



## Language and Text-to-Speech technologies for culture-embedded language learning

Gelan, Anouk

E-mail: [anouk.gelan@uhasselt.be](mailto:anouk.gelan@uhasselt.be)

Hasselt University (Belgium)

### Abstract

*In this paper, we will report on the results of the European project TST-ID – “Speech technology integrated learning modules for Intercultural Dialogue”. The European Union promotes language and culture learning and stresses the importance of linguistic diversity. However, the range of attractive e-learning materials is limited mainly to English, French, German and Spanish. The objective of this project was to increase the availability and quality of learning opportunities for less widely-used and less taught European languages using a barrier-free and highly accessible e-learning environment.*

*Today, a very wide spectrum of tools for CALL exist, a lot of them web-based. The internet offers a lot of possibilities for communicative culture-embedded language learning. We will discuss the tutorial software for multimedia CALL used in the TST-ID project, based on Otto & Pusack’s model of six characteristics of good foreign language courseware authoring tools. Specific functionalities such as the electronic dictionary and the intelligent correction for writing exercises are described. Special attention will be given to the Text-to-Speech technologies, which enable the conversion of (fragments of) texts into natural sounding speech. These were integrated with a double goal: on the one hand increase the accessibility of e-learning packages for visually-impaired learners, learners having difficulty reading (e.g. dyslexic learners) or having a preference for auditory learning; on the other hand exploiting some new didactic possibilities of this technology, in particular for communicative second-language learning. We will then describe the didactic methodology and contrastive approach used in the e-learning modules developed by the TST-ID project giving some concrete examples. Finally, we will highlight some interesting future developments for tutorial software and in particular the exploitation of speech technologies, that, although not much didactic use has been made of them so far, have reached a level of readiness for CALL applications.*

### 1. Introduction

#### 1.1 European language & culture policies

Multilingualism and linguistic diversity are one of the pillars of the European Union. 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. But the range of foreign languages spoken by Europeans is narrow, limited mainly to English, French, German and Spanish. 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<sup>1</sup> 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.

## **1.2 Potential of ICT for foreign language & culture learning - CALL**

Today's language teaching puts real-life communication skills first. Since communicating in a foreign language 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 at the same time different language aspects (grammar, syntax, spelling, vocabulary, functional expressions...), 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. Ever since the 60s, several forms of Computer-Assisted Language Learning (CALL) have come into being evolving together with up-coming language teaching methods and of course technological advances. But 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 direct access to reality, the target culture and authentic communicative situations
- 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
- 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.

## **2 E-Learning modules for culture-embedded language learning**

### **2.1 The technology: 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 (specialised language programs / generic software, content creation by teachers/trainers / by students), 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", developed as from 1998 by the Centre of Applied Linguistics of Hasselt University in collaboration with technological spin-off Brainlane, belongs to the 1st 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 (Levy 1995; Levy 1997; Hubbard and Siskin 2004; Garrett 2009).

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.

---

<sup>1</sup> This project was funded under the Lifelong Learning Programme of the European Commission, Transversal programme, Key Activity 2 Languages.



- (1) Suitability: the Lingu@Tor program offers a wide set of 20 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 one's 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.

- (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, chiasitic 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 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.

- (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.

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 for authors is available 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.

## 2.2 Didactic methodology and contrastive approach

The e-courses developed by the TST-ID project build on short dialogues and authentic texts representing real-life communicative situations in the target culture important to FL A1-A2 learners. TTS synthesis is used as a dialogue simulator (male and female synthetic voices); the learner gets audio input and at the same time scans the text together with its synchronous translation in the mother tongue. As Butzkamm (Butzkamm 2003) puts it, word-for-word translations are a good way to help learners see through unaccustomed and odd-sounding FL structures without resorting to grammatical terminology. In the same exercise, users can practise pronunciation selecting specific words, word combinations, entire sentences and choosing the speed of speech. The exercise material of the modules characterises by a didactic progression, and presents links to feedback screens, translation of new words, a dictionary and theory pages accentuating important linguistic characteristics, offering help for comprehending grammatical, lexical, syntactic aspects of linguistic input and providing communicative and cultural contrastive information.

Since the modules target adult FL learners, they focus on contrastive aspects, taking into account positive and negative transfer from the learner's mother tongue to the target language. As an example, not the common word order patterns of French and Dutch are emphasised but rather those divergent



patterns that have proven to cause difficulties for French-speaking learners of Dutch. In order to avoid misunderstandings or superficial grasp of meaning, all activity instructions as well as linguistic and cultural explanations are presented in the mother tongue. Another advantage of using the mother tongue is the possibility to use richer, more authentic texts sooner and the fact that more complex linguistic structures don't always need to be avoided.

Moreover, the modules deal with cultural elements (sometimes influencing language use) in a contrastive way and incorporate these into the language learning. FL learners should also become culturally competent and understand and learn how to handle the "culturally connotative meaning of words" (Liaw 2006). A concrete example: the French use a lot of abbreviations. "Un smicard" is nominalised from SMIC ("salaire minimum interprofessionnel de croissance", a specifically French concept) and refers to a person living on minimum wages.

Finally, as put forward by Chapelle (Chapelle 1998) describing 7 hypotheses relevant for developing multimedia CALL, the learner needs to have several opportunities to produce target language output. This is more precisely the case in the form of writing and dialoguing exercises, for which the learner receives detailed feedback from the system allowing (or forcing) him to notice and correct errors. An important didactic plus-value of TTS in this respect 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.

### 3. Conclusion

To conclude, existing tutorial software presents some interesting features for good FL learning with the help of the computer. 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 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). 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 for semi-autonomous learning, within structures of blended learning (with an e-coach or classical feedback sessions with a teacher). In such a system, the computer as tutor taking on all repetitive and mechanical teaching tasks complements the human tutor taking on the role of 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 LMS 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).

In this paper, some challenging uses of TTS technology for CALL have also been highlighted, 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.



## References

- [1] Butzkamm, W. (2003). "We only learn language once. The role of the mother tongue in FL classrooms: death of a dogma." Language Learning Journal 28(1): 29-39.
- [2] Chapelle, C. (1998). "Multimedia CALL: lessons to be learned from research on instructed SLA." Language Learning & Technology 2(1): 22-34.
- [3] 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.
- [4] Garrett, N. (2009). "Computer-Assisted Language Learning Trends and Issues Revisited: Integrating Innovation." Modern Language Journal 93: 719-740.
- [5] 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.
- [6] Handley, Z. (2009). "Is text-to-speech synthesis ready for use in computer-assisted language learning?" Speech Communication 51(10): 906-919.
- [7] Hubbard, P. and C. B. Siskin (2004). "Another look at tutorial CALL." Recall 16(02): 448-461.
- [8] Kilickaya, F. (2006). "'Text-To-Speech Technology': What Does It Offer To Foreign Language Learners?" CALL-EJ Online 7(2).
- [9] Levy, M. (1997). The tutor-tool framework. Computer-Assisted Language Learning: Context and conceptualisation, Oxford University Press.
- [10] Leyy, M. (1995). Integrating CALL: the tutor and the tool. EuroCALL '95, Technology Enhanced Language Learning: Focus on Integration Valencia: Universidad Politécnic de Valencia.
- [11] Liaw, M. L. (2006). "E-learning and the development of intercultural competence." Language Learning & Technology 10(3): 49-64.
- [12] 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.
- [13] Otto, S. E. K. and J. P. Pusack (2009). "Computer-Assisted Language Learning Authoring Issues." Modern Language Journal 93: 784-801.