# Language and Text-to-Speech technologies for highly accessible language & culture learning

*Gelan, A.*

Centre Applied Linguistics, Hasselt University

**Key words:** *CALL, Text-to-Speech, tutorial software, e-learning, languages*

**ICL conference 2010 - Hasselt University**

## Abstract:

*This contribution presents the results of the "Speech technology integrated learning modules for Intercultural Dialogue" project. The project objective was to increase the availability and quality of e-learning opportunities for less widely-used and less taught European languages using a user-friendly and highly accessible learning environment. The integration of new Text-to-Speech developments into web-based authoring software for tutorial CALL had a double goal: on the one hand increase the accessibility of e-learning packages, also for learners having difficulty reading (e.g. dyslexic learners) or preferring auditory learning; on the other hand exploiting some didactic possibilities of this technology.*

# 1   Introduction

## 1.1  *European policies on languages & cultures*

Multilingualism and linguistic diversity are one of the pillars of the European Union. The European Union considers itself as "a common home in which diversity is celebrated, and where our mother tongues are a source of wealth and a bridge to greater solidarity and mutual understanding (E.C. 2005)".  In March 2002, the Barcelona European Council defined the objective of each European citizen learning at least two foreign languages from a very early age. Therefore, the Commission supports universities in the development of distance learning tools and the provision of quality education through cooperation and exchange. But the range of foreign languages spoken by Europeans is narrow, limited mainly to English, French, German and Spanish. However, learning one "lingua franca" alone is not enough (E.C. 2003). Translingual/ transcultural competence in the globalised world of today is crucial. To be successful professionals, people must function comfortably in cultures other than their own (Garrett 2009). While quite a lot of e-learning materials exist for the "larger" European languages, there is an important gap as to the less-widely taught European languages (LWULT).
The TST-ID project[1] provided an answer to the need for attractive learning material for other European languages such as Romanian, Slovak, Turkish and Dutch. Moreover, this e-learning material had to be accessible to as many potential learners as possible, also to disadvantaged users, traditionally excluded from learning opportunities and depending on the help of others.

---

[1] This project was funded under the Lifelong Learning Programme of the European Commission, Transversal programme, Key Activity 2 Languages.

1(8)

### 1.2   Potential of e-learning for foreign languages & cultures - CALL

Over the years, language teaching has become ever more skill-based and less knowledge-based. While behaviouristic approaches based on drill & practise dominated in the years 60 and 70, the communicative approach in language learning won ground as from the 70s and 80s (Davies 2007, 2005). Communicative language teaching put authentic communication first; not the theoretical foundations of a language were important but the uses of the language that learners can make in diverse communicative situations. Today, integrated approaches to language learning are adopted. Since real-life communication requires putting into practise a combination of language skills (listening, reading, speaking, writing, interacting) and cultural skills (knowledge of culture-specific behaviour, values and customs, skills for dealing with cultural differences) and covers different language aspects (grammar, syntax, spelling, vocabulary, functional expressions…) at the same time, language is taught on the basis of explicitated communicative situations grounded in the cultural context of the foreign language.

For simulating communication in real-live situations, modern ICT are exceptionally well-suited. The first forms of Computer-Assisted Language Learning (CALL) originated in the 60s (stimulus-response learning in the language lab) and evolved together with up-coming language teaching methods and of course technological advances. The biggest revolution in CALL was caused by the advent of the internet. Some of the advantages of using the web for language learning include:

- the use of hypertext
- the link to reality and authentic communicative situations
- direct access to the culture linked to the foreign language
- the combination of (textual, visual, auditory) media enabling the integration of all language competences into one learning activity
- the anywhere-anytime access
- interactivity and the possibility of automatising specific inevitable language teaching routines (computer as tutor)
- the learner taking control over the path that he/she follows through the learning materials
- suitability for different learning preferences (e.g. visual, auditory learners),
- collaborative learning thanks to web 2.0 technologies, etc.

### 1.3   CALL today

Despite the potential of multimedia ICT for language learning, the impact of CALL in foreign language education has been rather modest so far. Several types of barriers to the integration of e-learning into language education seem to exist, in parallel to the current status of e-learning in education in general. As described by our e-readiness measurement model (Schreurs 2009), four groups of barriers to e-learning can be identified: learner characteristics (ICT skills and experience of learners, attitude, motivation), availability of technological facilities for e-learning (infrastructure, internet connectivity, flexible LMS, tracking of results, accessibility), organisational support (willingness of organisations such as schools or governments to invest financially, in teacher training and in learning time for teachers) and the process of e-learning and solutions offered (administration, didactical support, teachers' access to tools, available course contents and materials, personalisation).

