i.e., not showing a loss) as a function of time beginning with 100% of the cohort surviving. The event history model assesses whether loss is a possible function of time between test occasions. For an event phenomenon

impervious to time, the rate of loss would be expected to remain close to zero across all observed time lags. However, the prediction based on previous studies (e.g., Nagasawa, 1999; Reetz-Kurashige, 1999) would be for the rate of loss to increase as time lag increases, which was in fact found for listening ratings. Event history analysis revealed that the projected survival rate for listening was fairly long, with over 80% of cases found to maintain or improve listening ratings with three years' lag between test occasions (see Figure 2).

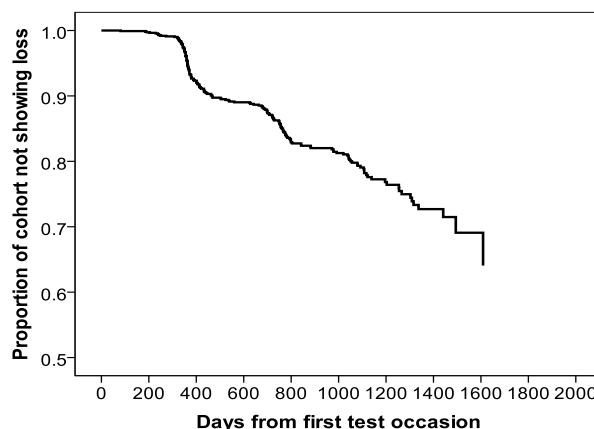


Figure 2. Loss in listening ratings over time

Version change was a marginally significant covariate in the event history analysis ($B = 0.32$, $p < 0.07$), with the predictable impact of increasing the rate of loss for listening: the odds of showing loss were projected to be roughly 1.37 times higher for people experiencing a version change than those who did not experience a version change. In addition, whether or not the person engaged in formal language training was entered as a covariate into the analysis. Participation in training did not significantly affect rate of loss for listening ILR ratings ($B = -0.15$, $p = 0.51$). However, the initial proficiency of the individual, coded as *low* for those with an initial ILR rating of 2 or lower and as *high* for those with first ILR rating of 2+ or higher, was a significant covariate ($B = 0.71$, $p < 0.01$, see Figure 3).

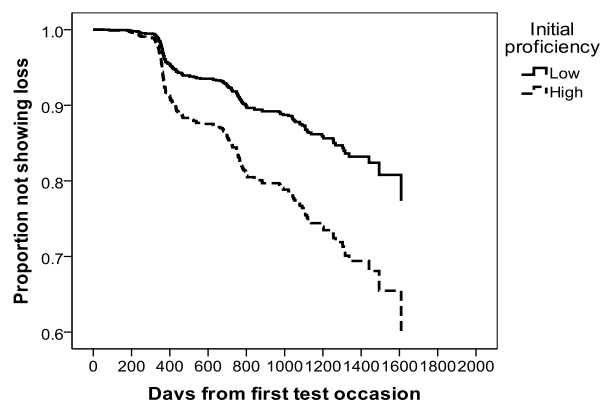


Figure 3. Initial proficiency and loss in listening ratings

⁴ Cases where the first rating on record was 0 (providing no opportunity to see a pattern of loss over time) and those where two test ratings were listed for the same date (suggesting the person was tested twice in one day) were excluded from event history analyses.

Surprisingly, however, those with higher scores exhibited a faster rate of loss, with the odds of this group showing loss projected as roughly 1.99 times higher than the odds for those with lower initial ratings. This finding is contrary to what has been found in previous studies (Clark & Jorden, 1984; Gardner et al., 1987; Kaufman, 1995; Nagasawa, 1999), and will be discussed in the Summary below.

Across all sets of reading test records, 17.6% of the cases were coded as showing a loss compared with the first score. As for listening, the incidence of loss increased with the amount of time between test occasions for reading tests scores (see Figure 4). Event history analysis projected the survival rate for reading to be comparable to that seen for listening: about 80% of cases were found to maintain or improve listening scores with three years' lag between test occasions.

Consistent with the analyses completed for listening scores, test version change was entered as a covariate into the event history analysis for reading scores. This factor failed to approach significance for reading scores, however ($B = 0.13, p = 0.48$). Whether or not the individual engaged in formal language training also failed to approach significance as a covariate for loss of reading ratings ($B = 0.12, p = 0.56$). However, as for listening ILR ratings, the first rating on record, coded as *high* = 2+ or higher and *low* = 2 or lower, was again a significant covariate ($B = 0.55, p < .05$; see Figure 5). The difference between the high and low proficiency groups was similar to that seen for listening scores, with the higher proficiency group showing a faster rate of loss than the lower proficiency group.

