A spoken dialogue program for second-language learning

General Background

Communicative language teaching (CLT) is widely accepted as an effective language teaching approach in the past decades. CLT is based on the theory that the primary function of language use is communication (Celce-Murcia, 2001) ; thus, the primary goal of the language learners is to learn communicative ability. In other words, language learners should make use of real-life conversational situations to facilitate language learning. Within the CLT paradigm, the recent Task-based language teaching (TBLT) approach has developed as a leading pedagogical and syllabus design approach. TBLT utilizes and emphasizes authentic, communication-driven tasks that provide task-related focus on form congruent with the learner's own internal syllabus that aims to prepare students to use their linguistic skills in meaningful interactions outside the classroom (Long, 2014). Examples of language learning tasks include filling out a form, buying a pair of shoes, making airline reservation, borrowing a library book etc. (Long, 1985).

With the rise of computer technology, the field of computer-assisted language learning (CALL) has grown and language educators have incorporated more computer-mediated language learning into classroom and curriculum in the last two decades (Grgurovic, Chapelle, & Shelley, 2013). TBLT has become widely accepted as the major framework that can help technological designs in CALL systems. In reviewing the relationship between technology and language learning tasks, Ziegler (2016) summarized that technology not only facilitates and supports the development of

linguistic and pragmatic skills, but its benefits go beyond improved quantity and quality of language production, with positive effects found for affective, social and cultural factors. For example, González-Lloret and Ortega (2014) found that technology-mediated tasks could reduce learners' anxiety and increase their motivation and creativity, as well as promote more engagement and participation in language learning.

Speech Recognition

Automatic speech recognition (ASR) technology opens up possibilities for the training of oral skills in CALL systems. In a recent review of technology and foreign language learning, Golonka et. al (2014) summarized that although ASR accuracy is not 100%, students typically reported positive experiences when using the software, increased motivation to practice the language and an increased confidence in their ability to use the language.

Although ASR technology has achieved great improvement in recent years and is widely adopted by many CALL systems, its major function is the training of pronunciation. As Golonka et. al (2014) summarized that in the field of language learning, ASR is used as a part of CAPT software to improve learners' FL pronunciation. Using ASR for training conversations that allows users to speak freely within a task design is still in a very early developmental stage. As a result, although visual aids have been used in videos and multimedia curriculum, most of the visual aids are delivered passively as medium to present input, very few visual aids are used interactively that are capable of providing feedback according to the users' oral productions. Such limitations greatly undermined the effects of training oral skills for CALL systems.

Visual Aid

Visual aid is another major component in CALL systems. Visual aids in CALL used in language teaching combine verbal and graphical materials promotes language learning. Studies show people learn better from better from words and pictures than from words alone (Butcher, 2014; Mayer, 2009). Paivio's dual coding theory suggests that the cognitive process of the human brain proceeds in relation to the processing of information derived through interplay of both verbal and visual elements (Paivio, 1986). Danan (1992) explained that dual coding theory emphasizes the importance of providing input for both the verbal and the visual representational system and for learners, visual traces are remembered better than are verbal components and also have an additional effect when items are encoded dually.

Problem Statement

Although oral proficiency is the focus of CLT approach and conversation that simulates real-life situation is the major theme for TBLT curriculum, the training of oral skills is often neglected in classroom instructions even under a TBLT context. One of the major challenges for oral skill training is that conversational practice is very hard to implement in classrooms especially when the student-teacher ratio is high. CALL systems provide a viable solution to this problem because students can access preprogramed oral input in L2 and practice individually with their own pace and as often as they want. However, most CALL systems that help students practice conversation have limitations.

CLT research suggests learners should be encouraged to express their own meaning as early as possible (Omaggio-Hadley, 2001), and such oral production practices need to happen under real

conditions of communication so the learner's linguistic knowledge becomes automatic (Ellis

1997). However, such "real conditions of communication" are not easy to implement in language

classes due to constrains of limited teaching resource, and learners in English as foreign

language environments lacking the opportunity to engage in real conversation with native

speakers. Thus, CALL systems that provide training for conversation can fill this gap for second

language learners. However, many current CALL systems that provide training modules for

conversation often restrict users from what they can produce. For example, although equipped

with an ASR function, the interactive language learning APP Duolingo only allows users to

choose from one of several pre-written answers in conversation practices. Such limitations

restrict learners from producing their own L2 languages and might impede their oral skill

development.

Proposed Solution

The program is coded with JavaScript and run on the Chrome browser. Focusing on the

limitations of spoken task designs in current CALL systems, the present project integrates ASR,

Speech synthesis and animations into spoken tasks to help the language learners practice L2

conversations in real-life situations. The language task of the program design is a simulation of a

real-life conversation at a fast-food restaurant and the user needs to complete the task by

successfully ordering a meal by giving speech commands to the program.

With JavaScript coding in the browser API, the present project uses Chrome speech recognition

to get users' L2 production as speech input and matches speech input with different keywords for

different dialogue situations. This will allow the users to express L2 sentences freely with their

own meanings as long as the L2 production contains a target keyword. When the target keyword

is detected in the users' language production, the program will generate appropriate replies from

the pre-stored rely bank and plays back an appropriate reply with Chrome speech synthesis. The

reply serves as input for the language user and carries the conversation forward. At the same

time, the program displays pictures of the keywords with pre-programed animation that serves as

visual aids for the conversation. Another benefit of using computer programing to facility the

teaching of conversation is that it is easily replicable. Although the current program is designed

for language learners to complete a very specific task (ordering a meal), the main structure of the

program can be easily replicated and users only need to modify the data objects in the program to

generate new language tasks. The speech recognition keywords, reply sentences and visual aid

pictures can be changed and expanded according to the context of new language tasks.

References:

Butcher, K. R. (2014). The multimedia principle. In R. Mayer (Ed.). The Cambridge handbook of multimedia learning (2nd edn; pp. 174–205). New York: Cambridge University Press.

Celce-Murcia, M. (2001). Language teaching approaches: An overview. *Teaching English as a second or foreign language*, *2*, 3-10.

Danan, M. (1992). Reversed subtitling and dual coding theory: New directions for foreign language instruction. *Language Learning,* 42, 497-527

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer assisted language learning*, *27*(1), 70-105.

Grgurovi´c, M., Chapelle, C. A., & Shelley, M. (2013). A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL*, 25, 165–198.

Long, M.H. (1985) A role for instruction in second language acquisition: Task based language teaching. In Hyltenstam, K., Pienemann, M. (Eds.), Modeling and assessing second language development (pp. 77–99). Clevedon: Multilingual Matters.

Long, M. (2014). *Second language acquisition and task-based language teaching*. John Wiley & Sons.

Mayer, R. E. (2009). Multimedia learning (2nd edn). New York: Cambridge University Press.

Omaggio-Hadley, A. (2001). Teaching language in context (3rd ed.). Boston: Heinle & Heinle

Paivio, A. (1986). Mental Representations: Dual–coding approaches. New York: Oxford University Press.

Ziegler, N. (2016). Taking technology to task: Technology-mediated TBLT, performance, and production. *Annual Review of Applied Linguistics*, 36, 136–163.