According to Bax (Bax 2003), CALL hasn't reached the stage of "normalisation" yet. This is "the stage when a technology is invisible, hardly ever recognised as a technology, taken for granted in everyday life and used every day by language students and teachers as an integral part of the education process. Where in the past the expectations of what technology could do for language education were too high and exaggerated (the computer can do anything and will replace the teacher), today more realistic ideas on the role of ICT in education exist. As stated by Bax, today "most people in language education would recognise that CALL does have a relative advantage, but we are still at a stage where the majority of teachers are nervous of it" (Bax 2003).

# 2   Technologies for CALL

## 2.1   Complexity of CALL research

Being a interdisciplinary field, CALL is rather complex: it combines disciplines underlying second language learning theories and practises with technical domains such as computer science, instructional design and human-computer interaction (Hubbard 2003).
When studying the history of CALL, not only the didactical evolution in language teaching but also technological advances, their adoption by the community and by the world of education and users' technical skills are to be taken into consideration.
When designing language learning applications or tools, didactic and technological expertise need to coincide in a complementary way, aiming at the enhancement of CALL didactics and at the application of new technical possibilities to effective didactic uses.

## 2.2   Case: authoring software for tutorial CALL

Today, a very wide spectrum of tools for CALL exist, which can be classified into 3 large groups: learning content creation tools (whether specialised language programs or generic software, whether enabling content creation by teachers/trainers or by students: *see (Otto and Pusack 2009) for an interesting overview and evaluation of tools*), tools for communication and tools enabling access to and use of authentic material (Garrett 2009).

The authoring platform for web-based language learning "Lingu@Tor" belongs to the first group. It can be defined as a software package for tutorial CALL as described by Hubbard and Siskin 2004, starting from Levy's dichotomy of the computer as tool versus the computer as tutor (Leyy 1995; Levy 1997; Hubbard and Siskin 2004; Garrett 2009). As from 1998, the collaboration between a partner specialised in language didactics - the Centre of Applied Linguistics of Hasselt University – and a technological spin-off of the same university, named Brainlane, led to the development of an authoring software package for language learning, with the support of the European Union.

Lingu@Tor enables language teachers who have no knowledge of computer programming whatsoever to develop interactive, multimedia e-learning exercises and courses can be developed for various languages (UNICODE) and cultures.
In the following, we will try to apply Otto & Pusack's six characteristics of good foreign language courseware authoring tools (Otto and Pusack 2009) to the Lingu@Tor software and the e-learning modules developed in the course of the TST-ID project. Particular attention will be paid to the integration of the Text-to-Speech technology.

(1) Suitability: the Lingu@Tor program offers a wide set of 20 different exercise templates for teaching languages and cultures. Among those are some playful vocabulary training exercise types but also listening and dialoguing exercises with native speaker recordings, dialoguing exercises with recording of the own voice, clickable hotspots and full sentence writing exercises with intelligent correction. Thanks to the hypertext system, which links sets of exercises to thematic tailor-made theory pages, individual learners can choose to run through the course thematically rather than to follow the chronologic order as suggested by the course author. The student controlling the pace of learning and making choices in what and how he learns feels more confident in his learning, which has a positive influence on motivation.

In the TST-ID project, a Text-to-Speech engine with SAPI 5 voices for all (available) languages of the project was integrated into the authoring platform.
Text-to-Speech can be defined as the process of synthesising natural sounding speech from any text via special software (Kilickaya 2006).
TTS is an interesting tool for users with reduced sight or with auditory learning preferences, who can use the TTS as reading machine (Handley 2009) (listen to instructions, theory, answer possibilities), and the technology also offers many didactical benefits for foreign language learning. However, these have been not much exploited so far. One of the few CALL applications that use TTS today are talking dictionaries.
In the TST-ID project, TTS synthesis is additionally used as a dialogue simulator (several male and female voices per language are put at the disposal) and as pronunciation model: users can practise pronunciation (selecting words, word combinations, entire sentences and choosing the speed of speech). But an important didactic plus-value lies in the fact that proper typed texts can be spoken out loud, which in some cases reveals typing errors not perceived by the students visually. This way, together with the Lingu@Tor intelligent correction system, students get very rich spelling feedback from the system.

