A Bilingual Translation System in Foreign Language Teaching

Viatcheslav Yatsko, Maxim Kozlov

Computational Linguistics Laboratory
Katanov State University of Khakasia, Russia
eng@khsu.ru

Abstract

The paper suggests a classification of computer systems used in foreign language education and focuses on the description of TITE – a distributed bilingual translation system designed for out-of-class activities. Detailed description of the architecture and functions of TITE is given.
The paper emphasizes a need to develop curriculum–integrated educational computer systems and formulates some requirements for such systems.

1. Introduction

Computer technologies have had a considerable impact on foreign language learning and instruction. Many experts [1; 2; 3] point out such their advantages as 1) fast and impartial grading of students' language proficiency; 2) instant feedback and effective control over students' progress; 3) rich opportunities for interactive audio and video presentation of language materials; 4) the opportunity to provide a richly supportive language-learning environment in which students are helped individually to develop, expand, and refine their own expressive and communicative abilities in a foreign language.
Contemporary computer systems used in foreign language learning can be adjusted to meet the needs of the individual learner who can proceed at his/her own pace depending on the level of language proficiency. "Individualization" of learning process has always been considered as one of the computer's important advantages, and most of contemporary computer systems are designed to facilitate language self-study rather than to be used in classroom environment. Hence a wide spread opinion that the role of the teacher is that of facilitator and that teacher-led view of learning is being gradually replaced by a learner-centered model [4;5]
The aim of this paper is to suggest a classification of contemporary computer systems used in foreign language learning and instruction and to provide evidence that an important part of computer-assisted language instruction is computer systems that can be used in classroom to facilitate interaction between the teacher and the learner and that can be integrated into existing curricula. Special emphasis will be made on the description of TITE (Translation in Teaching English) system – a bilingual translation system developed at Computational Linguistics Laboratory (CLL), Katanov State University of Khakasia.

2. Classification of computer systems used in foreign language education

2.1. Technological criteria for classification of educational computer systems

It's obvious, that educational computer systems (ES) can be classified according to two main criteria: technological and didactical. According to a technological criterion we distinguish local and network ES. Local ES are designed for users who have access to the computer where the given ES is installed. Usually such ES are installed on home computers and are used by one or several family members for language self study. ES of this group comprise numerous software products distributed on compact discs. One of the recent (and perhaps the best) systems of this type is *Talk to me* series developed by *Auralog* company to study English, Spanish, French, German, and Italian languages [6]
Network ES are designed for all users registered in the given computer network. Basing on criteria adopted in computer science we can distinguish local-area (LAN), metropolitan (MAN), and wide-area (WAN) ES.
LAN ES are designed for a group/groups of students studying at a given educational institution. We can consider as an example a Bulletin Board System (BBS), implemented at one of Japanese Universities [7]. The system was used by a group of 10 students studying English as a foreign language. By means of BBS the students passed messages in the target language to each other and to the teacher thus discussing problems of language use, homework, and preliminary debating questions suggested by the lecturer. Main debate was conducted during in-class activities. Thus BBS was used as a tool for preparation for classroom activities.
BBS is an example of a LAN ES with an undistributed access designed to facilitate communication between the lecturer and the students during out-of-class activities. LAN ES with a distributed access allow the lecturer to organize educational process and exert effective control over students' progress. An example is PASS (Partially Automated Symmetric

Summarization) system developed at CLL [8]. PASS is a distributed system that has a teacher's application (TA) and a student's application (SA). TA module allows the lecturer to download assignments for students and get statistic data about their errors, while SA module is used by students for fulfilling lecturer's assignments and saving results of their work.

