The role and shape of speech technologies in well-designed language learning environments.

Jozef Colpaert

IOIW, Universiteit Antwerpen, Belgium

jozef.colpaert@ua.ac.be

Abstract

No technology carries an inherent, direct, measurable and generalizable effect on learning. Nor does speech technology. Its assumed added value is not a starting point for design, but a hypothesis resulting from design. Role and shape of speech technologies are simply logical and natural consequences of the specification of an optimal language learning environment for a specific context and for specific subconscious personal goals. These ecological and psychological paradigm shifts will be illustrated by a discussion of the design concept behind a project on pronunciation training. Challenges for CALL research on this topic will be briefly discussed.

Index Terms: pronunciation training, courseware design, learning environments

1. Introduction

When we look back on thirty years or so of speech technology, we observe an impressive technological evolution, in hardware and software, in concepts, routines and models. On the other hand, we have to admit that this speech technology has not yet permeated our daily life and certainly not language education. There was indeed no pot of gold at the end of the voice rainbow. The reason for this, in our view, had less to do with technological limitations than with underestimated psychological aspects [1].

CALL is the field par excellence where speech technology could play an obvious role: on the level of the interface (navigation), for practicing oral communication (simulation of natural interaction) or for specific pronunciation training (remediation of specific topics). Despite their technological complexity and power, despite their pedagogical usefulness, despite their theoretical soundness, speech systems have not yet been able to show their effectiveness on the basis of evidence in terms of usage and satisfaction. The problem is a design problem.

2. Educational Engineering

Educational Engineering is our Instructional Design model for guiding the design, development, implementation and evaluation of educational artefacts for learning, testing and teaching. These educational artefacts can be documents, tools, content, concepts, models and solutions such as textbooks, syllabi, lesson plans, curricula, graded readers, exercises, tests, applications or electronic learning platforms.

The term engineering does not necessarily refer to technology, but it primarily denotes the typical actions we have to undertake when not enough knowledge is available for attaining our goal. Engineering is not only about solving practical problems by applying scientific knowledge, it is also about building knowledge through real-world implementations, in a systematic and verifiable way, using working hypotheses that should be empirically and theoretically validated.

Educational engineering is needed because there is not enough knowledge available for creating perfect artefacts. By its very nature, education can and will never be perfect. It will always be l’art du possible. Educational Engineering is geared towards obtaining the best possible results, applying the best possible methodologies, taking into account as many actors and factors as possible.

Educational Engineering can be considered a design-process oriented model. Generally speaking it distinguishes itself from other design-process oriented approaches on several points. First, Educational Engineering focuses on a larger process than on design alone. It also embraces and clearly specifies other stages such as Analysis, Development, Implementation and Evaluation (ADDIE). The Analysis stage focuses on the identification of elements of the learning context which are amenable to improvement, or which should be taken into account during design, but does not state anything about the eventual design. Design focuses on the conceptualization, specification and possible prototyping of educational artefacts. Result of the design stage is a mental representation, a virtual construct, a blueprint or even a metaphor. Actual development is clearly left for its proper stage, followed by Implementation and Evaluation. Educational Engineering states that more time, energy and effort should be put in Analysis and Design as these stages are crucial for the eventual quality and effect of the targeted product.

Educational Engineering is based on real-world iteration: its starting point is a concrete problem, but it does not focus on the problem alone. It focuses on how to reach an optimal solution which in most cases cannot be realized in one step, due to resistance, financial limitations, technological challenges or practical constraints. The long-term engineering process goes through a series of ADDIE lifecycles (Fig.1), and each of these cycles formulates a very precise and justifiable intermediate change on the pathway to the optimal solution. The evaluation stage of every cycle not only validates the suggested changes, but also confirms or readjusts the concept of the optimal solution along the way. This concept serves as a lighthouse for all actors involved. The lighthouse metaphor is not accidental here: a lighthouse shows direction to boats, but it never is their final destination. So will the eventual solution be different than the initially conceived optimal solution.
Contrary to Rapid Application Design (RAD) or Rapid Instructional Design (RID), Educational Engineering does not insist on much iteration during design itself, as it considers its typical real-world iteration as the most important source of information. While RAD and RID consider real-world iteration too slow and time-consuming – which in a large number of cases can be a valid argument – the goal of Educational Engineering is more than creating optimal solutions: it also intends to validate research hypotheses, build knowledge and share expertise in an academic context.

