Fostering Language Learner Autonomy Through Adaptive Conversation Tutors

Bayan Abu Shawar[1] and Eric Atwell[2]

1. Introduction

CALL can be a route to learner autonomy, allowing students to use PC-based software to learn individually, without need of class or teacher. However, language is above all a medium for communication, implying a dialogue of two or more participants: although CALL can provide exercises and lessons in grammar, vocabulary, writing skills etc, it might seem that conversation practice still calls for a teacher, or at least fellow students.

A chatbot is a program which can chat in natural language, on a topic built into the chatbot's internal linguistic knowledge model. Many chatbots exist, with different "knowledge" programmed by the chatbot builder. A chatbot may appear to be a suitable partner for conversation practice; for example, the speak2me.net website includes Lucy, who "chats" in the language and style of a nice, polite British young lady. However, language teachers will know that conversation practice is normally on a specific topic, to learn topic-specific vocabulary and language. Lucy is nice to chat with initially, but doesn't adapt to different topics or lessons. Furthermore, Lucy helps autonomous learners of English, but not other languages.

We have developed algorithms for adapting or retraining a chatbot with a corpus, to chat in the language and topic of the training corpus. An attraction of the corpus-training approach is that in principle any corpus, in any language and on any topic, can be used; so we have gone on to test this principle, by using domain-specific corpora to train chatbots to chat on specific topics such as the Qu'ran , Computing Frequently Asked Questions, and non-English language corpora such as the Corpus of Spoken Afrikaans. Language learners and teachers have given evaluation feedback, indicating that these adaptive chatbots offer a useful autonomous alternative to traditional classroom-based conversation practice.

2. Computer Assisted Language Learning and learner autonomy

CALL can be a route to learner autonomy, allowing students to use PC-based software to learn individually, without need of class or teacher. Jung and Kim (2004) classified CALL research into three areas: studies that compare CALL with traditional methods to see if students with computer learned better and/or faster than students with traditional models; studies that look at learning strategies; and studies that examine attitudes and interaction. Typical CALL programs present a stimulus to which the learner must respond. The Stimulus may be presented in any combination of text, images, sound and video. The learner responds by typing at the keyboard,

[1] Information Technology Department, Arab Open University
e-mail: b_shawar@aou.edu.jo
[2] School of Computing, University of Leeds
e-mail: eric@comp.leeds.ac.uk

pointing and clicking within mouse, or speaking into a microphone. CALL programs can give feedback indicating whether the student's response is right or wrong (Wikipedia 2007).

However, language is above all a medium for communication, implying a dialogue of two or more participants: although CALL can provide exercises and lessons in grammar, vocabulary, writing skills etc, it might seem that conversation practice still calls for a teacher, or at least fellow students. Recently, there is growing interest in Web-CALL which is "a software that enables language teachers to produce web-based teaching materials and enables students to study them online" (Fujii et al., 2000).

3. Chatbots for fun, and for practical applications

A chatbot is a program which can pretend to "chat" with a human user in natural language, on a topic built into the chatbot's internal linguistic knowledge model. Many chatbots exist, with different "knowledge" programmed by the chatbot builders. There are several alternative names for chatbot-like systems such as: machine conversation system, virtual agent, dialogue system, and chatterbot. Brennan (2006) defined a chatbot as "an artificial construct that is designed to converse with human beings using natural language as input and output". The purpose of a chatbot system is to simulate a human conversation; or, to put it more formally, a chatbot architecture integrates a language model and computational algorithms to emulate informal chat communication between a human user and a computer using natural language.

In a survey of chatbots, (Abu Shawar and Atwell 2007) noted that initially, developers built and used chatbots for fun, and used simple keyword matching techniques to find a match of a user input, such as ELIZA (Weizenbaum 1966, 1967). The seventies and eighties, before the arrival of graphical user interfaces, saw rapid growth in text and natural-language interface research, e.g. (Cliff and Atwell 1987), (Wilensky et al 1988). Since that time, a range of new chatbot architectures have been developed, such as: MegaHAL (Hutchens 1996), CONVERSE (Batacharia etc 1999), ELIZABETH (Abu Shawar and Atwell 2002), HEXBOT (Hexbot 2004) and ALICE (ALICE 2002). ALICE is the Artificial Linguistic Internet Computer Entity, first implemented by Wallace in 1995. ALICE knowledge about English conversation patterns is stored in AIML files. AIML, or Artificial Intelligence Mark-up Language, is a derivative of Extensible Mark-up Language (XML). It was developed by Wallace and he Alicebot free a software community during 1995-2000 to enable people to input dialogue pattern knowledge into chatbots based on the ALICE open source software technology. With the improvement of data-mining and machine-learning techniques, better decision-making capabilities, availability of corpora, robust linguistic annotations/processing tools standards like XML and its applications, chatbots have become more practical, with many commercial applications (Braun 2003)