The next type is MAN ES that provide access to educational resources for students and teachers in a given city. An example is *TeleNex* network designed to provide continuous professional support for secondary level English teachers in Hong Kong [9]. *TeleNex* provides teachers with access to three databases. *TeleGram* is a database of information about English grammar and usage customized for the Hong Kong teaching context; *TeleTeach* is a database of graded teaching materials which are designed to be printed out and used in the classroom; *TELEC Secondary Learner Corpus* (*TSLC*) is a corpus of English texts of various genres (personal and business letters, compositions, examination scripts, etc.) produced by Hong Kong secondary school students who study English as a foreign language. Analyses of *TSLC* revealed interesting patterns of usage characteristic of Chinese learners, such as underuse and overuse of certain lexical units, most frequent errors of various types (lexical, collocational, syntactic). These patterns are presented in *Student's Problems* files in *TeleGram* database so that any teacher can familiarize himself/herself with most frequent errors made by students. Using references to *TeleTeach* database the teacher can at the same time get materials for correcting students' mistakes.

WAN ES comprise various resources in the Internet available for an unlimited number of users. These resources must be classified according to some didactic criteria.

2.2. Didactic criteria for classification of educational computer systems

According to didactic criteria we think it possible to distinguish two main types of ES. 1) ES designed for language self-study; 2) ES designed to facilitate student-teacher interaction. Using traditional terminology we can say that ES of the first type belong to the domain of *Computer Assisted Language Learning* (CALL), whereas systems of the second type pertain to the domain of *Computer Assisted Language Instruction* (CALI).

CALL ES can be divided into systems that involve controllable and uncontrollable language learning; ES with uncontrollable language learning can be divided into systems providing learners with a certain learning environment and systems without such environment. To ES with learning environment belong Multi-user Object Oriented Domains (the so-called MOOs) comprising various virtual objects that the learner can visit and fulfill certain actions. A typical example is SchMOOze virtual university created by a group of educators from Hunter College, City University of New York [10].

Systems without learning environment include numerous chat rooms scattered all over the Internet. Getting connected to a chat the learner can communicate with native speakers to improve his/her language proficiency.

ES with controllable language learning are represented by the mentioned above systems distributed on compact discs. All systems of this type involve assessment of student's progress at various stages of language learning.

ES in CALI we divide into integrated and non-integrated. Integrated ES are systems integrated into a curriculum or a course studied by students at an educational institution. For example PASS system described earlier was integrated into *Text Theory* course studied by foreign language learners at Russian Universities. Another integrated system developed at CLL is TITE designed for students who study theory and practice of translation according to the curriculum of *Translation Studies* speciality introduced at Russian Universities several years ago. PASS and TITE differ in one essential respect: PASS is designed for in-class activities whereas TITE – for out-of-class activities. Thus, integrated ES can be divided into two corresponding types.

Non-integrated ES do not involve integration a curriculum or academic course. An example is commercial Internet site *Englishtown* [11]. The site suggests to the user testing English language proficiency and creating a study plan (depending on the purpose of his/her language learning) as well as participating in teacher-led discussion and practice interactive lessons that focus on all basic language skills. At the end of the course learners get a diploma certified by Suffolk University. Actually *Englishtown* offers a kind of distance language coaching.

Figure 1 represents our classification of ES.

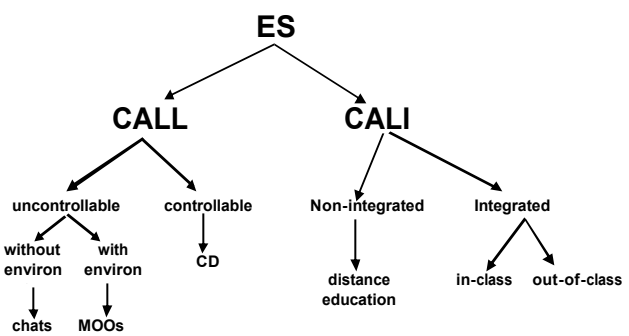

Figure 1. *Classification of computer systems used in foreign language education*

