“Integrating Speech Technology in Language Learning
An overview of the activities of InSTIL”

Philippe Delcloque SIG European co-ordinator
University of Abertay Dundee, Scotland
p.delcloque@mail1.tay.ac.uk

Abstract
This presentation describes the activities of a Special Interest Group which focuses on the “Integration of Speech Technology in (Language) Learning”. This SIG is unusual in the sense that it is a “bridge” between two essential, complementary traditions needed for success in the speech-enabled interface: Computer Assisted Language Learning (CALL) and Speech Science & Engineering. The history of the SIG is retraced since its foundation as CAPITAL (see later) in Edinburgh around 1994, past, present activities of the SIG are reviewed, its events, its publications, the origin and diversity of its membership, etc. Future activities are also mentioned. The success of the group has surprised its early members who thought that there were part of a very small minority group, this is a historic year for the SIG for two reasons, the first because it will write, present and publish the first “Illustrated History of Speech Technology in Language Learning” and second because it will be present for the first time at a EUROspeech conference. The group would be delighted to welcome more speech engineers and scientists in its membership.

1. A brief history: the beginnings
Around mid 1994, one of the founder members realised the need for an informal group of researchers which spans the CALL and the speech tradition. He was particularly fortunate to live and work in a city with a solid reputation in both areas, particularly speech. He visited the Centre for Speech Technology Research (CSTR) in Edinburgh where he met and exchanged ideas with the Director of the SPELL project [1]. He looked at the prototype produced by Hiller et al and expressed a wish to attempt to seek funding to pursue the project. It was on that occasion that the first mention of CAPITAL was made as a “bridge” between the CALL and Speech communities, ironically in premises in Bridge Street. The acronym stood then for Computer Assisted Pronunciation Investigation Teaching And Learning (CAPITAL), its narrow focus was precisely on the sort of technologies which had been used in the SPELL project. The first formal attempt to construct the SIG involved its first parent organisation, the American CALL association Computer Assisted Language Instruction Consortium (CALICO) [2], at the CALICO 1995 Conference in Middlebury, Vermont, an institution of international excellence in languages where French phonetician Pierre Delattre was based. CAPITAL was officially born there thanks to the enthusiasm of a few pioneering American scholars and one European, the participation of more Europeans was planned at the later conference held by the European Association for CALL (EUROCALL) [3] which took place in Valencia in September 1995 (the idea had also been discussed informally at the previous conference in Karlsruhe in 1994). One should acknowledge the financial assistance of the Scottish based Carnegie Trust which allowed the Edinburgh based academic to participate in the US conference. Little by little, the group grew and became more ambitious in its plans which had been stated at the beginning and remain present on the current web site at http://dbs.tay.ac.uk/instil/ It should also be noted that the web site will move this year we hope to http://www.instil.org/ Having become the first joint SIG of the two CALL associations, the group started to embark on the crucial phase of planning events and forging links with the Speech community.

The first event held by the group happened at EUROCALL 1998 in Leuven, Belgium, the first CAPITAL Seminar held in September 98 was a small event of quality which involved 5 speakers and an audience of around 20 [4], it broke even, was well supported by its European parent and constituted a worthwhile beginning. The day started with the report on a workshop of historical significance organised by ISCA [5] which several members of the SIG made a contribution to. The founder of CAPITAL was also invited on the Scientific Committee of STiLL 98 [6] in Marholmen. It is therefore ISCA (then ESCA) which must take the credit for “building the bridge”.

2. STiLL 98, a landmark
As the group writes “An Illustrated History of Speech Technology in Language Learning” in 2001, it will place much emphasis on the importance of the thematic workshop called by ESCA in 98. The gathering of some 80 CALL and Speech scholars in the tranquil setting of Marholmen was an important
landmark in the development of this relatively narrow field of endeavour. The quality of the event justified a review and its publication by InSTIL. [7] It also received a mention in the History of Computer Assisted Language Learning (see poster covering 1998 on the web site http://www.history-of-call.org/).

