Towards a linguistically scaffolded curriculum. How can technology help?

Ton Koenraad
Maaike Hajer

Research Centre for Teaching in Multicultural Schools
Utrecht University of Applied Sciences
Archimedeslaan 16
3584 BA Utrecht
The Netherlands
lectoraatmsc@hu.nl

Geke Hootsen
Rintse van der Werf

Department of Intercultural Communication,
Faculty of Arts, Tilburg University
Warandelaan 2
5037 AB, Tilburg
The Netherlands
r.vanderwerf@uvt.nl

Towards a linguistically scaffolded curriculum. How can technology help?

Abstract

This paper reports on the rationale for and development of the CATS project[i].
The main objective of this project (2006-2007) is to develop a model for curriculum redesign to optimize the development of the academic and professional literacy skills of Dutch students in Higher Education. The theoretical underpinnings for and the developmental process of a content and language integrated approach facilitated by Information and Communication Technology (ICT) will be presented.
We describe the design of language corpus-based software tools aimed at the promotion of student autonomy in academic reading and writing skills development and the use of a streaming video application to facilitate feedback procedures to workplace related specialised discourse.
We will outline the design of the first experimental implementations in three Universities of Applied Sciences in The Netherlands. In the conclusion the implications of the interim evaluation results for the further development of the project are discussed.

1. Introduction

As reported in a number of publications (Crul & Wolff, 2003; Severiens, Wolff, and Rezai, 2006) many, especially first year, students in Dutch Higher Vocational Education appear not to have the adequate language proficiency levels in Dutch to complete theirs studies successfully and to function well as a professional. The adoption of competency-based pedagogical models in Dutch Higher Professional Education, the increased intake of students studying subject matter through a second language and the entry of new target groups such as Secondary Vocational Education students have been suggested as possible explanations for the increase in the numbers of students involved.

The limited effectiveness of the current provision of support in Dutch language proficiency and study skills for students at Hogeschool Utrecht (HU) and the need to maintain educational quality levels motivated the University Board to address this problem. The Research Centre for Teaching in Multicultural Schools at the HU Faculty of Education was invited to come up with ideas for improvements. Inspired by local implementations of content based teaching in primary and secondary schools (Hajer, 2005) and by similar models for higher education (Crandall & Kaufman, 2002) a new approach was defined. The Centre has developed plans that offer alternatives to the limitations of the current practice such as: hardly any transfer of skills due to the remedial model applied and no focus on domain specific language possible because of mixed discipline student groups. To realize improvements interventions are proposed at four levels: university administration, faculty management, lecturers and students. Key pedagogic strategies in this integrated curriculum model are the raising of awareness of students' personal proficiency levels through diagnostic procedures, definition of linguistic demands of curriculum tasks, extension of subject teachers' repertoire of related didactical skills, empowerment of student autonomy and peer tutoring and assessment. For more information see Gangaram Panday, Beijer, & Hajer (2007).

The heterogeneity of student cohorts in terms of language proficiency levels and the need for further individualisation due to the transformation to demand driven curricula call for solutions with high levels of flexibility. To cope with these demands the possibilities of ICT-support for an integral approach to language development were researched. In Beijer, Hajer, & Koenraad (2004) technologies are identified that can support the main pedagogical features of the Content Based Approach (CBA)-based model presented in this study: provision of (individualised) domain specific language input, facilities for active content processing and language production tasks involving functional focus on form and options for peer and tutor feedback.

On the basis of these ideas the CATS project was defined and received funding from the Digital University, a Dutch consortium of organisations for Higher Education.

To realise the project's goals (i.e. providing support for the objectives at student level in the wider language development project (Gangaram Panday et al., 2007: 4) run by the Centre) and related deliverables specific work packages were defined. They respectively addressed the selection of ICT-applications and the development of additional technologies, design of curriculum experiments and the production of the related materials and methodologies. Evaluation of the project and the use of the tools in the various experimental courses was planned to lead to the realisation of the key deliverable of the project: an overview of the implications of pedagogical re-engineering of (a part of) the curriculum and practical guidelines to help other organisations with the implementation.

In section 2 we will first describe the software that is being used and developed in the project and discuss the related design principles and selection criteria. Then, in section 3, we briefly outline the settings in which the software and materials were piloted. In section 4 we summarise the evaluation research design and we conclude with some reflections on the interim experiences and results in section 5.

