SELF-STUDY WITH LANGUAGE LEARNING SOFTWARE IN THE WORKPLACE: WHAT HAPPENS?

Katharine B. Nielson

University of Maryland

Many language training software packages are intended for self-study and marketed as complete language learning solutions; however, little is known about how well they work or under what conditions they should be used. This article reports on a research study conducted at the University of Maryland Center for Advanced Study of Language that explores the way adult learners use Rosetta Stone and Auralog's TELL ME MORE—two popular, commercially available, technology-mediated self-study packages. Volunteers from different United States government agencies agreed to use these programs according to each manufacturer's usage guidelines and to complete regular assessments to document their language proficiency throughout the study. The most striking finding was severe participant attrition, which was likely due to a variety of technological problems as well as the lack of sufficient support for autonomous learning in the workplace. This lack of compliance with self-study suggests that despite the logistical ease of providing language learning software, more resource-intensive types of language training are more likely to be effective.

Keywords: Computer-Assisted Language Learning, Distance Learning, Learner Autonomy, Online Teaching and Learning

Computers and the Internet have made foreign language self-study materials increasingly easy to access and use, and there are now many software applications marketed as complete language learning solutions, from free self-study courses such as the BBC's online language offerings or LiveMocha®, to for-pay options through companies like Rosetta Stone®, Auralog®, and Transparent Language®. These organizations advertise their products for self-study, and the commercial products are especially appealing for novice learners, stating, for example, that they will help "you achieve your language learning goals faster than you ever thought possible" (Rosetta Stone), or that their program is "is the most advanced language training program available" (TELL ME MORE ®). However, to date, there has been no independent empirical research to support these claims, and little is known about learner experience or learning outcomes when individuals use computer assisted language learning (CALL) products as stand-alone resources. After a discussion of the existing research on self-study, distance language learning, and language learning software, the current article reports on a study designed to explore how independent language learners use stand-alone language training programs. The research presented here investigates learner use of two popular commercially available products: Rosetta Stone and Auralog's TELL ME MORE.

PREVIOUS RESEARCH

There is no existing empirical research on learning outcomes from foreign language self-study using commercially available, stand-alone CALL materials. There is, however, research from related areas that suggests the most effective learning is not achieved by learners working alone, and that any materials designed as stand-alone, self-study solutions will have to compensate for this lack of interpersonal interaction. For example, researchers investigating learner autonomy, or "the ability to take charge of one's own learning" (Holec, 1981, p. 3), make it clear that achieving autonomy—a condition argued to be beneficial to the language acquisition process—does not necessarily come about as a result of self-study. In fact, according to Benson's (2007) literature review on autonomous learning, "learners do not develop the ability to self-direct their learning simply by being placed in situations where they have no other

option" (p. 22). That is, autonomy is learner-internal, and not a situational condition. This acknowledgement that self-directed learners require interpersonal support is made explicit in Fernández-Toro's (1999) training manual for foreign language self-study:

> For the learning experience to be successful, learners require appropriate support, not only in the form of learning materials (many of which are produced by teachers), but also advice and training. No resource centre can operate effectively without the backup of adequate human resources (p. 7).

In other words, learners engaged in self-study require more than just access to resources if they are to succeed.

This claim is supported in the research on self-access centers, which are resource centers (usually affiliated with a university or other institution) that provide independent language learners with self-study materials. The research on these centers generally suggests that self-directed programs require support systems in order to be effective for language learning. For example, in her review of European universities with self-access language centers, Littlemore (2001) determined that these centers function best when they provide learners with well-thought-out support, guidance, and training. She found that when self-access language centers were implemented by universities as a method to cut costs, learner satisfaction was lower than when the centers were well-staffed and designed to foster the language learning process through advising, peer work, and other guided practice. In her overview of the history of the self-access language center at the University of Hull, Mozzon-McPherson (2007) supported Littleton's finding by detailing the significant extent to which language advisors contribute to the autonomous language learning process.

In addition to the research on learner autonomy, a discussion of online, computer-mediated self-study should include some of the research findings from distance language instruction, which also indicate that interaction is critical to the success of online learners. While distance learning research has largely considered instructor-mediated distance courses, some of the findings are relevant to computer-mediated self-study. For example, a "sense of community" is critical to the success of any online learner (Liu, Magjuka, Bonk, & Lee, 2007; Sadera, Robertson, Song, & Midon, 2009; Rovai, 2002), and especially foreign language learners, who require an online community not only to engage them in learning but also to foster the second language acquisition (SLA) process (Compton, 2009; Fleming, Hiple, & Du, 2002; Hampel & Stickler, 2005; Lie, 2009; Murphy, 2008; White, 2006). In other words, any program intended to facilitate online foreign language acquisition must not only follow principles of effective online course design but also incorporate the elements of effective instructed SLA, including opportunities for output, interaction, and appropriate feedback, which generally require interpersonal communication (Blake, 2008, 2009; Lai, Zhao, & Li, 2008; Nielson & González-Lloret, 2010; White, 2006).

Since commercially available packages like Rosetta Stone and Auralog claim to work as all-in-one solutions for language learning, the software should be designed to establish the conditions conducive to SLA. However, there is very little empirical evidence about how to do this. This lack of research on materials is not unique to CALL. As Chappelle (2010) points out, "[t]he amount of published work on materials evaluation is surprisingly small in view of the impact that materials have in the instructional process" (p. 67). One of the few research studies on independent learning that cites specific foreign language software is Ulitsky's (2000) examination of the learning strategies of highly motivated, experienced second language (L2) learners using one of two multimedia software packages intended for autonomous use. The software packages were custom-made by instructors at the University of Albany, who used the Annenberg video series *Destinos* and *French in Action*, along with interactive exercises and quizzes based on the videos' contents. The participants in the study were experienced language learners who were either at the intermediate or beginning level in French or Spanish (the learners studied the

language with which they had the least experience). Ulitsky discusses various strategies employed by these experienced learners as they worked through the language learning materials, but important to the discussion here were the strategies of seeking out native speakers and using outside resources. All of the 26 participants in Ulitsky's study sought outside resources to complement the self-study process with these particular materials.

Another study that considered independent use of specific software was Murray's (1999) experiment with French learners using a simulated immersion software package, *Á la rencontre de Philippe*. This program allowed students to follow a pre-planned story line in a number of different directions, responding to questions and comments from the main character (Philippe) by selecting multiple choice text responses. As in Ulitsky (2000), participants in this study were highly motivated, autonomous learners who used a variety of outside resources with the software program, and there were no learning outcome measures to gauge the software's effectiveness.

