TOLD LIKE IT IS! AN EVALUATION OF AN INTEGRATED ORAL DEVELOPMENT PILOT PROJECT

David Barr, Jonathan Leakey, and Alexandre Ranchoux
University of Ulster, Northern Ireland.

ABSTRACT

Much established pedagogical and CALL (computer-assisted language learning) research advocates an integrated constructivist approach to the use of technology in language learning. This paper reports on a pilot project delivered to first year undergraduate French students. The project aim was to deliver a blend of collaborative and individual learning through a combination of CALL programs and online activities alongside traditional face-to-face conversation classes. Using quantitative analysis of a pre- and posttest and a variety of questionnaires, this project assessed student progress in developing oral skills across two groups, one (the treatment group) using technology and the other (the comparison group) being a traditional conversation class. Each group covered the same content and underwent the same assessment procedures. In addition, through qualitative analysis measures, the project evaluated the role played by additional variables in the learning process, as well as student and staff reactions to the two approaches. The study concludes by showing that while progress was made by both groups, the progress made by those not using technology was significantly greater than that made by students using technology over a short-term study. It also highlights the need for developing pedagogy to ensure that CALL-based teaching goes beyond rehearsal activity to achieve message-orientated communication.

INTRODUCTION -- BACKGROUND TO THE RESEARCH

A glance through relevant literature since 2000 shows very little reference has been made to the role of computer technology in oral language development.[1] More frequently, the role of oral work is hinted at in titles referring to content looking at "CMC" (computer-mediated communication), "multimedia," "e-learning," or "online interaction." On closer inspection, however, these latter, more general, references tend to deal with written communication forms such as tandem learning via e-mail, discussion forums, or chat-rooms.

Levy's CALL survey of 1990 has two interesting findings in this respect. Of 17 categories of teaching approaches, the "communicative" approach was the most popular, with the "oral" approach and "direct method" coming in seventh and eighth, respectively (Levy, 1997, p. 123).

However, of 17 categories of CALL software development activity "Speaking" appears in the 15th position, just ahead of "Vocabulary" and "Other," the top three items on this list being "Reading," "Writing," and "Gap-filling" (Levy, 1997, p. 143). Clearly, developmental work in those "early days" lagged behind the pedagogically desirable and normal classroom practice.

Most CALL-based teaching and learning has tended to focus on non-oral activities such as software or Web-based reading, writing, or gap-filling type activities. The conversation class, pair and group role-plays, and discussions have for the most part taken place in ordinary classrooms. Felix (2001) lists "lack of speaking practice" first on the students' list of disadvantages of using Web-based programs for language, along with "distraction," "no interaction with peers," "inadequate feedback," and "absence of teacher" (p. 47). According to Boullier (2000),

> The combination of various media and techniques has always existed: oral practice did not disappear with the advent of writing, the same applies for hand-written practice with printing, newspapers with TV, post with telephone … each media and technique re-defined its sphere of actions according to its specificity. (p. 145; author's translation)

The teaching of all the language skills has gone through just such a phase of re-definition in response to the latest computer-based media; in the case of the oral skill, it is merely taking a little longer.

Technology and oral language development have been rare bedfellows and for one main obvious reason: The technology for oral language development has posed the greatest challenge to both hardware and software developers. More specifically, one of the main challenges has been to create speech recognition software that is sensitive and accurate enough for language learning (Hincks, 2003). Speech recognition software lends itself mainly to drill-type activities; further logistical and technological hurdles must be crossed if one aims to get beyond purely text-based tandem exchanges and encourage real-time oral communication across campuses or across national boundaries.

There have been experiments using telephone and video-phone link-ups where there are gains in smoothness of data transfer but loss in terms of cost, which has usually been prohibitive. The main hurdles in any attempt at Internet-based link-ups appear to be the challenge of overcoming time-zones and timetabling differences, ensuring adequate bandwidth at both ends, and the cost of calls. Goodfellow, Jeffreys, Miles, and Shirra (1996), in their study of a "video-conference try-out," draw several conclusions that will still apply however good the technology becomes. They state that the "language interaction [video-conferencing] supports is in many ways different from the 'face-to-face' equivalent" and cite, amongst others, such aspects as the restrictions the video-conference imposes on the teacher's moderation of group interaction, the distortion of the normal use of body language to manage interaction, and problems of managing camera viewpoints as challenges imposed by the technological context (p. 16). They conclude by saying that "we have to plan for it [i.e., videoconferencing] and adapt our teaching and learning methodologies accordingly. That way we will be able to enhance, rather than merely repackage, the educational service we offer" (p. 16). An overall pedagogy for successfully integrating technology into oral work, be it in local or remote mode, still seems to be lacking along with a holistic approach to technology-based assessment of oral development. The initial TOLD (Technology and Oral Language Development) project at the University of Ulster has focussed on communication within the classroom/e-lab, but the aim in future realisations of TOLD is to explore a workable and cost-effective link-up with a francophone university. Such a collaboration will aim to prepare students at both ends to work in a structured way and to integrate the content of live exchanges with ongoing curriculum work.

Research at the University of Ulster has looked closely at the issues of student resistance (Gillespie & McKee, 1999) and staff reluctance (Gillespie & Barr, 2002) with regard to CALL. From a psycho-linguistic perspective it would seem likely that of all communication mediated by technology, using technology for face-to-face speech would generate the greatest affective hurdles in the minds of teacher and student since it is the skill that least needs technology. As Levy puts it, "Face-to-face speech is the only technology-free mode of communication, aside from sign language" (2000, p. 184).

The TOLD Project has been designed to examine the value of technology-mediated oral communication by piloting a blended learning approach making full use of a recently installed multimedia laboratory and measuring student progress over three years. Perhaps the greatest challenge for the project was to ensure the principle of the *real need to communicate* applied in our use of technology for oral communication. By integrating technology into perhaps the most authentic and meaningful form of communication -- oral conversation -- this project goes some way to addressing the criticism of Warschauer (1996), who highlighted a weakness of much multimedia based language teaching: "Using multimedia may involve an integration of skills (e.g., listening with reading), but it too seldom involves a more important type of integration -- integrating meaningful and authentic communication into all aspects of the language

learning curriculum" (p. 6). While this pilot project falls short of oral communication over the Internet or by video conferencing, the aim is, in its next stages, to apply the lessons learned in-house and trial a pedagogical approach to remote oral communication as well.

Our approach drew primarily on communicative and constructivist theories of second language acquisition. In exploring these theories, we found that the following elements were particularly pertinent to oral development. According to the communicative approach to language learning, activities are best geared to flow in an ordered progression from rehearsal to meaningful performance. Carl Dodson (1978), for example, advocates a two-staged training approach for communicative acts: the *rehearsal* stage and the *performance* stage. Dodson emphasises the importance of allowing the student to pass from "medium-orientated communication" to "message-orientated communication" (p. 48).

According to constructivist theories, learning should be multi-modal, task-based, and content-based (Warschauer & Healey, 1998). There should be a strong focus on learner-centredness, with the teacher acting as a facilitator and the learner free to make his or her own interpretations. According to Driscoll (1994), student ownership of learning should be fostered through reflection. Examples of this might include learner logs and goal-setting, reflection, and monitoring of progress. Self-awareness of knowledge construction should also be encouraged (Driscoll, 1994). Another aspect of constructivist learning is Vygotsky's zone of proximal development (1978), which involves developing a scaffolding for knowledge construction through student collaboration. All of these features were considered important in underpinning the pedagogy behind the TOLD project.

In constructing a project looking at assessment and aiming also to measure progress, the TOLD Project team was also mindful of past criticism levelled at weaknesses of much research involving quantitative measurement of student progress using CALL. Due to various constraints (e.g., pressures of timetabling, small class sizes, problems with the obtaining of comparison groups, unfamiliarity with statistical analysis) it has often been difficult for language teachers to address charges of small sample sizes, faulty statistical analysis, and inadequate length of treatment to measure educational outcomes (Reeves, 1993; Schmitt, 1991;). Salaberry (1996) has called for a careful combination of qualitative and quantitative analysis that addresses these various design problems and that is founded on a sound basis of pedagogical theory.