2. Software Selection and Development

Processing of subject content can be a challenging task for students initially studying a new discipline as words used in academic (written) texts are often less frequent and less generic and therefore more difficult than those in (spoken) everyday language. Also, words in academic language often have a different or a more specific meaning than (the same) words in everyday language (Hajer et al., 1995; Schleppegrell, 2001). Students in tertiary education have to learn specific terminology and concepts to be able to complete their education. Students from ethnic minority groups and students from families with a low socio-

economic status face a double problem when they enter higher education. Their vocabulary is often smaller than that of Dutch L1 students. Therefore, they might have more difficulty understanding the reading materials. In addition, they have to learn great numbers of new words through reading those texts.

A first concern in the CATS project was therefore to develop tools that could support the efficient provision of domain specific content that meets demands like authenticity and currency (Mishan, 2005) and facilitate ways to present the learner with enhanced input e.g. by marking, modification and/or elaboration (Chapelle, 2003), and so would allow processing of subject content with a focus on meaning by providing glosses for obligatory vocabulary elements and flexible just-in-time lexical support for vocabulary acquisition in general.

To this end a number of software deliverables have been designed, developed and implemented in experimental sessions. The software applications are to a large extent inspired by work in the research project 'Models for Adaptive Second Language Acquisition' (MASLA) at Tilburg University . The project is targeted at the construction, implementation and empirical evaluation of models for computer supported second language acquisition (SLA). In these models factors such as the stable and dynamic characteristics of the users, human-computer interaction and interface design and content and situations for SLA are integrated. Applications are developed that focus on vocabulary acquisition. Students work online on lessons that are personalised and dynamically generated. A lesson is generated from a sequence of meaningful, comprehensible, texts that are enriched with meaning focussed annotations and learner tasks. For more information see Werf, Hootsen, Vermeer, & Suijkerbuijk (2005). These texts are selected from a large corpus of reading materials based on the vocabulary used in those texts. This approach is both in theory and practice very usable in designing tools for integrating content processing and language development. In CBA the focus is on domain specific meaningful input and annotations and tasks based on this input. Also, the 'open content' and automated material analysis tools using a large

general corpus of Dutch, the CELEX-corpus[1], as applied in MASLA help in building domain specific implementations. This is described in further detail in the next paragraphs.

Translating CBA principles to software requirements

The CBA characteristics and the general issues in supporting learners in a competence based curriculum have resulted in some essential requirements for the software that was developed. Next to the requirement that solutions developed in this project are sufficiently generic to make transfer to other disciplines feasible, *scalability* and *learner autonomy* are leading principles in this project and have led to the following specifications:

1. there should be a strict separation between content and support functionality based on this content;
2. users (both teachers and students) can add domain specific content, such as authentic texts;
3. for each form of learner support based on the content there must be an authoring function so that domain specific implementations can be designed. Yet, this authoring functionality should be the same over domains;
4. domain specific implementations should 'grow' both with the amount of content and teacher and student usage;
5. usable support should be reached with a minimum of time investment from both teachers and learners.

In order to meet these requirements a corpus-based approach was selected as a feasible solution. Corpus technology takes a collection of texts, a corpus, as its central unit of analysis and builds upon this corpus with analysis tools that provide information relevant for language pedagogy. Examples are frequency lists of the vocabulary used in the texts and common language patterns that can be identified in the texts. Student can 'query' a collection of texts and receive relevant examples

---

[1] http://www.ru.nl/celex/subsecs/section_source.html

of language use while carrying out pedagogic tasks such as the production of a professional document or the delivery of 'evidence based' peer feedback.

Deliverables: corpus software, annotation tools and tests

In the CATS software, corpus building functionality is the starting point for each domain specific implementation. First, users join an existing course, or create a new course. In a course, users can then upload and edit text materials, create annotation tasks for these texts, create tests, and sequence content blocks as lessons. The texts can be described using metadata field-value pairs, for example ' text type-reader' , and ' text goal-writing' . Based on these metadata fields users can create 'on-the-fly' subcorpora as a sub-selection of the complete corpus. An example would be a corpus of texts that contain only texts with a texttype 'reader', or texts that have been written by students only. Using subcorpora, users can compare language use in different text types, for example point out differences in patterns used by teachers or students, or syllabi and books. Also, the language used in a domain specific corpus can be compared to everyday language. Users can analyse these (sub)corpora using corpus linguistic technology: word frequency lists, concordancing and collocation tools.