None of the research on self-study with specific software packages examined the use of the "all-in-one" commercial solutions (e.g., Rosetta Stone, TELL ME MORE, Pimsleur, etc.) marketed to independent learners. There are several possible reasons for this. First, because these packages are intended to be used in lieu of in-person instruction rather than as a supplement to instructor-mediated classes, many universities, which tend to be the setting for CALL research, do not provide students with access to them, instead offering supplemental CALL materials. In addition, because of the shift in SLA to an interactionist approach (Gass & Mackey, 2007), much CALL scholarship concerns how learners use technological tools for interaction, for example, the wealth of research on computer-mediated communication (Thorne & Payne, 2005; Sykes, Oskoz, & Thorne., 2008), or how specific CALL applications can promote interaction (González-Lloret, 2003) rather than evaluations of software packages. That said, there are a few references to either Rosetta Stone or Auralog in CALL scholarship. Godwin-Jones (2007) mentions Auralog's TELL ME MORE and Rosetta Stone in his review of trends in self-paced instruction, stating that "these programs are built around pre-planned lessons with distinct goals prescribed in a linear, guided path," and he goes on to say that the audio, graphics, video, and speech recognition software make these products potentially very powerful (p. 11). Another paper that mentions commercially available resources is Lafford, Lafford, and Sykes (2007), which evaluates software for Spanish lexical acquisition. Lafford et al. (2007) uses theories of SLA to evaluate the potential for several different language learning products, including those provided by Auralog and Rosetta Stone. Overall, the authors found that most of the Spanish CALL products they reviewed fall short of creating the environment necessary for lexical acquisition, pointing out specifically that:

> *Rosetta Stone Spanish Latin America*, and Auralog's *TELL ME MORE Spanish* have outstanding technical infrastructure (e.g., excellent graphics, videos, pictures, and speech recognition software), but these products do not incorporate a number of the aforementioned research-based insights (e.g., the need for culturally authentic, task-based activities) that informed SLA scholarship might have given them. (p. 516)

Another mention of Rosetta Stone is in Saury's (1998) presentation on creating a psychological foundation for evaluating language learning software, in which she claims that Rosetta Stone is "one of the few software programs designed with an awareness of how multimedia can capitalize on the psychological processes of language learning" (p. 6). Saury then commends Rosetta Stone for having "consciously decontextualized" content so that learners can build an internal model of how language works. However, this claim was made in 1998 and is based on Krashen's theories of natural learning, which have fallen out of favor in more recent SLA research. While the boundary between explicit and implicit learning and the subsequent development of automaticity has not yet been clearly established, the idea that acquisition occurs with language out of context is not supported.

There is ample evidence that language learners require support, interaction, feedback, and appropriate materials to benefit from self-study. However, none of the studies on independent or online learning specifically looks at currently available commercial products that tout themselves as complete solutions. Although products are occasionally mentioned in the SLA literature, there are conflicting assessments of their potential for success. Despite the logical reasons for this lack of research (e.g., many universities do not provide access to stand-alone software packages, most CALL research is focused on interaction, self-study is not an easy area to research given its isolated nature), the fact that there has been no investigation into how these products work is something that merits attention. With for-profit companies marketing themselves to individuals, corporations, and non-profits as the panacea for language training, consumers, theorists, and language teachers deserve to know how well these products work and under what conditions they should be used.

THE CURRENT STUDY

While commercially available, stand-alone language products are not generally found in schools or universities, there is a large population of learners currently relying on these CALL resources: employees of the United States government (USG). Although there is a great deal of high-quality, face-to-face language training available to some USG employees—for example, agency-based language training centers such as the Foreign Service Institute, and the Defense Language Institute Foreign Language Center—there are agencies that lack in-house language training as well as the resources to send all interested parties for off-site instruction. Products from companies like Rosetta Stone, TELL ME MORE, and Transparent Language have been adopted by agencies across the USG as they attempt to find a scalable solution to make language training available to individuals for whom in-person instruction is impractical or impossible. The article will report the results of a two-part research study conducted by the University of Maryland Center for Advanced Study of Language (CASL) intended to explore how adult learners in USG workplaces use technology-mediated self-study packages.

There are many factors that could contribute to the success of self-directed learners, such as learner-internal beliefs and practices (Bown, 2006; White, 1999), the support available to learners (Murphy, 2008), the way the resources are used (White, 2005), and the content of the resources themselves (Lafford et al., 2007). Because of the number of potentially confounding variables with self-study, and because media or method comparison studies do not typically provide robust findings (Blake, 2009; Surry & Ensminger, 2001), the CASL researchers did not set up an experimental protocol to compare self-study packages—either to one another or to more traditional methods of language training. Instead, the study was designed to examine self-directed learner use and outcomes under typical agency workplace conditions in two different phases. The first phase (P1) examined the use of Rosetta Stone (RS) in Arabic, Mandarin, and Spanish throughout different government agencies, while the second phase (P2) examined the use of Auralog TELL ME MORE (ATMM) in Spanish by the employees in the U.S. Coast Guard (USCG). The conditions throughout the studies were designed to follow manufacturer's guidelines and to replicate the conditions under which self-study packages are currently used throughout the USG in order to answer the following research questions:

1.   Is self-study with stand-alone CALL materials an appropriate solution for USG employees?

2.   What language gain can be expected from 0-beginners who use CALL materials according to the manufacturer's instructions?

3.   Does RS work equally well with Arabic, Chinese, and Spanish?

4.   Is ATMM an effective tool for learners with previous Spanish training who want to improve their foreign language proficiency?

5.   Should supervisors and/or trainees be given any guidelines for self-study?

Based on the review of the literature on self-study and foreign language learning, the research team hypothesized that self-study CALL could be appropriate for some learners. Based on the content, scope, and sequence of both RS and ATMM described in the next section, the team predicted that 0-beginners would not gain much communicative competence but would master some vocabulary, and that learners with some prior language training would improve their Spanish proficiency after using ATMM. Finally, given the research findings from previous studies of autonomous learners, the final research hypothesis was that both supervisors and trainees would require guidelines for self-study.

METHOD

Participants

All participants in this research study were USG employees working for agencies that provide self-study CALL as a method of language training. They were recruited through announcements via e-mail and agency Web sites, which were generated by the USG clients who commissioned this research. In the P1 stage, employees from any level or position within participating agencies were given their first choice of target language (TL) as long as they were 0-beginners and therefore did not have any prior experience with the TL. Phase 2 was open to any USCG employee from any level or position who wished to study Spanish, so unlike the P1 participants, some of this group had prior experience with the TL. All volunteers who met the participation criteria were included in the study (P1: $n = 150$; P2: $n = 176$). All participants were motivated adult learners who sought out the opportunity for language training and readily agreed to the language study protocols required by this study. The Institutional Review Boards for the University of Maryland and the USG approved this research, and all participants gave informed consent.

Materials

Self-Study Materials

Participants in P1 used RS Version 2 (V2) in Arabic, Chinese, and Spanish, and P2 participants used ATMM Version 9 (V9) in Spanish; all software was accessed via the Internet—while both packages can be accessed via CDs, the procedure at the USG agencies participating in this research was to provide online access, so the CDs were not available to participants. Both RS and ATMM contain reading, writing, listening, and speaking exercises, and each includes voice recognition tools so learners can provide spoken responses (to match pre-determined scripts) in addition to comparing their recordings to those of a native speaker. Rosetta Stone is based primarily on exercises that require learners to match images with written or spoken language, while ATMM has a combination of vocabulary and grammar drills in a variety of formats, from fill-in-the-blanks to word scrambles, and also includes exposure to scripted dialogues as well as reading and listening passages. Rosetta Stone does not contain grammatical explanations or a glossary, while ATMM does provide these features. Both programs appear to have content sequenced in terms of grammatical complexity, with the units and levels in each program based on themes. Both programs track learner progress and recycle material based on performance. Both ATMM and RS provide limited exposure to vocabulary in context and genuine discourse. ATMM provides some specific cultural information.

Assessments

Participants in P1 took oral achievement assessments that were administered over the phone. Because the content of RS is heavily weighted toward vocabulary acquisition, and because participants were 0-beginners, existing general proficiency tests were unlikely to effectively measure learner progress with RS. Designed by CASL researchers, the over-the-phone achievement tests were criterion-referenced and dichotomously scored assessments, which were created specifically to assess the information taught throughout RS Level 1. The tests involved asking participants to go to a Web site to view several picture-

based prompts and then asking them to describe each image in the target language; after the description, learners were asked three specific questions about each image.[1] Three experienced language instructors (one each for Arabic, Chinese, and Spanish) scored the achievement tests. Each RS language course contained identical images and content, with equivalent descriptions of each image provided in each language, so the same achievement test was used for Arabic, Chinese, and Spanish, with the intention to compare success between groups. Participants in P1 were asked to take a telephonic American Council for the Teaching of Foreign Languages (ACTFL) Oral Proficiency Interview (OPI) as an exit test because this test was available in all languages, could be administered over the phone, and is intended to assess general proficiency.

