

# Providing Language Instructor with Artificial Intelligence Assistant

K. Pietroszek<sup>1</sup>

<sup>1</sup> Wilfrid Laurier University/Department of Languages and Literatures, Waterloo, Canada

**Abstract**—This paper presents the preliminary results of developing HAL for CALL, an artificial intelligence assistant for language instructor. The assistant consists of a chatbot, an avatar (a three-dimensional visualization of the chatbot), a voice (text-to-speech engine interface) and interfaces to external sources of language knowledge. Some techniques used in adapting freely available chatbot for the need of a language learning system are presented. Integration of HAL with Second Life virtual world is proposed. We will discuss technical challenges and possible future work directions.

**Index Terms**—artificial intelligence, intelligent tutoring systems, tele-learning.

## I. INTRODUCTION

In almost every area of science and art there is an increasing interest in computing technologies. The computer is already well known to be a powerful teaching tool for both instructors and students. However, despite the enormous research efforts in the area of Computer Assisted Learning and Instruction (CALI), the existing technology is used mainly in relatively passive way: the computer serves as a tool rather than an instructor's assistant. Nevertheless, recently there has been an increasing interest by the CALI community in applications of artificial intelligence natural language processing in intelligent tutoring systems [8]. Recent results [11] show that it is already possible to create an AI assistant capable of performing some teaching tasks such as:

- reacting to and correcting common mistakes in exercises; providing the student with tailored feedback on her/his performance,
- delivering knowledge incrementally and in an organized way, presenting it in an interesting, navigable form and supporting efficient information search,
- answering unambiguous questions immediately without the need to bother the instructor (for example providing a grammar rule explanation),
- reacting to the individual needs of a student, his or her abilities and weaknesses and adjusting the teaching techniques appropriately; dynamically keeping track of those aspects of the subject which the learner has mastered and in what areas the learner is deficient.

After substantial initial development effort all the above and many other tasks can be performed by an AI assistant in interactive fashion, without needing to involve the instructor. Moreover, the student may interact with the assistant not only through the traditional peripherals such as a keyboard or a mouse. It is possible to generate speech

from AI assistant's textual responses. The quality of currently available voices, for example developed by Natural Reader and others [12], are in our opinion more than satisfactory. Some advancement has also been done in the field of speech recognition and pronunciation error detection [13]. Both-ways voice human-computer interfaces [14] such as Dragon Naturally Speaking are already available on the market.

The use of AI assistants could potentially contribute to every area of teaching, and not only at the University level. Its consequences may be compared to the recent introduction of automated marking (e.g. online quizzes) or virtual learning environments such as WebCT. The AI assistant may play the role of an upper-year colleague, who knows the material of the course and is willing to help at any time. In addition the AI assistant will reduce the workload of teachers and teaching assistants, such as the number of students' emails, which place great demands on the instructor's time. Of course the AI assistant abilities are, and for long time will be, limited. It is unlikely that AI assistant will replace a living teacher or teaching assistant in the near future.

In the recent years many intelligent tutoring systems have been developed [11]. However, we are not aware of any projects of similar scope (see Section 6 for Related Work) and we believe that our work is both novel and an important step towards better understanding of potentials and capabilities of AI language assistants.

The remainder of this paper is structured as follows. Section 2 is a brief description of a HAL<sup>1</sup> for CALL project. A more detailed description of a central module of the AI assistant, the Mind, and its anticipated capabilities and foreseen weaknesses is given in Section 3. Section 4 elaborates on the technologies used in prototype implementation of the system. Section 5 sheds some light on future work. We survey some related work in Section 6. Finally, Section 7 concludes the paper.

## II. HAL FOR CALL PROJECT IN A NUTSHELL

HAL for CALL is a prototype of an AI assistant meant to serve as a language instructor assistant. It benefits from the results of research in natural language processing, intelligent tutoring, text-to-speech generation, voice recognition and virtual reality. It consists of several subprojects: a Mind, an Avatar, a Voice and a World Interface.

---

<sup>1</sup> HAL is a name of artificial intelligence from the Space Odyssey 2001. Along with Alice (from Alice in Wonderland) it is a popular name for chatbots.

The Mind module is a chatbot, an artificial intelligence algorithm designed to simulate an intelligent conversation (see Section 6). It is not only capable of conducting a conversation, but also of correcting some grammar and spelling mistakes as well as providing definitions and translations of words and phraseological units.