However, a corpus and corpus inspection software alone are not enough for autonomous and effective use of these tools in a normal curriculum (Braun, 2006). Additional form and meaning focussed input, tasks and assignments have to be added in order to make instruction more effective. For this purpose the CATS project delivered an 'open annotation framework for teachers and learners'. In this framework, (parts of) corpus texts can be annotated with extra linguistic and/or contextual information. Both teachers and students can create annotated versions of corpus texts. Annotations and annotated texts have different roles in student tasks. Firstly, these annotated texts can be viewed as part of a reading exercise focussing on specific parts of the texts. Secondly, the annotations can be queried. For example, a student

working on an essay can query the annotation database for a list of examples of linguistically correct abstracts. Thirdly, students can be asked to create a specific type of annotated text themselves, focussing their attention on the structural features of a text. Both the annotations themselves and the results of student tasks based on them are valuable input to the corpus as a whole and serve as input for the creation of new content and tasks.

Teachers can create lessons that consist of sequences of authentic (annotated) texts, tests and tasks. Students can also build their own corpus by uploading their own writing materials or editing directly in an online text editor. These materials can be checked by a teacher and be published in the corpus being used in a particular course. To exploit the potential of the corpus tool even further with additional software components, a cloze test application was developed and used in combination with an online multiple choice vocabulary test module for the assessment of general and domain specific language proficiency. These also served to realise the evaluation activities as described below in the next section.

3. Experiments

Focus on reading and writing

To test the annotation and corpus inspection tools an experimental course study will be run at the Faculty of Economics & Management in Utrecht. The study will take place in February 2007 during regular class hours in two to three consecutive weeks. Students' look-up behaviour when reading obligatory reading content and their comprehension of key concepts will be checked. Additional tasks involve writing a short summary of each of the texts, writing a number of annotations themselves, and answering several comprehension questions.

The technology should also facilitate support for the development of student *writing skills*. The more so because in competency-based curricula assessment and testing methods increasingly involve the production of products that are current in the practice of the the students' future professional lives. This involves mastering the discourse of the professional community of the discipline one specialises in. A process that according to Wilkinson (2003) is slow and partly subconscious for mother tongue speakers. To develop domain-specific writing expertise sufficient input from reading of discipline specific content and adequate feedback on productive assignments are needed. For Second Language (L2) users the acquisition of this expertise is more complex. Here additional guidance in writing is needed and can lead to improved quality if closely integrated with the domain (Wilkinson, 2003).

Two contexts for piloting the corpus informed approach to the development of domain-specific writing competency have been selected. One in the Faculty of Education at the Rotterdam University of Professional Education and one at the Institute for Law at the Utrecht University of Applied Sciences. Studies at the latter organisation prepare for the qualification Bachelor of Law and lead to careers such as assistant in a solicitors' firm, court of law or community social services department. The experiment here aims to support students in learning how to write formal, professional letters. Activities involve reviewing structural elements, locating and annotating evidences of good/bad practice and providing peer feedback. Simons, Koenraad & vd Werff (2007) provide a more detailed description of these experimental designs.

Experimenting with software for orals skills

For the instrumentation of experiments with a focus on oral skills the online video application 'Digital Video for the Digital University' (DiViDU)

was selected as it offered functionalities to help raise awareness of personal language performance and could facilitate monitoring and feedback processes. The application was designed to support learning from professional practice as it forms the basis for competence-based learning. It addresses the complex problem of the absence of frequent and meaningful interaction between practical experience and theory by providing a video-based learning environment which facilitates attention to theory, peer learning and coaching in the workplace. Central to a learning activity in DiViDU is a situation in professional practice that has been recorded on video. This could be a recording made by a student of his own or somebody else's (peer, expert) functioning in a professional context. The application supports task types for learning to analyze, learning to reflect and learning to demonstrate to what extent one has mastered a professional competency. As these play a crucial role in any tertiary professional education program, they allow the tool to be used generically (Kulk, Janssen, Gielis, & Kösters, 2006).