3. TITE – an integrated bilingual translation system

3.1. The idea

TITE (translation in teaching English) is a bilingual translation system to be used in foreign language teaching during out-of-class activities. General idea underlining TITE is the following. At the beginning of the term a student is given a number of texts to be translated from his native tongue into the target language or vice versa. These texts fall into two groups: limited

and unlimited. A limited text has a time limit: the student must translate it in the time allotted by the teacher (for example 15 minutes). If the student fails to keep within the time limit all the text translated by him is deleted. The translation stops as soon as he/she makes a mistake (enters an incorrect symbol). Unlimited texts are larger than limited ones and don't have a time limit; the student can save the results and proceed with the translation later. While working with TITE students are supposed to memorize texts under translation well, especially words and constructions, on which they spend much time as well as the context of their use. Since students get statistic data about their mistakes they are expected not to make most frequent mistakes again thus improving their command of the target language.

As soon as the student finishes translating the text a record is made in the log on the server so that the teacher can check student's progress. The teacher is also provided with statistical data about student's mistakes. By the end of the term the student must translate all texts assigned to him/her so as to get credit.

3.2. The architecture and algorithms of TITE

TITE is a distributed system realized in Delphi 7 that comprises Student's Application (SA) and Teacher's Application (TA) linked by an SQL server. The system works on a relational database managed by SQL Server 2000 and requires Windows 2000/2003 Server operating system.

By means of Teacher's Application the teacher can 1) assign attributes to students; 2) download texts and assign attributes to them; 3) assign texts for translation to the students; 4) The functioning of TA involves the following database operations and algorithms.

1) Download texts for translation. Using "equal texts" option (see Figure 3 below) the teacher must download two (or more) texts in different languages for translation. Usually one text is in the target language, and the other one is its equivalent in the student's native tongue, though the number of downloaded texts is unlimited. The teacher can download equivalent texts in Russian, English, Spanish, and German assigning then Russian and English texts for translation. Later he can assign for translation English and German texts from the same group.

Depending on the direction of translation one of the texts becomes a reference source, while the other one – a text to be translated. For example if the student is given the assignment to translate from Russian into English the Russian text uploads for translation, while the English one becomes a reference source (not seen by the student) with which the translation done by the student is matched.

Downloading is realized by means of "copy-paste" procedure. Files in any text format supported by Windows (.txt, .doc, .rtf, .pdf, html) can be downloaded via clipboard. As soon as the text is downloaded it is converted into internal system's format for further processing. TITE can process any texts in any language.

define translation rules; 5) get statistics about students' mistakes.

By means of Student's Application the students can: 1) get connected to the SQL server; 2) enter translation; 3) save the results of translation of unlimited texts; 4) get statistic data about his/her mistakes. Figure 2 represents TITE's generalized architecture

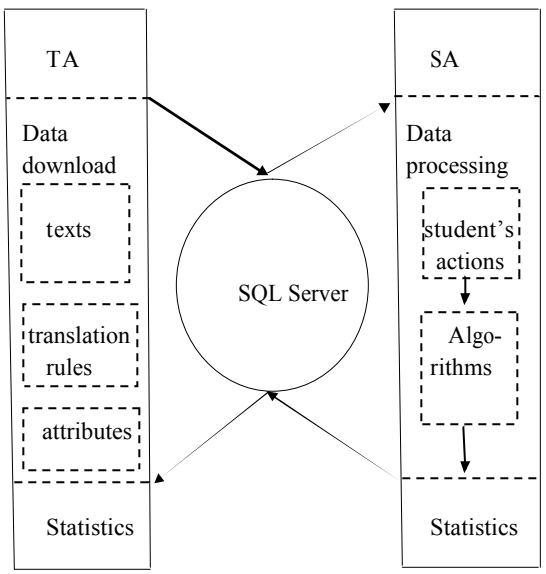

Figure 2*: The architecture of TITE*

2) Assign attributes to texts. Downloaded texts must be assigned attributes such as "title", "language", "limited", "unlimited", "direction of translation", and "translation order" Two equivalent texts must have the same title but different "language" attributes. Each text must have either "limited" or "unlimited" attribute, and direction of translation for each group of 2 texts must be specified. Apart from that the teacher can specify sequence in which texts must be translated. In this case the student can proceed to the translation of the next text, only if he has translated the previous one.

