

MEASURING ORAL PROFICIENCY IN DISTANCE, FACE-TO-FACE, AND BLENDED CLASSROOMS

Robert Blake, [University of California – Davis](#)

Nicole L. Wilson, Pearson & [University of California – Santa Cruz](#)

María Cetto, [University of California – Davis](#)

Cristina Pardo-Ballester, [Iowa State University](#)

Although the foreign-language profession routinely stresses the importance of technology for the curriculum, many teachers still harbor deep-seated doubts as to whether or not a hybrid course, much less a completely distance-learning class, could provide L2 learners with a way to reach linguistic proficiency, especially with respect to oral language skills. In this study, we examine the case of *Spanish Without Walls (SWW)*, a first-year language course offered at the University of California - Davis in both hybrid and distance-learning formats. The SWW curriculum includes materials delivered via CD-ROM/DVD programs, online content-based web pages, and synchronous bimodal chat that includes sound and text. The contribution of each of these components is evaluated in the context of a successful technologically assisted course. To address the issue of oral proficiency, we compare the results from both classroom and distance-learning students who took the 20-minute *Versant for Spanish* test, delivered by phone and automatically graded. The data generated by this instrument shows that classroom, hybrid, and distance L2 learners reach comparable levels of oral proficiency during their first year of study. Reference is also made to two other ongoing efforts to provide distance-learning courses in Arabic and Punjabi, two languages where special difficulties in their writing systems have an impact on the design of the distant-learning format. The rationale for offering language courses in either a hybrid or distance-learning format is examined in light of increasing societal pressures to help L2 learners reach advanced proficiency, especially in less commonly taught languages (LCTLs).

INTRODUCTION

"Do you use technology in your foreign-language classroom?" Few language teachers would dare to answer "no" to this question for fear of being classified as out of touch with best practices. Not surprisingly, most teachers routinely employ videos or CDs/DVDs, and many now take advantage of web pages to distribute syllabi, assignments, cultural material, and even lecture notes. But when foreign-language faculty are polled (at least in the case of faculty from U.S. research universities) about accepting credit for language courses delivered in a hybrid or completely distance-learning (DL) format, the positive sentiments toward technology begins to fade.¹ Some foreign language (FL) teachers tend to harbor deep-seated doubts as to whether or not a DL course could ever provide L2 learners with a way to gain linguistic proficiency, especially when oral language skills are in question. Perhaps others secretly worry that these new DL classes might displace them.

Reasons for Implementing DL or Hybrid Course Formats

In all fairness, a DL format is not the appropriate learning environment for everyone; the format self-selects for those who have both the ability and preference to work more independently. In particular, a completely virtual course appeals to people who work full-time and, therefore, need flexible access to instruction. The popularity of a hybrid course (which offers reduced face-to-face meetings along with other DL teaching techniques) stems from these factors, but this format also attracts students who want fewer days of class. Many students belatedly find out that both hybrid and DL language classes require a

high degree of self-motivation and independent work skills. This accounts for the routinely high dropout rate for the DL learning environment for all disciplines (Carr, 2000; Dreyer, Bangeni & Nel, 2005).

Technologically enhanced formats are increasingly being explored as a means to meet student demands for less commonly taught languages (LCTLs) such as Arabic, Punjabi, or Farsi. Frequently, there are no course offerings for these languages available locally, and the DL format may be the only way to fulfill both interest and demand. Nevertheless, no one would dispute that to reach ACTFL's (American Council on the Teaching of Foreign Languages) level of superior proficiency (which corresponds to the U.S. government's ILR (Interagency Language Roundtable) rating of 3+ or level C in the Common European Framework), L2 students need to interact face-to-face with native speakers, preferably in the target country. Davidson's (2007) longitudinal study of Russian L2 learners shows quite conclusively that study abroad—and, preferably, a yearlong program as opposed to a semester program—is a *sine qua non* for reaching these more advanced levels.² Likewise, the basic facts concerning L2 development have not changed: the road to advanced proficiency is arduous, requiring anywhere from 600 to 2,200 hours of instruction (Bialystok & Hakuta 1994, p. 34).

In the present study we seek more data to evaluate whether DL courses can be a valid way to start down this long road without disadvantaging DL students with respect to oral proficiency. Accordingly, this investigation addresses beginning and intermediate Spanish language instruction and poses the following question: Can the hybrid or distance-learning students keep pace in terms of oral proficiency with the outcomes normally found with traditional L2 classroom learners that meet five days a week? Although researchers have frequently observed that DL courses help develop strong literacy skills because so much of this learning environment is text-based (Warschauer, 1997), we wanted to investigate the effect of DL courses on the development of oral proficiency.

Previous DL studies

In her thorough review of four decades of research in the DL language field, White (2006) points to the progress and promise that the DL environment holds for the FL profession. She also cautions the field that small-scale studies need to be converted into fully implemented DL courses:

The innovations reported in the literature owe much to the work of early adopters—who pursue those innovations, usually in small-scale pilot studies, find ways through the barriers that emerge, and report their findings. An important area for inquiry and critical reflection concerns the process of migrating these innovations to contexts for mainstream course delivery. (p. 259)

Blake and Delforge (2007) have reported on several fully implemented DL courses (Cahill & Catanzaro, 1997; Chenoweth & Murday, 2003; Soo & Ngeow, 1998), but none of these studies assessed the effects that the DL format itself might have had on fostering oral proficiency, a major concern of the FL field with its current emphasis on communicative competence. A notable exception can be found in the DL courses in French, German, and Spanish offered through the Open University where researchers are investigating in earnest the issues concerning oral proficiency assessment and the effect of the DL format (Hampel, 2003; Hampel & Hauck, 2004; Lamy, 2004). Lamy (2004), in particular, has tried to define what constitutes conversational (if not oral) proficiency in an online learning context.³ But only Volle (2005) in her implementation of a DL course in central Texas has actually attempted to measure online articulation (with drills, free conversation, and reading), grammatical accuracy (adapted from Weir's, 1990, protocol), and conversational competence or *proficiency* in response to a series of modified OPI (Oral Proficiency Interview) prompts (e.g., "greet someone, answer questions about school or routines, talk about the past, describe specific items about family or clothes"). Her study was designed to gauge student improvement in these three areas after having completed a one-semester online course. However, lack of control over several variables makes the results uninformative.