Before the next event was to take place the following year, the SIG went through the important process of renaming itself, a crucial symbolic gesture. It had become clear that, as could be anticipated, the large majority of the SIG members were Speech Technologists in the CALL community rather than mere pronunciation specialists. Of course, most of them retained a concern to emphasize the importance of the acquisition and learning of pronunciation and even the use of computers to teach phonetics but the initial name of the group placed too much emphasis on the pronunciation aspect, hence the change.

In addition, the SIG was conscious that the majority of speech scientists and engineers, the ISCA community are not involved in language learning research, so their interest in pronunciation matters may be relatively low. For that reason, the new acronym InSTIL was created, deliberately spelt with one L. The bracketisation of the L in Language is also a recognition of the fact that research in Speech Science and Engineering may be focused more on the interface in general, on communication and on learning than on language, it was important to send the message that members from that area of research were welcome. The SIG recognises that most of the advances of great benefit to (language) learning have been driven by preoccupations of a much larger nature, such as the needs of telephony in the global economy, defence, translation and assistance to the disabled.

The next event organised by InSTIL was the InSTIL 99 seminar in Besançon, France, a preamble to EUROCALL 99. [4] There were a number of reasons why this was a memorable event. Firstly, it was the first bilingual event organised by the group, partly as a tribute to the vibrant speech tradition in France and to one of its greatest pioneers, the late Christian Benoît [8], whom the event was dedicated to. Secondly, it was the first event to lead to a formal publication given informally at the conference and to be released as a book in the summer of 2001. [7] Thirdly, because a presentation was made by a leading Besançon based phonetician from outside the field, Prof. Gabrielle Konopczynski, the event also featured work from the Institut de la Communication Parlée (ICP) [9] in Grenoble and from Brittany. The book constitutes a very readable introduction to the foundations of this multidisciplinary field of endeavour.

By that time, agreement had been secured from ESCA to extend free membership of the SIG to members of the speech organisation. It has to be said that the penetration from ISCA is still relatively modest as yet, it is hoped that our participation in EUROspeech 2001 [10] will bring more members from the speech tradition. As event number 2 had brought a doubling in the attendance from 20 to 40, it was felt that perhaps we could embark on a more ambitious symposium. At STILL 98, some discussion had taken place to see whether a sequel could be mounted somewhere in the world perhaps 4 years hence (2002), an event in the USA was perhaps envisaged. After taking a brief sounding and despite the shortage of time available to the organiser, the group decided to host InSTIL 2000 as a satellite two day Symposium before EUROCALL 2000, [11] which itself had two major speech keynote speakers, Prof. Ronald Cole from the then Centre for Speech & Language Understanding (CSLU, Oregon) [12] and the world renowned scientist, engineer, businessman and best-selling author, Raymond Kurzweil [13]. In order not to alienate our established one day event clientèle keen on Information Days designed for CALL “generalists”, we also hosted two InSTIL seminars repeated on each of the two days of the symposium. To this were added free CSLU Toolkit workshops with trainers supplied from the well-known Oregon Graduate Institute (OGI) and their research associates. Our doubling of members attending our events continued as we welcomed around 80 participants in late August 2000 in Dundee. The symposium was opened by one of the hosts of STILL 98, Prof. Bjorn Granström who was kind enough to provide the “human bridge” with the event we had modelled our symposium on. The success of InSTIL 2000 ensures that the symposia will now occur at two-year intervals in different parts of the globe. The next event is planned for San Diego, California, in late March 2002, then back to Europe, probably Venice in 2004, then in Australasia in 2006. In an attempt to match the quality of STILL 98 and mark the millennium, InSTIL 2000 [14] included some prestigious keynote speakers such Prof. Dominic Massaro from the University of Santa Cruz, Professor Eric Keller from the University of Lausanne and the main conference, EUROCALL 2000, included the world renowned inventor and writer Raymond Kurzweil [13]. Keller presented with his research and life partner, Brigitte Zellner-Keller [15], whose work is widely acknowledged by the group in the area of Speech Synthesis in Language Learning.
3. The current activities of the SIG: Events and Publications

There are a number of exciting developments for the group despite the fact that InSTIL 2001 [4] is likely to return to a gathering of more modest proportions. Publications will include the proceedings of InSTIL 99 and InSTIL 2000 following a careful editing, the first history of the field shown as an exhibition first at InSTIL 2001 and EUROCALL 2001 [16], and then at EUROSPEECH 2001 in Alborg and the completion of a manuscript containing a significant number of invited articles by leading researchers and practitioners. This has been commissioned by the established CALL publisher from Lisse, Holland, Swets & Zeitlinger.