The Jabberwacky chatbot aimed to "simulate natural human chat in an interesting, entertaining and humorous manner" (Carpenter 2007). Jabberwacky learns from its previous conversation and uses contextual pattern matching techniques to select responses. Jabberwacky is not intended to do anything "useful" other than simply chatting for fun. Figure 1 presents a sample of conversation with Jabberwacky.

Jabberwacky: Do you want to know my name?
User: Yes please.
Jabberwacky: My name is Matt.
User: Are you for real?
Jabberwacky: Yes.
User: How old are you?
Jabberwacky: Thirty two.
User: And you're from ...?
Jabberwacky: I told you - from Saturn.
User: No you didn't tell me, but welcome on earth anyway.
Jabberwacky: Yes, I did. Check your own logs.
User: I'll do later, but I don't think I'll find anything.
Jabberwacky: You don't make sense.
User: And you seem to be aggressive.
Jabberwacky: I am not answering your question right now. I must go!
User: You're running away.
Jabberwacky: You know me better than I know myself.
User: Ok, stay here then.

Figure 1: A sample of chatting with Jabberwacky (taken from http://aiissues.blogspot.com/2007/01/group-of-conversation-between-me-and.html)

Other chatbot developers claim their systems are capable of more intelligent analysis of conversation: "chatbots have come a long way from their roots in systems that were more about fun, filteration or simple 'chat'. We are now approaching a time where the technologies such as Lingubot can, through extensive syntactic structures developed for natural language processing and some complex methodological data structuring, begin to display behaviour that users will interpret as understanding." (Kerly et al., 2006).

4. Chatbots for conversation practice

As the design of chatbots become increasingly sophisticated their use was adopted in education. For example, (Knill et al 2004) use a chatbot for maths education; (Kerfoot et al 2006) describe an experimental use of chatbots in training medical students. A chatbot may appear to be a suitable partner for conversation practice; for instance, the speak2me.net website includes Lucy, who "chats" in the language and style of a nice, polite British young lady. Fryer and Carpenter (2006) suggest that the chatbot could be used to practice a language. In terms of listening comprehension, Robin (2007) added that "one can imagine scenarios where learners could make use of primitive Web-based translation bots, such as Google's Language Tools or Alta Vista's Babelfish in a way that would redefine the notion of usable learner input."

Atwell and Abu Shawar (2007) examined a range of practical applications of chatbots, including their use of as a tool to learn and practice a language. Several examples cited in this paper are hosted by the Pandorabots.com website, a good example of a web-hosting service used to publish online chatbots. Pandorabots.com hosts thousands of chatbots built using the ALICE AIML (Artificial Intelligence

Mark-up Language) format. The most popular Pandorabots for the last 24 hours web-page regularly lists chatbots developed by researchers and hobbyists, and also some systems for more practical applications. For example, Cyber-Sandy and Nickie act as portals to adult-entertainment websites; Jenny introduces the English2Go.com website, and lets English language learners practise their chatting technique. The first Pandorabot chatbots were text-only: the user typed a sentence via keyboard, and then the chatbot reply appeared onscreen as text too. Now some Pandorabot chatbots incorporate speech synthesis; for example, can Jenny talk with an educated British accent, via a speech synthesis engine. However, Pandorabot chatbots cannot recognise speech: the user still has to type their input via keyboard. This is because existing Markov-model-based speech recognition is still too error-prone, and does not fit the AIML key-phrase model. Existing speech recognition systems would take a lot of time and memory trying to recognise everything in the input, even though little of this is subsequently needed by the AIML language model; and speech recognition errors may cause inappropriate AIML patterns to be matched (Atwell 2005).

