Finding a Voice: Learning pronunciation in a second language using a dedicated speech technology

Thomas Kerr
The Learning and Teaching Centre
Macquarie University

This paper reports on a recent experiment that used Wimba Voice Board (WVB), an online asynchronous recording and playback utility, for teaching basic pronunciation rules to a group of novice learners of Spanish. The experimental design used a pre/post-test format with an intervention where participants in the experimental group were given access to a built-for-purpose Blackboard online unit and encouraged to engage with the learning materials in self-directed study (Victori and Lockhart, 1995; Lee, 1997). Participants were also able to voluntarily upload their own attempts at pronunciation to a WVB module accessed by all group members, listen to attempts uploaded by peers, and provide optional feedback. Control group members were taught the same content in a traditional teacher-led classroom setting. Two focus groups were conducted with members of the experimental group. Analysis of the collected data showed that WVB was able to produce results that were equivalent to those achieved by members of the control group.

Keywords: second language learning, speech technologies, Wimba Voice Board

Introduction

Text-based asynchronous discussion boards have been used in second language (L2) learning settings as long as the World Wide Web has existed, at least since the early 1990s (Warschauer, 1996). In teaching approaches that emphasise communicative competence (Hymes, 1967; Canale, 1983; Swain, 1985; Littlemore and Low, 2006), the issue of correct pronunciation becomes critical if the goal of effective communication is to be achieved. Speech technologies such as Wimba Voice Board (WVB), an online asynchronous recording and playback utility (Wimba.com, 2010), offer educators the ability to provide opportunities for groups of learners to exchange recorded attempts at pronunciation, thus promoting peer and self-directed learning.

A recent study at a a metropolitan Australian university attempted to teach a group of novice Spanish learners the basic pronunciation rules that govern correct speech production using a speech technology (WVB), and without explicit instruction by a teacher. Ethics approval was obtained for the study as it involved quantitative analysis of the participant's test scores and qualitative analysis of their opinions about the software used. Participants in an experimental group were given access to a built-for-purpose Blackboard unit containing text-based instructions on the basic rules of Spanish pronunciation. They were also able to access model recordings of all the words and phrases used. Two WVBs were included, allowing personal and group asynchronous practice. An equivalent control group was taught the same rule set by a native Spanish-speaking tutor in a traditional classroom setting.

The research question had two components; first, it asked if a speech technology such as WVB could be used to improve pronunciation in a second language, and second, it asked if the learning was effective when compared to traditional ways of learning the same material. The choice of the learning task was kept deliberately simple, as it was the efficacy of the speech technology itself in a specific learning task that was under investigation, rather than the technology's effectiveness as a language-learning tool. Scanlon and Issroff (2005) argue that basing the evaluation of a learning technology solely on maximised learning outcomes is too limiting in attempting to understand how the technology influences learning, given the complex issues (Oliver and Harvey, 2002) involved (p. 431). They advocate the use of an approach based on an extended version of Activity Theory that accounts, in part, for this complexity. Bearing this in mind, the researcher chose to extend the reach of the study by incorporating a qualitative analysis of the software use based on focus groups conducted with some of the participants exposed to the speech technology.

A primary aim of the research was that the results obtained and their subsequent analysis would inform both the future design of online language courses and those educators seeking to incorporate speech technologies in their teaching. This potentially contributes to the conference theme of 'learning for the future' and would enable L2 courses be delivered in an alternative mode in subsequent offerings.

Experimental study design and deployment

The study design used a mixed methods approach (Johnson and Turner, 2003), where both quantitative and qualitative data were collected and analysed. Quantitative data consisted of results obtained from pre and post-tests administered to two groups. Qualitative data was derived from analysis of two focus groups conducted at the conclusion of the study's experimental phase.

The experimental approach used was based on a comparative pre-test post-test control group design as described by Shadish, Cook and Campbell (2002). Participants were randomly assigned to either an experimental or control group and further randomized for gender. The final makeup of the groups was nine participants in each, with genders approximately equally distributed between groups. Participants were recruited from staff at the author's institution. In order to establish a common baseline for previous language-learning experience, all participants were asked if they had any previous experience in learning Spanish. Only respondents with no previous experience were subsequently chosen as participants in the study.