Figure 1. Engineering cycles

The optimal solution is a hypothesis, but so are the consecutive intermediate loops. Every hypothesis is based on theoretical findings, practical experience, and the outcome of previous loops. The role of theory is to feed the process with as many useful, relevant and substantiated findings and concepts as possible, in order to increase the efficiency of the process, to guarantee the effectiveness of its product and to reduce the risk of failures. In the case of language learning and teaching for example, theories to be taken into account pertain roughly speaking to the following fields: pedagogy, psychology, technology, linguistics and specific sub-disciplines such as Human-Computer Interaction (HCI), Second Language Acquisition (SLA), Computer Mediated Communication (CMC), Motivation Theory, Activity Theory and Cognitive Multimedia Theory. The integration of this theory typically happens during two stages. During the Analysis phase, the educational engineer checks whether or not enough theoretical knowledge and findings are available for carrying out the project, and (s)he has to make the inventory of all required knowledge for designing the optimal solution. Secondly, during Design, the final shape will to a great extent, but not exclusively, be determined by theory. So theory is not directly applied to nor translated into the solution (like applying 4C/ID as such in a concrete learning situation), but it serves as one of the premises of a logical reasoning that forms the hypothesis.

Educational Engineering intends to be a universally applicable model (statement to be validated), but it does not state anything about the eventual shape of the solution (as this mainly depends on the context), nor about which theories are relevant, useful and/or applicable. The product should not be evaluated as such on its features, and this for two reasons: a) a product is by definition always an intermediate solution and b) as the product will always depend on the local context. Applying the same model leads to polymorphous results. Design should not be confused with shape. While we can easily observe shape, good design often remains invisible. Because design refers to the work behind the shape.

3. Distributed Design

This optimal solution, the virtual product of the design process, will always be a hypothesis, but the proposed process itself will also remain a hypothesis to be continuously subject to theoretical and empirical validation. Seven hypotheses, paradigm shifts really, grown out of the confrontation of practical experience with theoretical findings, have led to our own specific design model:
- The ecological paradigm shift
- The psychological paradigm shift
- Focus on the process
- Distributed design
- Ontological specification
- Generic Content
- Teacher Support

We will now briefly discuss the first two paradigm shifts, leaving the others for upcoming publications.

3.1. The ecological paradigm shift

Research into the learning effect created by a single educational artefact, such as a new technology (tablet, serious game, IWB …) often leads to the No Significant Difference Syndrome (the more difficult to prove the effect, the more complex the statistics), or at least to non-generalizable results. The main tenet of Distributed Design is that the added value of a particular educational artefact is proportional to the extent to which it contributes to the creation of an optimal learning environment (OLE).

The term learning environment in its traditional acceptance refers to a collection of components such as actors (learner, teacher, parent, policy maker, content provider …), content, infrastructure, technology and models (for teaching, learning and evaluation). The Distributed Design approach defines the learning environment more as a self-regulating system, a learning ecology, where more attention goes to the interplay between the components of the environment, the context and the rationale behind its design. It focuses on the possible effect on learning of this entire ecology, and tries to research to what extent this ecology can be optimized, in other words leading to better results for all actors involved, both in quantitative and qualitative terms.

An OLE is a blueprint of an ideal learning environment which by definition will (perhaps) never exist. As already stated, its function is that of a lighthouse: it shows direction. In the same vain, an OLE should perhaps never be realized as such, but its main purpose will be to guide the decision process along the way. An OLE also has its specific scope. This scope is determined by the users of our educational artefacts: it can be a class, a grade or degree, an institution, a country or even the entire world (e.g. Open University).

An OLE cannot be realized in one step, but it should inspire small changes to be undertaken in the existing learning environment, typically every year. Every redesigned learning environment should always be seen as an instantiation of the OLE. This instantiated learning environment or ILE should be
specified in detail. The purpose of an ILE is to test a hypothesis, and after evaluation and validation, formulate a new hypothesis leading to a new ILE along the pathway to the OLE. The number of changes in the design of a new ILE, compared to the previous one, depends on available resources, on resistance to be expected, on the research-oriented nature of the activity etc. Hypotheses are based on previous experience (evaluation of previous ILEs, exchanges with colleagues worldwide …) and on theory. The reasoning leading to a new hypothesis should be based on a sound construct based on substantiated evidence. It is obvious in this respect that also the design of the OLE can and should be adjusted along the way on the basis of these intermittent evaluations.

3.2. The psychological paradigm shift

An OLE should be designed with a clear focus on a particular pedagogical goal. Pedagogical goals are mostly well documented, easy to find, explicit and detailed. Their formulation largely depends on the scope of the OLE, and they range from lesson plans over course goals (“At the end of this course you will be able to …”) and grade descriptors (French 101 or Common European Framework for Languages), to country level (official learning programmes). The term optimal learning environment refers to its very _raison d’être_, i.e. to offer the best possible guarantee that the set pedagogical goals can be realized as efficiently and effectively as possible.

However, especially in cases of lesser motivation, it is counterproductive to focus exclusively or too directly on the realization of these pedagogical goals. It is far more efficient to focus on personal goals first. Personal goals can be considered subconscious volitions that hinder or stimulate the learning process. The problem is that these goals are quite difficult to elicit [2].

The starting point – or angle of attack – of Distributed Design is the point where personal goals and pedagogical goals conflict such as in cases where students have to learn French but may not be motivated to do so, where students have to learn to be autonomous but they may prefer strong guidance, or where students should learn how to cope with chaos but they may prefer a strong structure.