3) Assign attributes to students. Each student must be assigned such attributes as first mane, last name, group number, year of study, e-mail address. These data are required to get information about student's activities and to get statistics about his/her mistakes made during translation.

4) Define translation rules. TITE provides the teacher with a number of options to adjust the difficulty of translation to the level of students' language proficiency.

"Synonyms" option allows the student to get synonyms help. When this option is checked synonyms of the word that is being translated appear in the appropriate section of student's application (see Figure 4 below).

Using "time limit" option the teacher sets time limit for the translation of a particular text if this text has a "limited" attribute.

"Punctuation" option when checked allows the student to translate texts without paying attention to punctuation marks and symbols of text formatting such as dots, question marks, colons and semi-colons, hyphens, comas, white spaces that appear in the text automatically as soon as the preceding word has been translated. The first version of TITE didn't have this option; we decided to add it after system's testing when it turned out that students wasted much time on punctuation which isn't of great importance for their English language proficiency.

"Dictionary +" option allows the teacher to make up a list words that will appear in the text for translation, and the student won't have to translate these words.

"Dictionary –" option allows the teacher to make up a list words that won't appear in the text for translation, i.e. the rest of the words will appear in the text and the student will have to translate only the words from the list.

Two latter options are useful when the teacher wants students to memorize vocabulary on a certain topic or to drill the use of certain lexical and grammatical units, for example prepositions used after verbs.

5) Get statistics about students' mistakes. Using this option the teacher can get information about the number of mistakes made by a specific student/students/group of students in the given word. The information is presented in the table form, words arranged according to the number of mistakes. This option provides the teacher with important feedback about gaps in students' knowledge.

6) Get information about students' activities. Each connection of each student to the server is registered in a log, records from which are available for the lecturer. These records have information about the connection date, time spent for translation on this day, and texts, whose translation was finished. Such feedback makes it possible for the teacher to control student's activities.