The pre-test was administered to individual members of both groups in an isolated test space chosen for the purpose. All participants were given a list of 20 words to pronounce in Spanish, each with an accompanying contextual phrase containing the word. The list consisted of previously unseen words and associated contextual phrases in Spanish known to be challenging for new learners of the language ((Diaz-Campos, 2004). Words such as *nicaragüense* (Nicaraguan) and *rehenes* (hostages) are not normally encountered in introductory Spanish courses, but were included to provide a high level of pronunciation contrast in order to test the efficacy of the two teaching methods used. Participants were asked to work through the list making their best guess as to how to pronounce each word and phrase. All attempts were recorded using the Wimba Voice Board technology and coded for subsequent rating by raters who were native speakers of Spanish. In the intervention phase of the study all participants were then either taught in a teacher-led face-to-face traditional classroom or given access to an online unit for self-directed study.

The control group received instruction in the basic rules of Spanish pronunciation in a traditional classroom setting. Two sessions were run, each lasting approximately 45 minutes and attended by groups of 4 and 5 participants. The instructor used a series of slides titled "Basic Spanish Pronunciation" (Fig.1) that included basic rules of pronunciation with example words and sentences, and a list of the same 20 words and contextual phrases used in the pre-test.
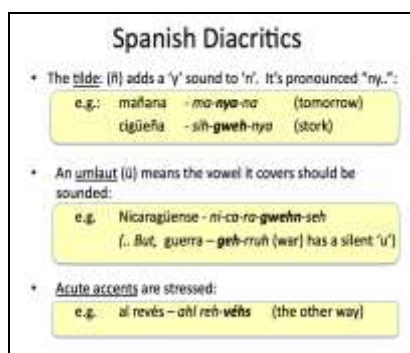
Figure 1: Example slide from the Basic Spanish Pronunciation slideshow used with the Control group

The teaching approach used was listen-and-repeat where the instructor provided a model sound and the participants attempted to repeat it. The instructor also provided some additional phrases not included in the slide materials in order to provide additional contextual examples of use. Participants were asked to make both group and individual responses when pronouncing words and phrases from the list. At the conclusion of the session, the participants were offered a printout of the instructional slides and told that they could refer to them if they wished. There was no instruction to either collaborate with their peers or to avoid comparing their learning experience. That is, participants were free to practice their pronunciation as they wished, with or without the support of their co-participants.

The experimental group did not receive any direct instruction apart from basic instruction on how to record, upload and play back personal and peer attempts at pronunciation. All of their learning was conducted online within a dedicated Moodle unit. They were able to access the same set of slides used with the control group, with the difference that all model sounds for the words and phrases were pre-recorded by the instructor and

made available through embedded playback consoles on each page. The actual content of the Basic Spanish Pronunciation module was the same for both the control and experimental groups. In addition to the pronunciation module, the experimental group was able to access two dialogue modules. The Private Practice Board gave users the opportunity to record, upload and play back their attempts at pronouncing the word list, knowing that access was restricted to the online moderator and themselves. The Pronunciation Discussion Board was designed to allow opportunities for peer-supported learning. Participants were encouraged to make a minimum of two uploads of to the board over a 48 hour period and to comment on at least one other learner's attempt.

All participants then completed a post-test consisting of recordings of the same list of 20 Spanish words and phrases, but in randomized order in order to avoid the possibility of threats to validity through memory or sequenced learning effects that might arise if the lists were exactly equivalent.

Within one week of the administration of the post-test, two focus groups were run with five volunteer members of the experimental group. Both sessions were recorded and transcribed for further analysis. The focus groups were conducted with two and three volunteer participants from the experimental group only, as the qualitative enquiry related to use of the WVB speech technology used by them for completing the pronunciation task. Prompt questions used by the focus group leader were designed to elicit responses relating to technology use from both cognitive and emotional points of view. Participants were asked pragmatic questions such as whether they thought the use of the WVB software had helped them in the learning task, ("Do you think the VB software helped your pronunciation of some Spanish words? If so, how do you think it helped?"), how useful they thought interacting with other learners was, and how this experience compared with interacting with a tutor. They were also asked to compare use of the software for learning with learning in a classroom or tutorial group.