Most of the effort in Distributed Design goes into trying to reconcile these conflicting goals into a strong concept. The concept underlying an eventual OLE can be expressed as a metaphor (such as a city, a space station, a forest or a power plant). This concise representation makes sure that all actors involved (designers, developers, users and stakeholders) carry more or less the same mental image. This is important for the design team, but also for the teachers and learners.

4. Designing for pronunciation training

4.1. Computer Assisted Pronunciation Training

Computer Assisted Pronunciation Training (CAPT) is a fairly recent discipline. As its history is well known by the readers of this volume, we will not try to provide another overview here, but we just wish to refer to the many publications on the topic by a.o. our eminent colleagues Catia Cucchiarini and Helmer Strik, such as [3]. There are, however, a couple of aspects that are worth mentioning at this stage. Epistemologically speaking, we see four challenges. The first one is a _technological_ one: the use of spectral analysis versus speech recognition routines. Each approach has its limitations and affordances. The second one is related to the _feedback_ problem: how to make sure that visual or auditory feedback is meaningful for the learner. Thirdly, the effect of _auditory discrimination training_ on pronunciation: it makes little sense to train pronunciation if the learner does not hear the difference between two sounds. And finally, the issue of _washback_: how can the analysis and evaluation of CAPT system usage impact on the way pronunciation is taught in language education?

4.2. The DISCO project

The DISCO project originated within the framework of the Dutch-Flemish STEVIN programme of the _Nederlandse Taalunie_ and aimed at developing and testing a prototype of an ASR-based CALL application for Dutch as a second language (DL2). Radboud University Nijmegen focused on the linguistic-didactic aspects of pronunciation, while the University of Antwerp focused on design, usage and usability issues.

While the initial design [4] reflected a more traditional user-driven and lexis/grammar-oriented approach (with the three main program components phonology – morphology – syntax), we very soon became aware that a different interface was needed in order to guarantee acceptance and willingness in the users’ mind. With respect to pronunciation, we aimed at the achievement of intelligibility, rather than accent-free pronunciation. The target user group consisted of highly educated immigrants in Flanders and the Netherlands who wanted to be able to communicate with locals in the most natural way (Fig. 2).

![Interaction example](image)

_Figure 2: Interaction example_

Our analysis in terms of personal goals ([2]) indeed pointed out that the targeted learners wanted to feel integrated and accepted as quickly as possible, without losing or betraying their own identity and culture. They also preferred a learning process that would give them a feeling of freedom, respect and
autonomy, without a school-like approach with strong guidance nor exaggerated focus on metalanguage or --again-- civic duties. The system had to be usable as a part of the learning environments at various DL2 centers in Flanders and the Netherlands [5].

The design we eventually came up with was based on the following premises: a/ not the learner, but the system has to make the choice between phonology, morphology or syntax as exercise items, based on an intelligent analysis of user behavior and performance; b/ the user has to interact with an agent or a persona in the program who is interested in the user and who wants to help him/her; c/ the user has to interact in a natural way with this agent, in a simulated branched conversation through oral selection without being aware that his/her pronunciation is constantly being analyzed; d/ only after detecting systematic pronunciation errors, the system will interrupt the interaction with the appearance of doctor Spraak (Fig. 3), who will invite the user to work on some specific listening and speaking tasks.

Figure 3: Doctor Spraak

5. Discussion

The current system affords the most advanced and up-to-date functionalities in pronunciation training. The design stage considerably simplified the interface and the system architecture in general. The goal of the project was not to deliver a market-ready product, but we would have reached this stage, and we would have been able to test it extensively, if one of the project partners had been able to deliver the system on time.

Challenges for further CAPT research include a deeper psychological study of feedback scenarios, the role of recast and auditory discrimination and a thorough analysis of user behavior and performance before formulating our next hypothesis.

6. Conclusions

Speech technology in general, and CAPT in particular, seem to have the advantage that they can be used in many language learning and teaching situations, without having to be adapted significantly to the local users and their context. This feature is quite deceptive as we state that design, meaning adapting a system construct to the ecology and the psychology of the user, not only leads to more and better use, to more learning results, but also to a less complex system architecture and a simple interface.

Having made an analysis of a target group (DL2) and of a number of similar learning contexts (university context) we were able to design a specific learning environment. The role and shape of the system we needed for pronunciation training are completely determined by that learning environment and are not transportable as such to other environments.

The role and shape of speech technologies in general should not be determined by affordances, expectations, hypes nor perceptions. They should be the result of a methodological and justifiable design process. Their eventual role and shape depend on the local context and will always be polymorphous.

In order to end on a less positive tone: what we have learned on the level of the design process is, again similar to our experience with other projects, that most of the resistance against ecological and psychological design seems to come from the actors involved themselves: project partners, learners, teachers, policy makers, publishers etc. We will come back to this issue in a forthcoming publication.

7. References