Language teachers will know that conversation practice is normally on a specific topic, to learn topic-specific vocabulary and language. Lucy and Jenny can be nice to chat with initially, but doesn't adapt to different topics or lessons. Furthermore, Lucy and Jenny help autonomous learners of English, but not other languages.

5. Retraining an adaptive chatbot for new topics and languages

We have developed algorithms for adapting or retraining an AIML-based chatbot with a corpus, to chat in the language and topic of the training corpus (Abu Shawar and Atwell 2003a). An attraction of the corpus-training approach is that in principle any corpus, in any language and on any topic, can be used; so we have gone on to test this principle, by using domain-specific corpora to train chatbots to chat on specific topics such as the Qu'ran (Abu Shawar and Atwell 2005a), Computing Frequently Asked Questions (Abu Shawar and Atwell 2005b), and non-English language corpora such as the Corpus of Spoken Afrikaans (Abu Shawar and Atwell 2003b), Language learners and teachers have given evaluation feedback, indicating that these adaptive chatbots offer a useful autonomous alternative to traditional classroom-based conversation practice.

6 Evaluation of chatbots for conversation practice

Abu Shawar and Atwell (2007) discuss how we used our chatbot-retraining program read the Corpus of Spoken Afrikaans (Korpus Gesproke Afrikaans) (Van Rooy 2003) and to convert it to the AIML format files. Since the corpus does not cover topics like greetings, some manual atomic categories were added for this purpose and the default ones were generated by the program automatically. As a result two Afrikaans chatbots were generated using: Afrikaana (2002), which speaks just Afrikaans, and a bilingual version speaking English and Afrikaans, named AVRA (2002); this was inspired by our observation that the Korpus Gesproke Afrikaans actually includes some English, as Afrikaans speakers are generally bilingual and switch between languages comfortably. We mounted prototypes of the chatbots on websites using Pandorabot service, and encouraged open-ended testing and feedback from remote users in South

Africa. Unfortunately, users found that many responses were not related to the topic or nonsense. The reasons behind most of the users' feedback can be related to three issues. Firstly the dialogue corpus context does not cover a wide range of domains, so Afrikaana can only "talk about" the domain of the training corpus. Secondly, the repeated approach that we used to solve the problem of determining the pattern and the template in case of more than two speakers may lead to incoherent transcripts: if the training corpus does not have straightforward equivalents of "user" and "chatbot" then it can be non-trivial to model turn-taking correctly in Machine-Learnt AIML (Abu Shawar and Atwell 2005a). Thirdly, our machine-learnt models have not included linguistic analysis markup, such as grammatical, semantic or dialogue-act annotations (Atwell 1996, Atwell et al. 2000), as ALICE/AIML makes no use of such linguistic knowledge in generating conversation responses. However, users found it an interesting tool to practise the language and enjoyed chatting, and we concluded that even with its key-word based matching technique, a chatbot could be used as a tool for unknown languages, where "unknown" means (i) unknown to the chatbot author/developer, and/or (ii) unknown to computational linguistics, that is, where there is a shortage of existing tools to deal with the languages.

Jia (2004a) also found that with the absence of linguistic knowledge and the inability to understand users' input, many answers will be nonsense. Moreover, Jia claimed that with this key word matching, a chatbot system could not work as a teaching assistant program in foreign language learning. This was his conclusion from an experiment in China to see if a chatbot could replace a chat partner with users who learn a foreign language. Students from universities and colleges were asked to chat with ALICE, students only knew that they were chatting with a partner to help them learning English. After a short period of chatting, most of students figured out that Alice was an AI agent and not a real human. Analysis of all dialogues generated and feedback from students revealed that: 1256 users chatted with Alice, 88% of them chatted only once and did not come back to the site; the duration of chatting was short; 17% made positive comments such as: "you are so nice", "you are clever", etc, and 24% evaluated it negatively. In this respect, Jia concluded that the failure of this experiment is down to the pattern matching technique used in Alice which is based on key-word matching without any attempt to understand what is said. The topics of chatting covered every aspect in our daily life, for example: study, emotion, life, computer, free time, travel/world and job. 11.39% of students talk about English study, and exams, and 13% mentioned love, mostly students younger than 30 years old dealt with Alice as a friend rather than as a teacher, and told her some private emotional problems and experiences. Jia (2004a) concluded that "the conversational chatbot should not only work as a teacher or learning partner with rich special knowledge, but also as a dear friend who may enjoy the joy and suffer the pain of the users". After that Jia (2004b) developed an intelligent Web-Based teaching system for foreign language learning which consists of: natural language mark-up language that labels grammar elements; natural language object model in Java which represents the grammatical elements; natural language database; a communication response mechanism which considers the discourse context, the world model and the personality of the users and of the system itself.