This and the fact that the web application had been successfully used for the development of reflective practice in a variety of professional contexts, among which teacher education and dentistry, led to the adoption in the CATS project as it was expected that the educational potential could be enhanced by including a focus on the realisation of discourse in complex professional contexts. Noticing, for instance, could be stimulated with the help of analysis tasks of model performances.

The Utrecht and InHolland Faculties of Health Care have been experimenting with software that offers a focus on orals skills in the Oral Hygiene curriculum. For this, an element aimed at the development of interactional skills to promote attitudinal and behavioural change in dental hygienic habits has been selected. Students will be asked to analyse and react to recordings of model and peer sessions with patients and select fragments from personal recorded sessions that in their view demonstrate their competent professional behaviour and the related, specific linguistic aspects involved.

4. Project Evaluation

During the academic year 2006-2007, several experiments will be done to test the effectiveness and usefulness of the ICT tools that were developed. These will be measured in terms of improvement of students' language proficiency in general and in terms of academic language, and also in terms of the ability of students and teachers to use the tools and their appreciation of the tools.

Firstly, both at the beginning and at the end of the academic year, students will be given a number of tests to determine their language proficiency and their insight into their own proficiency. In addition, at the beginning of the academic year they are asked to fill in a questionnaire about their language background. The first test is a ninety-item general multiple choice vocabulary test, since vocabulary is thought to be one of the most important factors in language proficiency (see e.g. Laufer, 1992; 1997; Nation, 2001). Furthermore, students are given two cloze tests, one general and one subject-specific test. In these tests, both word knowledge and knowledge of spelling rules are tested. Finally, students are given questionnaires in which they estimate their own language proficiency. The results of these tests at the beginning of the academic year are analysed to determine which students fall within the category of low proficiency students and therefore need extra support. Furthermore, the results at the beginning of the academic year will be compared with those at the end of the academic year to test students' progress.

Secondly, to test the tools for each part of language proficiency (reading, writing, and oral proficiency), data will be collected from the various experiments described before try and find answers to the following questions:

At student level:
- Do students use the support that is offered to them?
- Do students find the materials useful and easy to use?

- Do students improve their language competencies for study and profession?
- Does the autonomy of students increase: do they have more insight in their own language proficiency and in the way they can improve this?

At the level of the teacher:
- Are teachers better able to analyse the problems students have with language and to describe the competences students need in their study and work?
- Have teachers improved their didactic skills to support students' development of language skills?
- Do teachers find the materials and tools that are offered useful and easy to use?

Though the experiments described here have not all been carried out yet, some preliminary results will be presented in the next section.

5. Discussion

Our experiences with the use of ICT in a linguistically scaffolded curriculum strengthen our motivation to further research its potential. However, several conditions have to be fulfilled to make this pioneering stage to a success. The first experiments show that preliminary activities such as student needs analysis and the specification of linguistic demands in curriculum tasks need expertise, time and attention. Also, as seen in many innovative ICT projects, the professional development of language experts and content teachers is crucial for the adoption of the new methods in the curriculum. Both faculty staff and students need training in software use and –equally important- need to understand the related pedagogical concepts and methodology. All involved need to develop additional 'linguistic' information literacy skills. In this respect involvement of content teachers in software development and related task design appears a good strategy to reach these goals. For more detailed information on the actual experiments and an analysis of the evaluation data, including user

appreciation of the software tools, see (Simons, Koenraad & van der Werf, 2007).

Material collection for corpus development has proven to be harder than expected. Publishers, NGOs and commercial firms appear very reluctant to provide content and authentic documents in digital form. This despite our written testimonies guaranteeing fair use only and the application of anonymisation protocols. Resource development so far has therefore been largely dependent on products produced by senior students at the faculties involved. But since we expect useful content to become increasingly available on the web (open source educational content, professional e-zines, community of practice portals, relevant social software content) we also plan to do some follow-up experiments using RSS-feeds and dedicated webcrawling tools for corpus development such as BootCaT (Baroni, Kilgarriff, Pomikálek, & Rychlý, 2006). The hope is that the examples of pedagogically informed corpus use as developed in the CATS project will help in building a case for future open content initiatives.