The Avatar module is a three-dimensional (3D) model serving as an incarnation of the Mind. Its function is not merely a visualization of chatbot. According to [15] talking with 3D avatar is more engaging and motivating for the student than chatting with bodiless application. The reason is that chatting with avatar gives an impression of conversation with a real human being.

The user experience is further enhanced, if avatar is equipped with speaking ability. This leads to the next module of the system, the Voice, which is basically a text-to-speech engine interface enabling chatbot to speak with close to a natural human voice.

Finally, the World Interface module consists of a set of interfaces to external components which can be used both by the Mind, the Avatar and the Voice. Through World Interfaces the Mind gains access to all kinds of dictionaries and encyclopedias as well as to a search engines such as Google. These sources of language knowledge may be predefined or filtered by the language instructor in a manner similar to the professor advising his teaching assistant to use a particular dictionary or website. The World Interface can also run an external application such as a pronunciation error detection engine.

While the Mind module uses the World Interface mostly to gain external knowledge or abilities, the Avatar module uses it to put both itself and the student into the Internet-based virtual worlds such as Second Life [16]. This feature, although still under development, will open a whole new perspective of controlled learning experience in the environment closely resembling real life situations and real cultural and social context (see Section 5.1).

In the reminder of this paper we will discuss mainly the Mind module capabilities. The Avatar, the Voice and World Interface modules are beyond the scope of this paper and are just briefly referenced in Future Work.

### III. THE MIND MODULE CAPABILITIES

#### A. Chatting

The basic goal of the Mind module is to involve a student in a chat. The chatbot uses the simplest yet one of the most powerful natural language processing (NLP) models developed by Richard Wallace and first implemented as A.L.I.C.E. chatbot [17]. HAL's brain is a clone of A.L.I.C.E.'s brain with the addition of patterns specific for language learning.

A current version of the Mind module offers two simplistic learner models: beginner and advanced. If the chatbot is set to deal with a beginner, it talks with simpler vocabulary and grammar structures. Moreover, it often provides translations and gives explanations and definitions in student's native language. For example, if a French native speaker were chatting with HAL on the beginner level, a typical session could look as follows:

STUDENT: Hi!

HAL: Hi there!

STUDENT: How are you?

HAL: I am doing very well. How are you?

STUDENT: Good. What means "very well"?

HAL: "very well" means "très bien" in French.

When the learning level is set to "advanced", the chatbot moves into more grammar structures, richer vocabulary, some forms of slang or idiomatic expressions etc.

STUDENT: Hi! How are you?

HAL: Dude, my logic and cognitive functions are normal since the time I came into existence<sup>2</sup>. What about yours?

#### Limitation of the Current Approach

We are aware that simple division into difficulty levels is not a sufficient and effective form of learner model. A fully developed version of HAL for CALL should include a dynamic learner model, recording the student's misconceptions and language knowledge gaps and directing the conversation accordingly (e.g. [18], [19]).

#### B. Identifying and correcting grammar mistakes

The Mind module is capable of identifying some violations of grammar rules such as incorrect word order, lack of inversion in questions, use of a wrong verb form etc. In case of mistakes, the module either informs the student that it doesn't understand or, more often, provides a correction and an explanation. It is not guaranteed that the proposed correction will preserve the original meaning of the sentence. This shouldn't be surprising for even in normal conversation misuse of grammar structures may lead to misunderstanding. Moreover, the corrections and explanations are provided in a fashion which does not interrupt the flow of dialogue. The response to a mistake always falls into the following structure: alert, correction, explanation, and reference to original sentence. Let us look at the annotated example below. Please note that the annotations are only for reader's reference and do not appear in the real conversation.

STUDENT: Hal, have you ever think how wonderful it would be to fly?

HAL: [alert] Oops, I think I caught a grammar mistake! [correction] I guess you wanted to say: "have you ever thought". [explanation] You should have used a Past Participle form of the verb! [reference to original] Coming back to our conversation... Yeah. I don't think I have ever thought how wonderful it would be to fly. What's it like?