Participants in P2 were asked to take the ATMM placement test, which comes with the software. In addition, participants in P2 took the Versant for Spanish oral proficiency assessment, which is a standardized test administered over the phone and computer-scored, using a scale from 20 to 80. Scores on this assessment correlate highly with other measures of oral language proficiency (e.g., a correlation of .88 with the ACTFL OPI and a correlation of .92 with OPI scores on the Interagency Language Roundtable (ILR) scale by government-trained raters [Pearson, 2008]). See http://www.ordinate.com/products/spanish.jsp or Fox and Fraser (2009) for more information on this assessment. The Versant Spanish test has been used in other empirical studies to measure learner proficiency after Spanish language training (Blake, Wilson, Cetto, & Pardo-Ballester, 2008; Burwell, González-Lloret, & Nielson, 2009). During P2, the Versant test was used as a pre-test for learners with some Spanish proficiency at the outset of the research and as a post-test for all participants.

Hardware

In addition to PCs, all participants were required to obtain a headset with a microphone to use the self-study materials and complete the assessments.

Procedures

The procedures for each phase were slightly different because of the nature of the language learning software and the research questions. In addition, because P1 was completed before P2, some of the findings from P1 influenced the protocol for P2.[2]

Phase 1 (June 2008): Participants completed a pre-study survey detailing prior language study/use and agreed to use RS for 10 hours per week for 20 weeks—following the manufacturer's guidelines that it would take 200 hours to complete level 1 of the program. They also agreed to study for no more than three hours in a particular day; this study protocol was included as a mechanism to keep study time distributed throughout the week. In addition, participants agreed to keep a weekly "learner log," in which they could record the time they spent working, technical difficulties, and whether or not they consulted any additional language resources. All P1 participants agreed to complete a listening/speaking assessment over the phone every five weeks or after 50 hours of study, and they agreed to complete an ACTFL telephone OPI at the end of 200 hours of study.

Phase 2 (November 2009): Participants completed the pre-study survey detailing prior language study/use that was administered during P1 and agreed to use ATMM for at least five hours per week for 26 weeks. This protocol was established to mimic the training requirements established by the USCG for ATMM participation. Each week, participants were given three hours of release time from their job duties to use the software at work, and they were asked to use the materials for an additional two hours per week on their own. Phase 2 included students who had some Spanish proficiency at the start of the research; these participants agreed to take the Versant for Spanish test as a pre-test as well as the ATMM placement test to determine starting level. In addition, all participants agreed to keep a weekly "learner log" as well as to complete two exit tests: the ATMM proficiency test and the Versant for Spanish computerized oral proficiency test.

The purpose of this research was to determine how well self-study works under typical agency workplace conditions. When resources for independent learning are made available at the agencies employing participants in this study, they are generally provided with little guidance or supervision. However, previous research on autonomous learning made it clear that learners require support. For this reason, the lead researcher on this project maintained frequent contact with participants, working with agency personnel to provide software licenses, and serving as the point of contact for any issues encountered during the study. In addition, she provided encouragement, monitored progress, and followed-up individually with all participants who were not regularly using the software. The research team scheduled all assessments for both P1 and P2 by working with participants to identify convenient times.

RESULTS

Software Usage

The most striking finding of both P1 and P2 was severe attrition in participation. Despite initial participant interest as well as active researcher involvement and encouragement, participants in both phases of the study spent very little time using the CALL materials before stopping completely, if they used the materials at all. In fact, many participants never managed to access the software—51% ($n = 77$) of the P1 participants and 41% ($n = 73$) of the P2 participants never logged in to their accounts. Only one participant in P1 and four participants in P2 completed the full study protocol. Tables 1 and 2 provide participation records.

Table 1. *Record of Participation in P1 (RS)*

Participant Activity	Arabic	Chinese	Spanish	Total
Volunteered and signed consent forms	50	50	50	150
Obtained RS accounts	50	37	33	120
Actually accessed accounts	38	19	17	73
Spent more than 10 hours using *Rosetta Stone*™	18	13	5	32
Completed the first assessment (50 hours of use)	13	5	3	21
Completed the second assessment (100 hours of use)	5	0	1	6
Completed third and fourth assessments and OPI (200 hours of use)	1	0	0	1

Table 2. *Record of Participation in P2 (ATMM)*

Participant Activity	0-beginner	Non-zero beginner	Total
Signed consent forms and received accounts	82	94	176
Took Auralog Placement Test	52	51	103
Used Auralog for .5 to 5 hours	32	29	61
Used Auralog for 5.5 to 10 hours	7	9	16
Used Auralog for 10.5 to 15 hours	5	12	17
Used Auralog 15.5 to 25 hours	2	7	9
Used Auralog for more than 25 hours	0	7	7
Took Versant Post-test	3	19	22
Took Auralog Exit Test	0	4	4

Assessment Outcomes

The attrition from both P1 and P2 meant that the assessment outcomes data were extremely sparse. There were 21 P1 participants (Arabic = 13, Chinese = 5, Spanish = 3) who took the first assessment, which was designed to test learners on mastery of the content presented in the first 25% of RS Level 1. There were five Arabic participants and one Spanish participant who took the second assessment; none of the Chinese participants took the second assessment. There was only one Arabic participant who completed the full study protocol and took assessments three and four. The single successful Arabic participant received a 96% on the first interim assessment, and then perfect scores on the final three assessments. Because of low *n*-sizes, it is not statistically sound to compare scores between language groups. In general, using the software prepared some students to describe and answer questions about the images on the assessments.[3] See Figure 1 for a graph of the student scores on the first assessment by the number of hours they spent using RS.
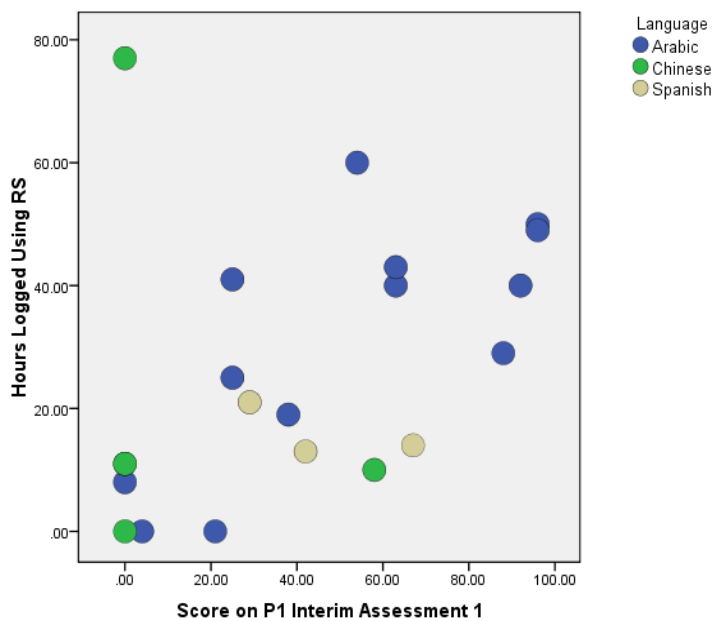

Figure 1. Scores on P1 Interim Assessment #1 by number of hours logged by each participant using RS.