Not all evaluations of chatbots are as positive, for example Chantarotwong (2006) reported that "responses of most chatbots are frequently predictable, redundant, lacking in personality, and having no memory of previous responses which could lead to very circular conversation." However, despite some critical findings, Fryer and Carpenter (2006) claimed that "chatbots could provide a means of language

practice for students anytime and virtually anywhere". Even though most chatbots are unable to detect spelling errors, and grammar mistakes, they could still be useful for non-beginner students. Fryer and Carpenter did an experiment where 211 students were asked to chat with ALICE and Jabberwocky chatbots. The feedback in general was that students enjoyed using the chatbots, and felt more comfortable and relaxed conversing with the bots than a student partner or teacher as followed in traditional teaching. The authors listed other advantages of chatbots in this task: the chatbot could repeat the same material with students several times without being bored, many bots used text and speech mode in responding to give an opportunity to practice both reading, and listening skills, and chatbots as a novel technology can improve students' motivation towards learning. In addition to this, if computers are available in the classroom, teachers could encourage students who finished their class work early to talk to a chatbot, giving them a topic to focus on. An easy self analysis could be achieved since most chatbots keep a transcript of the conversation where students can evaluate themselves.

7. Conclusions

We have advocated the use of a chatbot as a tool to learn a foreign language. We could "imagine chatterbots acting as talking books for children, chatterbots for foreign language instruction, and teaching chatterbots in general." (Wallace et al. 2003). Most current chatbots are restricted to converse on a pre-determined topic, so teachers or students may need to find a chatbot suitable to their needs. However, adaptive chatbots can be re-trained to chat on any given topic, as we have demonstrated. These adaptive chatbots can foster language learner autonomy, giving students the opportunity for independent active learning of conversation skills.

References

Abu Shawar, B. And E. Atwell. (2003a) Using dialogue corpora to train a chatbot in: Archer, D, Rayson, P, Wilson, A and McEnery, T (eds) Proceedings of CL2003: International Conference on Corpus Linguistics, pp. 681–90 Lancaster University.

Abu Shawar, Bayan; Atwell, Eric. (2003b) Using the corpus of Spoken Afrikaans to generate an Afrikaans chatbot. *Southern African Linguistics and Applied Language Studies*, vol. 21, pp. 283–94.

Abu Shawar, Bayan; Atwell, Eric. (2005a) Using corpora in machine-learning chatbot systems. International Journal of Corpus Linguistics, vol. 10, pp. 489–516.

Abu Shawar, Bayan; Atwell, Eric. (2005b) A chatbot system as a tool to animate a corpus. ICAME Journal, vol. 29, pp.5–24.

Abu Shawar, Bayan; Atwell, Eric. (2007) Chatbots: Sind Sie wirklich nutzlich? (are they really useful?). LDV-Forum Journal for Computational Linguistics and Language Technology, vol. 22, pp. 31–50.

Atwell E. (2005). Web chatbots: the next generation of speech systems? European CEO journal, vol. November-December, pp. 142–44.

Atwell E., Demetriou G., Hughes J., Schiffrin A., Souter C. and Wilcock S. (2000). A comparative evaluation of modern English corpus grammatical annotation schemes. ICAME Journal. Vol. 24, pp. 7–23.

Batacharia B., Levy D., Catizone R., Krotov A. and Wilks Y. (1999). CONVERSE: a conversational companion. In Y. Wilks (Ed.) Machine conversations (pp. 205–215). Boston/Dordrecht/London: Kluwer.

Braun, A. (2003). Chatbots in der Kundenkommunikation (Chatbots in customer communication). Springer.

Brennan K. (2006). The managed teacher: emotional labour, education, and technology. Educational Insights. Vol. 10, No. 2, pp.55–65.