This paper examines the following questions:

1) Does computer technology enhance significantly progress in students' oral language development?
2) What factors may affect students' oral language development when using computers?
3) How do staff and students react to the use of computer technology for oral language development?

Our null hypothesis is that a CALL environment makes **no** difference to learning gains in oral language development. The alternative hypothesis (i.e., that it does make a significant difference, be it positive or negative) will be gauged by a configuration of data collection methods (Levy, 2000, p. 180) including quantitative analysis of learning gains.

With only one semester of a three-year quasi-experimental study[2] completed, our sample size of 29 drops just short of enabling us to assume normality and we have had to run appropriate analyses for small samples (see the later description of the Mann-Whitney test, Wilcoxon's matched pairs test, and the Spearman Rho test) alongside the usual parametric tests. Nevertheless, the project team aims to repeat the study over the next six semesters with first year students so as to ensure a larger sample size and will also follow their progress over the course of their degree programme, with the aim of obtaining increasingly reliable data.

METHODOLOGY

Context of Project

1) Hours -- 1 hour per week over a 12-week semester, the standard time allocated to French conversation classes
2) Language groups -- The students were divided into four small conversation classes of between 5 and 11 students, along course lines. This was determined mainly by timetable considerations. Two foreign language assistants each taught two classes.
3) Comparison groups -- Two of the four groups were "comparison" groups taught in the "traditional" manner (i.e., usually in the classroom, sometimes also in the [analogue] language laboratory; never in the multimedia lab). For the purpose of this paper we will refer to these two groups as one entity: "the comparison group."
4) Treatment/test groups -- two of the four groups were "treatment" groups and were taught every week in the multimedia laboratory. For the purpose of this paper we will refer to these two groups as a single entity: "the treatment group."
5) The comparison groups were made up of Combined Arts students and the treatment groups made up of Applied Languages students. Our analysis of student ability found broad comparability between the two groups.
6) Resources available to the two treatment groups were
    - FLA;
    - digital multimedia laboratory; and
    - the software TellMeMore® (Version 5) by Auralog. This software includes speech recognition technology with a variety of colourful interactive activities rehearsing all the four main language skills and provides an interactive glossary and grammar.

Data Collection and Evaluation Methods[3]

1) ***Language Experience Questionnaire.*** This asked for data such as language qualifications, number of foreign languages studied, amount of time spent in the country of the target language, and student confidence and fluency levels across the range of language learning skills. An overall percentage rating was obtained for each student. This was correlated with student progress (posttest percentage less pretest percentage).
2) ***ICT-use survey***. This collected information on student access to and use of a range of common ICT (information and communications technology) applications, e-mail, and the Web for personal and study purposes. It also gathered data on student use of mobile phones. An overall percentage rating was obtained for each student. Again this was correlated with student progress. Both surveys support the theory of Driscoll (1994), advocating self-awareness of knowledge construction.
3) ***Journals***. Students were given a paper-based log at the start of the project and were asked to list their goals for progress in their oral language development from a checklist of oral skills and then, for each session, to record their impressions of the lesson. They were asked to comment on what they thought had worked well and make suggestions as to how it could have been improved. This supports Driscoll's theory about student ownership of learning.
4) ***Pre- and Posttest.*** All students sat the same pre- and posttest under the same conditions. The setting for the tests was a traditional (analogue) language laboratory. There were five sections to the test: a pronunciation task; some personal questions; a listening comprehension exercise presented initially without transcript of the text and questions, and then with these, for which the students recorded oral answers; and lastly, an oral résumé of an extract of a television documentary.

PROJECT ENVIRONMENT, CONTENT, AND DELIVERY

The Environment

The Faculty of Arts at the University of Ulster is spread across multiple campuses. Our project utilized the language resources available to us on the Coleraine campus. The facilities at hand include a Multimedia lab computer-based and an audio-visual laboratory.

The multimedia classroom possesses a comprehensive video and audio switching system for the management of the classroom. *SmartClass*® Plug and Play computer supervisory system enables the teacher to maintain audio and visual contact with each student and the computer screen they are using. The system has a teacher console and 16 workstations each networked and equipped with *Robotel* technology, in particular the *SmartClass*® multimedia language-learning environment (see Figure 1). The language tutor can monitor and communicate with individual students or the whole class, broadcast material from and to any workstation in the network, or take control of any student's screen, keyboard or mouse functions at any time.

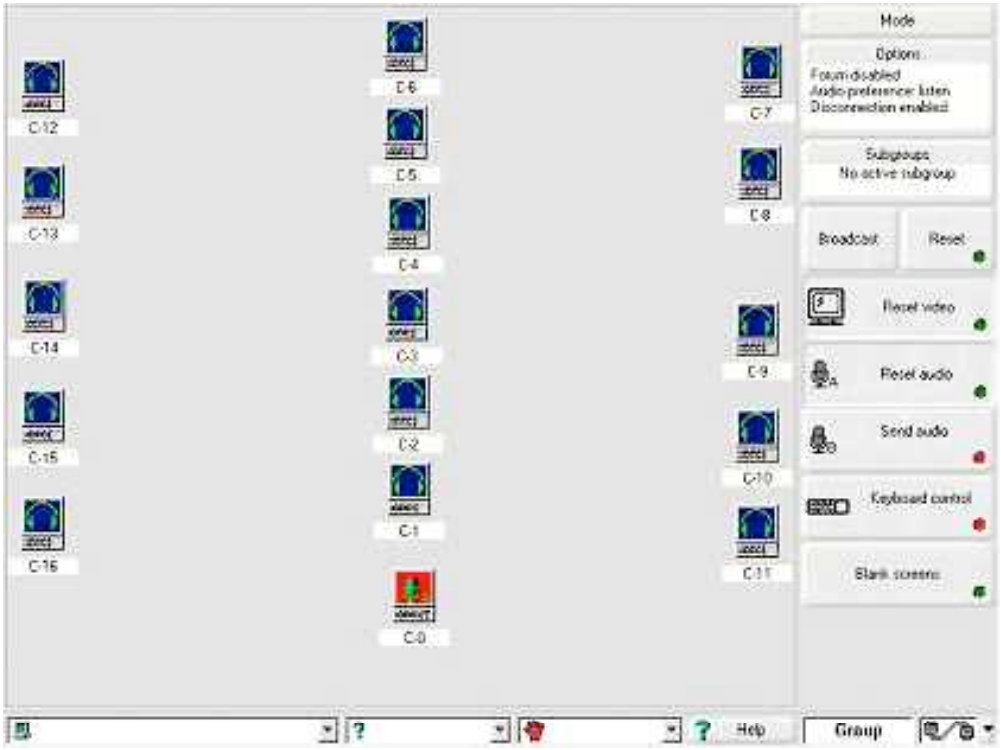

Figure 1. Controlling the multimedia classroom: the *SmartClass*® visual interface

The hardware was used in conjunction with *TellMeMore*® (version 5, Auralog), including speech recognition technology S.E.T.S. (spoken error tracking system) which automatically detects pronunciation errors. The *TellMeMore*® software includes a variety of colourful interactive activities rehearsing all the four main language skills as well as providing an interactive glossary and grammar.

Content

Our pedagogical approach was informed by a desire to draw on the effectiveness of both traditional teaching and learning methods as well as the benefits of a technological environment. An effective traditional method is the small group discussion with its emphasis on human interaction in a meaningful context, which is difficult to replicate in a technological setting. One of our aims in this project was to replicate this small group intimacy factor within the potentially impersonal e-lab. An example of an effective technological method might be the use of CALL software for pronunciation drilling with

individualised instant feedback. We felt it important to make the most of this feature in the e-lab, and for the sake of comparability we introduced a pronunciation-drilling element into the comparison (Non-Tech) group's routine. Thus, to a large extent our pedagogical approach was determined by our desire to keep the content similar across the two groups.

Planning

Our lesson planning was informed by the desire to blend the best of both approaches. Activities were also designed to flow in Dodson's (1978) progression from rehearsal to meaningful performance, allowing the student to pass from "medium-orientated communication" to "message-orientated communication" (p. 48).