The relative dearth of mainstream DL language course offerings means relatively fewer students to sample from and, consequently, more difficulties in comparing DL student outcomes with those of students found in more traditional classroom formats. Not only are there fewer fully implemented DL classes—and, therefore, students—to sample from, but, in our experience, DL students tended to respond only to emails that directly relate to their own progress or grade. By way of contrast, classroom students can be cajoled in person to complete research questionnaires. Finally, the burden of isolating the experimental treatment so as to enable comparisons between DL and classroom student outcomes presents formidable challenges, as will be discussed below in more detail.

Not all of these issues can be resolved at once, but the need to provide more data on oral proficiency in the DL context looms first and foremost. The present study examines oral proficiency data from students in three types of first-year Spanish courses:

- a hybrid course (two days instruction per week plus seven hours of study via technology)
- a completely virtual DL course
- a control group of traditional classroom students.

Oral proficiency was measured in this study by using *Versant for Spanish* (Pearson), an instrument administered by phone and scored automatically by means of speech recognition and parser software (see description below). Our intent is not to test whether the DL or hybrid format produces better results than the classroom format, but rather to ascertain if students in those technologically supported learning environments can keep pace with the oral progress demonstrated by students in face-to-face learning environments.

METHODOLOGY

During 2005-2007, the *Versant for Spanish* instrument was administered to 233 students enrolled in traditional Spanish classes and 85 students enrolled in non-traditional (hybrid, distance) course formats at the University of California - Davis. All students were enrolled in the first seven levels (ten-week quarters numbered here as 1 through 7) of Spanish language instruction, which comprises the first- and second-year college Spanish curriculum. All students took the Spanish test over the phone sometime during the final two weeks of the course in order to give a more accurate representation of their proficiency at that particular level. The performances of the two groups, first-year Spanish (SPA 1-3) and second-year Spanish (SPA 4-7), were then compared. It was predicted that course format (traditional, hybrid, or distance) would not affect outcomes at the lower-division level with respect to oral proficiency.

In addition, 15 heritage learners (HL) who enrolled in a lower-division traditional course for native speakers also took the test. HL students constitute an extremely heterogeneous group: while most HL students have distinct oral advantages over non-heritage learners, they also tend to lack competence in academic Spanish or the more formal registers (Blake & Zyzik, 2003). Nevertheless, it was predicted that the HL students would outperform the non-HL students on the *Versant for Spanish*, a prediction that was born out by our findings (see results below). Heritage learners were included in the study in order to test whether the *Versant* test was sensitive enough to differentiate between advanced learners and heritage or home language learners. Heritage learners are not the same as native speakers and represent a continuum of linguistic proficiency. Could *Versant* separate them successfully from advanced L2 learners?

***Spanish Without Walls* and its Hybrid Cousin**

Spanish Without Walls is a first-year DL Spanish curriculum divided into three quarters. It combines multimedia language materials from three sources:

- *Tesoros*, a five-disk CD-ROM or one-disk DVD detective story (Blake, Blasco & Hernández, 2001);

- Content-based web readings and Flash activities⁴;
- A collaborative CMC tool running on the Flash communications server that allows for both asynchronous and synchronous textual communication (both controlled character-by-character and by carriage return) in addition to Voice-over Internet protocol (VoIP) sound exchange.

The CDs/DVD serve as the course textbook.⁵ The remaining online materials are packaged into a course management system (Moodle) designed to teach first-year Spanish grammar and vocabulary, provide exercises, conduct testing, present authentic Spanish-language readings, and enable oral communication with teachers and peers.

Students alternate between use of the CD-ROMs/DVD and the *SWW* website in order to cover the scope and sequence of a normal university first-year Spanish language course. They are held accountable for the CD-ROM/DVD material by means of online exams that cover the vocabulary, storyline, and grammar presented in *Tesoros* (Blake & Delforge, 2007). Students are also required to chat live using both text and voice with their instructor in groups of no more than three at least once a week for one hour and several more times with their assigned partners, as their mutual schedules permit, in order to complete the collaborative content-based tasks. For example, one student might research the capital cities of four Latin American countries, while his/her partner would investigate the same type of information for four other countries. During the chat, the students share their results with each other in jigsaw fashion.

The hybrid courses (SP2V and SP3V) are the equivalent of the last two thirds of an introductory college-level Spanish curriculum (SP2-3) delivered in a hybrid course format. These courses utilize the same *Spanish Without Walls* curriculum, but students meet with their instructor only two days a week for three hours and then perform web activities, work with the CDs/DVD, and carry out CMC synchronous dialogues using Adobe's *Breeze* (which provides text chat, Voice over IP, and whiteboard/character-by-character text exchange). The course materials and exercises focus on developing the same array of skills promoted in the classroom courses, quarters 2 and 3, and provide students with exposure to a variety of Spanish accents from Spain and Latin America, as well as extensive reading of authentic texts, guided writing practice, the second part of a first-year grammar and vocabulary sequence, and cultural awareness of the Spanish-speaking world. The DL and the hybrid formats of *Spanish Without Walls* differ only in a few respects. The hybrid version has two regular class meetings per week, while the virtual course has no face-to-face encounters. The materials and the approximate number of hours the students are asked to invest studying are roughly equivalent, after making adjustments for the different formats.