Due to the low *n*-size as well as the fact that the interim assessments were designed to be achievement tests rather than proficiency tests, it is difficult to generalize about potential proficiency outcomes after using RS based on participant performances; however, there are some observable trends in the available data. For example, hours logged using RS in Arabic correlates with performance on the interim assessment ($r = .72$, $p = .005$). In addition, learners studying Spanish were able to score an average of 46/100 points on the interim assessment after an average of just 16 hours. Learners studying Chinese appeared to have a harder time than either the Arabic or the Spanish participants with the interim assessments—four of these participants received a score of 0 on the test, with hours logged ranging from 0 to 77. Additionally, it is worth noting the large standard deviation values for both the test scores and the hours logged, especially for Chinese and Arabic, indicating that the means do not reflect the distributions well and that there were huge variations in both the number of hours logged and performance on the achievement tests. See Table 3 for the mean hours logged and scores earned for the P1 interim assessment.

Table 3. *Descriptive Statistics for P1 Interim Assessment 1*

Language	Data	*N*	Min.	Max.	*M*	*SD*
Arabic	Score	13	0	96	51.15	34.91
	Logged Hours		0	60	31.08	19.51
Chinese	Score	5	0	58	11.60	25.94
	Logged Hours		0	77	21.80	31.20
Spanish	Score	3	29	67	46.00	19.31
	Logged Hours		13	21	16.00	4.36

The proficiency outcomes data from P2 are very similar to P1—sparse. The instrument used to measure learner proficiency in P2 was the Versant for Spanish oral proficiency exam. Subjects who complete this test receive an overall proficiency score as well as a score for each of the following categories: sentence mastery, vocabulary, fluency, and pronunciation. For more information on how the test is scored, see Blake et al. (2008). There were only three 0-beginners who took the post-test Versant Spanish assessment. The 0-beginners were not pre-tested because they identified as being absolute novices with no previous exposure to or instruction in Spanish. Table 4 shows the P2 0-beginner group's ATMM placement test scores (scored from 0 to 10), their Versant post-test scores (scored from 20 to 80), and the number of hours they logged using ATMM prior to the assessment. It is striking that these participants hardly used the software at all—the longest a 0-beginner used ATMM was an hour and twenty minutes[4]—and that their Versant test scores, for participants claiming to be 0-beginners with under two hours of instruction, were so high. Their placement test scores also indicate that they were not absolute beginners (particularly participant 265, who tested at the level of an intermediate student); however, they are included as a separate group here because they self-identified as novices and, for that reason, they were not asked to complete a pre-test.

Table 4. *P2 Test Scores for 0-beginners*

Subject Number	ATMM Placement Test Score	Time Logged (h:mm)	Overall Score	Sentence Mastery	Vocabulary	Fluency	Pronunciation
P206	2.5	1:19	24	20	20	21	42
P279	1.1	1:42	23	20	20	32	36
P265	4.4	0:20	40	28	20	55	55

There were 19 non-0-beginner participants in P2 who took the Versant for Spanish exit test. Based on their pre-test scores, these participants were divided into three groups: Novice (n = 12), Intermediate (n = 4), and Advanced (n = 3).[5] Table 5 provides the descriptive statistics for these groups.[6] The novice and intermediate learners spent significantly more time using ATMM (novice mean hours logged = 14:06; intermediate mean hours logged = 15:45) than either the 0-beginners (mean hours logged = 1:07) or the advanced learners (mean hours logged = 3:20).

Table 5. *P2 Mean Test Scores for Participants with Starting Spanish Proficiency*

Proficiency Level	N	ATMM Placement Test Score	Time Logged (h:mm)	Overall Gain Score	Sentence Mastery Gain Score	Vocabulary Gain Score	Fluency Gain Score	Pronunciation Gain Score
Novice	11	3.12	14:06	2.55	2.90	0.09	2.18	2.64
Intermediate	4	4.25	15:45	1.25	6.00	-1.25	0.75	-0.50
Advanced	3	6.60	3:20	0.00	0.00	0.00	-4.60	-1.00

Note. One novice participant who logged 134 hours with ATMM is excluded here.

The results for the novice group and the intermediate group were similar in some respects. Both groups improved their overall scores on the Versant test, averaged roughly the same number of hours with the software (novice = 14:06 and intermediate = 15:45), and improved their sentence mastery scores (which is a measure of grammar/syntax) and their fluency and pronunciation scores (which measures timing, rhythm, and phonology). Neither group improved its vocabulary scores (the intermediate scores dropped by 1.25 points). The advanced group did not show any improvement in its overall scores, and in some areas, the advanced participants' scores went down after using the software. This decrease for the advanced group is likely a ceiling effect from the Versant for Spanish proficiency instrument. The learners tested at the very top of the scale for the pre-test, and the categories for which they lost points were fluency and pronunciation, which the advanced learners could have inadvertently manipulated by deliberately speaking slowly. As a group, the three advanced learners hardly used the software at all, so their decrease in proficiency likely has nothing to do with their use of Auralog.

Figure 2 depicts the change in scores on the Versant test during P2 by the hours logged by all of the participants for each of the three proficiency levels—this figure does not include the 0-beginners because they did not take a pre-test. As suggested by the descriptive statistics, many of the scores are clustered around 0, with 9 of the participants using the software for up to 20 hours and having their scores stay the same. There were 5 participants who used the software from 10 to 28 hours who had gain scores of 3 to 7 points.[7] There is one outlier in the data, the single novice participant who used the software for 134 hours (close to the total number of hours required by the P2 protocol) and increased her overall Versant proficiency score by 10 points.
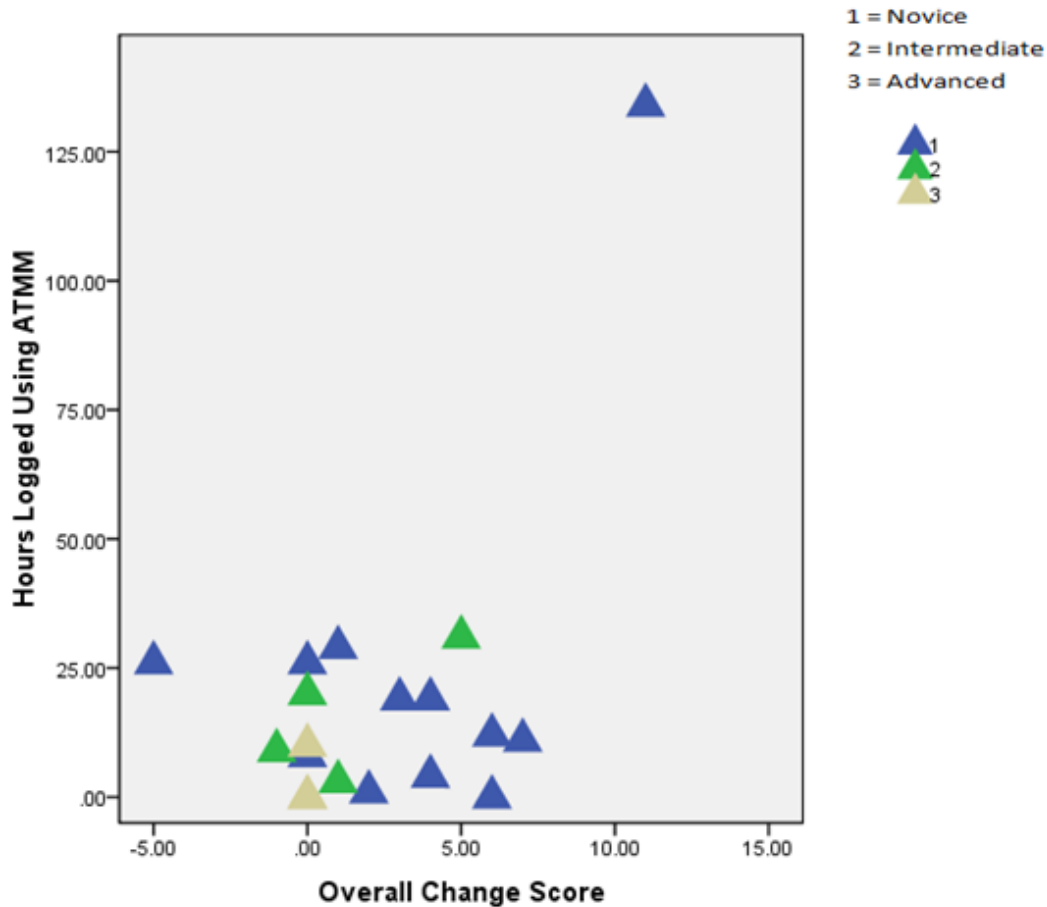
Figure 2. Overall P2 Versant gain score by hours logged using ATMM.

