And Then There Were None? Measuring the Success of Commercial Language Courses

Jeff McQuillan

Introduction

Self-study language courses have a long history in L2 education, but they have failed to capture the interest of researchers in applied linguistics (see Krashen, 1991, and Krashen & Kiss, 1996). These courses have been produced by several companies, including Berlitz, Pimsleur, Living Language, the Rosetta Stone and more recently, by independent producers via podcasts and websites (McQuillan, 2006). Some of the courses are limited to books and CDs, while others include an internet component or are delivered exclusively online. Lessons tend to focus on expressions related to travel or daily conversation, and are often guided by a grammatical syllabus.

Most companies that produce these courses have not released sales figures, but if the sales rank on the e-commerce site Amazon.com is a valid indication of popularity, at least some of the courses are among the most popular educational items sold. At the time this article was written (summer, 2018), for example, the book *Easy Spanish Step by Step* by Bregstein had an Amazon.com sales rank of 978, making it one of the top 1,000 books sold on the site. By comparison, books written by two recent U.S. presidential candidates Trump and Clinton had sales ranks of 8,548 and 3,990 respectively.

While popular, few of these courses have been evaluated for their effectiveness in promoting foreign language fluency. The only exception is Harry Winitz's *Learnables* (2003), which has been evaluated numerous times (Winitz & Reed, 1973; Winitz, 1981, 1996).

A rough measure of gauging the success of a self-study course is to look at perseverance in study. Do the students manage to reach the intermediate and advanced levels of the course? One of the earliest studies of persistence in foreign language study was conducted by Dupuy and Krashen (1998). The researchers collected background data and observed the classroom behaviour of a group of intermediate and advanced level college students. Their main interest was to document the characteristics of those who had "survived" the lower-level courses, and had advanced to the upper-division classes. They concluded that only a very small percentage of lower-division students did in fact reach what they refer to as the "Promised Land" of upper-division courses. Those who did advance in their studies had extensive exposure to the language outside of the classroom: 84.5 per cent had participated in a study abroad program.

Data from other sources confirm Dupuy and Krashen's findings on the high attrition rate in language courses. Table 1 summarizes data on foreign language course enrollment in high school over a period spanning 75 years. Coleman (1930) reported on statewide high school foreign language enrollments by level, for an unnamed northeastern U.S. state in 1925. Draper & Hicks (2002) provide more recent data from the year 2000, covering all 50 states. In addition to the raw figures, I have calculated the percentage of the students who "survived" each passing year of study, dividing the number of students at each course level by the total beginning (Level I) enrollment.

Table 1
Foreign Language Enrollments by Level in High School in 1925 (1 State) and 2000 (50 states)

Level	Total Enrollment, 1925	% Survivors	Total Enrollment, 2000	% Survivors
I	3,594	/	1,133,626	/
II	2,839	79%	797,800	70%
III	364	10%	346,200	31%
IV	160	4.5%	201,805	18%

Note. For the year 2000, the data includes any higher levels (e.g. Spanish V or VI) plus Advanced Placement courses. Data from Coleman (1930) and Draper & Hicks (2002)

Note that there are similar declines from Level I (freshman) to Level IV (senior) in both sets of data. Also, the attrition rate is particularly steep after the second year. In 1925, 95.5 per cent of students had dropped out of study before reaching Level IV. In 2000, it was marginally better at 82 per cent, but still high.

The situation does not improve at the college level. Furman, Goldbert, and Lusin (2007), report that of the 1,536,614 undergraduates enrolled in the top 15 foreign languages in the U.S. colleges in 2006, only 17 per cent were enrolled in upper-division courses. This attrition rate is similar to what we find at the high school level.

There are probably a variety of reasons as to why students fail to advance in formal language study. Many students take these courses as a requirement for graduation and therefore stop at the lowest class necessary to reach that goal. Ramage (1990) conducted a survey across a group of high school students (N=138). He asked them to indicate their agreement to statements related to their reasons for foreign language study on a 3-point Likert scale, in which "3" meant the subject agreed. A clear majority of the students indicated that one of the reasons they were taking the class was to fulfill a graduation requirement (mean score of 2.59). Yet, a significant number of students also said they genuinely wanted to learn the language. More than half of the students said they were "interested in learning to read and write" (mean score of 1.69), and nearly as many said they had a "particular interest in Spanish/French culture" (mean score of 1.59).