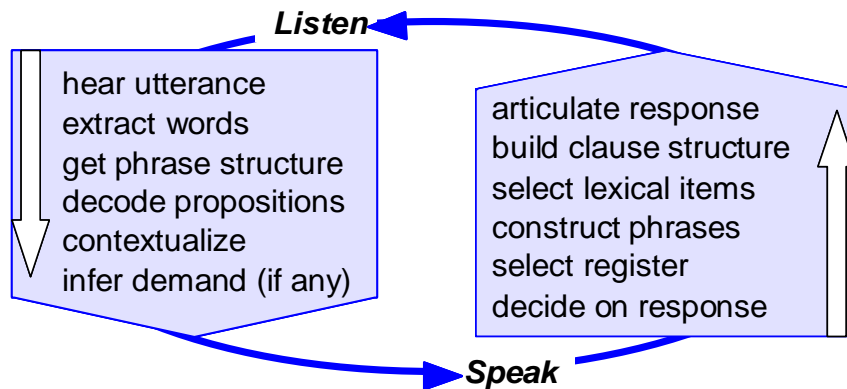
The hybrid format is designed for students with some limited high-school experience in Spanish, while the virtual course is targeted for pure beginners. In reality, both true beginners and *false* beginners (i.e., students who have studied Spanish in high school) routinely enroll in both traditional and DL/hybrid. It is rare in California to find students without some previous formal exposure to Spanish in the schools. There is no reason to assume that these false beginners are not distributed normally in all delivery formats. The false beginners in the DL classes, however, given their slightly older age range, would have left off taking high school Spanish many years before the classroom student counterparts with all the implications concerning language attrition. Eliminating these students would have drastically reduced the number of students in the sample.

In contrast to the hybrid and DL students, the L2 learners in the traditional classroom environment attend class five days a week and use *Dos Mundos* (Terrell, Andrade, Egasse, & Muñoz, 2002) instead of *Tesoros* (Blake, Blasco, & Hernandez, 2001) as a textbook. The scope and sequence for grammar and vocabulary topics found in both these language textbooks are approximately the same. It is expected that traditional students will invest one additional hour of home study for every hour of class, bringing their weekly language study average to ten. Hybrid students spend three hours in class per week and should also study at home or chat for an additional seven hours. DL students have no class meetings, but are asked to work for ten hours with the combined curriculum of CD/DVD activities, web pages, and chat

assignments. There were no controls, however, for the actual amount of out-of-class time spent by each student studying Spanish in any of these formats, including the traditional one.

Measuring Proficiency: *Versant for Spanish*

The *Versant for Spanish* test published by Pearson was designed utilizing principles from Levelt's (1989) model of speech production and comprehension. In this model, there are many basic elements required to participate in a spoken conversation: a person must track what is being said, extract meaning as speech continues, and then formulate and produce a relevant and intelligible response. These component processes of listening and speaking are schematized in Figure 1, adapted from Levelt.



Adapted from Levelt, 1989

Figure 1. Conversational processing components in listening and speaking.

Versant for Spanish requires participants to read aloud, listen and repeat, say the opposite, answer short questions, build sentences from jumbled-up word combinations, answer open-ended questions, and retell stories. Because the test is designed to elicit responses from the test-taker in real time, it estimates the test-taker's level of automaticity with the language. Automaticity is the ability to access and retrieve lexical items, to build phrases and clause structures, and to articulate responses without conscious attention to the linguistic code (Cutler, 2003; Levelt, 1989).

The responses from test-takers are scored automatically by means of a speech recognition and parser program based on a probabilistic grammar built from a large corpus of native speakers ($n = 435$) from a variety of Spanish-speaking countries as well as another corpus of L2 learners of Spanish ($n = 579$) (Bernstein, Barbier, Rosenfeld, & de Jong, 2004). The computer algorithm has been "trained" by expert human raters. Separate values for vocabulary, sentence mastery, pronunciation, and fluency are weighted and then combined in order to come up with an overall score on a scale from 20 to 80.

Results from *Versant for Spanish* have high correlations with the ACTFL, ILR, and CEF proficiency tests ($r = .86$, $r = .90$, and $r = .92$, respectively; Wilson, Bernstein, & Rosenfeld, 2006). Because each item in *Versant for Spanish* contains multiple sources of information (e.g., pauses between and within words, response latency) that are fed into the scoring algorithm, a 20-minute phone sample provides sufficient data to identify the proficiency level of the speaker.

The purpose of the present study, however, is not to argue for the merits or statistical validity of *Versant for Spanish* versus other human-scored oral proficiency tests (see Bernstein et al., 2004). *Versant for Spanish* is particularly advantageous for assessment research with DL students, who are only accessible by email contact, because they can take this test from anywhere in the world at any time using only a

telephone. The fact that *Versant for Spanish* takes only 20 minutes to complete further heightens its attractiveness as a research instrument in the DL context. Consequently, in using this instrument, we were able to generate a large amount of data with minimal expense and effort in order to establish a reliable basis of comparison for any data generated by DL students.

RESULTS AND DISCUSSION

Figure 2 below displays *Versant for Spanish* scores by level for 248 traditional classroom learners as a function of their respective cumulative densities: in other words, each *Versant for Spanish* test score is matched to the proportional number of students from that group who have reached that score or below. Given the range of language proficiency among these eight levels, any particular percentage value should theoretically correspond to radically different *Versant for Spanish* scores, which is exactly what a glance at Figure 2 will confirm.