The design of the CALL software used (*TellMeMore*®) seemed to reflect an awareness of these stages (e.g., start with pronunciation drilling and progress to simulated interactive role-plays). However, we were conscious that the software gave no opportunity for face-to-face communication between the students and they saw this too. For this reason, our lesson design included progression beyond the software to the message-orientation phase of more traditional group discussion and presentation.

Planning and Content Features Common to Both Groups. Lesson plans shared a similar structure involving progression from pre-communication to rehearsal, information retrieval, assimilation, and final meaningful production in the target language. All groups shared the following 11 oral language development skills targets:

- pronunciation
- accent and intonation
- fluency
- one-to-one with a French person
- one-to-one with an English speaker
- responding spontaneously in a conversation
- responding to visual or aural input (e.g., from TV/Radio)
- taking an active part in a structured group discussion
- taking an active part in an unstructured group discussion
- giving a group presentation
- giving a presentation on my own

These skills informed lesson planning and student goal setting and were integral to monitoring, feedback, and students' reflection on their own learning. This was drawn from social constructivist theory, for example Vygotsky's Zone of Proximal Development (1978), which involves developing a scaffolding for knowledge construction through collaboration.

Delivery

For the complete list of lesson plans and topics, please refer to http://www.arts.ulster.ac.uk/lanlit/french/research/told/

Although classes for the treatment group took place in the multimedia teaching lab, while students in the comparison group did not have these facilities at their disposition, the lesson topics, texts, comprehension questions, and pronunciation drills remained common to both groups. Students in the treatment group recorded their pronunciation of a passage or respond to a series of pre-recorded questions digitally, while the comparison group recorded themselves using an analogue recorder. An example of a pronunciation-comprehension-oral production activity can be found in Appendix A.

Multimedia was a feature of all the oral classes for the treatment group, and not just an add-on. The only time that students regularly broke from interaction with the computers was for the purpose of group discussion or conversation. Given the restrictions of the software and hardware resources available to us

and the fact that we were not using CMC (computer mediated communication) for this project we concluded that the multimedia lab was best suited for the tutorial and rehearsal and assessment phases of oral work. We found a strong case can be made for the use of technology for these phases of oral skills teaching and learning.

Although our pedagogical approach reflects a combination of objectives and requirements set out a priori, in practice, each tutor enjoyed a degree of flexibility in terms of his/her content delivery, presentation, and style. Also, the original pedagogical plan evolved during the course of the semester, in light of ongoing feedback from tutors, students, as well as research staff. The original goals given to the students remained unchanged, as did the guiding principles of message-orientated communication via the process of pre-communication, rehearsal, information retrieval, assimilation, and final meaningful production.

QUANTITATIVE ANALYSIS

Research Question 1: To evaluate the learning gains in oral language made by the treatment and the comparison groups and test for significance

For an analysis of **individual** learning gains see Appendix B, which presents the collated raw data, showing individual scores for the predictor (or independent) variables of attendance, language learning experience, and ICT-use, and the individual scores for the outcome (or dependent) variables of the pre- and posttest, as well as the difference between the two. These are all expressed as percentages. What follows below is primarily an analysis of the group mean scores taken from the raw data.

An Analysis of Group Learning Gains

Figure 2 and Table 1 show the mean improvement, or group learning gain, from the pre- to the posttest for the treatment group (Tech) and the comparison group (Non-Tech). At this stage we are looking at global mean scores, not the learning gains in each of the oral skills tested for.

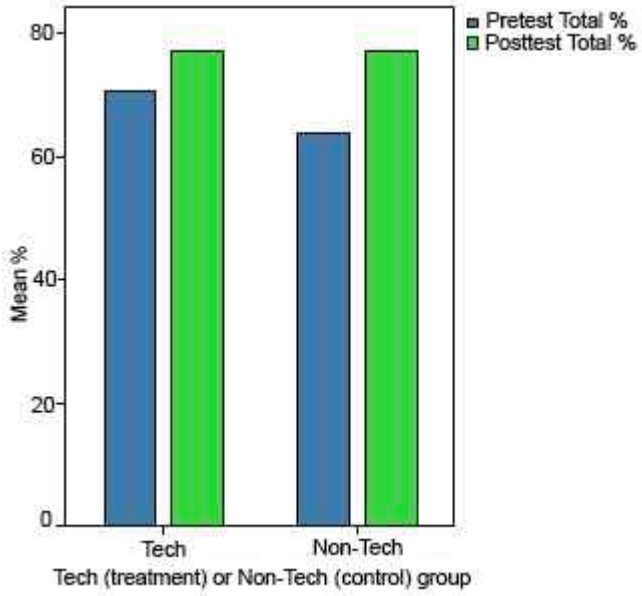

Figure 2. Mean improvement from the pre- to the posttest for the treatment group (Tech) and the control group (Non-Tech)

Group Statistics for the Means of Pre-/ Post-Test Totals (see also Figure 2)

Table 1. Mean Improvement From the Pre- to the Posttest for the Treatment Group (Tech) and the Comparison Group (Non-Tech)

	Tech or Non-Tech	*N*	*M*	*SD*	Std. Error Mean
Total %	Tech	15	70.60	13.222	3.414
	Non-tech	14	63.64	15.184	4.058
Ptotal	Tech	15	77.20	7.408	1.913
	Non-tech	14	77.21	10.519	2.811

From Figure 2 and Table 1 we can see that over the course of the semester and between pre- and posttest, both groups on average made progress, while the Non-Tech group made more progress than the Tech group. This quantitative data raises several qualitative questions:

- Why did the comparison group (those not exposed to technology), while starting from a lower mean starting point (64%) make a 13% leap up to 77%, compared with the treatment group's 6% rise from 71% to 77%?
- Was the technology a hindrance, and if so at what point and for how long? And would this hurdle over a longer study period be reduced as the students got used to the technology?
- How much did the time spent coaching in how to use the multimedia lab and associated software, and dealing with technical glitches, affect the progress of the treatment group?
- Can we triangulate data from the qualitative data (the student and staff logs and evaluations) that would clarify the quantitative data?
- If we look closer at the individual tasks and skills do they reveal if the progress is being made in some tasks and skills more than others?
- Are other variables (such as language learning experience, ICT-use, attendance) influencing the data?

Having looked at the global group learning gains we can now compare progress in each of the different tasks and skills tested for in the pre- and post-test. In Table 2, the first 11 rows cover the *pre-test* scores for the five tasks and five skills assessed and the total (Total%); the next 11 rows cover the equivalent scores for the *post-test.* (Note: Tasks 1 and 3 became unusable for the purposes of comparison because the wrong text/procedure in the pre-test was mistakenly used in one of the pre-test sittings.)

The following explanations refer to the abbreviations we have used to denote different sections of the pre- and post-tests:

Task 1,2,3,4,5	refers to the *pre-test* score
ptask 1,2,3,4,5	refers to the *post-test* score
Task 1 and ptask1	Pronunciation task -- students read from a text selected to assess command of a range of phonemes that typically pose a challenge to anglophones, and are marked for Pronunciation, Accent/Intonation, and Fluency.
Task 2 and ptask2	Students answer questions (x5) about themselves (e.g., how long they've been learning French/ where they come from/ describe their personality/ interests/ plans for the summer). They are marked for Pronunciation, Accent/Intonation, Fluency, Content, and Grammar.
Task 3 and ptask3	Students answer, orally, questions on a heard text (no transcript). They are marked and graded by the teachers.
Task 4 and ptask4	Students answer, orally, questions on the same text as above, but this time seen as a transcript; the questions are also seen this time.
Task 5 and ptask5	Students view a short (3 minute) video clip (from a documentary about France 2) and give an oral résumé of the content (after 2 viewings).

	Students were not given freedom to replay the video.
Total and Ptotal	refer to the respective totals from the two tests
Pronunciation and ppron, etc.	likewise refer to the respective (pre- and post-test) scores for each *skill* assessed in the different tasks

Teachers conducted all marking after the tests were completed. All answers were recorded on audio-cassette.