This quite sophisticated feedback is a result of simple pattern matching and searching. In the above case the "have you ever think" sentence part matches "have you ever [Verb->PPVerb]" correction pattern. It triggers the Mind module to search through the Past Participle verb table. "Think" verb is not found there. Therefore, the Mind needs to verify if "think" is a verb at all. When "think" is found in the general table of verbs it is concluded that the wrong verb form was used and the alert, correction, explanation sequence is generated. Then, the chatbot comes back to the original topic of the conversation and responds to the corrected student's utterance.

<sup>2</sup> Most of HAL responses come from original A.L.I.C.E. chatbot or are their modifications. Grammar explanations, spelling corrections, dictionary and encyclopedic definitions are added to HAL by the author.

### Limitation of the Current Approach

The simple pattern matching approach to grammar correction works in our case, because we can safely assume that the student is trying to say something meaningful. However there are some limitations to this technique. If a student is saying something which does not conform to any single grammar pattern known by the system, the chatbot gets confused. For example, let us look at a student sentence "Have you ever deeply think?". It is so incorrect that it "confuses" the Mind. This is because "have you ever [Adverb Verb->PPVerb Adverb]" does not exist in the database of HAL correction patterns.

The other limitation is that the system does only one correction pattern match per utterance. In the case of many mistakes in a single utterance the Mind will ignore all mistakes except the first one matched. For example,

STUDENT: Hal, have you ever think how wonderful would it be to flew?

HAL: [alert] Oops, I think I caught a grammar mistake! [correction] I guess you wanted to say: "have you ever thought". [explanation] You should have used a Past Participle form of the verb! [reference to original] Coming back to our conversation... Yeah. I don't think I have ever thought how wonderful would it be to flew. What's it like?

As we see the Mind module didn't correct the second part of the sentence. Even worse, it repeated the wrong sentence in the response. We realize that it is a strong limitation. However, even state-of-the-art grammar checkers, such as the one used in Microsoft Word, have difficulties to catch that kind of grammar mistakes. When a more advanced grammar checking engine becomes available, it will be integrated with the HAL for CALL system<sup>3</sup> via the World Interface module. On the other hand, even if the problem were overcome, a sentence with many mistakes would lead to a very long and elaborate chatbot divagation on grammar. It would contradict the main goal of the Mind which is to keep the conversation interesting and engaging for the student.

### C. Spelling mistakes recognition and correction

The Mind module identifies and corrects some spelling mistakes. Similar to the grammar correction approach, the response to a misspelled sentence consists of alert, correction, explanation and reference to original sentence. However, the explanation is optional and in most cases omitted. For example,

STUDENT: I dont know.

HAL: [alert] Spelling mistake? [correction] I assume you meant "I don't know". [reference to original] Well, I don't know, either.

The analysis of spelling is done in similar fashion as in many contemporary word processors; when a mistake is found the system searches for similarly looking words. However, instead of presenting a list all possible corrections, the Mind module randomly chooses the one which results in a match with a known pattern. This can be done efficiently, because the list of possible spelling corrections is usually quite short. In the example, "dont"

replaced by "don't" results in an "I don't know" pattern with a standard response "I don't know, either".

### Limitation of the Current Approach

As in case of state-of-the-art spell checkers, if the word is in the dictionary, it won't be identified as misspelled, even though it doesn't fit the sentence semantically. Moreover, such a meaningless sentence often leads to chatbot confusion as in the following example.

STUDENT: I thing you would love apples but I don't now.

HAL: Right now? That is interesting: Student said he thing you would love apples but he or she do not.

This cannot be easily prevented, inter alia because among all possible combinations of words it is much easier to find a meaningless sentence than a meaningful one.

### D. Tricks to Simulate Intelligent Conversation

The previous example shows that the Mind module is not able to respond meaningfully to the sentence with just two spelling mistakes. However, such a sentence would not pose a problem for most people. Therefore, it should be stressed that the chatbot module does not attempt to understand the ongoing conversation. Instead, it uses several tricks to simulate an intelligent conversation.

The first trick is to look for cue words or phrases from the students and reply with a fully or partially pre-calculated response<sup>4</sup>. If the database of possible patterns and responses is big enough the conversation "seems" natural without any attempt at understanding on the side of chatbot. For example,

STUDENT: What do you think about apples?

HAL: An Apple: 1. Temptation of Adam. 2. Inspiration for Newton. 3. Empire for Steve Jobs. Would you like to hear a dictionary definition for apple?