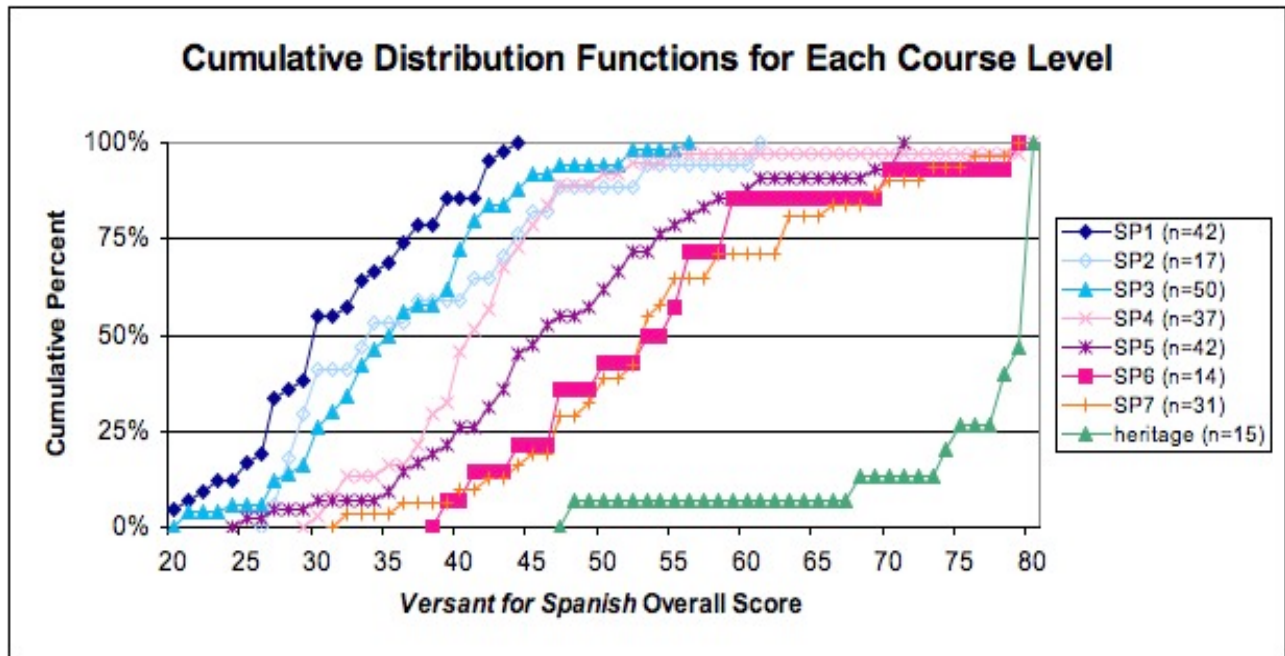


Figure 2. Cumulative density function for classroom students

The results from plotting the cumulative density function suggest that certain language levels as defined by the number of quarters of language study should be clustered together. This visual impression was statistically tested by means of a Tukey Honestly Significant Difference (HSD) Test, the most appropriate statistical measure to employ when no *a priori* hypotheses exist with respect to groupings. The HSD results are displayed in Table 1. Three statistically distinct groups emerge: (1) a first-year+ group that includes SP1, SP2, SP3 and, somewhat more weakly, SP4; (2) a second-year+ group consisting of SP5, SP6, SP7 (although note that quarter 7 is, technically speaking, the beginning of the third year); and (3) the heritage speakers group.

The HSD values for the SP4 group deserve further comment. While from the point of view of the SP1 students, the SP4 students were statistically different, there were no significant differences among the scores of the SP2, SP3, and SP4 students. Clearly, the SP4 group represents a transition class after which these L2 Spanish students experience a demonstrable jump in linguistic capabilities to the next plateau (SP5, SP6, SP7). The score ranges, eliminating the 10th and 90th percentiles, can be summarized as follows:

Course Number	<i>Versant for Spanish</i> Score
Courses 1, 2, & 3	25 - 40
Course 4	31 – 49
Courses 5, 6, & 7	37 - 68

These facts are captured visually in [Figure 3](#), where the cumulative density functions have been plotted after pooling the data into three levels in order to show the differences in oral proficiency that occur after the fourth quarter of language study. In terms of oral proficiency, the heritage students, not unexpectedly, outperformed all of the non-heritage groups. [Figure 3](#), then, represents the normed values for placement scores for oral proficiency according to the *Versant for Spanish* instrument. We then proceeded to test the hybrid and DL students and use their scores as a way of determining if they have met the normal expectations for oral proficiency as a function of time on task (i.e., quarters of language study).

Table 1. Tukey Honestly Significant Difference (HSD) Test for *Versant for Spanish* scores by Quarter (*p*-values)

Quarters	1	2	3	4	5	6	7	HL
1	--	.97	1.0	.43	.00	.00	.00	.00
2		--	1.0	.97	.00	.00	.00	.00
3			--	.71	.00	.00	.00	.00
4				--	.09	.00	.00	.00
5					--	.77	.36	.00
6						--	1.0	.00
7							--	.00

[Figure 4](#) plots the cumulative density function for all DL learners in both hybrid (SP2V and SP3V, $n = 65$) and virtual classes (*SWW*, $n = 21$) along with the scores from the traditional first- and second-year+ classroom learners ($n = 233$) and the heritage learners ($n = 15$) for comparative purposes (total $n = 334$). [Figure 4](#) shows that the DL student performance clearly follows the trends set by the first-year+ group of classroom learners. The strongest hybrid learners (top 50% of the class) are performing slightly poorer than the strongest students from the first-year+ group, but the latter group includes students from quarter 4 as well. The higher performing students from the purely DL group (*SWW*) scored slightly higher than the first-year+ control group (see [Table 1](#)), but, again, in statistical terms, there are no aggregate group differences among students from all three formats.