There is little data available on the motivations of independent adult second language students who pursue study outside the formal classroom. Presumably, some students have short-term objectives such as travel to another country, and seek only a very basic level of proficiency. However, as in the case of Ramage's (1990) subjects, many adults would no doubt like to reach higher levels of competency in the language.

Although there are few independent evaluations of commercial language programs, we do have some evidence on persistence rates within and across language course levels. I will summarize the results from three such studies: McQuillan (2008), who used unobtrusive or "non-reactive" methods of examining a set of course books used by adult acquirers; Nielson (2011), who conducted a quasi-experimental study of a group of mostly government workers who were given access to two commercial language courses;

and Ridgeway, Mozer, and Bowles (2017), who reported attrition rates based on more than 125,000 users of one of the largest self-study language courses, *Rosetta Stone*.

McQuillan (2008): Use of Library Self-Study Books

McQuillan (2008) attempted to gauge the persistence of independent language students who used print versions of self-study language books that they borrowed from a large, urban library system. He created an "unobtrusive measure" (Webb, 1966), called the "Wear and Tear" Index, to measure roughly how much of a book was read by the patrons. Similar indices have been used by previous researchers such as Debois (1963) and Moestller (1955) (as cited in Webb, 1966), to determine which parts of a library reference book were most frequently consulted, and which newspapers advertisements were seen by readers.

These indices often use a combination of markers indicating both "erosion" (physical degradation, such as tears in a book page or bent back page corners) and "accretion" (added dirt, dust, and smudges). McQuillan's Wear and Tear Index (2008), included both types of measures:

1. The separation of the pages close to the binding.
2. Fingerprints or smudges on the pages or the corners.
3. Worn or wrinkled corners likely caused by page turning.

He examined a set of 10 self-study language books, representing six different languages. He recorded the highest page number in the book that showed some evidence of one or more measures in the Wear and Tear Index. To ensure that there had been a sufficient amount of patron use of the books, only those books that had been in circulation for at least one year were examined. Table 2 (adapted from McQuillan's Table 1) reports the name of the book, the last page used, the total number of pages, and the

Table 2

Use of Commercial Language Courses by Library Patrons

Course Title (Publisher, Year)	Last Page Used	Total Pages*	Percent Read/Used
E-Z Spanish: A beginner's course (2001)	35	328	11%
Teach Yourself: Spanish (2003)	86	318	27%
Learn Spanish the fast and fun way (1997)	37	264	14%
Rush Hour Spanish(2003)	21	104	20%
Spanish Now! Level 1(2005)	80	505	26%
Portuguese for Dummies(2006)	20	304	7%
Learn German the fast and fun way (1997)	15	252	6%
Chinese for Dummies(2005)	38	314	12%
Teach Yourself Beginner's Italian (1999)	44	193	23%
Teach Yourself Cantonese (1995)	56	254	22%
		Average	16.8%

*Note. * = excluding glossaries or bilingual dictionaries at the end of the volume; adapted from McQuillan (2008)*

estimated percentage of use and progress in the course. Note that in no case did the average percentage of course book use exceed 27 per cent, with an average of 16.8 per cent read.

McQuillan's study may actually have understated the dropout rates, since he measured only the highest page number showing evidence of some use by patrons. This is not the same as the *average* use of the course book. One or two outliers could have used the book more extensively, skewing his estimates. For a more accurate measure of persistence, we need to measure course use more directly, as was done in the next two studies.