Table 2. Task-by-Task and Skill-by-Skill Comparison of Pre- and Posttest Scores for Both the Treatment and Control Groups

	Tech or Non-Tech	*N*	*M*	*SD*	Std. Error Mean
Task 1%	Tech	6	69.67	11.639	4.752
	Non-tech	4	69.00	5.477	2.739
Task 2% (PRE)	Tech	14	70.57	14.053	3.756
	Non-tech	14	62.14	22.017	5.884
Task 3%	Tech	6	71.67	13.110	5.352
	Non-tech	3	58.00	12.288	7.095
Task 4%	Tech	14	75.57	15.820	4.228
	Non-tech	13	72.15	16.411	4.551
Task 5%	Tech	15	65.47	17.900	4.622
	Non-tech	14	60.00	14.655	3.917
Pronunciation %	Tech	15	65.60	13.087	3.379
	Non-tech	14	58.21	10.504	2.807
Accent/Intonation %	Tech	15	67.80	12.531	3.236
	Non-tech	14	57.36	10.135	2.709
Fluency %	Tech	15	75.60	14.966	3.864
	Non-tech	14	68.36	18.333	4.900
Content %	Tech	15	73.60	15.099	3.898
	Non-tech	14	70.14	18.691	4.995
Grammar %	Tech	15	72.53	15.514	4.006
	Non-tech	14	61.64	18.612	4.974
Total %	Tech	15	70.60	13.222	3.414
	Non-tech	14	63.64	15.184	4.058
ptask1	Tech	15	77.80	7.103	1.834
	Non-tech	14	71.71	10.730	2.868
ptask2 (POST)	Tech	15	79.53	9.219	2.380
	Non-tech	14	77.57	11.863	3.170
ptask3	Tech	15	72.27	10.720	2.768
	Non-tech	14	74.71	20.435	5.462
ptask4	Tech	15	83.53	8.383	2.164
	Non-tech	14	78.43	20.114	5.376
ptask5	Tech	15	71.73	10.899	2.814
	Non-tech	12	68.50	11.180	3.227
ppron	Tech	15	74.40	7.405	1.912
	Non-tech	14	71.93	10.269	2.745
pAccInt	Tech	15	78.60	10.642	2.748
	Non-tech	14	78.64	15.653	4.183

pfluency	Tech	15	81.60	9.132	2.358
	Non-tech	14	81.50	13.178	3.522
pcontent	Tech	15	80.33	7.247	1.871
	Non-tech	14	82.29	12.356	3.302
pgrammar	Tech	15	74.93	9.830	2.538
	Non-tech	14	72.43	10.471	2.799
PTotal	Tech	15	77.20	7.408	1.913
	Non-tech	14	77.21	10.519	2.811

In Table 2, the highlighted Task 2 boxes (in which students answered questions about themselves) show, as an example, progress in both the treatment and the comparison groups. For this skill greater progress was made by the comparison group.

In order to show whether there are any significant relationships between compared data, the following comparisons were of interest:

1) Comparing the students' pre- and post test total scores for a broad gauge of progress over the semester. Had progress been made by everyone irrespective of which group they were in?

2) Comparing the students' pre- and post-test scores by the independent variable of which teaching group they were in. Here we were interested in comparing progress made across the two groups, comparison and treatment, to see if significantly more or less progress was made by one or other.

The above comparisons were carried out for each of the five tasks in the test and for each of the five skills under analysis (pronunciation, accent/intonation, fluency, content, and grammatical accuracy) to see if more progress had been made in certain skills rather than others and whether there were any reasonable assumptions possible as to the causes for this. Finally, in order to assess whether it was possible that factors, other than which group they were in, were influencing students' progress, an analysis of the other variables we had tested (attendance, language learning experience, ICT-use) was made.

The most useful tests for our project were those which

1) Compare the means in the pre- and posttests to see if there has been any general improvement over the semester in the whole cohort ("within-subjects" analysis).

2) Compare the means of two or more independent samples (groups of individuals), in our case the treatment group and the comparison group, to see if one group has made significantly more progress than the other ("between-subjects" analysis).

For both of the above, an independent samples *t*-test (for parametric samples) was carried out, and its non-parametric equivalent (Mann-Whitney Test). A parametric sample is one where the sample is sufficiently large (30 or more records) for one to be able to make reasonably safe generalisations. Non-parametric analysis is required for small samples such as ours where normality, or safe generalisability, cannot be assumed (i.e., the total sample size was N=29 and the independent samples were no larger than N=15). Given that the results of the two types of analysis were very similar, for the interests of brevity we have only included here the parametric data and tables for this and all remaining analyses.

Test Differences of the Mean From Two Sets of Observations From the Same Group of Individuals. Here we were interested in comparing one group's performance under one set of conditions (i.e., the pretest by overall totals, or by task and skill) with their performance under another set of conditions (i.e., the posttest, again looking at overall totals and separate tasks and skills).

Our null hypothesis is that exposure to technology in oral language development makes no difference. We can accept the alternative hypothesis (i.e., that exposure to technology, even over the course of one semester, does make a difference to progress in oral language development) if our p value in our

compared means tests for pre- and posttest scores across the two groups is less than or equal to 0.05 (i.e., is at a 95% or higher level of confidence).

Test the Degrees of Relationship or Correlation Between Variables. This test was to see if there was a positive or negative correlation between the variables attendance, language learning experience, and ICT-use, and to see how strong this correlation was. It is worth noting that correlation does not imply causation. As with any correlation, there could be a third variable which explains the association between the variables we measured. So in the case of the TOLD project, even if we showed that there was a strong positive correlation, say, between ICT-use score and progress in the Tech group, a third variable such as "positive exposure to something new" (the so-called "Hawthorne effect") may be playing a significant role, especially as these were undergraduates in their first weeks of experiencing a new multi-media lab.

For this analysis, a Pearson's Product Moment Correlation Coefficient (Pearson's *r*) Test was carried out and its non-parametric equivalent (Spearman's Rho Test). Tables 3 and 4 confirm the above conclusions by means of a paired *t*-test showing that improvement took place in both groups. This paired *t*-test shows, by group, the differences in the compared means of the pre- and posttests. A minus sign before the figure in the mean column actually denotes an increase (not a decrease) from pre- to posttest. A comparison of the *p*-values (final column) shows that the confidence levels that improvement was not down to chance are high for both groups, but highest for the Non-Tech (comparison) group. The significance values (*p* or Sig.) are both less than 0.01 (i.e., above the 99% confidence level). We therefore reject the null hypothesis (that in this case the sample means of the pre- and post-tests are equal across both groups) and accept the alternative hypothesis that improvement did take place. However, we must conclude that because the comparison group also made significant progress the improvements cannot be attributed to technology.

Table 3. Paired *t*-Test Showing, by Group, the Differences in the Compared Means of the Pre- and Posttest Totals for the Control Group (Non-Tech)

	Paired Differences					*t*	*df*	Sig. (2-tailed)
	M	*SD*	Std. Error Mean	95% Confidence Interval of the Difference				
				Upper	Lower			
Pair 1 Total % = PTotal	-13.571	8.925	2.385	-18.724	-8.419	-5.690	13	.000

Table 4. Paired *t*-Test Showing, by Group, the Differences in the Compared Means of the Pre- and Posttest Totals for the Treatment Group (Tech)

	Paired Differences					*t*	*df*	Sig. (2-tailed)
	M	*SD*	Std. Error Mean	95% Confidence Interval of the Difference				
				Upper	Lower			
Pair 1 Total % = PTotal	-6.600	8.069	2.083	-2.131	-11.069	-3.168	14	.007

Table 5. (Parametric) Independent Samples Test -- Comparing the Means of the Two Groups Against Each Other

		Levene's Test for Equality of Variances		t-Test for Equality of Means						
		F	Sig	*t*	*df*	Sig (2-tailed)	Mean Differ-ence	Std. Error Differ-ences	95% Confidence Interval of the Difference	
									Lower	Upper
Total %	Equal varliances assumed	.331	.570	1.318	27	.198	6.957	5.277	-3.871	17.785
	Equal variances not assumed			1.312	25.877	.201	6.957	5.303	-3.946	17.860
PTotal	Equal varliances assumed	1.144	.294	-.004	27	.997	-.014	3.360	-6.908	6.879
	Equal variances not assumed			-.004	23.205	.997	-.014	3.400	-7.045	7.016

In the parametric test described in Table 5 the *p* (Sig.) values highlighted for both the pre- and posttest across the two groups are greater than 0.05, that is, falling below acceptable confidence levels for us to infer a significant difference between the two groups sets of results. We, therefore, cannot accept the alternative hypothesis (H_1) that exposure to technology over the course of one semester makes a significant difference to progress in oral language development.