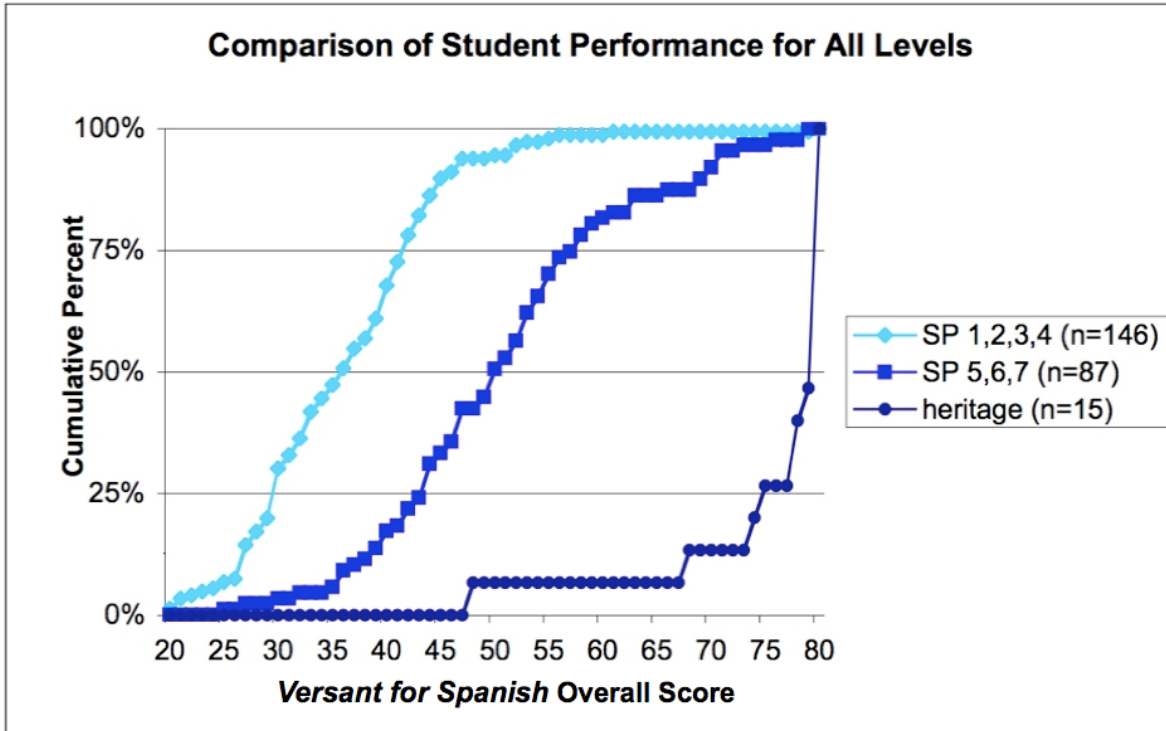


Figure 3. Cumulative density function for classroom student by clusters

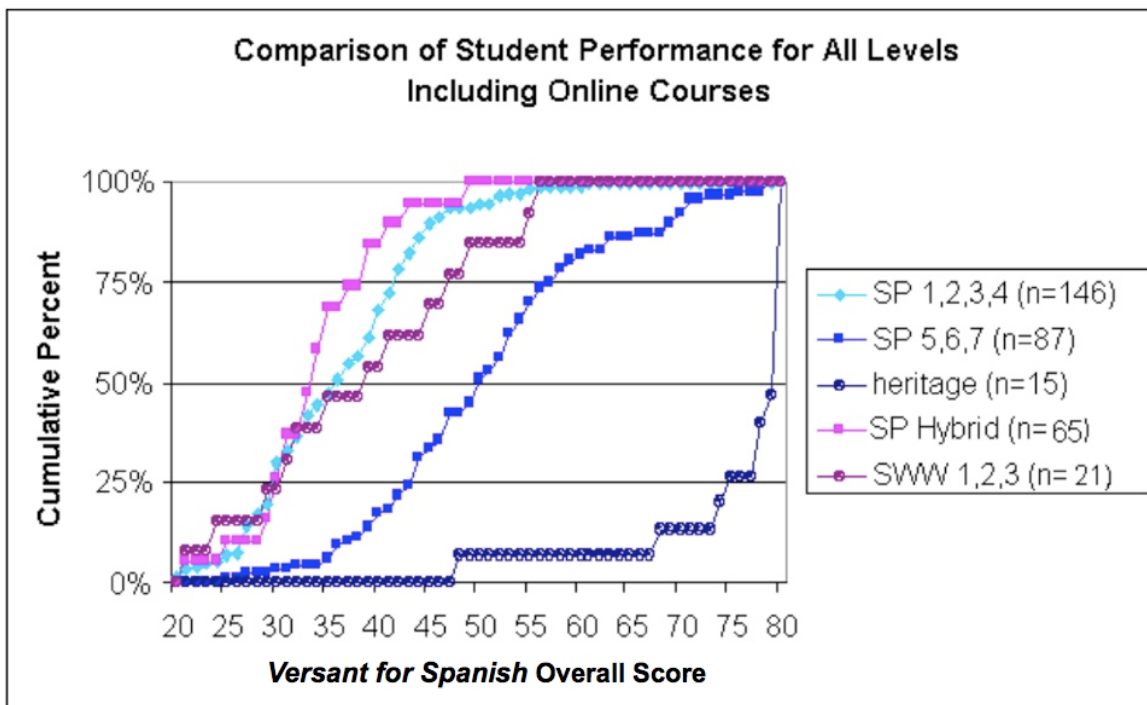


Figure 4. Cumulative density function for classroom and DL students

More importantly, the individual t-tests given below in [Table 2](#) reveal no significant differences among hybrid, DL or their classroom counterparts (i.e., DL 1 compared to classroom learners from level 1, DL 2 compared to classroom 2, etc.).