(2) Interactivity: Lingu@Tor enables familiar types of interaction through Multiple Choice and true-false,  cloze and several variants of (audio/image/text) drag & drop exercises, but also has exercise types with an intelligent correction system marking faulty segments (missing parts, typing errors, chiastic structures). The correction is based on variant lists of correct sentences for each task sentence.
It also presents features for author-generated feedback for correct answers/for errors (whether anticipated or general errors) or links to the web.
Another important form of feedback for the student is the availability of an electronic dictionary which can be tailor-made for a specific course. This system was upgraded in the course of the TST-ID project: the student selects at random a word from the texts or exercises and clicks to get access to the dictionary information (translation, audio recording, context sentences, grammatical info, antonyms, synonyms) and pronunciation by a synthetic voice. The dictionary application uses stemming (inflected words are mapped to their stem or root form), a.o. on the basis of suffix stripping.
The level of interactivity also increases thanks to the TTS technology, which allows the generation of audio from non-planned texts. It gives the student the freedom to experiment at his own pace with fragments of speech independently from target language native speakers or from a teacher.
.

The offer of interactive web-based exercise formats and TTS enables teachers to put their creative ideas into *e*-practise, to use a maximum of authentic materials (created by native speakers for an audience in the target language) and to generate varied educational contents that are motivating for their students, which is, in the end, the real challenge of CALL learning material (Garrett 2009).

(3) Media: the authoring platform supports different kinds of media such as text, pre-recorded native speaker audio or authentic mp3 files from the web, avi. or mov. video fragments, recording of proper voice, live-generated synthetic speech (using a Text-to-speech engine), still images and links to web pages.
No chatting or voice-over-ip communication has been integrated into the platform so far. Even though it is possible for the learner to record his/her own voice, no computerised correction or feedback is possible.
The training of speaking skills remains one of the most difficult parts of CALL and, considering today's evolution in Speech recognition technologies, will undoubtedly be one of the important challenges of the coming years.

(4) Record keeping: within the "Mentor on-line" environment developed specifically for Lingu@Tor courses, all results of logged in users are tracked and saved to a database. Students can obtain scorebooks; teachers/coaches can view students' performances and all exercise and student statistics (scores, time of action, time spent, copy of exercise with given answers). Thanks to the storage of the full context of the exercises with of all of the students' intermediary answers, students' real efforts can be measured from these user data, which make misinterpretation or cheating difficult.
In the asynchronous coaching zone, coaches can send additional remediation in text or audio format and assign new exercises to students. Students view the new information as soon as they login again.
So far, no progress report listing finished and remaining exercises or exercises marked for rehearsal or commented by the learner/teacher, can be presented within specific courses. Since this is a typical feature of Learning Management Systems, this feature should coincide with the embedding of Lingu@Tor course material into other existing LMS.

(5) Ease of use: when devising the authoring platform, an important prerequisite was the user-friendliness, making the platform available to teaching staff and developers having no programming skills. All components of Lingu@Tor e-courses can be created on the basis of an Office-like interface with WYSIWUG editing fields. Exported contents can be viewed with most known web browsers, upon installation of a web plug-in. A users' manual was written and is available to authors in 4 languages.

(6) Accessibility: using the web for e-learning ensure great accessibility in terms of numbers and computer skills of users and anywhere and anytime access for independent learners, as well as for learners assisted by a teacher.
One of the ambitions of the TST-ID project was to create accessibility of e-learning packages also for visually impaired learners or learners having difficulty reading. Therefore, a Text-to-Speech engine was implemented into the existing software, allowing learners to have synthetic voices pronounce all selected text fragments of the module.
To date, there has not been much dedicated research and development in curriculum-focussed design or development for dyslexic students (Greene 2006). TTS technology offers a lot of possibilities for these kinds of learners who have difficulties retaining

grapheme-phoneme links. With the TTS, they can have constant audio-visual revision, clicking on a word or word combination to hear what it sounds like. Another tool put at the disposal of these users is screen magnification.

# 3   Conclusion

## 3.1   Challenges for content creation software for language learning and for CALL

To conclude, existing tutorial software presents some interesting features for good foreign language learning with the help of the computer. Self-paced exercise-based courses are most suited for out-of-class learning and practising activities (self-study, remediation, rehearsal). Although they function well for independent e-learning, it is widely accepted that maximum return is still obtained when used in combination with more traditional teaching practises with a human tutor. Indeed, where in the past exaggerated expectations on the role of ICT in language education existed, today technology is generally considered to be only one of the factors crucial to good foreign language learning.

e-Learning contents are used for semi-autonomous learning, within structures of blended learning (with an e-coach or classical review sessions with a teacher). In such a system, the computer as tutor taking on all repetitive and mechanical teaching tasks does not replace but complements the human tutor taking on the role of learning advisor, providing more specific feedback and setting up class activities such as role-play.