We must now check, task by task and skill by skill, whether there are any exceptions to this inference that might show that significant progress was made only in the treatment group. Tables 6 and 7 show the results of submitting the data to a paired samples *t*-test in which mean progress in each task and skill in the pre-post test is tested for significance. We were surprised to see that most results seemed to favour the comparison group.

Table 6. Task-by-Task and Skill-by-Skill Paired Samples *t*-Test for the Treatment Group (Tech)

		Paired Differences					*t*	*df*	Sig. (2-tailed)
		M	*SD*	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Total % - PTotal	-6.600	8.069	2.083	-11.069	-2.131	-3.168	14	.007
Pair 2	Task 2% - ptask2	-8.429	11.817	3.158	-15.252	-1.605	-2.669	13	.019
Pair 3	Task 4% - ptask4	-7.357	13.703	3.662	-15.269	.555	-2.009	13	.066
Pair 4	Task 5% - ptask5	-6.267	18.219	4.704	-16.356	3.823	-1.332	14	.204
Pair 5	Pronunciation % - ppron	-8.800	9.615	2.483	-14.125	-3.475	-3.545	14	.003
Pair 6	Accent/Intonation % - pAccInt	-10.800	10.930	2.822	-16.853	-4.747	-3.827	14	.002
Pair 7	Fluency % - pfluency	-6.000	11.458	2.958	-12.345	.345	-2.028	14	.062
Pair 8	Content % - pcontent	-6.733	13.128	3.390	-14.004	.537	-1.986	14	.067
Pair 9	Grammar % - pgrammar	-2.400	11.224	2.898	-8.615	3.815	-.828	14	.421

Table 7. Task-by-Task and Skill-by-Skill Paired Samples *t*-Test for the Control Group (Non-Tech)

		Paired Differences					*t*	*df*	Sig. (2-tailed)
		M	*SD*	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Total % - PTotal	-13.571	8.925	2.385	-18.724	-8.419	-5.690	13	.000
Pair 2	Task 2% - ptask2	-15.429	18.912	5.054	-26.348	-4.509	-3.053	13	.009
Pair 3	Task 4% - ptask4	-7.000	27.139	7.527	-23.400	9.400	-.930	2	.371
Pair 4	Task 5% - ptask5	-7.917	14.311	4.131	-17.010	1.176	-1.916	11	.082
Pair 5	Pronunciation % - ppron	-13.714	11.317	3.024	-20.248	-7.180	-4.534	13	.001
Pair 6	Accent/Intonation % - pAccInt	-21.286	14.824	3.962	-29.845	-12.726	-5.373	13	.000
Pair 7	Fluency % - pfluency	-13.143	13.132	3.510	-20.725	-5.561	-3.745	13	.002
Pair 8	Content % - pcontent	-12.143	12.347	3.300	-19.272	-5.014	-3.680	13	.003
Pair 9	Grammar % - pgrammar	-10.786	15.338	4.099	-19.642	-1.930	-2.631	13	.021

This data can also be displayed as bar charts (see Figures 3 and 4). Confidence levels are marked with asterisks: * = confidence from $p = 0.05$ (i.e., at the 95% level) up to, but not including, $p = 0.01$ (i.e., at the 99% level); ** = confidence from 99% to 100% confidence level, or $p = 0.01$ or less.

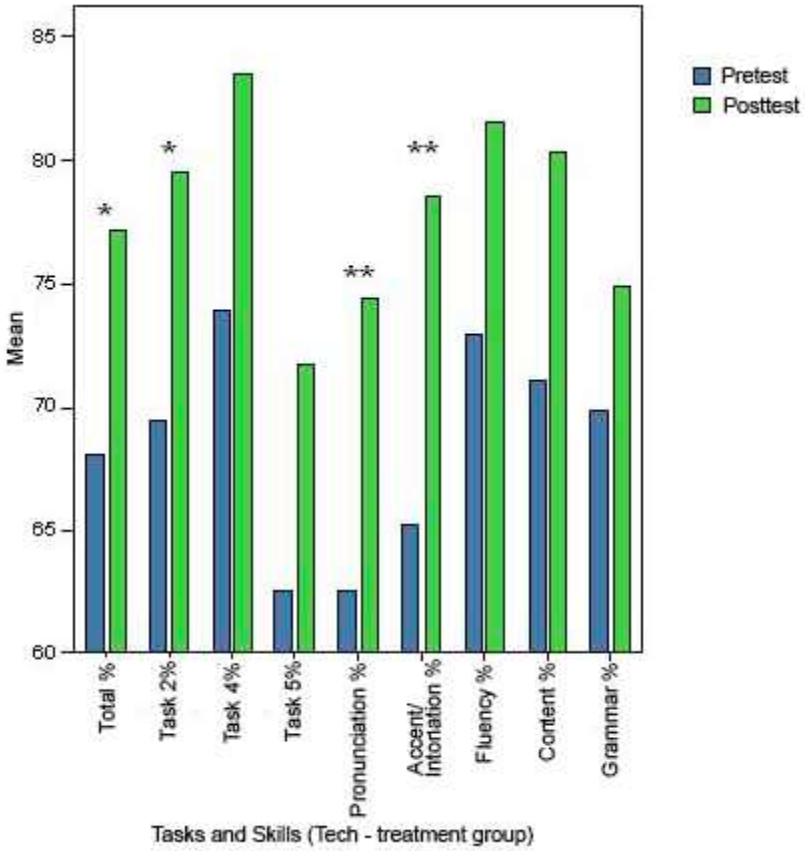

Figure 3. Progress for the Tech (treatment group) with significance values (*/**)

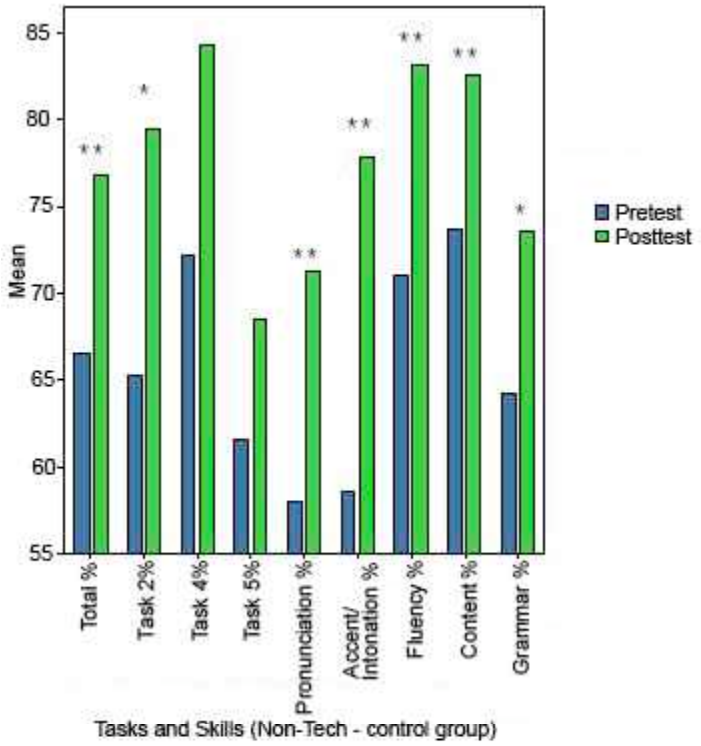

Figure 4. Progress for the Non-Tech (comparison group) with significant values (*/**)

Research Question 2: What factors may affect students' oral language development when using computers?

Tables 6 and 7 and Figure 3 and 4 reflect the progress made in both groups with the comparison group actually showing more areas (overall total, one task, and all five skills) than the treatment group (overall total, one task, and only two skills) of significant progress that cannot be attributable to chance. Neither of the remaining tasks (Task 4, the hear-and- respond comprehension; and Task 5, the résumé of a video extract) showed significant improvement for either group. Such results must be left to one side.