Learner Feedback

In keeping with their limited usage of the self-study CALL materials, the participants in P1 were similarly non-compliant with the learner logs, which were intended to collect information about learner use of the software, technical problems, and the use of outside materials and resources. In the first week of the study, only 46 of the 73 participants completed the learner log. By week four, 19 participants used the learner log, and by week 10, there were just three. Figure 3 shows P1 learner log usage throughout the 20 week study, with a rapid decrease in the first half of the study (weeks one to 11). Unfortunately, there are no learner log data from the P2 participants; the USCG agreed to distribute a weekly survey to all study participants in order to collect their learner use data, but was never able to do so because of internal technical problems. Learner feedback from the participants in P2 came from personal e-mails to the CASL researchers and an end-of-study anonymous survey completed by 30 P2 participants (17%).
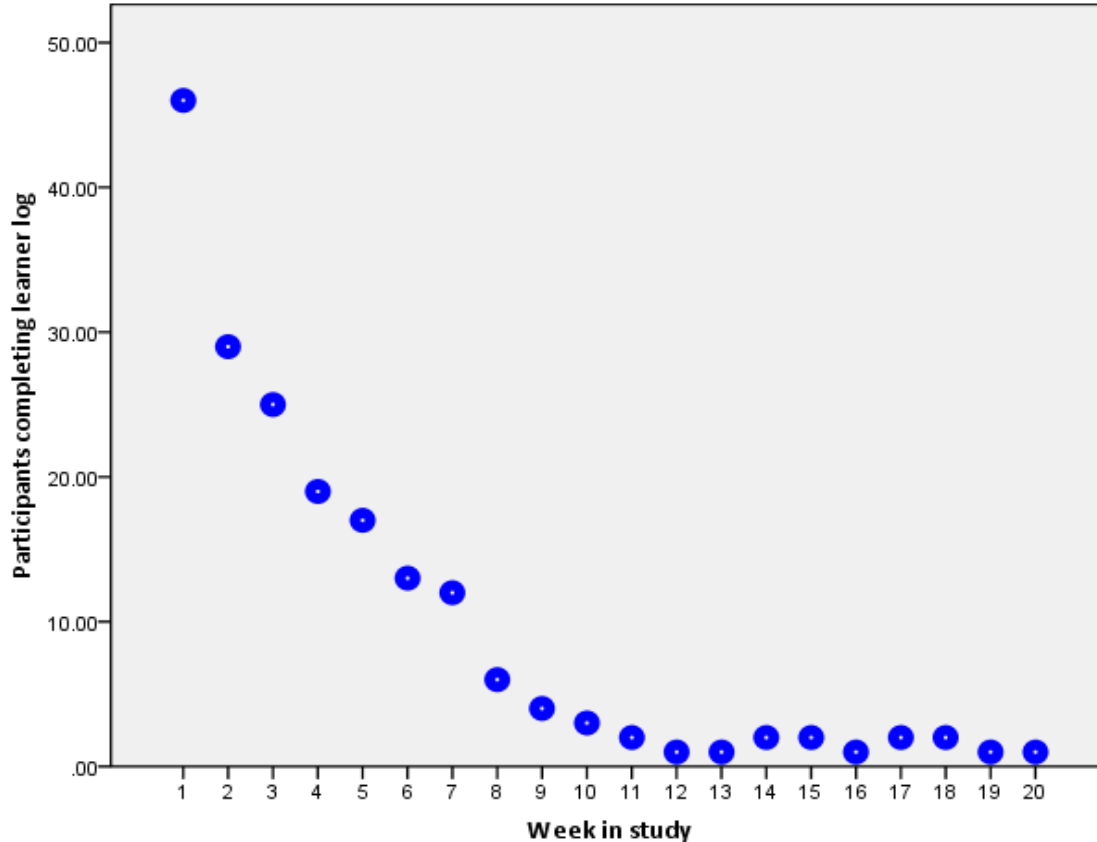
Figure 3. P1 learner log usage in terms of number of participants by week of the study.

User feedback from the P1 learner logs suggests that, at least initially, technological problems were a great source of user frustration. During week one, there were complaints from 20 different participants who had technical difficulties while using the CALL materials. Among other issues, participants had trouble with system crashes, were unable to use microphones in their workplaces, were unable to access the materials using wireless connections, and could not download required plugins to their work computers. There were six questions on the P2 end-of-study survey, and 30 (17%) participants responded; of the 180 total responses, 48 identified technical problems with the software. The chief complaint was that the software did not ever work on the USCG computers, so participants had to use the materials at home. There were also complaints that the software didn't work at home, and that microphones and speakers did not work. Finally, there were comments from participants in both P1 and P2 that when they were required to travel for work, they were unable to access the CALL materials remotely. These technological complaints in both P1 and P2 persisted despite continuous support from the lead researcher as well as the agencies' technology specialists.

In addition to technological complaints about using the software, there was also negative user feedback about the content of the CALL materials. The P1 participants who continued with either Arabic or Chinese until the first interim assessment commented that it was difficult to learn the non-Roman characters without any explicit instruction; eight P1 participants mentioned needing additional resources to understand the RS content, from online grammar guides and dictionaries to native speakers. Nineteen of the 30 P2 participants who completed the exit survey reported seeking outside resources to supplement the ATMM materials, including eight learners who sought out native Spanish speakers for conversational

practice. There were 35 negative comments about either the content or the structure of the ATMM materials. These comments largely concerned the lack of job-specific vocabulary, the lack of a guided plan to the materials, the lack of explicit instruction for 0-beginners, and the lack of conversational vocabulary and practice. There were four comments from P2 participants that the software was useful.

The final source of learner feedback was through personal communication with the lead CASL researcher during P1, usually to explain the reasons for attrition. Of the 119 participants in P1 who dropped out of the study, 72 did not provide a reason, even when asked directly (46 did not begin using the software, and 26 e-mailed the researcher to drop out without stating why). Thirty-five people dropped out because they were sent overseas for work, their work situations changed, or they simply did not have enough time. Nine people dropped out for personal reasons, five dropped out because they could not access the software, and two dropped out because they did not like RS. Participants in P2 did not provide specific reasons for attrition; their responses to the exit survey suggest that technological problems, lack of job-specific content, and lack of time were the chief obstacles to participation.

DISCUSSION

The discussion of the results is organized by the research questions. Q1 asked whether or not self-study with CALL materials was an appropriate solution for USG employees. Given the high attrition rate, and the limited usage of the materials by those participants who did not immediately drop out, self-study with CALL does not seem to be a particularly promising solution for adults in the workplace. Aside from the outlier in each of the two phases of the study, 0-beginners do not appear to be a good fit for this training method. In addition to the very low participation in P1 (which consisted exclusively of 0-beginners) and the extremely high attrition of 0-beginners in P2—only five of the 82 absolute beginners who began the study persisted with the ATMM software for more than 10 hours—there were several telling comments from 0-beginners in both P1 and P2, who wanted more guidance with the materials, more explicit information about how the languages worked, more help navigating through the self-study packages, and more content relevant to their job needs. These requests for more support are not surprising in light of the research findings discussed at the outset of this article, which indicate that independent learners using CALL require support (Littlemore, 2001; Mozzon-McPherson, 2007) perhaps because, as suggested by Jones (2001), "CALL should not be too closely associated with self-access or autonomy, and … teachers are needed to drive the CALL process" (p. 360).