Some upcoming challenges for tutorial software and for CALL in general have been cited, among which a very important one: standardisation, further embedding generated SCORM compatible course contents into the existing Learning Management Systems which are currently widely used, mostly by education institutes. These platforms can host very different types of learning contents and offer more possibilities of on-line communication. In this sense, they should also allow for the future pedagogical exploitation of current Web 2.0 possibilities for collaborative learning (e.g. Wiki, social bookmarking, Slide Share, VoIP).

Globally, an important challenge for CALL will be to study the best ways of eliminating gradually the barriers currently hindering the integration of ICT in language learning. To this end, it will have to pay particular attention to the learner, but also to available technologies, their adoption rate and possible applications. It will have its role in sensitising and correctly informing educational stakeholders, and in the further development of CALL didactics, training and tools.

## 3.2   New technologies

CALL will continue to evolve together with trends in language pedagogy and technological advances. Since the way we communicate and find and process information influences the way we learn, the evolution of these two will continue to be intrinsically linked.
In this paper, some challenging uses of TTS technology for CALL have been highlighted within the context of the European TST-ID project, addressing issues of accessibility as well as the improvement of CALL didactics. It is clear that although over the years TTS technology has faced a lot of difficulties acquiring the complexity of naturally sounding speech, it has now reached a level of readiness for CALL applications despite imperfections

remaining at the level of accuracy, naturalness and expressivity (Handley 2009). Research exploring further potential didactic uses of the technology and the evaluation of their specific plus-value is needed.

Finally, the TST-ID project hopes to contribute to the dissemination of these new functionalities to a wide audience. We believe that people should be informed about this technology, its possible uses (e.g. reading on-line newspapers), advantages and limitations. For a higher accessibility, it is important that a maximum of applications and websites offer TTS to their users.

## References:

[1]   Bax, S. (2003). "CALL--past, present and future." System 31(1): 13-28.

[2]   Davies, G. (2007, 2005). Computer Assisted Language Learning: Where are we and where are we going? UCALL Conference; E-learning and Japanese Language Education: Pedagogy and Practice. University of Ulster at Coleraine; Oxford Brookes University.

[3]   E.C. (2003). Promoting Language Learning and Linguistic Diversity: An action plan 2004-2006, Communication from the Commission to the Council, the European Parliament, the European economic and social committee and the committee of the Regions, Commission of the European Communities: 21.

[4]   E.C. (2005). A new framework Strategy for Multilingualism, Communication from the Commission to the Council, the European Parliament, the European economic and social committee and the committee of the Regions, Commission of the European Communities: 17.

[5]   Garrett, N. (2009). "Computer-Assisted Language Learning Trends and Issues Revisited: Integrating Innovation." Modern Language Journal 93: 719-740.

[6]   Greene, C. N. (2006). Computer-assisted language learning (CALL) for dyslexic students. 10th International Conference on Computers Helping People with Special Needs. Lecture Notes in Computer Science 4061., Linz, Austria, Springer Berlin / Heidelberg.

[7]   Handley, Z. (2009). "Is text-to-speech synthesis ready for use in computer-assisted language learning?" Speech Communication 51(10): 906-919.

[8]   Hubbard, P. (2003). "A Survey of Unanswered Questions in CALL." Computer Assisted Language Learning 16(2): 141 - 154.

[9]   Hubbard, P. and C. B. Siskin (2004). "Another look at tutorial CALL." Recall 16(02): 448-461.

[10] Kilickaya, F. (2006). "'Text-To-Speech Technology': What Does It Offer To Foreign Language Learners?" CALL-EJ Online 7(2).

[11] Levy, M. (1997). The tutor-tool framework. Computer-Assisted Language Learning: Context and conceptualisation, Oxford University Press.

[12] Leyy, M. (1995). Integrating CALL: the tutor and the tool. EuroCALL '95, Technology Enhanced Language Learning: Focus on Integration Valencia: Universidad Politécnica de Valencia.

[13] Luppi, E., Primiani, R., Raffaelli, C., Tibaldi, D., Violi, A.M. (2009). Net4Voice – new technologies for voice-converting in barrier-free learning environments: Development of innovative learning methodologies, experiment and results. eLearning papers 13.

[14] Otto, S. E. K. and J. P. Pusack (2009). "Computer-Assisted Language Learning Authoring Issues." Modern Language Journal 93: 784-801.

[15] Schreurs, J., Gelan A., Sammour, G. (2009). "E-Learning Readiness in Organisations – Case Healthcare." International Journal of Advanced Corporate Learning (iJAC) 2(2).

## Author(s):

Anouk Gelan, researcher
Hasselt University, Centre Applied Linguistics
Campus Diepenbeek
Agoralaan, gebouw D
3590 Diepenbeek
Belgium
anouk.gelan@uhasselt.be
www.uhasselt.be/ctl