However, for the skills of fluency, content, and grammar, the comparison group did make significant progress while the treatment group did not. What factors were at play here? One is not so surprised that fluency and content improved more in the comparison group as more time was spent in this group on *meaningful communication.* What is more surprising is the fact that the treatment group, which had access to grammar drilling software with built-in feedback, did not progress more in the area of grammar. One would expect, however, that a group with access to CALL grammar reference and drilling tools as well as Web-based topic-based content would benefit in the longer term.

The one task that both groups had in common in terms of significant progress made was Task 2 (the five personal questions) and the skills they had in common were pronunciation and accent/intonation. Progress in Task 2, one could say, is not a surprising outcome given that students are likely to feel most comfortable when talking about familiar topics such as themselves whatever the context (technology based or not). This tends to be the area that FLAs use to start weekly conversation classes. From a skills point of view, one would expect pronunciation and accent/intonation would be the skills that a conversation class with a native speaker would develop the most, again regardless of the environment.

A configuration of our data sources might help to highlight those factors contributing to learning gains and differences across the two groups. The student logs, for example, point to the human element or lack of it as being a key factor influencing student (and to some extent staff) reactions. The qualitative analysis

below will assess this. Further research including targeted questions in a Focus group might also have got closer to the factors thrown up by the quantitative findings. A larger sample size would also have enabled multivariate analysis which would have highlighted possible other factors involved.

SUMMARY OF QUANTITATIVE FINDINGS

In answer to our first research question as to whether computer technology significantly enhances students' progress in oral language development we have drawn the following conclusions:

1) The parametric and non-parametric results were very similar for all the tests analysed for this paper, implying that the sample size for these tests (N=29) was sufficiently reliable data from which to make inferences. Nevertheless, a larger sample size would be preferable from which to make generalizable statements.

2) The language learning experience survey showed that both the treatment and the comparison groups were, when viewed as a whole, starting from the same ability/experience benchmark. This went some way towards countering the skewing effect that might have been caused by the fact that these groups were not randomly selected but self-selecting according to course.

3) Given that our data showed that the CALL environment did lead to improvement we might have been tempted initially to reject our null hypothesis (that the CALL context does not make a significant difference to oral language development). However, in the light of the between-subjects comparison (paired *t*-tests) we have seen that the difference/improvement cannot, over the short term in any case, be attributable to a computer-based environment since both groups progressed, and if anything the comparison group made more progress. We controlled as much as we could also for differences attributable to teacher input, by ensuring that the two native-speaker tutors both taught in each context and that the subject matter was similar.

4) While both groups were shown to have made significant progress over the semester, the comparison group (Non-Tech) generally made more progress than the treatment (Tech) group. This can be seen by comparing, skill by skill, the mean of each group's progress from pre to post-test (see Table 2). This shows that the comparison group made significantly more progress than the treatment group, which also made progress. The average percentage gain for the treatment group was 5.44, whereas that of the comparison group was 15.64.

5) It is now clear that two alternative hypotheses will now need to be tested in greater depth: H_1 = The CALL environment does make a positive difference in certain aspects (skills or tasks) of oral language development, such as accent and pronunciation/intonation; and H_2 = The CALL environment hinders certain aspects (skills or tasks) of oral language development.

In answer to our second research question which aimed at isolating the factors that may affect oral language development, the following conclusions were drawn:

1) According to our statistical findings, it appears that the CALL environment inhibited progress in oral language development more than the traditional non-CALL setting. While overall both groups progressed, the comparison group showed significant progress in more oral skill areas than the treatment group. Nevertheless, given the cohort size, it would be inappropriate to say conclusively that the technology hindered performance. A more likely explanation would be that more of the curriculum could be covered in the Non-Tech classes than in the Tech group because time did not need to be spent adjusting to the technology and this ensured good linguistic progress could be made by those students not using technology.

2) Factors that may have acted as inhibitors or brakes on progress may have been: the short length of the study so far combined with the fact that the treatment group lost some time actually getting used to

the computer-based environment and software, time which the other group were able to use on task; also the fact that the students and the teaching staff were novices in the CALL environment, which may have exacerbated resistance and reluctance on the part of some of them, just as much as for some it may have had a positive (Hawthorne) effect. We did not in this study isolate quantitatively which effect resulted in individual cases. However, it may be possible to map statements (positive or negative) from the student logs with individual scores to see if there was a correlation. Focus groups could have teased out the reasons on a skill by skill and task by task basis.

3) A negative correlation was found to exist between student improvement over the semester and their language learning experience in both the treatment and comparison groups. In other words both contexts most encouraged the weaker students. Given that we have shown that the ICT environment did not contribute significantly to progress, we must look elsewhere for an explanation of this finding. It might, for example, be due to the smaller group context (as opposed to the more threatening lecture/ large seminar environment) and the closer attention students received in the experimental situation, which may have raised the confidence of the weaker students.

4) No significant correlation was found to exist between ICT-use scores and progress and attendance and progress. The first variable (ICT-use) we attributed to the fact that this was a new ICT environment for most, if not all, and that new ICT skills were therefore being learned by all. The second variable (attendance) we felt was not significant given the short-term nature of the study. Both variables may well play a larger role in results over a longer-term study.

QUALITATIVE ANALYSIS

Even though the statistical evidence showed that the pedagogical benefit of using technology for oral work was unclear (thereby answering the first of our research questions), all the same, the effect of computer technology is not just measured in these terms. The views of students and staff towards the use of technology in oral language development also merit consideration to allow us to gauge the reaction of both groups to the technology and help us answer the third of our research questions. The qualitative evidence has been drawn from student and staff logs and reports as well as classroom observations.

Acceptance by Learner

The reaction of students towards the use of technology also helps us to consider some of the factors that may affect their oral development and therefore help us address the second of our research questions. In the treatment group, the initial weeks of the semester were spent learning how to use the multimedia lab technology. In the first few weeks, there were problems using the digital voice recorder. That said, however, what one student considers problematic, another student may find advantageous. This was the case with the use of headphones to record answers to a video comprehension. One student considered the headphones difficult to use, while another student welcomed using them as it meant answers to questions could be recorded without other students hearing what was said, which was less intimidating for the student. In addition, the information provided on the logs from students in the technology group made very little reference to problems using the technology and certainly the number of students who described these problems as affecting the efficacy of the classes was limited (one student). One possible explanation for this is that students saw beyond the shortcomings of the technology and felt it was making a difference in their learning because of the opportunities for practice through pronunciation exercises. This supports the findings of Burnage (2001), who discovered that the most popular CALL package among students at the University of Cambridge was a difficult-to-use DOS-based program, because students appreciated its pedagogical value (p. 169).

The statistical data have shown that students made progress in oral development in both treatment and comparison groups and that this cannot easily be attributed to the use of technology. These findings are supported by the results of the student logs, which showed that every student (in both the comparison and

treatment group) who took the time to complete a log felt that progress in oral work was made in several of the areas listed in the oral skills checklist outlined earlier in this paper. The fact that the use of technology in this experiment did not add anything significant to the learning process calls into question the need for using it -- an issue that was not lost on the teaching colleagues in their assessment of the project. It also highlights a wider issue in the area of CALL and ICT -- whether perceived pedagogical benefit of technology corresponds to the actual benefit derived. Students in the treatment group were willing to use technology and generally were very upbeat about its use. In fact, in some cases the use of computer technology was cited as the most positive aspect of the classes, making classes more interesting. This supports a widely held view in CALL research that technology motivates students: even critics of CALL technology concede this point (Ross, 1991, p. 65). That said, however, improving motivation on its own does not mean that CALL is an effective pedagogical tool. In the case of this project, CALL software and resources were used to maximise the learning opportunities for students; however students in the comparison group were exposed to similar learning opportunities in more traditional surroundings. As a result, the pedagogical impetus for using the technology was lessened.

Student logs also revealed that just under half the students in the treatment group (7 out of 15) described the group discussions and debates as the best aspect of the oral development classes. These activities were the least technological aspects of the oral development classes, although students in the treatment group used the Web to conduct research for ideas before starting the group debate. One of the main reasons given by students for this preference was that the debates allowed them to engage in meaningful and relevant, "message-orientated" communication. This feedback highlights an important question regarding the use of computer technology in oral development: Is message-orientated communication enhanced by technology? The technology may help in the development and practice of oral skills through drill and practice and pronunciation exercises (the rehearsal stage) but its role in the application of this practice (the performance stage) is not as clear. This differs from other language skills where technology can be used for both stages. Written skills, for example can be practised through CALL exercises, and it is possible to put these skills into practice through e-mail or other forms of text-based electronic communication.