Table 2. t-Test Results Comparing the DL and Classroom Formats

LEVEL (number of quarters)	SP 1 (n= 42) (mean=34.4, sd=5.3)	SP 2 (n = 17) (mean=37.6, sd=10.0)	SP 3 (n = 50) (mean=36.6, sd=7.6)
SWW 1 (n = 13) (mean=33.4, sd=8.7)	$t(53) = 1.03$	$p = 0.35$	-- --
SWW 2 (n = 7) (mean=41.1, sd=9)	--	--	$t(22) = 0.96$ $p = 0.35$ -- --
Hybrid 2V (n = 37) (mean =34.9, sd=6.9)	--	--	$t(52) = 1.63$ $p = 0.11$ -- --
Hybrid 3V (n = 28) (38.8, sd=6.9)	--	--	-- -- $t(76) = 1.00$ $p = 0.32$

The number of DL students ($n = 21$) who took *Versant for Spanish* exam is, again, limited for the reasons already discussed above; more data is being collected for a future study. This obvious shortcoming, however, is mitigated by the strength of the *Versant* baseline results. In essence, the large number of baseline data points from traditional students establishes a placement norm against which each DL and hybrid score can be compared and situated in its own outcome space. Looked at from this relative perspective, both types of DL formats and curriculum have allowed students to develop a level of oral proficiency that is comparable with the outcomes for oral proficiency exhibited by classroom learners, allowing for individual differences. In other words, DL students are not falling behind on oral proficiency, as judged from the *Versant for Spanish* results.

Giving all 334 students a pre-test to establish their initial Spanish capacity would have been ideal, but difficult to manage during the add-drop period of each quarter's beginning. Again, we were not so much interested in this study in charting the gains (i.e., a comparison of pre- vs. post-test results) that the DL students might have made with respect to measures of oral proficiency after one quarter of instruction, but finding out whether or not they were falling behind their classroom counterparts. In other words, do the DL students enjoy the same placement results? Accordingly, we have used the large number of data points from the classroom learners as a way of establishing placement norms. Future studies, however, should include a pre-test to verify that all students in these different learning formats are starting at approximately the same linguistic level. Likewise, a pretest for working memory, such as that used by Payne and Whitney (2002) and Payne and Ross (2005), would be extremely useful to see if the strong text-based support and less pressured chat environment routinely cited as advantages of the DL learning environment have a positive effect on the weaker students. These working memory scores might also be correlated with specific *Versant* items that focus either on vocabulary or on syntax in order to yield a new set of interesting comparisons.

Actually pinpointing which factors are responsible for putting the DL students on an equal footing with their classroom counterparts with respect to oral proficiency is not an easy task. On the one hand, the SWW students were older, more mature, and more self-motivated than their younger classroom counterparts. False beginners in the DL course, however, have had more time to forget whatever knowledge they had gleaned from high school courses. More importantly, students who enjoy working in a virtual learning environment tend to self-select, as has been noted above. These learners appear to be, from our teaching experience, more efficient and responsible L2 learners since they pay tuition in advance for the DL course and have real career objectives tied to their learning of Spanish. One might expect that

hybrid students should follow suit, but many students enrolled in the hybrid course appear to be looking for reduced class time and less work, according to our instructors' direct observations and informal conversations with students. On the other hand, the scope and sequence of the course materials is virtually the same in all formats. The DL and hybrid environments principally differ from the classroom courses in terms of the nature of the chat activities, the amount of individual attention received from instructors during chat sessions, and the amount of out-of-class study time. Again, administering a pre-test to all students would have strengthened our ability to compare more effectively the effect of these different learning environments.

No doubt, the required synchronous chat sessions for both DL and hybrid students, which offered both text and audio exchanges, make a major contribution to the level of individual practice and the extent of instructor attention, which might even exceed what can be found in traditional classrooms given their burden of 25 to 30 students in a 50-minute period. While the positive effects of audio practice on oral proficiency should be transparent, Payne & Whitney (2002) and Payne and Ross (2005) have also shown that keyboarding exerts a positive affect on oral proficiency as well. The present study is limited in not being able to tease out the differential effect of oral vs. text chatting.

Whatever the case, the *Breeze* bimodal chat sessions for the DL and hybrid students appear to help them perform on the *Versant for Spanish* test at similar levels as students studying in the face-to-face classrooms. The FL profession is just beginning to study how to most effectively utilize these bimodal CMC tools (Blake 2005; Erben, 1999; Fleming, Hiple & Du, 2002; Hampel, 2003; Hampel & Hauck, 2004; Holmberg, Shelley, & White, 2005; Jepson, 2005; Lamy, 2004; Thompson & Hiple, 2005; Thorne & Payne, 2005; White, 2006; Zähler, Fauverge & Wong, 2000).

Finally, available oral-proficiency testing instruments (e.g., ACTFL OPI, ILR OPI, CEF, *Versant*) are not sensitive to the relatively small gains in speaking ability made by beginning learners. These gains consist mostly of learned material that learners can produce in predictable situations but are not yet able to adapt to "real life" use. Despite efforts to measure these gains by introducing "plus" sublevels (ILR), and "low," "mid," and "high" sublevels (ACTFL), tests of communicative proficiency are generally not well-suited for beginning L2 learners who do not yet possess speaking skills that could be described as communicative.

CONCLUSIONS AND FUTURE PROSPECTS FOR DL LANGUAGE LEARNING

The potential for instruction in LCTLs in places where face-to-face learning is not available has triggered great interest in DL recently. This interest in the DL format must be situated within a proper L2 developmental context: the process of L2 acquisition will take students numerous hours of instruction and, eventually, will require study abroad. A DL course is only one piece in this equation, but a very viable means to afford students access to introductory instruction in LCTLs, especially when local classroom options are lacking or do not fit into personal schedules. The present study offers the FL profession some sense that adding the DL component to the foreign-language curriculum is a responsible and reasonable option, with palpable benefits for oral proficiency. Clearly, more assessment studies from mainstreamed DL language courses are needed to overcome the limitations of the present study as discussed above.