Nielson (2011): Users of Rosetta Stone and Tell Me More Software

Nielson (2011) looked at two groups of U.S. government employees working in agencies that provide self-study language training. The Rosetta Stone group (N = 150) consisted of employees from a number of different agencies. All the employees were "absolute beginners" in the language they had chosen to study (Arabic, Chinese, or Spanish), i.e. they had no previous coursework in the language. The Tell Me More group (N = 176) consisted of students employed by the U.S. Coast Guard. The Tell Me More group only studied Spanish, but unlike the Rosetta Stone group, the students were at various proficiency levels. All subjects were volunteers who had sought to participate in the study, and Tell Me More students had been given time off their regular duties to do so (up to three hours per week).

The Rosetta Stone group used a popular internet-based software program designed for self-study of languages. While the program is available on CD-ROM, participants could only access an online version, as per the procedure of the participating government agencies. The Tell Me More group used Aurolog's *Tell Me More* software, also available only online.

Rosetta Stone students agreed to use the course materials online for 10 hours per week for 20 weeks, giving them time to complete the recommended 200 hours for Level I of the courseware. Tell Me More students agreed to use the courseware for at least five hours per week for 26 weeks.

To measure the program's effectiveness, Rosetta Stone students were given proficiency interviews over the phone, in which they were asked to identify and describe pictures similar to the ones that appeared in their course. The tests were administered after the completion of each 50-hour segment of the 200-hour long study period. They were also given an ACTFL Oral Proficiency Interview (OPI) as an exit test.

The Tell Me More students took the program's placement and exit tests, as well as the Versant for Spanish oral proficiency assessment, which correlates highly with the OPI exam (Fox & Fraser, 2009). Tell Me More students who already knew some Spanish were given the Versant as a pre-test, and all students were to be given it as a post-test. Students in both groups were asked to keep a "learner log" to track how much time they had studied the materials.

Although Nielson's (2011) intention was to measure the effectiveness of the programs in promoting language acquisition, she found that "the most striking finding [for both groups]…was severe attrition in participation" (p. 116). I have summarized the attrition data from her study in Table 3. In both the Rosetta Stone group and the Tell Me More group, there were steep dropout rates, as indicated by the "percentage of survivors" column (the number of students reaching that level divided by the total number of students enrolled in the program at the outset). Nielson used different categories to report the data for the Rosetta Stone and Tell Me More groups (as noted in Activity column of Table 3). For the Rosetta Stone results,

Nielsen used signing into an online account as the third milestone, whereas for Tell Me More the third milestone was "Used TMM for 5 hours." Other differences are noted in column 1 of Table 3. Despite these differences, the pattern is very clear.

Of the students who signed up for the Rosetta Stone courses, only 21 per cent completed even

Table 3
Attrition in Participation Using the Rosetta Stone and Tell Me More Software Programs

Course Progress	Rosetta Stone (RS)	% Survivors*	Tell Me More (TMM)	% Survivors*
1. Volunteered and signed consent forms	150	/	176	/
2. Obtained RS account / Took TMM placement test	120	80%	103	58.5%
3. Actually signed into RS account / Used TMM for 5 hours	73	49%	61	35%
4. Spent more than 10 hours using the course	32	21%	17	9.6%
5. Completed lst RS assessment (50 hours) / Used TMM 1525 hours	21	14%	9	5%
6. Completed 2nd RS assessment (100 hours) / Used TMM more than 25 hours	6	4%	7	5%
7. Completed final RS assessments and OPI / Took TMM exit test	1	.6%	4	2.2%

Note. From Nielson (2011) Tables 1 and 2, pp. 116-117
** % Survivors = Number of students reaching level/Total number of student who signed up for the course*

10 hours of the 200-hour course (5 per cent of the total). Only one of the 150 volunteers made it to the end. For the Tell Me More course, less than 10 per cent made it to the 10-hour mark (13 per cent of the way through the course), with a mere four completing the final assessment. The attrition rate from beginning to end was 99.4 per cent for Rosetta Stone, and 97.8 per cent for Tell Me More.

Although the language proficiency assessments were taken by only a fraction of the participants, Nielson found that more hours spent on the course did produce better scores on the interim assessments. She concluded, however, that the number of subjects who took the exams was too small to be of much use in evaluating the effectiveness of the programs.