Acceptance by Tutor

Discussions with the tutors showed that they were not opposed to the technology in itself, but felt it did not always fit in with the aims of the oral classes. One colleague remarked that the use of headphones in the multimedia lab created a gap between the students and her. The staff feedback in general pointed to a dehumanisation of oral classes when technology is introduced, and this was supported from classroom observations. In the early classes, computer technology was used for many aspects of student-staff oral communication, even though everyone was situated in the same room: in those circumstances, technology was creating an artificial barrier for all concerned. Such a barrier made it difficult for the tutors to appreciate the potential of using computer technology -- a theory that is supported by Partee (1996, p. 79). These views showed that unlike students, they had more reservations about using the technology because it did not seem to contribute to the learning outcomes of the oral classes. As a result, we found the tutors reaction to be one of pragmatism, in other words, only using the technology when it makes a difference to the learning process, a view confirmed by Gillespie and Barr (2002, p. 131).

See Appendix C for a tabular comparison of the benefits of a CALL and non-CALL environment for oral language development classes drawn up as a result of our experiences with the TOLD project.

CONCLUSION

The results of this paper are inconclusive in proving whether computer technology makes a significant difference in enhancing students' oral language development. Clearly, this study was limited by time, which meant that the experiment was really only effective over a 10-week period. Conducting the research over a longer period might reveal more definitive conclusions.

Despite the limitations of the project, our findings show that one main factor affected such development. The content of the classes meant that the use of technology in this project was not always relevant in achieving learning outcomes. This may also help us to explain why the students who did not use technology seemed to perform better: the students and the tutor in the comparison group did not need to spend valuable time in class to learn how to use the technology. The time taken up learning and using the technology in the treatment group was used for valuable oral practice in the comparison group. Other possible factors may be prior use and experience of ICT and attendance, although the results of this paper do not demonstrate a significant correlation between student performance and these factors and more evidence is required. Other factors, such as student learning styles may well be a significant variable and this is being looked at as part of a longer-term study.

Despite the inconclusive findings of the paper, we have discovered that students in general welcome the use of computer technology to enhance oral skills. In general, students saw the benefits of using computer technology for drilling oral skills such as pronunciation, although when it came to using these skills for meaningful communication, the traditional approach of class discussions and one to one conversations without any technology proved more successful. These findings are broadly confirmed by the reaction of the class tutors who believed that oral communication skills in a "real" context are best developed through spontaneous contact with human tutors and classmates. Under the current arrangements, technology is perhaps better kept out of free conversation and integrated more into pronunciation drilling and the development of associated skills as well as opening up the possibilities for video and audio conferencing, including tandem oral work with students from francophone universities. This would require a redefinition of the format of free oral conversation classes at Ulster.

FUTURE DEVELOPMENTS

After running the project for one semester, a number of lessons have been learned and the experiment will be repeated subject to the following changes:

1) carry out the project over a longer period (one academic year as opposed to one semester and to repeat the project with the following year's cohort of first year students);
2) re-visit the variables studied in the pilot project (ICT-use, language learning experience and attendance) to see whether over the longer period of the study their correlation with progress is more significant;
3) explore more comprehensively the influence of different variables such as gender and learning style;
4) develop individualized learning paths based on diagnostic surveys and tests, to cater for all levels of student ability;
5) modify the pedagogy used in the pilot project to make the technology more relevant to the learning experience (e.g., explore the possibilities of message-orientated communication locally, for example, turning the multimedia lab into a newsroom, and remotely, including cross-campus and international multimedia video conferencing); and
6) pay heed to a recent study by Hubbard (2004, p. 165) of over 90 research articles which sounds a note of caution regarding the overdependence on using for research data what he calls "neophytes" to the

CALL environment. He found that a high percentage of CALL research appears to have the following characteristics:

- Research subjects, whether students or teachers, are novices to CALL.
- They are also novices to the task or application under study.
- They are often studied exclusively during their initial experience.

APPENDIX A. Sample Lesson Plan With Follow-Up Activity

Activity -- Week 3
"Une loi contre les signes religieux ostensibles à l'école"

Première partie: Compréhension orale

2 écoutes successives du texte (écoute a froid, sans aucun support)
1ere écoute des questions (idem)
Re-écoute du texte
2ème écoute des questions
Ecoute découpée du texte

Deuxième partie: Production orale

3ème écoute des questions	Réponses des étudiants pour chacune des questions
1er travail sur la prononciation	Répétition de mots et d'expressions entendues
2ème travail sur la prononciation	Lecture du texte par les étudiants

Questions posées aux étudiants

1) A quelle date a été remis le rapport?
2) Qu'est-ce qui va être interdit à l'école?
3) A qui le rapport a-t-il été remis?
4) Combien y a-t-il de propositions dans le rapport?
5) Quel adjectif caractérise les signes qui vont être interdits dans les écoles?
6) De quelles religions sont les fêtes qui vont devenir jours fériés?
7) Quel code va être modifié?
8) A quelle date le Président de la République s'est-il prononcé sur le rapport?
9) Est-il favourable ou non à ce rapport?
10) Qu'en est-il du projet actuellement ?

Mots & expressions à répéter

La commission -- La commission Stasi sur la laïcité -- les signes religieux ostensibles -- le principe de laicite -- une appartenance religieuse ou politique -- la santé publique -- les jours fériés -- le Code du travail -- un jour de fête religieuse -- le credit de jours fériés -- favorable au projet -- entre les mains des parlementaires.

French Oral Skills Checklist

• Pronunciation	√
• Accent and intonation	√
• Fluency	√
• One-to-one with a French person	√
• One-to-one with an English speaker	

• Responding spontaneously in a conversation	
• Responding to visual or aural input (e.g., from TV/radio)	√
• Taking an active part in a structured group discussion	
• Taking an active part in an unstructured group discussion	
• Giving a group presentation	
• Giving a presentation on my own	

Follow-up: Listen-Watch-Respond-Discuss Activity

In another class, students in both groups watched a video extract on the theme of "laïcité." This was a news clip from a TV5 satellite broadcast, which had been recorded in analogue format (for the comparison group) and then digitised for the treatment group. Each class was divided into two groups (one group for and one group against) in order to prepare answers to these questions and to lead into a discussion of the topic of wearing of Islamic head-dress in French educational establishments.

APPENDIX B. Descriptive Statistics -- The Collated Raw Data

The following is an analysis of **individual** learning gains.

Here we present the collated raw data showing, individual scores for the predictor (or independent) variables of attendance, language learning experience, and ICT-use, and the individual scores for the outcome (or dependent) variables of the pre- and posttest, as well as the difference between the two. These are all expressed as percentages.

In the following table, *N*=30 (10 males, 20 females; however 1 female in the non-tech group was absent for the post-test, reducing *N* to 29 for most comparisons). Also,

Column 1	1 = Tech group (the treatment or test group) = 15; 2 = Non-Tech group (the comparison group) = 15
Column 2	1 = male student; 2 = female student
Column 3	Attendance score; maximum possible was 12 hours out of 12
Columns 4 & 5	Percentage scores for the identical pre- and posttests
Column 6	Denotes the difference between students' pre- and post-test scores
Column 7	The language learning experience survey score. This score represents a combination of objective measures (e.g., qualifications/ amount of time spent in the foreign country) for which we devised our own scoring system, and subjective measures (e.g., students' assessment of their own confidence and fluency levels in their language skills) for which we asked them to give a score on a scale of 1-5. The scores showed a measure of comparability between the groups in that the mean of each group's survey was nearly identical: 58.2% for the Tech group and 58% for the Non-Tech group. Of course, comparability student to student would be much more difficult to obtain.
Column 8	The ICT-use survey. This percentage score summarises a range of student responses covering their experience of using a range of applications (word-processing; spreadsheets; databases; WWW, etc.), their frequency of use of these applications, their access to computers in university and at home, and also their use of mobile phones. Most of these scores were given on a 1-5 scale.