The participation results from the non-0-beginners are slightly better; there were 92 participants in P2 who did not identify as 0-beginners, and there were 12 participants in this group who used the software for more than 10 hours. While this is slightly better than the five 0-beginners who persisted for more than 10 hours, this dropout rate of 93% is not encouraging. The severe technical problems experienced by all P2 participants may have caused much of this attrition. While this study did not provide sufficient data to support this conclusion, it is possible that self-study with CALL could be appropriate for learners with some prior language training who have adequate support, as suggested by Ulitsky (2000) and Murray (1999).

The overwhelming technological problems throughout the studies suggest that autonomous use of CALL materials requires more technical support than is currently available in the agencies participating in this study. In P1, many of the participants using RS worked in locations that did not permit them to download the applications necessary to run the software, and there were also many employees who were not allowed to use microphones at their workstations, which limited the usability of the RS software. As for P2, the language program manager at the USCG agreed to procure and distribute ATMM licenses to 250 employees, and he arranged for them to receive release time from their job duties to use the software at work. However, despite spending six months on the technical set up, the software was never successfully installed on the USCG computer workstations, so despite being told that they could use the software at work, many participants were required to use the materials offsite. Finally, the USCG intended to use

ATMM as a language maintenance program for its linguists and interpreters as well as a language training program for other interested employees. However, in general, the linguists and interpreters were unable to use the materials because they did not have reliable Internet access while at sea.

The second research question concerned the language gains that could be expected from 0-beginners who use CALL materials according to the manufacturer's instructions; however, due to severe attrition, there are very little data on which to base the answer to this question. Fifteen of the 19 P1 participants who persisted through the first assessment were able to ask and answer some questions about the images on the interim assessment. The performance of the 11 Arabic students did correlate with the number of hours they logged using RS. However, six of those 11 participants reported the use of outside resources as well, so it is difficult to know to what extent their proficiency can be attributed to the CALL materials. There were three self-described 0-beginners who took the Versant exit test during P2; however, those participants did not actually appear to be true absolute beginners. Regardless of this, none of them used the software for more than two hours, so no data exist on 0-beginners who used ATMM.

While it is impossible to make any sound claims about learner outcomes from the performance of a single participant, it is worth discussing the case of the single 0-beginner who completed the full P1 study protocol of 200 hours of self-study with RS. This participant received perfect or nearly perfect scores on all the interim assessments, yet his score on a post-study ACTFL OPI was a 0+ or Novice-High.[8] When reached for comment after having finished the study, he stated that "While Rosetta Stone does teach a lot of words, they are not always the words you need to have an actual conversation."[9] It is important to report that this participant was a career linguist who, before beginning this study, could already speak Korean, Spanish, Italian, and Portuguese to very high levels. His profile is not representative of the other USG employees who participated in this study and to whom these self-study materials are generally directed.

The third research question asked whether or not RS worked equally well with Arabic, Chinese, and Spanish. Again, due to the lack of proficiency outcomes data in P1 it is difficult to answer this question definitively. The three students who studied Spanish were able to score an average of 46/100 points on the first interim assessment after an average of 16 hours, rather than the 50 hours suggested as necessary to complete the first quarter of RS V2. The Arabic students, on the other hand, needed to spend the full 50 hours in order to do well on the first interim assessments. The five Chinese participants simply did not do well, regardless of the amount of time they studied, with 4 of them scoring no points on the first interim assessment. The learner logs contained many complaints about the program from those students studying Arabic and Chinese; participants commented that it was difficult to grasp the script without an instructor and that they had trouble figuring out which words went with which pictures. Six of the 13 participants who completed the first Arabic assessment reported seeking outside resources to help them with the program. Rosetta Stone includes no explicit instructions for how to read and write Arabic script or Chinese characters, or for how the Chinese tonal system works. This suggests that, at least initially, RS does not work equally well for all languages, and that the RS model of avoiding metalinguistic information is problematic for 0-beginners who need to learn to write in a TL with an unfamiliar writing system. Further, the CASL review of RS V2 in Arabic, Spanish, and Chinese revealed that the materials do not use any authentic discourse and that they are all translations of one another, with the software presenting the same vocabulary and structures regardless of language. This approach is at odds with recent theories of SLA which call for exposure to contextually appropriate target language discourse, both from an interactionist perspective (Gass, 2003; Gass & Mackey, 2007; Long, 1981, 1996) and from a sociocultural perspective (Lantolf, 2000).

Research question four addresses the effectiveness of ATMM for Spanish learners who were not absolute beginners. There were 19 learners with some starting proficiency who used the program over the six month study period; instead of the 130 hours indicated by the study protocol, 18 of these learners used the materials for an average time of 14:06 (novice), 15:45 (intermediate), and 3:20 (advanced). In addition,

some of these learners sought out additional language learning resources. There was some improvement by a few points on the Versant Spanish proficiency test for some of these learners in the sentence mastery category, the pronunciation category, and the fluency category, but there was no correlation between time spent using the software and gains on the Versant test. There was one participant in the novice group who used the software for 134 hours over the six-month period, adhering closely to the original study protocol. See Figure 4 for the Versant scores of this participant, which are slightly different from those of other novice learners, with improvements in vocabulary, pronunciation, and fluency, but no change in sentence mastery. Unfortunately, this participant did not provide the CASL research team with any additional information, so it was impossible to determine whether or not she undertook outside language learning activities along with her ATMM use; accordingly, we have no way of knowing whether or not we can attribute her gain scores on the Versant assessment to her use of ATMM. Based on the limited available data, some learners with some prior Spanish proficiency did improve their scores after a period of self-study with ATMM.
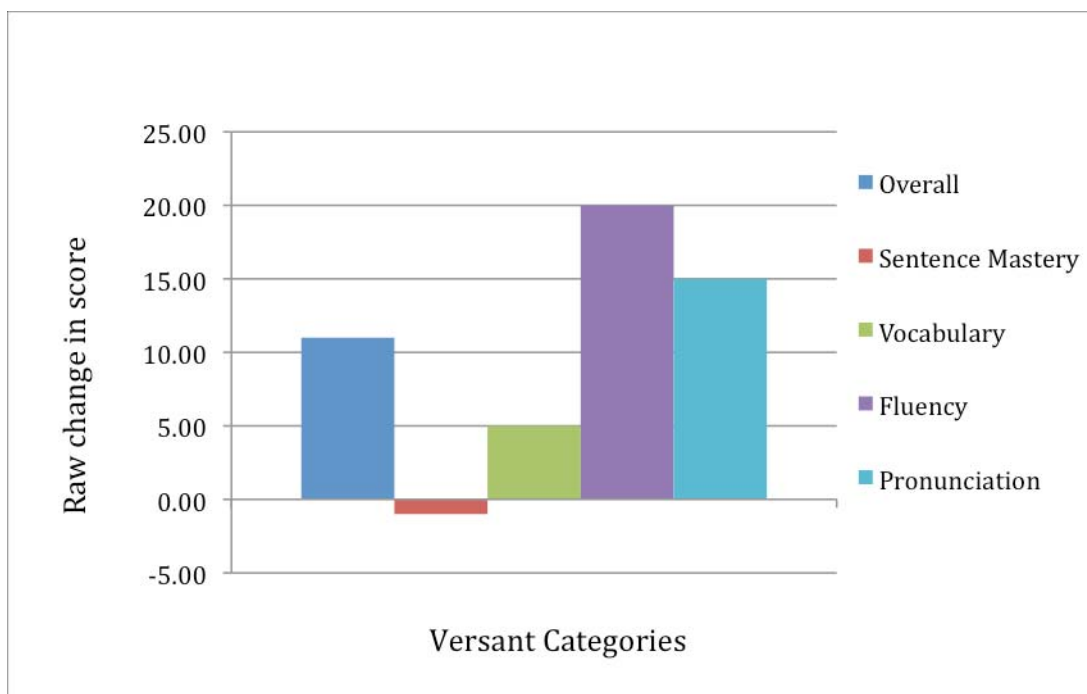

Figure 4. Versant gains by participant with 134 hours of ATMM software use.