Findings

The 18 participant recordings from the pre and post-tests were collected and de-identified with a series of code numbers. These were then rated by a native speaker of Spanish (Rater 1) using a rubric sheet with a five-point likert scale. Ratings included: "Needs practice" (1), "Below average" (2), "Average" (3),"Good" (4) and "Excellent" (5). To establish inter-rater reliability, 6 of the pre/post-test scores, randomly selected from both control and experimental groups were also rated by a second native speaker of Spanish (Rater 2) using the same rubric sheet, representing a 33% sample of the total population Statistical analysis of the two rater score sheets for the same six participants indicated that there was no significant difference between the two raters' scores. Rater 1's pre and post-test totals difference scores for all 18 participants were then compiled and ranked in ascending order then compiled in a chart for purposes of comparison (Fig. 2).
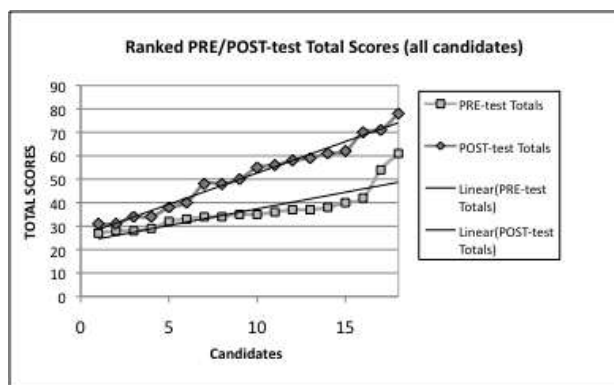
Figure 2: Ranked pre/post–test totals for all 18 candidates with trends indicated

Post-test total scores trend higher than the pre-test scores as would be expected, as all candidates had been exposed to the word list in the pre-test and then again in either the classroom session or in the WVB online unit. A Mean difference score was calculated for all 18 candidates resulting in a score of 14.67. On average, with the exception of a single outlier, all candidates improved their scores on any given word from the word list by an average score of 0.733 (Mean / word count).

The first part of the research question mentioned above suggests a hypothesis that an analysis of the data from the experimental group's pre and post-tests would show a significant improvement in performance indicated by the difference scores achieved; that is, a measurable, positive score difference in the post-test. The second part of the research question suggested a null-hypothesis; that is, that analysis would *not* show a significant

difference in pre and post-test difference scores between the experimental and control group participants. Pallant (2005) suggests that non-parametric analysis methods are indicated in small sample populations such as the WVB study, with non-normal distributions. To determine the validity of adopting this approach, a preliminary test series for normality of distribution was first run using Kolmogrov-Smirnov and Shapiro-Wilk tests. Establishing a normal distribution of difference scores requires a significance ('p') value of < 0.05. Both the Kolmogorov-Smirnov and Shapiro-Wilk measures (.200 and .991) were significantly higher, indicating that non-parametric analysis would provide the most authentic results.

Two measures were adopted to determine the significance of the effect of the intervention (the experimental group's use of WVB to complete the learning task);

- a non-parametric test (Related-Samples Wilcoxon Signed-Rank test) was applied to the experimental group's pre and post-test totals to determine the significance of the median of differences between them
- **a** Descriptives Table was produced including Mean, Standard Deviation and Variance measure**s.**

Results from both measures indicated that a significant, positive difference existed between the experimental group's test totals when pre-and post test were compared. That is, the post-test scores were significantly higher, supporting the hypothesis derived from the first part of the research question. In order to test the null-hypothesis derived from the second part of the research question, difference scores from the control and experimental groups were compared using an Exact significance measure for both 1-tailed and 2-tailed outcomes. A significant difference in the experimental and control groups' pre/post-test difference scores would be indicated by a significance level of < 0.05. Results obtained from the Asymptotic Significance 2-tailed score (.185) and the Exact Significance scores (0.190 and 0.197) for both 1 and 2-tailed outcomes all fell substantially above 0.05, indicating no significant difference detected when comparing the two groups' difference scores and supporting the null-hypothesis.