Nielson noted that not all of the attrition could be blamed on the programs. A significant percentage of the students apparently had a variety of technical problems with the software (browser plugins that would not load, system crashes, etc.). Some of the participants reported dropping out as they were assigned overseas during the course of the study, others cited not having enough time, or a change in their work situations. Most of them however, did not provide reasons for dropping out. In addition to the technological problems, there were also complaints about the content of the courses

themselves. Nielson (2001) concluded that the high dropout rates for the software programs meant that such self-study products are "unlikely to work by themselves", without proper support (p. 125).

Ridgeway, Mozer, & Bowles (2016): Institutional Use of Rosetta Stone

Ridgeway, Mozer, and Bowles (2016) analyzed a large set of data (N = 125,112) gleaned from students of the Latin American Spanish online course offered by Rosetta Stone. Rosetta Stone itself supplied its internal tracking data to the researchers; this data included "institutional" clients only (universities, governments and private companies). Ridgeway and colleagues looked at student enrollment and completion of the Level I, II, and III courses (beginning to advanced Spanish) between 2008 and 2014. Each level of the courses had 16 units, with a review/assessment activity at the end of the unit. Ridgeway et al. calculated the completion rates for each unit across the three levels, reporting the results by unit.

I summarized their findings by level, including the first unit, the 8th unit (mid-point), and the 16th and final unit, in Table 4. I estimated the approximate number of students on each level from Ridgeway et al.'s (2016) Figure 2 bar chart (p. 931). The data was reported by the researchers on a logarithmic scale, so my estimates are only approximate. I also calculated the "per cent of survivors" using the same approach as in my discussion on Nielson (2011) (number of students reaching that level divided by total enrollment in Unit 1).

Table 4
Attrition in Participation for Institutions Using Rosetta Stone

Course Progress	Level I Enroll.	% Survivors	Level II Enroll.	% Survivors	Level III Enroll.	% Survivors
Unit 1	86,000	/	12,000	/	6,000	/
Unit 8	25,000	29%	6,000	50%	2,000	33%
Unit 16	5,000	5.8%	1,000	16%	1,000	16%

Note. Data from Ridgeway et al. (2016).

The sharp decline in course completion is evident for all three levels. Dropout rates were highest in Level I, where only 29 per cent of the students completed the mid-point assessment, and just under 6 per cent made it to the end of the course. This is higher than Nielson's (2011) findings, although her data came from a small, more restricted sample. Level II students fared the best, with about one-half of them making it to the 8th unit, and 16 per cent completing the entire course. Level III students dropped out more quickly than in Level II, but had the same completion rate of 16 per cent. The overall dropout rate, calculating from enrollment in Unit 1, Level I, to completion of Unit 16, Level III, was similar to that reported by Nielson at 98.8 per cent.

Ridgeway et al. (2016) noted that there was a "sawtooth" pattern in the data, in that the number of students completing the final unit of Level I was lower than those starting Level II and the number of students finishing Level II was lower than those starting Level III. They attributed

this to the fact that "students tend to drop out within a level of a course, and new students join at the beginning of each of the three levels" (p. 931).

Discussion

Despite the use of very different methodologies and datasets, all three studies reviewed here reported similar outcomes: the number of students who make it to the end of independent self-study language courses is very small, falling somewhere between 1 and 16 per cent. At best, independent students appear to do no better, and usually worse than those enrolled in traditional language courses in high school and college.

Nielson's (2011) concluded from her data that independent adult language students would benefit from more "support" such as that offered in a traditional language classroom. However, attrition rates are only marginally better for students enrolled in "high-support" high school and college classes with live teachers. Indeed, it appears from Ridgeway et al. (2016), that those who enroll in the upper-level courses fare about the same as those in regular classrooms (roughly 16 per cent completion rate).

All three studies suffer from a potential design weakness, one that McQuillan (2008) noted: students who receive a free course may be less motivated than those who have paid for it. In both McQuillan (2008) and Nielson (2011), users of the course materials did not have to pay for the materials. Ridgeway et al. (2016), analyzed data from "institutional" users, where it was likely that the institution and not the individual user had purchased access to the software. Ridgeway and colleagues also note that "[s]ome institutions mandate the use of the software; others make the use optional," but "[w]e have no means of determining usage policy governing individual students" (p. 931).