The final research question was whether or not supervisors and trainees should be given any guidelines for self-study. Because research from the fields of autonomous foreign language learning and CALL suggests that learner support is necessary for effective self-study, the research design included a point of contact from the CASL research team to provide encouragement and assistance to participants. Despite this additional measure of support, which was not generally supplied by the participating agencies, there was very little use of the materials. Based on the significant pattern of attrition, including the large number of participants who never logged into the CALL materials at all, as well as the technology issues that plagued users in both P1 and P2, it is clear that both supervisors and trainees need guidelines for self-study. Supervisors must be prepared to provide learners with release time from their job duties, ensure CALL materials are suitable for their workplaces (in terms of both content and technological accessibility), and make sure that technical support is available.

In addition, supervisors should understand that merely providing access to self-study materials does not guarantee usage; additional resources to help users compensate for the inadequacies of stand-alone CALL

products as well as face-to-face support should be provided (Jones, 2001; Wegerif, 1998). Participants should anticipate the time commitment involved with self-study and be prepared for the difficulty associated with learning a language alone. The user manuals for autonomous learners (e.g., Fernandez-Toro, 1999) or language learning in general (e.g., Hurd & Murphy, 2005) might be a useful place to start. Further, while the CASL lead researcher did provide continuous support to participants in these studies, she was not a USG employee and had nothing to do with the participants' day-to-day job activities. It is possible that the same levels of encouragement and support from a direct supervisor would have resulted in more consistent learner engagement with the CALL materials. It is likely that a dedicated, online instructor would have been able to provide learners with crucial support; as discussed in the introduction, both instructor-mediated distance learning and supported self-access instruction are approaches to CALL that have been shown to be effective (Blake et al., 2008; Mozzon-McPherson, 2007; Murphy, 2008). Following this model, CALL products could be developed that provide learners with the means to interact with other learners or instructors, something which RS has begun with its new TOTALe product, and which other software developers offer as well (e.g., LiveMocha and WordChamp, among others).

Leaving aside the issue of interaction, participants in P1 and P2 commented that the content of the CALL packages was very generalized, with several P2 participants requesting a training program with nautical content so that the language would be job specific. These comments suggest that perhaps generalized self-study solutions—including ATMM and RS, as well as other products that claim to work for all learners— are not the best approach for working adults with specific language needs. As with all decision makers responsible for selecting foreign language training, supervisors in the USG should make sure to conduct a needs analysis before choosing a CALL program so that the content of the materials matches the job requirements of the learners (See Long (2005) for a discussion of the importance of a needs analysis in order to provide learners with relevant materials and tasks). Because learners were using these materials as part of their work-related activities there would likely have been more persistence if the training had included job-specific content (e.g., the Operational Language and Culture Series by Alelo or the USG-specific products created by Transparent Language).

CONCLUSION

This investigation of the use of self-study CALL materials by motivated USG employees in the workplace had one major finding: severe attrition. Despite beginning with large *n*-sizes, a wide range of enthusiastic participants from different positions within the USG, and researcher encouragement and support, this method of language training yielded very limited proficiency gains in only a handful of learners. These results are very important; this was the first empirical study intended to establish what happens when independent learners use commercial self-study materials in the workplace, and the lack of use suggests that this approach is not likely to be an effective approach for improving overall language proficiency in this context, especially for 0-beginners. Yet the USG continues to invest millions of dollars in self-study language products, and companies like ATMM and RS continue to advertise themselves as complete solutions to language learning, using endorsements from successful language learners and other clients to appeal to institutions and individuals eager to find a simple solution to their language training needs.

The truth is that learning a language is far from simple, especially for adult learners. Research from self-access centers and online learners indicates that independent language learners require support, guidance, and access to a wide-range of materials and resources in order to benefit from self-study. While CALL products offer increasingly sophisticated graphics and interfaces, they are not yet able to offer an alternative to human support or interaction. The participants in this study were adult learners eager for the opportunity to use these materials to study a second language. It is unlikely that any future study with these stand-alone products could find a more motivated group of people, and the results of both P1 and P2 indicate that these resources are unlikely to work by themselves. Managers and learners alike should

consider them as supplements to instructor-mediated training rather than stand-alone solutions. Future research with self-study materials would likely have more robust results if the products were selected according to learners' needs, if the learners were given specific (and measureable) learning goals, and if the participants began the study with some proficiency in the target languages.

NOTES

1. A Rasch analysis was completed on the first interim assessment to determine how well the test items fit the subject pool. The item-person map indicated a reasonable spread of difficulty across items as well as an expected pattern of ability across persons. The number of subjects who actually completed the first assessment was very low ($n = 21$), so absolute conclusions cannot be drawn, but the analysis did support the use of this assessment as a tool to measure learner acquisition of the RS content.

2. One of the findings of P1 was that participants had difficulty fitting 10 hours/week of self-study into their workdays, so the protocol for P2 was modified to require less work and to secure supervisor permission for participant release time from job duties. In addition, P2 also controlled for other obstacles to success identified during P1, such as technological problems and an exclusively 0-beginner subject pool.

3. A Rasch analysis of the test items and learner performances revealed that across all three languages, the images that were the most difficult to describe were those that depicted static images. This is most likely because RS focused on photos of people doing things and described all of the images in terms of actions (e.g., "the boy is jumping"). In addition, for participants who did not get perfect scores, the questions on the assessment that were the most difficult were those that requested specific information about pictures. The participants were better able to complete the monologic description task than to answer questions, most likely because the software did not prepare learners to answer questions.

4. While some 0-beginners actually used the software longer than that (see Table 2), none of those who did took the exit Versant assessment.

5. Participants were sorted into proficiency groups based on their initial Versant score—novice participants had scores from 20 to 39, intermediates had scores from 40 to 59, and advanced participants had scores from 60 to 80. Versant scores correlated with the ATMM placement test scores ($r = .86$).

6. One novice participant had a usage and performance pattern unlike the rest of the participants; her scores are not included in the descriptive statistics in Table 5, and are instead discussed separately in Figure 4.

7. Versant reports that the test/re-test scores ($N = 140$) changed an average of less than one point (Pearson, 2008).

8. See the following Web site for a description of learner oral proficiency at the 0+ level: http://www.languagetesting.com/actfl_guidelines8.html

9. It is important to note that this exploratory study of RS was conducted using Version 2 of the software, and while the study was underway, Version 3 (V3) was released. The changes made to the V3 product are not likely to have affected the outcome of the present study. For example, V3 continues to use the same images and content for every language taught by the program. In addition, there is no explicit information about grammar or a language's writing system. There is more vocabulary in V3 geared toward basic communicative tasks and greetings, but there is no chance to observe communication in a genuine context. Finally, V3 has newer photos and there is more variety in the drills, but the overall approach is the same. Learners use the materials to match images with sounds or words.

