Facilitating foreign language fluency through interactive conversational software

Elizabeth Roberts, Fall 2011

Despite the large variety of language study software available, there is still no effective way to practice conversational fluency in a foreign language without a fluent speaker of the language in question. Many programs, such as Rosetta Stone and Tell Me More, offer excellent alternatives to expand vocabulary, improve listening comprehension, etc., but still fall short when it comes to practicing conversational speech. In this project, I hope to develop software which permits flexible, sustained, and structured conversation practice in foreign languages at an elementary level. The description below reflects a fully-functional ideal of the software; I believe it is unlikely that I will be able to accomplish all of these goals in the course of a single semester, but I hope to complete the majority of them.

As presently conceived, the final form of this software will focus almost entirely on speaking and listening. The primary interface between the user and the program will be an aural/verbal interface rather than a visual interface, although a graphical interface will facilitate altering options within the program and provide written cues when the user is struggling with the dialogue. The user will begin by selecting a language and a module within that language (e.g. introductions, family, travel); before the conversation begins, they will be able to choose additional vocabulary outside of the core module which they would like to include. The program will provide an audio prompt, and the user may say one of several responses in return. The user will be able to click on a button to reveal the written form of audio prompt if necessary. Based on the clarity of the user's response, the software will either continue the conversation or prompt the user to try again, displaying the written forms of the user's choices. As the conversation progresses, a transcript of the dialogue so far will be displayed in the window. At the conclusion of the conversation, the user may choose to print the transcript, have another conversation in the same module, or return to the beginning of the program.

In order to test the production of conversation before the extensive and difficult implementation of the aural/verbal interface, I will first create a similar version of the software with relies on a strictly visual interface. As in the final form, the user begins by selecting a language and module; the computer will then provide a text prompt, and the user may select one of several responses by clicking the button with the appropriate phrase or sentence on it. A transcript will still be produced, and the program should function in all ways like the final form, excepting the differing nature of the user interface. Below are some examples of short sample conversations that might be produced by a module on introductions:
Computer: Hello.
User: Hello. How are you?
C: I am well. What's your name?
U: My name is Fred. What's your name?
C: My name is George. How are you?
U: I am well. Where are you from?
C: I am from China. Where are you from?
U: I am from America. Nice to meet you.
C: Have a good day!
U: You too!

Regardless of the details of the final implementation of this program, there are several important qualities which I believe the software must have. First, the conversations with the user must be focused on particular topics to effectively practice phrases and vocabulary; however, they must not be entirely static but should vary from use to use. Second, the conversation produced by the program and user must be logical and acyclic. In addition, there are other features which would greatly enhance the usability and usefulness of such software: the user should be able to customize the difficulty and extent of vocabulary included in the conversations; the software should be easily modifiable to provide support for additional languages; and the software should be easily expandable to cover more vocabulary within each language.

Based on the intended user experience and program requirements described above, I expect this software to have the following major components:

- a “conversation engine” which facilitates choosing the responses which will be offered to both the user and the program while conducting the conversation
- a set of questions and phrases in English which comprise the content of each module, as well as rules for transitioning between phrases
- translations of the module content into the chosen languages (most likely beginning with Korean and Spanish)
- a GUI to permit user interaction with the program
- a signal processing system to analyze the user’s attempts to speak and compare it to the acceptable options
- a bank of recordings of the module content in the target language by native speakers or advanced learners of that language

Depending on the difficulty of and time required to implement the above, there are additional components which could improve the usability and relevance of this software. One such component would be a grammar practice module, which operates in a manner very similar to the conversation modules but which offers variations on one response to the user, from which the grammatically correct answer should be chosen, rather than several “correct” responses of
which the user may select whichever he prefers. Another possible expansion of the program would be a pronunciation practice module that is based on the content of the conversation modules. If the user has difficulty saying the desired responses accurately enough for the computer to recognize a valid answer, this module would allow the user to select individual phrases from the conversation to practice and receive detailed information about his pronunciation inaccuracies.

In the process of preparing this project, there will undoubtedly be technical challenges which arise. The first challenges which I anticipate encountering include outputting characters which are not included in the ASCII encoding scheme and designing a conversation engine which allows for extensive flexibility without extensive memory usage. Additionally, developing an auditory signal processing and analysis system which can successfully match user input to recordings of native speakers, and possibly provide feedback about pronunciation, will require a significant amount of research and thought; this is exacerbated by the fact that ideally I will develop a system that functions more or less equally well on speech in any language.

The deliverables for this project will consist of four things. In addition to this proposal, I will provide all files relating to this software: this will include code, the text files with the module content, sample transcripts produced by the software, etc. At the conclusion of the project, I will also write a report on the process and results along with an abstract of the project and report.

Potential resources and references:

The Language Technology Institute at Carnegie Mellon University
Stephen D. Krashen, “Second language acquisition and second language learning”
Farzad Ehsani and Eva Knodt, “Speech technology in computer-aided learning: Strengths and limitations of a new CALL paradigm”
Ambra Neri, Catia Cucchiarini, and Wilhelmus Strik, “Automatic Speech Recognition for second language learning: How and why it actually works”