Nevertheless, we have shown that the *Versant for Spanish* instrument is capable of distinguishing different levels of oral proficiency that roughly correspond to first-year, second-year, and heritage students. Proponents of the face-to-face OPI may disagree with Levelt's construct of oral proficiency as operationalized in the *Versant for Spanish* exam, while others might state the obvious that there still exists a need for more data from DL studies. Although the findings presented here should, at the bare minimum, pique the interest of even the most cynical members of the FL profession, more evaluation studies are clearly needed from both a quantitative and qualitative approach. Nevertheless, taking individual L2 differences into account—which constitutes the true bane of all SLA studies--the first-year DL and hybrid students in this study approximate oral proficiency outcomes similar to those of first-year students

working in traditional classrooms. Most importantly, these students are not being disadvantaged by taking Spanish in a non-traditional format. No doubt, students will continue to self-select for the type of language instruction they prefer whenever given the chance. Accordingly, the profession should concern itself with providing legitimate options and increasing all avenues of access to language instruction, especially for the LCTLs. To echo White (2006), once again, there is an urgent need for mainstreaming DL materials into fully implemented courses, especially in the case of LCTLs.

The SWW project and its hybrid derivatives represent a successful experiment in providing the DL option for Spanish. The ten campuses that comprise the University of California system and the base for the UC Language Consortium (see <http://uccllt.ucdavis.edu>) are presently engaged in other ground-breaking projects for providing DL language instruction in Arabic (<http://169.237.245.74/aww/info.html>) from the Berkeley campus and Punjabi (<http://uccllt.ucdavis.edu/distancelearning.cfm>) from the Santa Barbara campus.⁶ Both projects have received funding from the Department of Education (the FIPSE and IRS programs, respectively). Unlike Spanish, both Arabic and Punjabi present writing challenges for students used to working only with Roman scripts. L2 learners of Japanese, Chinese, and Korean face similar issues (Fleming, Hiple, & Du, 2002). Using software and programming that utilize Unicode conventions simplifies most of the potential problems for conducting a DL course in languages with alternate scripts.⁷

The real challenge, however, lies in implementing DL language instruction within a sound pedagogical and truly conversational framework (Laurillard, 2002) that seeks "... to persuade students to change the way they experience the world through an understanding of the insights of others..." (p. 23), blending both direct experiential and formal or academic knowledge. This formidable task is academic in nature and most properly the business of the FL profession. The FL teacher, not the medium, will ultimately determine whether or not any given instantiation of a DL language course makes a positive contribution to the L2 student's long march to advanced proficiency. DL education is a worthy task for the FL profession.

NOTES

1. In the process of gaining acceptance from the UC Berkeley senate for credit for the DL format for Arabic Without Walls, the authors repeatedly encountered these skeptical attitudes from other language faculty, language departments, and course committees.
2. For earlier work on this topic, also see Brecht, Davidson, & Ginsberg (1995).
3. The way that Barr, Leakey & Ranchoux (2005) operationalized the notion of exposure to technology did not strike us as being comparable to the other DL studies reviewed above.
4. The SWW web pages were produced by María Victoria González Pagani, co-PI of the FIPSE grant, P116B000315.
5. *Tesoros* was first published in a five CD set; it has now been released as a single DVD.
6. Other teleconference efforts are well along in Danish, Swedish, and Filipino (see UCLA's World Language Center at <http://www.international.ucla.edu/languages/>). This study did not intend to address the teleconference model for DL.
7. For a basic primer to Unicode, see the entry in Wikipedia: <http://en.wikipedia.org/wiki/Unicode>.

ABOUT THE AUTHORS

Robert Blake (Ph.D. UTexas 1980) is professor of Spanish linguistics and director of the UC Language Consortium (<http://uccllt.ucdavis.edu>).

E-mail: rjblake@ucdavis.edu

Nicole Wilson (Ph.D. UC Santa Cruz 2005) is a research consultant for Pearson in Palo Alto, CA and a lecturer in psychology at UC Santa Cruz.

E-mail: nlwilson@gmail.com

Maria Cetto is a Ph.D. candidate in Spanish linguistics at UC Davis.

E-mail: macitto@ucdavis.edu

Cristina Pardo-Ballester (Ph.D. UC Davis 2006) is an assistant professor of Spanish linguistics at Iowa State University.

E-mail: cpardo@iastate.edu

REFERENCES

Barr, D., Leakey, J., & Ranchoux, A. (2005). TOLD like it is! An evaluation of oral development pilot project. *Language Learning & Technology*, 9(3), 55-78.

Bernstein, J., Barbier, I., Rosenfeld, E., & de Jong, J.. (2004). Development and validation of an automatic spoken Spanish test. In *InSTIL/ICALL 2004 Symposium on Computer Assisted Learning, NLP and speech technologies in advanced language learning systems*, Venice, Italy, June 17-19, 2004, ISCA Archive. Retrieved May 22, 2007 from <http://www.isca-speech.org/archive/icall2004>.

Bialystok, E., & Hakuta, K. (1994). *In other words*. New York: Basic Books.

Blake R. (2005). Bimodal CMC: The glue of language learning at a distance. *CALICO Journal*, 22(3), 497-511.

Blake, R., Blasco J., & Hernández, C. (2001). *Tesoros CD-ROM: A multi-media-based course*. Boecillo, Valladolid: Boecillo Editorial Multimedia (BeM) and New York: McGraw-Hill Publishers, Inc.