ACKNOWLEDGEMENTS

ABOUT THE AUTHOR

Katharine B. Nielson is a Principal Research Assistant at the University of Maryland Center for Advanced Study of Language (CASL) and a Ph.D. student in the Second Language Acquisition program at the University of Maryland. Her research interests include technology-mediated language training, SLA and cognition, and instructed SLA.

E-mail: knielson@casl.umd.edu

REFERENCES

Benson, P. (2007). Autonomy in language teaching and learning. *Language Teaching, 40*, 21–40.

Blake, R. (2008).The brave new digital classroom. Washington, DC: Georgetown University Press.

Blake, R. (2009). The use of technology for second language distance learning. *The Modern Language Journal, 93*(s1), 822–835.

Blake, R., Wilson, N., Cetto, M., & Pardo-Ballester, C. (2008). Measuring oral proficiency in distance, face-to-face, and blended classrooms. *Language Learning & Technology, 12*(3), 114–127. Retrieved from http://llt.msu.edu/vol12num3/blakeetal.pdf

Bown, J. (2006). Locus of learning and affective strategy use: Two factors affecting success in self-instructed language learning. *Foreign Language Annals, 39*(4), 640–659.

Burwell, G., González-Lloret, M., & Nielson, K. (2009). Assessment in a TBLT Spanish immersion course. Paper presented at 3rd Biennial International Conference on Task Based Language Teaching, Lancaster, UK.

Chapelle, C. (2010). The spread of computer assisted language learning. *Language Teaching, 43*(1), 66–74.

Compton, L. K. (2009). Preparing language teachers to teach language online: A look at skills, roles, and responsibilities, *Computer Assisted Language Learning*, *22(*1), 73–99.

Fernández-Toro, M. (1999). *Training learners for self-instruction*. London: Center for Information on Language Teaching and Research.

Fleming, S., Hiple, D., & Du, Y. (2002). Foreign language distance education at the University of Hawai'i. In C. A. Spreen (Ed.), *New technologies and language learning: Issues and options (Technical Report #25)*, (pp. 13–54). Honolulu, HI: University of Hawai'i, Second Language Teaching and Curriculum Center.

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

Gass, S. (2003). Input and interaction. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 224–255). New York: Basil Blackwell.

Gass, S. & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten, & J. Williams (Eds.), *Theories in second language acquisition* (pp. 175–200). Mahwah, NJ: Lawrence Erlbaum Associates.

Godwin-Jones, R. (2007). Emerging technologies tools and trends in self-paced language instruction. *Language Learning & Technology, 11*(2), 10–17. Retrieved from http://llt.msu.edu/vol11num2/emerging/default.html

González-Lloret, M. (2003). Designing task-based CALL to promote interaction: En busca de esmereldas. *Language Learning and Technology, 7*(1), 86–104. Retrieved from http://llt.msu.edu/vol7num1/gonzalez/default.html

Hampel, R., & Stickler, U. (2005). New skills for new classrooms: Training tutors to teach languages online. *Computer Assisted Language Learning, 18*(4), 311–326.

Holec, H. (1981). *Autonomy in foreign language learning* (first published 1979, Strasbourg: Council of Europe). Oxford: Pergamon.

Hurd, S., & Murphy, L. (2005). *Success with languages*. Abingdon, UK: Routledge.

Jones, J. (2001). CALL and the responsibilities of teachers and administrators. *ELT Journal, 55*(4), 360–367.

Lai, C., Zhao, Y., & Li, N. (2008). Designing a distance foreign language learning environment. In Goertler, S., & P. Winke (Eds.), *Opening doors through distance language education: Principles, perspectives, and practices* (pp. 85–108). Texas: CALICO.

Lafford, B., Lafford, P., & Sykes, J. (2007). Entre dicho y hecho …: An assessment of the application of research from second language acquisition and related fields to the creation of Spanish CALL materials for lexical acquisition. *CALICO Journal, 24*(3), 497–529.

Lantolf, J. P. (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

Lie, K. (2009). Virtual communication: An investigation of foreign language interaction in a Distance Education course in Norwegian. *Dissertation Abstracts International, 69*(9-A), 3484.

Littlemore, J. (2001). Learner autonomy, self-instruction and new technologies in language learning: Current theory and practice in higher education in Europe. In A. Chambers & G. Davies (Eds.), *ICT and language learning: A European perspective* (pp. 39–60). Lisse, The Netherlands: Swets & Zeitlinger Publishers.

Liu, X., Magjuka, R., Bonk, C., & Lee, S. (2007). Does sense of community matter? An examination of participants' perspectives of building learning communities in online courses. *Quarterly Review of Distance Education, 8*(1), 9–24.

Long, M. H. (1981). Input, interaction, and second language acquisition. *Foreign Language Acquisition: Annals of the New York Academy of Sciences, 379*, 259–78.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego: Academic Press.

Long, M. H. (2005). *Second language needs analysis*. Cambridge, UK: Cambridge University Press.

Mozzon-McPherson, M. (2007). Supporting independent learning environments: An analysis of structures and roles of language learning advisers. *System, 35*(1), 66–92.

Murphy, L. (2008). Supporting learner autonomy: Developing practice through the production of courses for distance learners of French, German and Spanish. *Language Teaching Research, 12*(1), 83–02.

Murray, G. (1999). Autonomy and language learning in a simulated environment. *System*, *27*(3), 295–308.

Nielson, K., & González-Lloret, M. (2010). Effective online foreign language courses: Theoretical framework and practical application. *EUROCALL Review, 17*. Retrieved from http://www.eurocall-languages.org/review/17/index.html#nielson_gonzalez

Pearson. (2008). Versant Spanish Test Description and Validation. Retrieved from http://www.ordinate.com/technology/VersantSpanishTestValidation.pdf

Rosetta Stone. (2011). How it works. Retrieved from http://secure.rosettastone.com/global/endangered/howitworks

Rovai, A. P. (2002). Building a sense of community at a distance. *International Review of Research in Open and Distance Learning*, *3*(1), 1–16.

Sadera, J., Robertson, J., Song, L., & Midon, N. (2009). The role of community in online learning success. *MERLOT Journal of Online Learning and Teaching*, *5*(2), 277–284.

Saury, R.E. (1998). Creating a psychological foundation for the evaluation of pre-packaged software in second language learning. Proceedings of ED-MEDIA/ED-TELECOM 98 World Conference on Educational Telecommunications, Freiburg, Germany.

Surry, D., & Ensminger, D. (2001). What's wrong with media comparison studies? *Educational Technology, 41*(4), 32–35.

Sykes, J., Oskoz, A., & Thorne, S. (2008). Web 2.0, synthetic immersive environments, and mobile resources for language education. *CALICO Journal, 25*(3), 528–546.

Tell Me More. Auralog difference. Retrieved from http://www.tellmemore.com/about/aboutus/auralog_difference.

Thorne, S., & Payne, J. S. (2005). Evolutionary trajectories, Internet-mediated expression, and language education. *CALICO Journal, 22*(3), 371–397.

Ulitsky, H. (2000). Language learner strategies with technology. *Journal of Educational Computing Research, 22*(3), 285–322.

Wegerif, R. (1998). The social dimension of asynchronous learning networks. *Journal of Asynchronous Learning Networks*, *2*(1), 34–49.

White, C. (1999). Expectations and emergent beliefs of self-instructed language learners. *System, 27*(4), 443–457.

White, C. (2005). Towards a learner-based theory of distance language learning: The concept of the learner-context interface. In B. Holmberg, M. Shelley, & C. White (Eds.), *Distance education and languages: Evolution and change* (pp. 55–71). Clevedon, UK: Multilingual Matters.

White, C. (2006). State of the art: Distance learning of foreign languages. *Language Teaching, 39*(4), 247–264.