The collection and/or development of a wide range of relevant discourse and textual resources (corpora, pedagogic annotation, video recording of expert behaviour, production of tasks) is a time consuming process. The project activities related to the applications used in this project were therefore deliberately developed to allow collaboration between institutions and teams. This is crucial for the critical mass needed to develop domain specific implementations of the content based framework. Content based curriculum development and (re)use is only feasible in a well organised community. The current project partner institutions are working towards a formal agreement to this end.

References

Baroni, M., Kilgarriff, A., Pomikálek, J., and Rychlý, P. (2006). WebBootCaT: instant domain-specific corpora to support human

translators. In *Proceedings of EAMT 2006*, Oslo: 247-252. Retrieved 13 November 2006 from: http://corpora.fi.muni.cz/bootcat/publications.cgi

Beijer, J., Hajer, M. and Koenraad, A.L.M. (2004). ICT-ondersteund Taalgericht Vakonderwijs in het Hoger Onderwijs. Vormgeving en bouwstenen in kaart gebracht. Utrecht: Digitale Universiteit. Available at: http://www.edia.nl/cats/documents.php

Braun, S., Kohn, K., and Mukherjee, J. (eds) (2006). *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods.* (English Corpus Linguistics 3, ed. Thomas Kohnen and Joybrato Mukherjee) Frankfurt am Main: Peter Lang.

Chapelle, C. A. (2003). *English Language Learning and Technology*. Amsterdam – Philadelphia. John Benjamins Publishing Company

Crandall, J. and Kaufman, D. (2002). *Content based Instruction in higher education.* Maryland: TESOL.

Crul, M. and Wolff, R. (2003). *Blijvers en uitvallers in het hoger onderwijs.* Utrecht: ECHO

Gangaram Panday, R., Beijer, J., and .Hajer, M. (2007). Challenges of integrating academic and professional language proficiency in the curriculum. Experiments at the Utrecht University of Applied Sciences. In *Proceedings of the ICLHE 2006 Conference.*

Hajer, M. (2005). Taalgericht vakonderwijs: tijd voor een nieuw vijfjarenplan. *Levende Talen Tijdschrijft*, 6-1: 3-11.

Hajer, M., and Meestringa, T. (1995). *Schooltaal als struikelblok: didactische wenken voor alle docenten.* Bussum: Coutinho.

Kulk, R., Janssen, J., Gielis, A. & Scheringa, E. (2005). DiViDU - Learning from Professional Practice through Online Video. In P.

Kommers & G. Richards (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005* : 3199-3204. Chesapeake, VA: AACE. Retrieved 13 November 2006 from: http://dividu.ilo.uva.nl/docs/Learningfromprofessionalpratice.pdf

Laufer, B. (1997). The lexical plight in second language reading. Words you don't know, words you think you know and words you can't guess. In Coady,J. & Huckin,T. (eds.). *Second language vocabulary acquisition*. Cambridge, New York, Melbourne: Cambridge University Press, 20-34.

Mishan, F. (2005). *Designing Authenticity into Language Learning Material.* Bristol. Intellect

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Schleppegrell, M.J. (2001). Linguistic features of the language of schooling. *Linguistics and education* 12 (4), 431-459.

Severiens, S., Wolff, R. and Rezai, S. (2006). *Diversiteit in leergemeenschappen. Een onderzoek naar stimulerende factoren in de leeromgeving voor allochtone studenten in het hoger onderwijs.* Utrecht: ECHO

Simons, M., Koenraad, A.L.M., and Werf, R. van der (2007). Technologies for a content and language integrated approach to dropout problems in Higher Education. Paper submitted for publication in the PALC 2007 Proceedings.
Available at: http://www.edia.nl/cats/documents.php

Werf, R. van der, Hootsen, G. J., Vermeer, A., and Suijkerbuijk, B. (2005). Automated user-centered task selection and input modification, Paper presented at the 2005 International Task Based Language Teaching conference. Leuven, Belgium

Wilkinson, R. (2003). Domain-specific writing: acquiring expertise in psychologic writing. In Leeuwen, C., van &, Wilkinson, R. (Eds*) Multilingual Approaches in University Education: Challenges and Practices.* Nijmegen, Valkhof Pers.

Notes

---

[i] An English version of this Dutch acronym would be: Computer Assistance for Language competences of Students