There is some evidence that many online courses suffer from high attrition. Jordon (2015) examined the attrition rates of 129 "Massive Open Online Courses" (MOOCs), free courses for adults in a variety of fields, offered through websites such as Coursera and Open2Study. She found that the average completion rate for the MOOCs was 12.6 per cent, with shorter courses (fewer than five weeks) and those with auto-graded assessments doing the best at retaining students. This completion rate is in the range of the best-case scenario for language courses, at least at the upper levels.

Many adults who begin their self-study language courses probably do so with the goal of being fluent, or at least conversant, in the language. The data reviewed here indicates that this rarely happens. One possible cause of low completion rates may be poor teaching methods. Krashen (2013) noted that the language instruction provided by the most popular self-study course, Rosetta Stone, was "not very interesting, and a long way from compelling" (p. 2). He concluded that the limited amount of evaluation data on the program provided "only modest support for its effectiveness" and that "studies do not agree on users' reactions" to the course (p. 2).

Similar problems of uninspiring language teaching have been reported in studies of traditional classrooms. Tse (2000) noted that research from the 1970s found that a large percentage of students found their foreign language classes "un-stimulating and uninteresting" (p. 72). McQuillan (1994) reported that one of the most common yet least effective second language classroom activities, grammar study, was judged to be far less interesting for undergraduate students when compared to a rarely used but more effective approach, sustained silent reading (Krashen, 2004). Yet all the courses in the three studies reviewed here relied largely on these same traditional teaching methods, including grammar study,

decontextualized vocabulary drills, exercises, and (in some cases) output "practice".

Little consideration so far has been given by companies producing self-study language materials to teaching methods based on "comprehensible input" (Krashen, 1982, 2003). Given the steep dropout rates in self-study language courses and the demonstrated superiority of comprehensible input methods to traditional instruction (e.g. Krashen, 2004; Mason & Krashen, 2004), these companies may wish at least to consider more effective methods, for the sake of both their customers' satisfaction and their own bottom line.

References

Amazon.com. *Instructional Bestsellers*. Retrieved August 15, 2018 at https://www.amazon.com/gp/bestsellers/books/11811/ref=pd_zg_hrsr_books_2_4_last.

Bregstein, B. (2006). *Easy Spanish step-by-step*. New York, NY: McGraw-Hill.

Coleman, A. (1930). A new approach to practice in reading a modern language. *Modern Language Journal, 15*(2), 101-118.

Draper, J., & Hicks, J. (2002). *Foreign language enrollments in secondary schools, fall 2000*. Alexandria, VA: American Council on the Teaching of Foreign Languages.

Dupuy, B., & Krashen, S. (1998). From lower division to upper division foreign language classes: Obstacles to reaching the promised land. *ITL: Review of Applied Linguistics*, 119-120, 1-7.

Fox, J., & Fraser, W. (2009). Test reviews: The Versant Spanish test. *Language Testing*, *26*(2), 312-322.

Furman, N., Goldberg, D., & Lusin, N. (2007). *Enrollments in languages other than English in United States institutions of higher education, Fall 2006*. New York, NY: Modern Language Association.

Jordan, K. (2015). Massive Open Online Course completion rates revisited: Assessment, length and attrition. *International Review of Research in Open and Distributed Learning, 16*(3), 341-358.

Krashen, S. (1982). *Principles and practice in second language acquisition*. New York, NY: Pergemon Press.

Krashen, S. (1991). How much comprehensible input did Heinrich Schliemann get? *System*, *19* (3), 189-190.

Krashen, S. (2003). *Explorations in language acquisition and use: The Taipei lectures*. Portsmouth, NH: Heinemann.

Krashen, S. (2004). *The power of reading* (2nd ed.). Portsmouth, NH: Heinemann.

Krashen, S. (2013). Rosetta Stone: Does not provide compelling input, research reports at best suggestive, conflicting reports on users' attitudes. *International Journal of Foreign Language Teaching*, *8*(1), 2-5.