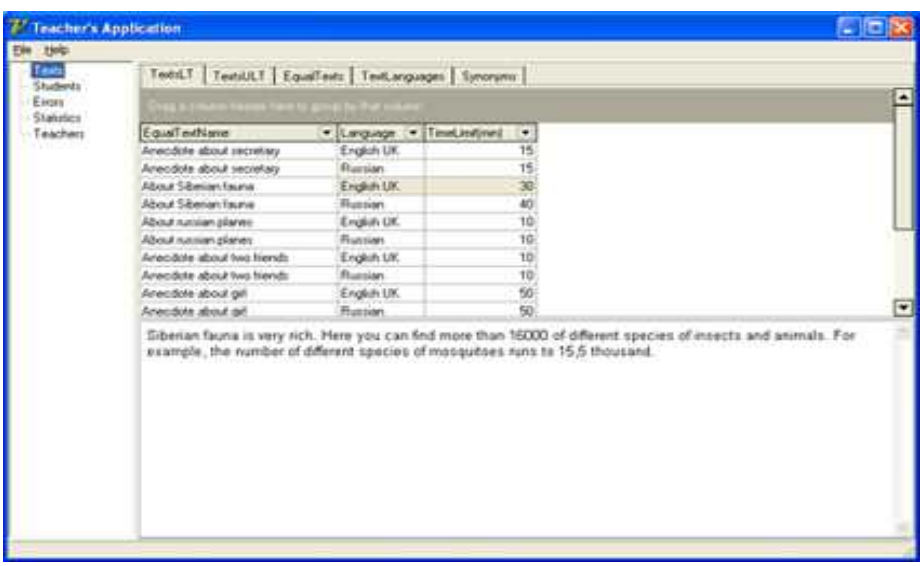

Figure 3*: A screenshot Teacher's Application*

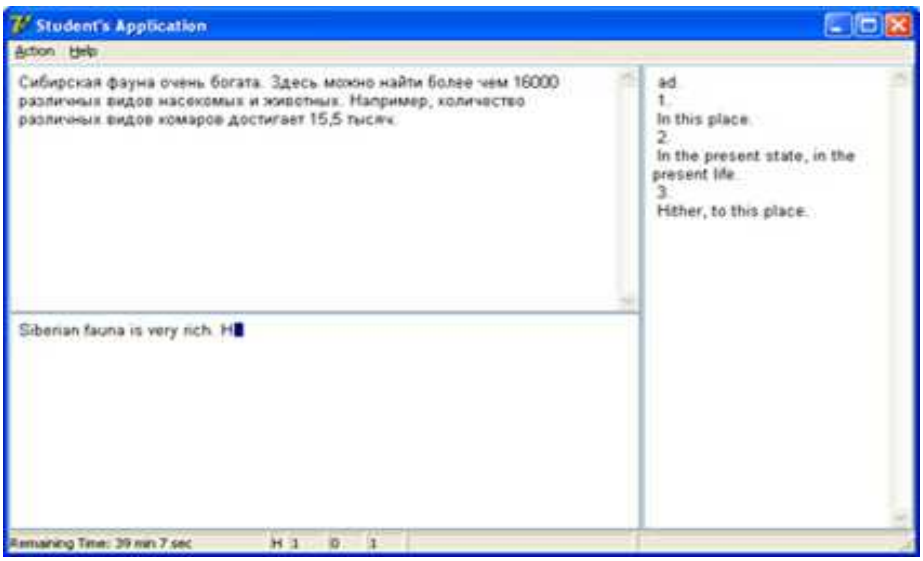

Figure 4*: A screenshot Student'ss Application*

The functioning of SA involves the following procedures and algorithms.

1) Start translation. The student presses "Start" button in the "File" menu item and the text assigned to him for translation appears in the upper section of Student's Application; the student enters his translation into the lower section. Correctness of the translation is checked by a simple matching algorithm: the first symbol entered by the student is matched with the first symbol of reference translation and so on. If the symbols are identical the student can proceed with the translation; if the symbols do not match the translation stops and the system produces a signal to inform the student about the mistake.

2) Remaining time (for texts with "limited" attribute only). As soon as the student has entered the fist symbol of his translation a reveres counting of the time left for translation starts. If the student fails to keep to the time limit all translation is deleted and he/she has to start it again.

3) Save translation (for texts with "unlimited" attribute only). The student can save his translation using the appropriate option in the "File" menu item so as to proceed with it later.

4) Statistics. The student gets statistics about mistakes in specific words in the same way as the teacher does. The difference is that the teacher has access to group statistics while the student can view his/her mistakes only.

3.3. Discussion

Currently we have passed the stage of beta-testing that was conducted on a group of 12 postgraduate students studying English as a foreign language at the Institute of Computer Science, Katanov State University of Khakasia. The testing produced the following results.

The system can be applied during out-of-class activities; our attempts to use it for in-class work failed because some students were rather annoyed with their texts deleted when the time limit was out and refused to proceed with the translation. Translation of limited as well as unlimited texts must be done in a more comfortable atmosphere, out of class. The lecturer must carefully calculate the time limit for each text taking into account the level of language proficiency of a specific student. The best result can be achieved if the text is deleted when little time (one minute or less) is left till the end of translation. In this case the student, on the one hand, is able to restore the text quickly and, on the other hand, memorize it well. Some students managed to finish translation at the third attempt but it turned out tiresome and produced negative effect.

Another difficulty is the choice of texts for translation. It's obvious that one and the same phrase or a sentence often admits of several variants of translation and it can take the student much time to find the correct variant. The teacher can avoid such translation ambiguity by using "Dictionary+" or "Dictionary –" options. Another variant is to adapt texts. Our test group translated adapted texts about programming languages, operational systems and other aspects of computer science.

