

〒Study on an Interactive CALL System for  
Pronunciation, Grammar and Composition  
Learning(発音・文法・作文学習のための対話型  
CALLシステムに関する研究) 〒

著者	権五 杓
号	3354
発行年	2004
URL	<a href="http://hdl.handle.net/10097/8626">http://hdl.handle.net/10097/8626</a>

氏名	グオン オ ピョウ 権 五 杓
授与学位	博士(工学)
学位授与年月日	平成17年3月25日
学位授与の根拠法規	学位規則第4条第1項
研究科, 専攻の名称	東北大学大学院工学研究科(博士課程)電気・通信工学専攻
学位論文題目	Study on an interactive CALL system for pronunciation, grammar and composition learning
指導教員	東北大学教授 牧野 正三
論文審査委員	主査 東北大学教授 牧野 正三 東北大学教授 澤谷 邦男 東北大学教授 鈴木 陽一 東北大学助教授 伊藤 彰則

## 論文内容要旨

### Chapter 1 Introduction

With the advent of globalization, in particular, the past decade has seen an influx of non-native speakers entering Japan for education, job, and shopping purposes. The ability to understand and communicate in other languages is increasingly important in our society. As a consequence, languages contribute more and more to personal fulfillment, mutual commercial success and international trade/communication. In order to master a foreign language, learners commonly need to acquire the skills from the following 4 exercises: Reading, Listening, Writing and Speaking.

It is best if a foreign language learner can be taught on a one-to-one basis. Active production of speech in such environment prepares learners to participate effectively in 'real conversation' later on. Unfortunately, a one-to-one basis learning is actually difficult for most learners, due to economic, time and place reasons. In reality, most learners attend classes in which they have to share their teacher with other learners. This reduces the amount of time each learner spends in speaking the foreign language. Conventional self-study methods that are commonly used include text materials, tape recorders, VCR language learning material, and composition textbook. However, the disadvantages are that the learner is not forced to say anything, and that the learner receives no corrective feedback.

In order to compensate for learner-teacher methods' and self-study methods' weak area, various intelligent and sophisticated diagnostic CALL systems have been proposed and developed so far. While the conventional systems are important step in the right direction, other more complex and ambitious CALL systems are conceivable and no doubt desirable. If I focus only on the speaking part: pronunciation and conversation exercise, there are still lots of problems. In pronunciation practices, the conventional methods cannot return the intelligibility score and instruction together. The disadvantage of the rule based methods is that the calculation of the intelligibility score is not

given and disadvantage of the model based methods is that the instruction is not given. In conversation exercise, most systems do not allow learners to produce various utterances actively as often as they do in learner-teacher interactive method. Most systems also cannot detect grammatical, lexical errors et al, if there are grammar and lexical errors in a learner's utterance. The other is that the path of the dialogue should be changed according to the learners' response utterances similar to the learner-teacher method.

## **Chapter 2 Pronunciation evaluation method**

In this chapter, I mainly discussed about pronunciation evaluation methods. I have briefly pointed out the disadvantages of 3 kinds of conventional pronunciation methods: Human perception methods, Model based methods and Rule based methods. While Human perception method returns proper feedback that includes the intelligibility and instruction, the disadvantage of this method is that it requires additional human perception if new minimal pairs are added. The greatest disadvantage of Model based methods is that it is difficult in returning proper instruction, because these methods are only concerned with how much the learner's pronunciation is similar to native's correct pronunciation. The disadvantage of Rule based methods is that it does not give the learner an intelligibility score that shows the percentage of native speakers who would understand the learner's pronunciation of each phone.