Krashen, S., & Kiss, N. (1996). Notes on a polyglot: Kato Lomb. *System, 24*, 207-210.

Mason, B., & Krashen, S. (2004). Is form-focused vocabulary instruction worthwhile? *RELC Journal, 35*(2), 179-185.

McQuillan, J. (1994). Reading versus grammar: What students think is pleasurable for language acquisition. *Applied Language Learning, 5*(2), 93-100.

McQuillan, J. (2006). Language on the go: Tuning into podcasting. *International Journal of Foreign Language Teaching, 2*(1), 16-18.

McQuillan, J. (2008). Does anyone finish the Berlitz tapes? A novel method of perseverance for commercial language courses. *International Journal of Foreign Language Teaching*, *4*(1), 2-5.

Nielson, K. (2011). Self-study with language learning software in the workplace: What happens? *Language Learning & Technology*, *15*(3), 110-129.

Ramage, K. (1990). Motivational factors and persistence in foreign language study. *Language Learning*, *40*(2), 189-219.

Ridgeway, K., Mozer, M., & Bowles, A. (2017). Forgetting foreign-language skills: A corpus-based analysis of online tutoring software. *Cognitive Science*, *41*, 924-929.

Tse, L. (2000). Student perceptions of foreign language study: A qualitative analysis for foreign

language autobiographies. *Modern Language Journal*, *84*(1), 69-84

Webb, E. (1966). *Unobtrusive measures: Non-reactive research in the social sciences*. New York, NY: Rand McNally.

Winitz, H. (2003). *The learnables: Book 1* (6th ed.). Kansas City, MO: International Linguistics Corporation.

Winitz, H. (1996). Grammaticality judgments as a function of explicit and implicit instruction in Spanish. *Modern Language Journal*, *80*(1), 32-46.

Winitz, H. (1981). A reconsideration of comprehension and production in language training. In H. Winitz (Ed.), *The comprehension approach to foreign language instruction* (pp. 101-140). Rowley, MA: Newbury House Publishers.

Winitz, H., & Reeds, J. (1973). Rapid acquisition of a foreign language (German) by the avoidance of speaking. *International Review of Applied Linguistics*, *11*(4), 295-317.

Jeff McQuillan worked at the School of Education at California State University, Fullerton and American Language School at the University of Southern California. He is a Senior Research Associate at the Center for Educational Development in Los Angeles, California.He has published widely in the areas of literacy development, first and second language acquisition, and bilingualism.

titleman@gmail.com

Call for Papers for the issues of LLT 2020

Language and Language Teaching (LLT) is a peer-reviewed periodical. It focuses on the theory and practice of language learning and teaching, particularly in multilingual situations.

We will henceforth invite contributions for papers for the forthcoming issues. The references must be complete in ALL respects, and must follow the APA style sheet. All papers must include an abstract (100-150 words) and a set of key words (maximum 6 keywords). Papers may address any aspect of language or language teaching. They MUST be written in a style that is easily comprehensible to school teachers, who are the primary target audience of this periodical. The articles may be centred around the learner, teacher, materials, teacher training, learning environment, evaluation, or policy issues. They must be original and should not have been submitted for publication anywhere else. A statement to this effect must be sent along with the article. Activities focusing on different aspects of language teaching are also invited.

The upper word limit (including the abstract, key words, references and a short bio-note) for each contribution in different sections of LLT is:

Article: 2000-2200 words (this could be extended to 3000 words if the article has some theoretical significance);

Interview: 2500-3000 words; Landmark: 2500-3000 words

Book Review: 1000-1500 words; Classroom Activity: 750 words; Report: 1000 words; Bio-Note: 30 words.

(You are requested to stick to the upper word limit.)

Papers must be submitted as a word document in MS Office 7. Please send the fonts along with the paper if any special fonts have been used. For images, please send jpeg files.

Last date for the submission of articles:

January Issue: July 15

July Issue: January 15

Articles may be submitted online simultaneously to the following email IDs:

agniirk@yahoo.com; amrit.l.khanna@gmail.com; jourllt@gmail.com