A second trick is to change the topic. For example, when the chatbot cannot find a match for the student's question, it proposes to teach the student a new meaning of the word used in the student's utterance.

STUDENT: Do you think eating fruits is healthy?

HAL: To tell you the truth, I don't think too much. It's tiring. Wait... you said "fruits"... Did you know you can say for example: "forbidden fruit"? The term "forbidden fruit" is a popular metaphor that describes any object of desire whose appeal is a direct result of the knowledge that cannot or should not be obtained.

As we see, the chatbot couldn't find the exact match for the question. The trick enables it to turn his ignorance into a teaching technique. In this case it randomly chooses a phraseological unit containing the word "fruit" and derives its definition from Wikipedia.

### E. Customization of the Mind module

The Mind module may be easily customized by the language instructor. Adding a simple pattern and

<sup>3</sup> We are currently testing an open source grammar checker, Link Grammar, developed in the School of Computer Science at the Carnegie Mellon University [25]

<sup>4</sup> This is a very old technique used by first chatbots such as ELIZA Joseph Weizenbaum and released in 1966.

responses does not require any programming skills. For example, adding a Portuguese translation of the word “fruit” comes down to writing a few lines in the translationToPortuguese.xml file:

```
<category>
  <pattern>How do you say fruta?
</pattern>
  <template>fruit</template>
</category>
```

Now, assuming that the student’s native language is set to Portuguese, the conversation could flow as follows.

STUDENT: How do you say fruta?

HAL: Fruit.

Another customization technique is to switch on or off parts of the HAL’s brain. As mentioned before, the major part of our chatbot brain content is a clone of A.L.I.C.E. and as such it is similarly structured. It consists of files categorized by topics. Any parts of the brain may be switched off at any time simply by removing a specific topic. This leads to a simple technique of preventing the vocabulary or topics from appearing in a conversation. Let’s assume that the instructor removed the “money” topic from the chatbot’s brain, because the vocabulary associated with money was not yet introduced in class. The Mind is then no longer able to give meaningful responses to any topic associated with money. Therefore, it must use the change topic tricks described in section 3.4. But responses generated via tricks tend to be less interesting. Additionally, the Mind could add on a discouraging utterance such as “Let’s change our topic. I’m not interested in money.” etc. In consequence, the student will naturally keep the money topic out of conversation. This way, he or she will be silently directed out of the topics which the instructor doesn’t want him to touch yet. This technique is based on a simple observation that is not engaging to chat with somebody who knows nothing about this particular conversation topic.

STUDENT: How much does it cost to buy an apple?

HAL: Oh, I don’t know what you mean by cost. Let’s change our topic. I’m not interested in cost.

#### IV. PROTOTYPE IMPLEMENTATION OF THE HAL FOR CALL

In the current stage of development the HAL for CALL system is just a proof of concept. It cannot in any way be considered as comprehensive. The prototype was created to show that such implementation is feasible with the use of existing technologies. The development of a full HAL system with all the possible functionalities would constitute a substantial effort for a small or middle size software company and is beyond the resources available for a single researcher.

The chatbot submodule of the Mind is implemented in Artificial Intelligence Markup Language (AIML) and processed by freely available AIML interpreter called Program D, written in Java [20]. The rest of the Mind module is implemented in Java. We are planning on trying AILM interpreter written in Linden Scripting Language (LSL)[21], which would allow HAL to access Linden’s Second Life virtual world (see Section 5.1). This leads to the Avatar module, which is being implemented

in LSL. The Voice module is still under development. We are testing MBROLA set of text-to-speech engines, mainly because it supports many languages [22].

#### V. FUTURE WORK

One of the goals of this paper is to trigger the discussion of possible future developments of HAL. Some feasible integration of HAL with other technologies are proposed below.

##### A. HAL for CALL Integration with Second Life Virtual World

Second Life is a constantly growing virtual world, whose citizens can meet, chat, buy land, build houses, do shopping – live their second life. The membership is free, but real moneys are spent on material, electronic or just virtual goods and services. In February 2007 this virtual world had more than two million citizens with an increase of one hundred thousand a month. Benefits from integrating an intelligent language tutoring system with this online community would be vast. It would enable learners to get into cultural and social environments almost undistinguishable from the real world. For example, students of English can teleport to New York, meet virtual incarnations of New Yorkers, go to the cinema to watch Hollywood movies – the possibilities are endless. All these activities could be supervised by HAL, which would serve as a guide, interpreter, encyclopedia or just a friend. The technology to do the integrations exists already and is constantly improving.