Analysis of the two focus groups showed that a significant level of anxiety was experienced by at least some members of the experimental group when making recordings for uploading to the Blackboard unit. Despite this, all participants reported that they thought that use of the WVB software had assisted them in a positive sense while attempting to learn correct pronunciation in the L2.

Conclusions

Results from the analysis of the pre and post-tests results for the two groups participating in this study indicate that 17 of the 18 participants showed a measurable improvement their ability to pronounce Spanish words correctly. This suggests that, excluding the control group results, use of WVB as the sole means of learning correct pronunciation in Spanish can produce measurable gains in novice learners. Results from the comparison of the experimental group's difference scores with those of the control group indicated no significant difference in the learning gains achieved. The conclusion suggested is that use of speech technologies such as WVB can produce equivalent results to those achieved through traditional forms of teacher-led classroom instruction in an L2. An interesting result of the focus group discussions was that at least two participants thought that, with appropriately designed content, prospective language students could use speech technologies such as WVB to preview a language course by attempting a simple learning task such as the basic rules of pronunciation before committing to an intensive study of the language itself. This suggests that speech technologies have a definite role to play in future language education and that the results of this study could inform the future design of online language courses in sometimes unexpected ways.

References

Canale, M. (l983), From communicative competence to communicative language pedagogy. In Richards, J., Schmidt, R. (Eds) *Language and Communication*. London: Longman, pp. 2–27.

Diaz-Campos, M. (2004). Context of learning in the acquisition of Spanish. In *Studies in Second Language Acquisition*, 26. Cambridge University Press, pp. 249-273.

Hymes, D. (l967). Models of the interaction of language and social setting. *Journal of Social Issues* 23(2), pp. 8–38.

Johnson, R. B., & Turner, L. A. (2003). Data collection strategies in mixed methods research. In A.Tashakkori, A. and Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioral research.* Thousand Oaks, CA: Sage, pp. 297–319.

Lee, I. (1997). Supporting greater autonomy in language learning. *ELT Journal,* Vol. 52/4. Oxford University Press, pp. 282-290.

Littlemore, J. and Low, G. (2006). Metaphoric competence and communicative language ability. *Applied Linguistics* 27(2), pp. 268-294.

Oliver M. & Harvey J. (2002) What does 'impact' mean in the evaluation of learning technology? *Educational Technology and Society* 5, pp. 18–24.

Pallant, J. (2005). *SPSS Survival Manual*. Sydney: Allen and Unwin.

Scanlon, E. and Issroff, K. (2005). Activity Theory and Higher Education: evaluating learning technologies. In *Journal of Computer Assisted Learning* **21**. Blackwell Publishing Ltd., London, pp. 430-439.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Swain, M. (1985) Communicative competence: some roles of comprehensible input and comprehensible output in its development. In Gass, M.A., Madden, C.G. (eds.) *Input in Second Language Acquisition*. Rowley MA: Newbury House, pp. 235-253.

Victori, M. and Lockhart, W. (1995). Enhancing metacognition in self-directed language learning. In System Volume 23, Issue 2, Elsevier, pp. 149-282.

Warschauer, M. (1996). Computer-assisted language learning: An introduction. In S. Fotos (Ed.), *Multimedia language teaching*. Tokyo: Logos International.

Wimba.com (2010). Wimba Voice version 6.0.3. Released December, 2010. Description available from: http://www.wimba.com/products/wimba_voice/

Author contact details:

Thomas Kerr, tom.kerr@mq.edu.au

Please cite as: Kerr, T. (2012). Finding a Voice: Learning pronunciation in a second language using a dedicated speech technology. Proceedings Ascilite 2012. (pp.484-488).