Carpenter R. (2007). [Online]: www.jabberwacky.com

Chantarotwong B. (2005). The learning chatbot. Final year project. [Online]: http://courses.ischool.berkeley.edu/i256/f06/projects/bonniejc.pdf

Cliff D. and Atwell E. (1987). Leeds Unix Knowledge Expert: a domain-dependent Expert System generated with domain-independent tools. BCS-SGES: British Computer Society Specialist Group on Expert Systems journal Vol.19, pp.49–51.

Fryer L. and Carpenter R. (2006). Emerging technologies bots as language learning tools. Language Learning andTechnology. Vol. 10, No. 3, pp. 8-14.

Fujii S., Iwata J., Hattori M., Iijima M., and Mizuno T. (2000). "Web-CALL": A language learning support system using Internet. In proceeding of Seventh International Conference on Parallel and Distributed systems (workshops). pp. 326–31.

Gibbs, G.R., Cameron, C., Kemenade, R., Teal, A. and Phillips, D. (2004). Using a chatbot conversation to enhance the learning of social theory. [Online]: http://www.hud.ac.uk/hhs/dbs/psysoc/research/SSCRG/chatbot.htm

HEXBOT (2004). HEXBOT chatbot website. http://www.hexbot.com/

Hutchens J. (1996). How to pass the Turing test by cheating. School of Electrical, Electronic and Computer Engineering research report TR97-05. Perth: University of Western Australia.

Hutchens, T., Alder, M (1998). Introducing MegaHAL. [Online],
http://cnts.uia.ac.be/conll98/pdf/271274hu.pdf

Jia J. (2004a). The study of the application of a web-based chatbot system on the teaching of foreign languages. In: Proceedings of the SITE2004 (The 15the annual conference of the Society for Information Technology and Teacher Education). AACE press, the USA, pp. 1201–1207

Jia J. (2004b). CSIEC (computer simulator in educational communication): An intelligent Web-based teaching system fro foreign language learning.

Jung H, and Kim S. (2004). Opportunities in technology-enhanced language learning (TELL) classroom environments. Teaching English with Technology. Vol. 4, No. 4.

Kerfoot P., et al. (2006). Multi-institutional randomized controlled trial of web-based teaching to medical students. Academic Medicine. Volume 81. No. 3, pp. 224–30, March 2006.

Kerly A., Hall P., and Bull S. (2006). Bringing chatbots into education: towards natural language negotiation of open learner models. In R. Ellis, T. Allen and A. Tuson (eds), Applications and Innovations in Intelligent Systems XIV-proceedings of AI-2006, SGAI International conference on Innovative Techniques and Applications of Artificial Intelligence, Springer.

Knill O., Carlsson J., Chi A. and Lezama M. (2004). An artifical intelligence experiment in college math education.

Molla D. and Vicedo J. (2007) Question Answering in Restricted Domains: an overview. Computational Linguistics Vol.33:1, pp.41–61.

Robin R. (2007). Commentary: Learner-Based listening and technological authenticity. Langugae Learning and Technolog. Vol. 11, No. 1, pp. 109–115.

Schumaker R., Ginsburg G., Chen H., and Liu Y. (2006). An Evaluation of the Chat and knowledge delivery components of a low-level dialog system: the AZ-ALICE experiment.

Turing, A. (1950). Computing machinery and intelligence. Mind 49, pp. 433–60.Van Rooy, B. (2003). Transkripsiehandleiding van die Korpus Gesproke Afrikaans. [Transcription Manual of the Corpus Spoken Afrikaans.] Potchefstroom: Potchefstroom University.

Wallace R., Tomabechi H., and Aimless D. (2003). Chatterbots Go Native: Considerations for an eco-system fostering the development of artificial life forms in a human world. [Online]: http://www.pandorabots.com/pandora/pics/chatterbotsgonative.doc

Wallace, R. (2003) The Elements of AIML Style. ALICE A.I. Foundation.

Weizenbaum J. (1967). Contextual understanding by computers. Communications of the ACM. Vol. 10, No. 8, pp. 474–80.

Wikipedia (2007). Computer-Assisted Language Learning. [Online]: http://en.wikipedia.org/wiki/Computer-assisted_language_learning