The Internet a "natural" channel for language learning

INUI Kentaro
Dept. of Computer Science
Tokyo Institute of Technology
2-12-1 O-okayama Meguro 152 Japan
e-mail:inui@cs.titech.ac.jp

The network as a motivational source for using a foreign language. Electronic networks can be useful in many ways for language learners. First of all, network facilities (e-mail, news, WWW home-pages) minimize not only the boundaries of time and space, but they also help to break communication bar-riers. They are a wonderful tool for USING a foreign language. E-mail, for example, can be used not only for interaction between teachers and students, but also for interaction among students (collaborative learning). Students can even ask for help from friends or "ex-perts" living elsewhere, on the other side of the globe.

There have been quite a few attempts to introduce these new tools into the classroom. For example, there are several well established mailing lists between Japanese and foreign schools. This allows Japanese kids to practice, let's say English, by exchanging messages with students from "abroad", chatting about their favorite topics like music, sport or any other hobby. Obviously, this kind of communication is meaningful for the student, since s/he can talk about things s/he is concerned with.

What role then can CALL system play in this new setting? Rather than trying to play the role people are very good at (answering on the fly questions on any topic, common sense reasoning, etc.), CALL system should assist people by providing the learner with information humans are generally fairly poor at . One way to help the user is by providing him with infor-mation (databases) he is looking for. For example, all language learners are concerned with lexicons. Having fabulous browsing tools, computers have a great advantage over traditional dictionaries. Also, people are not very good in explaining the contexts in which a word may be used, or in explaining the difference between two words. Last, but not least, existing NLP technology, such as parsing or machine translation, could be incorporated into the development of 'intel-ligent dictionaries'. However, before doing so, we have to consider several basic issues : what information is useful, that is, what in-formation should be provided to the learner, when and how? For example, rather than killing the user by an information overflow, -like these long list of translations that most electronic dictionaries provide, lists in which the user has to dig deeply in order to find the relevant word,- one could parametrize the level of detail, scope and grain size of trans-lations for a given text or text fragment. In sum, there should be a balance between the information provided by the system and the user's competence.

Following this line of reasoning we have started to work on a user friendly interface for a bilingual lexicon (English-Japanese). Two features of our prototype are worth mentioning: (a) the tool is implemented as a WWW application (http://tanaka-www.cs.titech.ac.jp/~inui/JLD.html), hence lexical information is pre-sented in a hypertext fashion, i.e. the user can jump from one information to another; (b) the system has a mechanism for example retrieval. The dictionary consists of a lexicon and a database of examples, the latter being a collection of collocations extracted from text corpora such as newspaper articles. When looking for a particular word, the user can input in English or Japanese any word that is semantically related with the one he is looking for. The system searches then its database, displaying those examples that exhibit this kind of relation. The examples are displayed in hyptertext format in terms of similarity, that is, the examples are grouped or ordered in terms of member-ship or proximity. Each example being linked to that part of the corpus from which it has been taken, each word in the example being linked to the corresponding dictionary entry.

The net as facility for resource sharing and development. The advantages of networks for sharing resources are obvious. For example, the CONJUGATE project between Melbourne universi-ty and Tsukuba University having started to put Ja-panese CALL software on an FTP server (http://www.intersc.tsukuba.ac.jp/conjugate.html). Al-though it is still at a preliminary stage, people from all over the world can access this information. This tendency will certainly increase with the development of platform independent languages such as JAVA. Yet, networks are also useful for resource development as they allow to reduce the gap between the developing team and the end user. The electronic dictionary of Japanese developed at Monash University is a good example.[i] The dictionary and reference tool are distributed as freeware (ftp://ftp.cc.monash.edu.au/pub/nihongo/). Note that the dictionary has currently more than 170,000 entries with English glosses and Kanji trans-criptions. What enabled this research group to build so quickly such a huge lexicon was the network. The users provided the developers with feedback, by adding new lexical items to the original dictionary. Due to this tight coupling, i.e. connection between the developers and the users, the dictionary grew

---

[i] Actually, MacJDic was developed by a graduate student from Harvard, whereas the body of the dictionary was first developed and maintained by the center administrator at Monash University. There have been also several groups of japanese teachers across the world who have contributed to the development. The final maintenance was under the control of Jim Breen.

very quickly. There is another reason to plead for this kind of feedback loop: software users are generally the ones who know best what their needs are, that is, what is useful.

Such schemes where learners directly participate in resource development allow for authentic communi-cation, hence there is a benefit for the learner, they also show the engineer the kind of information the learner is interested in, information which is usually hidden. In order to maximize the overall benefits we have decided to develop our Japanese lexicon tool as a WWW application. This allows for feedback from the users during the development phase. The next stage consist in moving away from a fully client-server set up, to a semi-stand alone implementation, based on the platform independent language, JAVA.