##### B. Speech Processing and Error Detection

A robust speech processing and error detection engine suitable for a highly variable learner speech would greatly contribute to the student experience of HAL for CALL system. There exist speech recognition and pronunciation error detection methods (e.g. [23], [13]) although substantial further effort is needed to incorporate them into the HAL for CALL project.

##### C. Dynamic learner model

Current HAL in CALL learner model is simplistic and needs to be enhanced. The learner’s knowledge and, even more importantly, misconceptions can be modeled dynamically and used to provide an adaptive interaction and tailored feedback. The current state of the model itself can be hidden from or opened to a student. “It has been argued that opening learner model provides opportunities for learner reflection and deep learning that enhances the learning experience [2] [3] [4] [5] [6]” [1]. Kerly et al. proposed that “natural language negotiation [of learner model] through a Chatbot may offer users the flexibility to express their views in a naturalistic and intuitive way” [1]. Incorporating a dynamic learner model would arguably improve the user experience of HAL for CALL.

##### D. A Need of Experimental Evaluation

Further in class experiments are needed to identify pros and cons of HAL for CALL. The early state of development of the prototype implementation does not allow for extensive experimentation yet.

Incorporation of other technologies such as emotion recognition systems, facial recognition systems, or attention detection systems would enhance the user

experience greatly, but are beyond the scope of this paper and current HAL for CALL development plans.

## VI. RELATED WORK

To our best knowledge the system most similar to HAL for CALL is the Tactical Language and Culture Training System (TLTS) originally developed by the University of Southern California's Information Sciences Institute under funding from the Defense Advanced Research Projects Agency (DARPA). It "helps learners acquire communicative competence in spoken Arabic and other languages. An intelligent agent coaches learners, assessing their mastery and providing tailored assistance. Learners then perform missions in an interactive story environment, where they communicate with autonomous, animated characters" [7]. However, TLTS concentrates on rapid acquisition of basic language skills useful during the military missions, while the HAL for CALL scope is more general and aimed at students of all language levels.

There exist many chatbots tailored for English as a Second Language. One example is DAVE E.S.L developed by Richard Wallace, author of A.L.I.C.E. It does not provide feedback on student performance, though, and has no connection to external sources of knowledge. However, there are chatbots which make use of external sources of knowledge. For example, Chomsky chatbot [24] builds its conversations with information retrieved from Wikipedia.

It is suggested by Kerly et al. [1] that technologies more advanced than AIML should be used in the development of educational chatbot. Lingubot technology is proposed as superior over A.L.I.C.E. In fact, in the recent years chatbots with built-in self-learning mechanisms tend to perform better than classical pattern-matching-without-reasoning algorithms<sup>5</sup>. However, in our opinion A.L.I.C.E. approach has one great advantage, namely its simplicity. Simple AIML brain contents can be created by nonprogrammers such as language instructors, while use of many other technologies demand in-depth programming and technical skills.