The results of the event history analyses reveal that the proportion of test cases in the dataset showing loss increased as the amount of time since the first test increased. However, training did not have a significant effect on the rate of loss for listening or reading. Because only a small portion (~14%) of the individuals included in the current analyses had formal language training on record, and only a subset of this group experienced a loss in ratings, it is possible that there was not enough power in these analyses to detect an impact of training on the rate of skill loss. Further, experiencing a test version change was a marginally significant covariate for the rate of loss of listening scores, but not for reading scores.

Summary and Conclusion

The findings from the current study are in line with previous results from studies examining change in foreign language skills in several ways. Consistent with previous research investigating the duration of the period of reduced input (e.g., Murtagh & van der Slik, 2004; Nagasawa, 1999; Reetz-Kurashige, 1999), the probability of loss increased as the amount of time since the first test increased for both reading and listening. These results indicate that, all other factors being equal (amount of foreign language use, motivation, etc.), ILR ratings for these individuals will tend to decrease as the amount of time between test administrations increases.

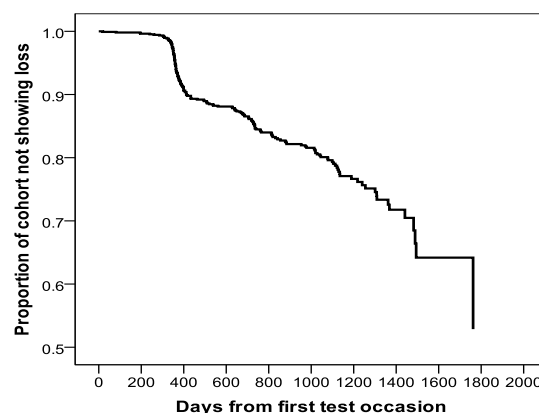


Figure 4. Loss in reading ratings over time

In contrast, the investigation of achieved proficiency (the degree of proficiency achieved in the language) and its effects on the rate of loss were not consistent with previous research findings. In the event history analyses, initial proficiency was a significant covariate for the rate of loss for all three skills, with higher proficiency individuals projected to have *faster* loss rates than lower proficiency individuals. This result is in the opposite direction from what is typically found when there is a difference in change for higher and lower proficiency learners (e.g., Clark & Jorden, 1984; Gardner et al., 1987). The current dataset contained many more individuals with *high* proficiency (a first score of 2+ or higher) than with *low* proficiency (a first score of 2 or lower), so it is possible that there was something unique about the individuals with lower initial ILR ratings that led to the relationship between initial proficiency and rate of loss. It is also possible that maintaining a higher rating is simply more difficult, with a smaller margin for error, leading to a faster rate of loss for this group. Further, there may be differences in motivation between the two groups, with the lower proficiency individuals more motivated to work to improve their language skills, and so slower to show loss over time. These possibilities will be investigated in future studies.

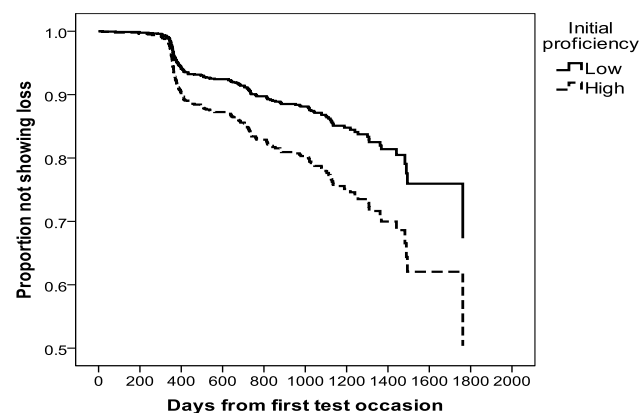


Figure 5. Initial proficiency and reading attrition

The current study also found no effect of whether an individual engaged in formal language training on the rate of loss for either listening or reading, though engaging in training did lead to a faster rate of improvement over time. This result runs counter to previous findings showing that the extent of target language use and exposure during the period of reduced input predicts retention (e.g., Clark & Jorden, 1984; Edwards, 1977; Murtagh & van der Slik, 2004). However, it is important to note that the information on formal language training was available for only a small subset of all people (14%), and that it is unclear why this subset of the sample participated in language training. It may be that these individuals were selected for training specifically because they struggled to maintain their foreign language skills. Further, the records of formal language training were the only information available about what the people included in the dataset might have been doing with their foreign language skills since the first language test in the dataset. It is very possible that large individual differences exist in the sample in terms of informal language training and other types of use and exposure (e.g., watching TV in the target language) during the period of time since the first test.