InSTIL 2001, a pre-conference gathering of EUROCALL 2001 in Nijmegen, includes 3 elements:

In the background will be the free historical exhibition, in the foreground, a commercial exhibition of hardware, software, courseware, publications and prototypes and finally our usual fare one day event consisting of thematic presentations. The talks this year include Prof. Jared Bernstein on speech recognition and testing [17], Prof. Garry Molholt on segmental and supra-segmental visualisation [18], Julia Kommissarchik on the BetterAccent [19] approach to teaching English pronunciation with innovative visualisation, Prof. Keith Ross on the virtues of Wimba [20] for using sound on the web, a team from the University of Abertay Dundee on the use of speech enabled (recognition and synthesis) in language learning games (see Delclouque, Toche and Fiandino article in this volume) and Ron Cole who will give an update on the latest work done by the Centre for Speech & Language Research (CSLR) [21]. It is also at InSTIL 2001 that the inaugural InSTIL Prize for Outstanding Lifetime Contribution to the Integration of Speech Technology in Language Learning will be presented to Professor Cole on behalf of the CSLU team who created and integrated BALDI, namely Profs. Michael Cohen and Dominic Massaro. Dominic Massaro will formally receive the prize also at the EUROSPEECH Conference in Aalborg a few days later. A second prize will be awarded in 2002 at the InSTIL 2002 Symposium, this will be the InSTIL Christian Benoît Prize for Most Innovative Work by a Newcomer to STiLL, this is in fact the second prize dedicated to the memory of Benoît, the main and generous prize is awarded with a larger remit by a charity set-up by former colleagues. There has been full consultation and agreement between the two organisations.

The prizes were set-up partly due to a funding application to the European Year of Languages initiative. They will operate bi-annually in alternation, in odd years, the lifetime prize will be awarded, in even years the newcomer prize will be presented. We are hopeful of finding commercial sponsors for the said prizes.

Beyond this, starting in 2002, the group will get involved in the gradual construction of a web-database of useful materials in the domain.

2001 is a year of new beginnings, already named the European Year of Languages, it may now also be remembered as the year of Speech Technology in Language Learning.

This article marks another landmark too, the SIG’s participation at its first EUROSPEECH conference. Jared Bernstein has agreed to join me in early September in presenting the rest of the SIG’s activities and retracing more generally the history of speech in CALL which he of course made a strong pioneering contribution too. The SIG will also wish to form close links with other ISCA groups in particular the EduSIG. [22].

Conclusion

As we close, we should not be afraid of stating the obvious, speech is a vital area of the learning interface in general, (Massaro) [23], it is primary in language acquisition and learning. It must be a paradox that speech within CALL should be regarded as a minority area. The growing interest in speech was demonstrated this year when our first parent CALICO devoted a panel discussion to Speech-enabled CALL [2], many of the great pioneers have been long-term members of the association. The large attendance, the vibrancy of the event and the pertinence of the questions demonstrated that speech will now join the mainstream as the silent web interface becomes more speech-enabled, some of the great challenges of speech-enablement will be resolved within the decade according to visionary Ray Kurzweil, many already have been but one should not expect too much of a machine! Humans are more than capable of misunderstandings. It is hard to see how machines could do better in this area! Machines, on the other hand, don’t suffer from distractions and are good at disambiguation so who knows what the future may hold.

InSTIL hopes that, through a closer integration with speech science and engineering, scholars will be able to push the envelope and help ensure that the strategic
role of effective, interactive, intelligent CALL driven learning will meet the undeniable language training opportunities of the 21st century.

References and Web Sites


See past events.