## REFERENCES

- [1] Kerly, A., Hall, P., Bull, S., Bringing Chatbots into Education: Towards Natural Language Negotiation of Open Learner Models. In R. Ellis, T. Allen & A. Tuson (eds), Applications and Innovations in Intelligent Systems XIV – Proceedings of AI-2006, the Twenty sixth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, Springer, London, 2006
- [2] Bull, S., Brna, P., Pain, H., Mr. Collins: A collaboratively constructed, inspectable, student model for intelligent computer assisted language learning. *Instructional Science*, 1995. 23: p. 65-87.
- [3] Dimitrova, V., STyLE-OLM: Interactive Open Learner Modelling. *International Journal of Artificial Intelligence in Education*, 2003. 13: p. 35-78.
- [4] Kay, J., Learner Know Thyself: Student Models to give Learner Control and Responsibility, in ICCE97 International Conference on Computers in Education, Z. Halim, T. Ottomann, and Z. Razak, Editors. 1997: Kuching, Malaysia. p. 17-24.
- [5] Morales, R., Exploring participative learner modelling and its effects on learner behaviour. 2000, Unpublished PhD Thesis. University of Edinburgh: Edinburgh.
- [6] Zapata-Rivera, J.D., Greer, J., Externalising Learner Modelling Representations, in Workshop on External Representations of AIED: Multiple Forms and Multiple Roles. International Conference on Artificial Intelligence in Education. 2001: San Antonio, Texas.
- [7] Johnson, L., Vilhjálmsson, H., Marsella, S., Serious Games for Language Learning: How Much Game, How Much AI?, AIED 2005
- [8] Pon-Barry, H., Schultz, K., Bratt, E.O., Clark, B., Peters, S., Responding to Student Uncertainty in Spoken Tutorial Dialogue Systems. In *International Journal of Artificial Intelligence in Education (IJAIED)* Volume 16, 171-194. Special Issue "Best of ITS 2004".
- [9] Eskenazi, M., Issues in the use of speech recognition for foreign language tutors, invited paper: *Language Learning and Technology Journal* (online) Vol. 2, No. 2, January 1999, pp. 62-76
- [10] Heilman, M., Eskenazi, M., Language Learning: Challenges for Intelligent Tutoring Systems, Workshop on Ill-defined Domains in Intelligent Tutoring 2006.
- [11] Murray, T., Authoring intelligent tutoring systems: An analysis of the state of the art. *International Journal of Artificial Intelligence in Education*, 10(1):98-129.
- [12] Speech synthesis links. Retrieved from <http://www.acoustics.hut.fi/~slemmet/speech.html>
- [13] Neri, A., Cucchiarini, C., Strik, H., Feedback in Computer Assisted Pronunciation Training: Technology Push or Demand Pull, ICSLP 2002
- [14] Bilmes, J.A., Li, X., Malkin, J., Kilanski, K., Wright, R., Kirchoff, K., Subramanya, A., Harada, S., Landay, J.A., Dowden, P., Chizeck, H., The Vocal Joystick: A Voice-Based Human-Computer Interface for Individuals with Motor Impairments, Human Language Technology Conf. and Conf. on Empirical Methods in Natural Language Processing, Vancouver, Canada, Oct. 2005
- [15] Sheth, R., Avatar Technology: Giving a Face to the e-Learning Interface. *E-Learning Developer's Journal*, Aug. 2003.
- [16] Second Life and Education. Retrieved from <http://secondlife.com/education>
- [17] A.L.I.C.E. Foundation. Official website: <http://www.alicebot.org>
- [18] Michaud, L.N., McCoy, K.F., Capturing the Evolution of Grammatical Knowledge in a CALL System for Deaf Learners of English. *International Journal of Artificial Intelligence in Education*, 2006. 16(1): p. 65-97.
- [19] Zapata-Rivera, J.D., Greer, J., Externalizing Learner Modeling Representations, in Workshop on External Representations of AIED: Multiple Forms and Multiple Roles. International Conference on Artificial Intelligence in Education. 2001: San Antonio, Texas.
- [20] Bush, N., Getting started with Program D. Retrieved from <http://www.alicebot.org/resources/programd/readme.html>
- [21] LSL Wiki. Retrieved from <http://lslwiki.net/lslwiki/wakka.php?wakka=HomePage>
- [22] The MBROLA Project. Official website: <http://tcts.fpms.ac.be/synthesis/mbrola.html>
- [23] Mote, N., Sethy, A., Silva, J., Narayanan, S., Johnson, L., Detection and modeling of learner speech errors: The case of Arabic tactical language training for American English speakers. *STIL Symposium on NLP and Speech Technologies in Advanced Language Learning Systems*.
- [24] Chomsky chatbot. Official website: <http://demo.vhost.pandorabots.com/pandora/talk?botid=b0dafd24e35a477>
- [25] Link Grammar. Official website: <http://www.link.cs.cmu.edu/link>

## AUTHOR

**K. Pietroszek** is with the Department of Languages and Literatures, Wilfrid Laurier University, Waterloo, N2L 3C5, Canada (e-mail: [kpietroszek@wlu.ca](mailto:kpietroszek@wlu.ca)).

<sup>5</sup> In 2005 and 2006 Jabberwacky-based chatbots won the Loebner Prize awarded to the most human-like computer.

Manuscript received 31 July 2007. Published as submitted by the author.