To overcome the disadvantage of the conventional methods, I proposed the evaluation method for Japanese pronunciations. The feature of the proposed method evaluates intelligibility scores and returns proper feedback together. To use expert knowledge about systematic pronunciation error made by Korean native speakers in Japanese, Japanese pronunciation problems were quantitatively analyzed using about 2800 utterances of 28 adult Koreans by 7 Japanese native speakers. Furthermore, pronunciation problems depending on the ability to pronounce Japanese were also analyzed. With our evaluation methods, a learner's pronunciation is evaluated according to its similarity to the native speaker speech for each phone. The correlation coefficient between the intelligibility scores and human scores is 0.72. It is verified at a significance level of 0.05. This method can also return proper instruction, because this method is based on the rule basis methods extended to the model basis methods.

I also incorporated an on-line speaker adaptation into the evaluation method to improve the evaluation accuracy. Comparison of evaluation results of Koreans' speech with human-evaluated scores and results of statistical analysis verified that the results of the evaluation were similar to those of native speakers.

### Chapter 3 Grammar error detection method

In this chapter, I mainly discussed about grammar error detection methods. In order to deploy ASR technology in our purposed voice-interactive CALL system, understanding of both the limitation of ASR technology and the limitation of learners is essential. Therefore, I defined the tasks for the limitation of ASR technology and pre-exercises such as basic vocabulary, grammar and conversation flow for the limitation of basic knowledge of foreign language learners. In defined tasks and pre-exercises, I investigated the tendency of grammatical and lexical errors. I also checked the influence of grammatical and lexical errors with/without pre-exercise. It is no doubt that pre-exercises can help learners to acquire basic knowledge for 'real conversation' with system, but the grammatical and lexical errors are decreased only from 0.223 / each sentences to 0.177. They tried to change the vocabulary and grammar, which are acquired during pre-exercise, with their limited knowledge of the vocabulary and grammar which are acquired in past. Error tendency are related to the particle, verb conjugation and lexicon. I classified the detected grammatical/lexical errors in the transcription and made grammar error rules.

In order to detect learners' grammatical and lexical errors, I proposed several methods and examine the recognition rate and error detecting rate. The feature of the proposed method 1 is that the original correct grammar which does not include any grammatical errors is used. The disadvantage of the manual method is that it is difficult to extend grammatical errors. To overcome this disadvantage, the concept of this method uses correct system grammar (which does not include grammatical and lexical errors) to recognize a learner's utterance and N-best recognition results are extended by grammar error rule to generate the predicted utterances and a learner's utterance is searched. The feature of the proposed method 2 is that general grammar which does not consider the relationship between verb, particle and each noun is used. To overcome the disadvantage of the manual method, this method uses a general grammar (which does not consider the relationship between verb, particle and each noun) to recognize a learner's utterances, whereby the N-best recognition results are calculated and the utterances of the learner are searched.

Comparing to the manual method (the best recognition rate / error detecting rate, but system grammar was constructed by hands), error detecting rate of the proposed method 2 is the same and the recognition rate is similar (It decreased about 2.7%, but the recognition improvement of the manual method was not significant at the significance level 0.05. Therefore, it can conclude that the recognition rates of both methods are the same). Considering of disadvantage of the manual method, I confirmed that the proposed method 2 is effective.

### Chapter 4 Voice-interactive CALL system

In this chapter, I mainly discussed about a voice-interactive CALL system. To realize such a system, I also discussed about system design. It is quite difficult for learners to try to produce

spontaneously what they want to say well with limited vocabulary and phrases. I introduced a pre-exercise to help learner to prepare 'interactive conversation exercise'. First, a learner does exercises on basic exercise (vocabulary, grammar and typical conversation) and acquires basic knowledge of a defined task. Then, they do applied exercise (pronunciation, grammar and conversation exercise).

To investigate the influence of a pre-exercise, I examined the number of interjection and out-of-grammar error from interview data. With the pre-exercise, the number of interjection per sentence decreased by about 0.122 per sentence from 0.787 to 0.645. Under the 'without pre-exercise' condition, 14% of total sentences include out-of-grammar errors, but 5.4% of the errors are committed under the condition 'with pre-exercise'. It is no doubt that pre-exercise can help a learner to produce their own sentences naturally.