Spreadsheet of Collated Raw Data for Full Cohort (Tech and Non-Tech), Showing *N*=29 in Most Cases (TOLD Project, Semester 2, April 2003)

Oral Group Tech (1); Non-Tech (2)	Gender M (1); F (2)	Attendance % classes (max = 12/12)	Pre-test %	Post-test %	Post less Pre %	Language learning experience survey %	ICT-use survey %
1	1	90	76	78	2	69	64
1	1	90	65	69	4	51	73
1	2	80	50	71	21	39	-
1	2	90	89	86	-3	58	48
1	2	80	67	74	7	53	71
1	2	50	45	64	19	49	
1	1	80	79	87	8	76	58
1	1	100	95	86	-9	73	71
1	2	80	70	86	16	57	
1	2	67	78	78	0	57	55
1	2	92	57	68	11	60	61
1	2	75	68	71	3	58	71
1	2	67	72	81	9	62	66
1	2	67	77	80	3	57	52
1	2	75	71	79	8	44	53
2	2	80	47	62	15	59	64
2	2	40	73	absent	-	60	-
2	2	70	76	89	13	69	-
2	1	50	61	70	9	64	-
2	1	50	77	79	2	77	-
2	2	90	63	84	21	44	68
2	2	100	81	89	8	53	50
2	1	70	74	82	8	60	-
2	1	30	50	75	25	65	-
2	2	100	69	86	17	54	59
2	2	100	64	72	8	67	53
2	2	70	72	87	15	50	66
2	1	90	77	74	-3	58	-
2	2	90	54	79	25	spoiled	-
2	1	70	44	53	9	54	46

APPENDIX C. Comparative Benefits of a CALL and Non-CALL Environment for Oral Language Development Classes

Benefits of Computer Technology in Oral Language Development	Benefits of Technology-free Oral Language Development classes
Monitoring	
• Tutors can monitor and intervene unobtrusively in students' activities in a number of ways that are not available in analogue language lab facility or traditional classroom context (keyboard, screen and mouse control).	• Monitoring and intervention is less discreet in traditional classroom context.
Pronunciation	
• Students can listen repeatedly to the recording of their own efforts against the standard of the native speaker • Students have individual access to resources on the Web, which give coaching in pronunciation, extending the boundaries of the classroom.	• Sometimes a student can go through a whole class without having spoken more than seconds/ a few minutes of French • Students cannot hear their own voice played back to compare against the native speaker or after correction • The tolerance threshold of the teacher is variable and can be more flexible than CALL packages.
Responding spontaneously in a conversation	
• Possibility for distance learning through computer-mediated video conferencing software, with target language institutions • Development of banks of role plays that are accessible on demand	• Traditional class lends itself better to this form of interaction
Responding to visual or aural input (eg. From TV/Radio)	
• A digital lab with streamed digital video/audio providing individual access and control of PLAY/PAUSE/ REWIND functions and the recording of student responses to stimuli or questions. Teacher can also take control of student consoles	• The traditional approach with one teacher and access to just one TV/Video/DVD player does not allow for individual control -- the advantage of this is that the teacher may not always want the students to have control.
Taking an active part in a group discussion	
• Possibility for distance learning through computer-mediated video conferencing software, with target language institutions	• Traditional classroom is better suited to this activity in the same room.
Giving a presentation	
• Advantage of a multimedia lab would mainly be for those presentations where the presenters wish for the audience to take an active part in looking at/hearing and responding to material on line • Best advantage at a distance	• More suitable where audience are in passive mode and where the presenters wish their faces to be seen by audience

NOTES

1. For example, an analysis of the titles of all of the articles that appeared in one of the leading CALL journals, *ReCALL,* between 2000 and 2004 reveals that the word "oral" appeared only twice.

2. A quasi-experimental study is different from an experimental study in that, while both have pre- and posttests and treatment and control groups, in the former there is no random assignment of subjects (see Nunan, 1992, p. 41). For this reason, too, we have used the term "comparison" group (rather than "control" group), as in our case the division into groups was not random but determined by the course and timetabling/teaching group considerations.

3. Project documentation such as the pre- and posttests, the surveys, and some lesson plans are available online at http://www.arts.ulster.ac.uk/lanlit/french/research/told/

ABOUT THE AUTHORS

David Barr teaches French at the University of Ulster. He completed his PhD in the area of computer-based language-learning environments and has published articles and a book in this area. He is a member of IALLT and EUROCALL and is the UK's national representative for the EUROCALL organisation.

Email: jd.barr@ulster.ac.uk

Jonathan Leakey is a Lecturer in French and German in the School of Languages and Literature at the University of Ulster, Northern Ireland. His principal research interest is computer-assisted language learning. He is currently undertaking doctoral research in the area of quantitative measurement of student progress in CALL.

Email: ja.leakey@ulster.ac.uk

Alexandre Ranchoux is currently a full-time PhD student at the University of Ulster, researching the area of pedagogy in CALL.

Email: a.ranchoux@ulster.ac.uk

REFERENCES

TellMeMore (Version 5) [Computer software]. Tempe, AZ: Auralog. Information is available at http://www.auralog.com/en/tellmemore.html

Boullier, D. (2000). La loi du support: Leçons de trois ans d'enseignement numérique à distance [The golden rules of pedagogical support: lessons learned from 3 years of distance digital teaching]. *Les cahiers numériques*, *1*(2), 145.

Burnage, G. (2001). Approaches to university network-based language learning. *ReCALL*, *13*(2), 167-178.

Dodson, C. J. (1978). *Bilingual education in Wales*. Schools Council Committee for Wales, pp. 5-11.

Driscoll, M. P. (1994). *Psychology of learning for instruction.* Needham Heights, MA: Allyn & Bacon.

Felix, U. (2001). The Web's potential for language learning: The student's perspective. *ReCALL, 13*(1), 47-58.

Gillespie, J. H., & Barr, J. D. (2002). Resistance, reluctance and radicalism: A study of staff reaction to the adoption of CALL/C&IT in modern languages departments. *ReCALL, 14*(1), 120-132.

Gillespie, J. H., & McKee, J. (1999). Resistance to CALL: Degrees of student reluctance to use CALL and ICT. *ReCALL*, *11*(1), 38-46.

Goodfellow, R., Jeffreys, I., Miles, T., & Shirra, T. (1996). Face-to-face language learning at a distance? A study of a videoconference try-out. *ReCALL*, *8*(2), 5-16.

Hincks, R. (2003). Speech technologies for pronunciation feedback and evaluation. *ReCALL, 15*(1), 3-20.

Hubbard, P. (2004, September). Some subject, treatment and data collection trends in current CALL research. In *11th International CALL Conference on CALL & Research Methodologies -- Proceedings*. Antwerp, University of Antwerp, 165-166.

Levy, M. (1997). Computer-assisted language learning: Context and conceptualization. Oxford: Clarendon.

Levy, M. (2000). Scope, goals and methods in CALL research: Wuestions of coherence and autonomy. *ReCALL*, *12*(2), 170 – 195.

Nunan, D. (1992). *Research methods in language learning.* Cambridge, England: Cambridge University Press.

Partee, M. (1996). Using e-mail, Web sites & newsgroups to enhance traditional classroom instruction. *T.H.E. Journal*, *23*(11), 79 – 82.

Reeves, T. (1993). Pseudoscience in computer based instruction: The case of learner control research. *Journal of Computer Based Instruction, 20*(2), 39-46.

Ross, M. (1991). The CHILL factor (or computer hindered language learning). *Language Learning Journal, 4,* 65-66.

Salaberry, M. R. (1996). A theoretical foundation for the development of pedagogical tasks in computer mediated communication. *CALICO Journal, 14*(1), 5-36.

Schmitt, R. (1991). Methodological weaknesses with CAI research. *Journal of Computer Based Instruction, 18*(1), 75 – 76.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

Warschauer, M. (1996). Computer-assisted language learning: An introduction. In S. Fotos (Ed.), *Multimedia language teaching* (pp. 3-20). Tokyo: Logos International.

Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching, 31,* 57-71.