Blake, R., & Delforge, A. (2007). Online Language Learning: The Case of Spanish Without Walls. In B. Lafford & R. Salaberry (Eds.), *The art of teaching Spanish: Second language acquisition from research to praxis* (pp. 127-147). Georgetown: Georgetown University Press.

Blake, R., & Zyzik, E. (2003). Who's Helping Whom? NNS/NS Networked Conversations. *Applied Linguistics*, 24(4), 519-544.

Blake, R., Blasco, J., & Hernández, C. (2001). *Tesoros CD-ROM: A Multi-media-based Course*. Boecillo, Valladolid: Boecillo Editorial Multimedia (BeM) and New York: McGraw-Hill Publishers, Inc.

Brecht, R., Davidson, D., & Ginsberg, R. (1995). Predictors of Foreign Language Gain during Study Abroad. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 37-66). Amsterdam/Philadelphia: John Benjamins.

Cahill, D., & Catanzaro, D. (1997). Teaching first-year Spanish on-line. *CALICO Journal*, 14(2-4), 97-114.

- Carr, S. (2000). As distance education comes of age, the challenge is keeping up with the students. *The Chronicle of Higher Education*, 46(23), A39-A41. Retrieved August 27, 2008, from <http://chronicle.com/free/v46/i23/23a00101.htm>.
- Chenoweth, N. A., & Murday, K. (2003). Measuring student learning in an online French course. *CALICO Journal*, 20(2), 284-314.
- Cutler, A. (2003). Lexical access. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science. Vol. 2, Epilepsy – Mental imagery, philosophical issues about* (pp. 858-864). London: Nature Publishing Group.
- Davidson, D. (2007). Study abroad and outcomes measurements: The case of Russian. *The Modern Language Journal*, 91(2), 276-280.
- Dreyer, C., Bangeni, N., & Nel, C. (2005). A framework for supporting students studying English via a mixed-mode delivery system. In B. Holmberg, M. Shelley & C. White (Eds.), *Distance education and languages: Evolution and change* (pp. 92-118). Clevedon, U.K.: Multilingual Matters.
- Erben, T. (1999). Constructing Learning in a Virtual Immersion Bath: LOTE Teacher Education through Audiographics. In R. Debski & M. Levy (Eds.), *WORLDCALL: Global perspectives on Computer-Assisted Language Learning* (pp. 229–248). Lisse: Swets & Zeitlinger.
- Fleming, S., Hiple, D., & Du, Y. (2002). Foreign language distance education: The University of Hawaii experience. In C. A. Spreen (Ed.), *New technologies and language learning: Cases in the less commonly taught languages* (Technical Report #25) (pp. 13-54). Honolulu, HI: Second Language Teaching & Curriculum Center.
- Hampel, R. (2003). Theoretical perspectives and new practices in audio-graphic conferencing for language learning. *ReCALL*, 15(1), 21-36.
- Hampel, R., & Hauck, M. (2004). Towards an effective use of audio conferencing in distance language courses. *Language Learning & Technology*, 8(1), 66-82.
- Holmberg, B., Shelley, M., & White, C. (Eds.). (2005). *Distance education and languages: Evolution and change*. Clevedon, U.K.: Multilingual Matters.
- Jepson, K. (2005). Conversations—and negotiated interaction—in text and voice chat rooms. *Language Learning & Technology*, 9(3), 79-98.
- Lamy, M.-N. (2004). Oral conversations online: Redefining oral competence in synchronous environments. *ReCALL*, 16(2), 520-538.
- Laurillard, D. (2002). *Rethinking university teaching*. London: Routledge.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Payne, J. S., & Ross, B. (2005). Synchronous CMC, working memory, and L2 oral proficiency development. *Language Learning & Technology*, 9(3), 34-54.
- Payne, J. S., & Whitney, P. J. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal*, 20(1), 7-32.
- Soo, K., & Ngeow, Y. (1998). Effective English as a second language (ESL) instruction with interactive multimedia: The MCALL project. *Journal of Educational Multimedia and Hypermedia*, 7(1), 71-89.
- Terrell, T. D., Andrade, M., Egasse, J., & Muñoz, M. (2002). *Dos mundos*. New York: McGraw-Hill Publishing Company.
- Thompson, I., & Hiple, D. (Eds.). (2005). *Selected papers from the 2004 NFLRC symposium: Distance Education, Distributed Learning and Language Instruction (NetWork#44)* [HTML document]. Honolulu:

- University of Hawai'i, National Foreign Language Resource Center. Retrieved March 6, 2006 from <http://nflrc.hawaii.edu/NetWorks/NW44>.
- Thorne, S., & Payne, J. S. (2005). Introduction to the special issue. *CALICO Journal*, 22(3), 369-370.
- Versant for Spanish [Computer software]. Available at <http://www.ordinate.com/versant/business/professional/samples.jsp>
- Volle, L. M. (2005). Analyzing oral skills in voice e-mail and online interviews. *Language Learning & Technology*, 9(3), 146-163.
- Warschauer, M. (1997). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal*, 13(2-3), 7-26.
- Weir, C. (1990). *Communicative language testing*. Englewood Cliffs, NJ: Prentice-Hall.
- White, C. (2006). Distance learning of foreign languages. *Language Teaching*, 39, 247-264.
- Wilson, N.L., Bernstein, J.E., & Rosenfeld, E. (2006, November). *Development and validation of an automatic spoken Spanish test*. The 4th Joint Meeting of the Acoustical Society of America and The Acoustical Society of Japan, Honolulu, HI.
- Zähler, C., Fauverge, A., & Wong, J. (2000). Task-based language learning via audiovisual networks: The LEVERAGE project. In M. Warschauer & R. Kern (Eds.), *Network-based Language Teaching: Concepts and Practice* (pp. 186-203). Cambridge: Cambridge University Press.