I also integrated the proposed pronunciation evaluation method in chapter 2 and grammar detection method in chapter 3 and developed a new concept of dialogue-based CALL system.

I also surveyed user trial for 6 Korean and administered a questionnaire. According to the results of questionnaire, most of learners answered that pronunciation evaluation and grammar error detection are very significant, useful and essential. However, the reason of grammar error and various examples of correct sentences are necessary in the technical point. They also pointed out that various grammar expression and dialogues are necessary in the contented point.

## Chapter 5 Conclusion

The long-term goal of our research is to develop a voice-interactive CALL system. I developed the new concept of an interactive CALL system. Using such a system, a learner can practice pronunciation exercise, grammar exercise, composition exercise and conversation exercise.

In this thesis, I discussed about the pronunciation evaluation method in chapter 2, grammar detection methods in chapter 3 and voice-interactive CALL system in chapter 4. I developed a new concept of dialogue-based CALL system and surveyed user trial. According to the results of questionnaire, pronunciation evaluation and grammar error detection are very significant and useful.

However, there still remain other subjects that need to be improved on. Even though I implemented a voice-interactive CALL system and surveyed user trial, I need long-term user trial quantitatively and survey the effectiveness of the system as a future work.

# 論文審査結果の要旨

近年、外国語を学習する人口の増加とともに、外国語を効果的に学習するための CALL (Computer Assisted Language Learning) システムが開発されてきている。現在の CALL システムの多くはコンピュータを利用したマルチメディア教材に過ぎず、外国語学習の際に問題となる発音学習や文法学習を積極的に支援するものではなかった。音声認識技術を利用して発音学習を支援する CALL システムも開発されているが、発音評価の精度が不十分であった。筆者は、韓国人学習者のための対話型日本語学習システムを題材として、高精度な発音評価および会話中の文法誤りの検出手法を研究した。本論文はこれらの成果を取りまとめたものであり、全編5章からなる。

第1章は序論である。

第2章では、高精度な発音評価手法の開発を行っている。筆者は、これまで提案されていた母語話者の発音から作成した確率モデルによって学習者の発音の評定値を算出する手法と、学習者の発音誤り傾向のモデルを利用して学習者の発音誤りを検出する手法の両立を提案した。さらに、母語話者と非母語話者の発音モデルと、発音誤りをモデル化した発音誤りルールを用いる手法を考案している。これは、発音評価手法を取り入れた CALL システムを実現する上で重要な成果である。

第3章では、学習者の発話の中から文法的な誤りを検出するための手法を開発している。従来は、代表的な文法誤りを受理する音声認識システムを個別に作成し、これを用いた認識結果から誤りを見つけるという方法がとられていた。筆者は、音声認識システムの文法とは別に、学習者の起こしやすい文法誤りをルール化した一般的な「文法誤りルール」をシステム内に保持し、これを利用して誤りの検出を行う手法を開発した。これは非母語話者の発話の中から一般的な文法誤りを検出する方法として初めてのものであり、高く評価できる。

第4章では、上記の成果を基に、日本語学習のための対話型 CALL システムを構築し、利用実験を行っている。学習者は最初に語彙、文法および会話例などについての学習（事前学習）を行う。その後、設定された場面の中で学習者とシステムが自由に会話練習を行う。このシステムにより、学習者は発音、文法および会話作文を総合的に学習することができる。被験者による評価により、システムが学習者にとって有用であることが示された。これは、高度な機能を導入した CALL システムの有効性を示した重要な成果である。

第5章は結論である。

以上要するに本論文は、発音誤りおよび文法誤りを検出する新しい手法を開発し、これに加えて、従来よりも教師との会話練習に近い対話型の外国語学習 CALL システムを実現したものであり、音声情報工学、教育工学ならびに電気・通信工学の発展に寄与するところが少なくない。

よって、本論文は博士(工学)の学位論文として合格と認める。