In conclusion, the current study introduced a method for investigating factors affecting change in adult foreign language skills in a longitudinal design, and described the results of an analysis of change in proficiency ratings for one group of foreign language professionals. Although this dataset was limited in several ways, including providing little information about use of the foreign language between proficiency tests, it did offer two or more data points for most individuals that were spread across a number of years. Further, the first ILR rating in the dataset offered a proxy for achieved language proficiency, so a given individual's current performance could be compared against his or her own previous abilities. A survey is currently being distributed to collect additional information about language learning history and current language use for professionals in this population to provide a more complete picture of the factors that affect change in foreign language skills.

Acknowledgements

This research was supported by the University of Maryland Center for Advanced Study of Language with funding from the Department of Defense.

References

- Bahrack, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General*, 113(1), 1-29.
- Bardovi-Harlig, K., & Stringer, D. (2010). Variables in second language attrition. *Studies in Second Language Acquisition*, 32(1), 1-45.
- Bollen, K. A., & Curran, P. J. (2006). *Latent curve models*: Wiley-Interscience.
- Carr, L., & Johnston, J. (2001). Morphological cues to verb meaning. *Applied Psycholinguistics*, 22(4), 601-618.
- Clark, J. L. D., & Jorden, E. H. (1984). A Study of Language Attrition in Former US Students of Japanese and Implications for Design of Curriculum and Teaching Materials. Retrieved August 4, 2011 from ERIC Document Reproduction Service, <http://www.eric.ed.gov>, No. ED243317.
- De Bot, K., & Weltens, B. (1995). Foreign language attrition. *Annual Review of Applied Linguistics*, 15(1), 151-164.
- Edwards, G. (1977). Second language retention in the public service of Canada. *Ottawa: Research Section, Official Languages Directorate*.
- Gardner, R. C., Lalonde, R., Moorcroft, R., & Evers, F. (1987). Second language attrition: The role of motivation and use. *Journal of Language and Social Psychology*, 6(1), 29-47.
- Hansen-Strain, L. (1990). The attrition of Japanese by English-speaking children: An interim report. *Language Sciences*, 12(4), 367-377.
- Interagency Language Roundtable. (2011). Descriptions of proficiency levels. Retrieved January 1, 2012, from <http://www.govtllr.org/Skills/ILRscale1.htm>
- Jordens, P., De Bot, K. D., & Trapman, H. (1989). Linguistic aspects of regression in German case marking. *Studies in Second Language Acquisition*, 11(2), 179-204.
- Kaufman, D. (1995). Where Have All the Verbs Gone? Autonomy and Interaction in Attrition. *Southwest Journal of Linguistics*, 14, 43-66.
- Lambert, R. D., & Freed, B. F. (1982). *The Loss of Language Skills*. Rowley, MA: Newbury House Publishers, Inc.
- Murtagh, L., & van der Slik, F. (2004). Retention of Irish skills: A longitudinal study of a school-acquired second language. *International Journal of Bilingualism*, 8(3), 279-302.
- Nagasawa, S. (1999). Learning and losing Japanese as a second language: A multiple case study of American university students. In L. Hansen (Ed.), *Second Language Attrition in Japanese Contexts* (pp. 169-212). Oxford: Oxford University Press.
- Oxford, R. (1982). Technical issues in designing and conducting research on language skill attrition. In R. D. Lambert & B. F. Freed (Eds.), *The Loss of Language Skills* (pp. 119-137). Rowley, MA: Newbury House.
- Reetz-Kurashige, A. (1999). Japanese returnees' retention of English-speaking skills: Changes in verb usage over time. In L. Hansen (Ed.), *Second language attrition in Japanese contexts* (pp. 21-58). Oxford: Oxford University Press.
- Russell, R. A. (1999). Lexical maintenance and attrition in Japanese as a second language. In L. Hansen (Ed.), *Second language attrition in Japanese contexts* (pp. 114-141). Oxford: Oxford University Press.
- Snow, M. A. (1982). *Graduates of the Culver City Spanish Immersion Program: A follow-up report*. Paper presented at the Sixteenth Annual TESOL Convention, Honolulu, HI.
- Snow, M. A., Padilla, A., & Campbell, R. (1988). Factors influencing language retention of graduates of a Spanish immersion program. *Applied Linguistics*, 9, 182-197.
- Weltens, B. (1987). The Attrition of Foreign-Language Skills: A Literature Review. *Applied Linguistics*, 8(1), 22-38.
- Weltens, B., & Grendel, M. (1993). Attrition of vocabulary knowledge. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 135-156). Amsterdam: Benjamins.
- Yoshitomi, A. (1999). On the loss of English as a second language by Japanese returnee children. In L. Hansen (Ed.), *Second language attrition in Japanese contexts* (pp. 80-111). Oxford: Oxford University Press.