It turned out that English lecturers were in no way prepared to operate the data base of TITE and to fulfill procedures described above. In this case there are two ways out: either to teach teachers or to impose database operations on the network administrator to whom the lecturers will submit all information needed by the database in familiar Word format.

4. Conclusions

In the field of foreign language education rapid development of computer technologies determined a shift to educational systems designed for language self-study rather than systems facilitating interaction between the teacher and the student. Admitting importance of the former type of ES we think that there is growing demand for curriculum-integrated systems that may be used in in-class as well as in out-of-class activities to develop language skills under teacher's control.

Basing on our experience we can formulate the following requirements for development of educational computer systems integrated into the process of foreign language teaching at university/college level.

Integrated educational systems (IES) must be distributed network systems. Network nature of IES makes it possible to involve in teaching process optimal number of students and teachers; distributed nature of such systems provides the teacher with a separate access to the database. Correspondingly, the architecture of an IES should comprise three main components: teacher's module, student's module and a database server. By means of teacher's module the teacher downloads assignments for students and operates the database; student' module is used by students to do teacher's assignments and save results of their work.

An IES must have options that allow the teacher to adjust the difficulty of assignments to the level of students language proficiency.

An indispensable feature of an IES is feedback about student's progress and activities which the teacher gets in the form of statistical data and logs on the server. Another possibility is timing of students actions.

An interesting direction of future work may be integration into IES the methodologies developed within the domain of computer learner corpora research [12].

5. References

[1] Popova, G.S., *Computer technologies in foreign language teaching*. (In Russian) http://iii04.pfo-perm.ru/Data2004/DConf04/PopovaGS.htm

[2] Garrett, N., "Computers in foreign language education: teaching, learning, and language-acquisition research", *ADFL Bulletin* 19, no. 3, p 6-12, 1988.

[3] Dobricheva L.M., *Computer technologies in foreign language teaching.* (In Russian) http://rsi.altai.fio.ru/section_b/2002/dobricheva.htm

[4] Peterson, M., "Computer assisted language learning (CALL) – moving into the network future", *Linguistic informatics – state of art and the future,* Amsterdam; Philadelphia: John Benjamins, p.248-257, 2003.

[5] Warschauer, M., Turbee, L., Roberts, B., "Computer learning networks and student empowerment", *System, vol.24, No 1, p. 1 -14, 1996*.

[6] *Auralog – innovation for language learning* . http://www.auralog.com/en/Individuals_talktome.htm

[7] Field, M.H., "Beyond the novelty – providing meaning in CALL", *Linguistic informatics – state of art and the future,* Amsterdam; Philadelphia: John Benjamins, p.258-278, 2003,.

[8] Yatsko, V., Shilov, S., Vishniakov, T., "Semi-automatic Text Summarization and Foreign Language Teaching", *Philologie im Netz , No 34, p. 48–59, 2005*. http://www.fu-berlin.de/phin/phin34/p34i.htm

[9] Allan, Q., G., "The TELEC secondary learner corpus: a resource for teacher development", *Granger S., Hung, J., Tetch-Tyson, S., ed-s, Computer learner corpora, second language acquisition and foreign language teaching* , Amsterdam; Philadelphia: John Benjamins, p.195-212, 2002

[10] Falsetti, J. *What the heck is a MOO and what's the s tory with all those c ows?* http://schmooze.hunter.cuny.edu/%7Eeflmoo/MOOarticle.html

[11] Englishtown http://www.englishtown.com

[12] Granger, S., "A bird's-eye view of learner corpus research", *Granger S., Hung, J., Tetch -Tyson, S., ed-s, Computer learner corpora, second langua ge acquisition and foreign language teaching* , Amsterdam; Philadelphia: John Benjamins, p. 3-36, 2002.