

論文 / 著書情報
Article / Book Information

論題(和文)	
Title(English)	Intelligent English Tutoring System: from the Analysis of Japanese Learner Corpus
著者(和文)	西原 明法
Authors(English)	Pakaporn THISAYAKORN, Akinori NISHIHARA
出典(和文)	日本教育工学会 第28回全国大会, , , pp. 993-994
Citation(English)	The 28th Annual Conference of JSET, , , pp. 993-994
発行日 / Pub. date	2012, 9
権利情報 / Copyright	本著作物の著作権は日本教育工学会に帰属します。 Copyright (c) 2012 Japan Society for Educational Technology.

Intelligent English Tutoring System: from the Analysis of Japanese Learner Corpus

Pakaporn THISAYAKORN, Akinori NISHIHARA

Department of Human System Science

Graduate School of Decision Science and Technology, Tokyo Institute of Technology

Abstract We propose a method of increasing error awareness of learners to achieve better proficiency of English. Our targets are Japanese learners, thus we analyze Japanese learner corpus to create quizzes and use to construct intelligent tutoring system. The system allows users to learn from conflict by error identification of erroneous sentences. This paper explains the process of corpus analysis and the extraction of quizzes. The selection of quiz to use is also discussed.

Keyword Educational Software, Computer Assisted Language Learning, Second Language Acquisition, English as a Foreign Language, Language Education, Learner Corpus

1. Introduction

It has been reported that computer assisted system is an effective tool that engages students to spend their time on lesson than traditional face-to-face approach does (US Department 2009). Learning with computer allows students to study at their own pace, and intelligent system can adaptively provide instruction to their ability levels.

For language study, computer assisted language learning (CALL) is widely used. Current pedagogy used is structure-based, where linguistic rules and correct examples are provided. However, according to reviews provided by Leeman (2007), several studies stated that learning from negative evidences plays an important role in Second language acquisition (SLA). MacWhinney (1987) also suggested that adult learners perform better when learning on conflict, thus not only should learners study valid forms, but also impossible patterns in target language.

In this study, we proposed a learning method through erroneous sentences via intelligent CALL by identifying and correcting errors. We aim to increase error awareness, thus will increase language proficiency of learners. Our method focuses on English as the target language studied by Japanese-native speakers. This paper explains first half of the system construction, which is how we acquired erroneous sentences and used them as quizzes in this intelligent tutoring system.

2. Error Treatment

To let students learn from erroneous examples effectively, type of errors to present must be considered. Various methodologies in error treatment have been proposed (Shahin 2011). Since the goal of language is to convey meaning, we focus on Hendrickson's (1987) suggestion that errors that block communication, so called communication errors, should gain priority for correction.

3. Corpus Processing

We choose NICT Japanese Learner Corpus created by Tono (2005) because it provides error tagging with corrections that we can use to create the system. The NICT Corpus divided errors into 43 categories by part-of-speech of error words and morphological, grammatical, or lexical rule violated.

Izumi (2007) investigated intelligible levels of errors in NICT corpus, from which we conclude that errors of categories with high frequency of unintelligible utterance are likely to be communication errors. We extracted only errors of categories with high frequency and created quizzes from them.

4. System Structure

Type of quizzes used in our system is error identification type, which students have to determine errors in erroneous sentences and correct them.

Following framework of the system suggested by Thisayakorn (2011), we constructed an intelligent tutoring system, which error identification quizzes are presented to students first to determine their weaknesses in each error category. Then students will learn on more exercises related to their weak points, with feedbacks when committing wrong answers. Their performance on exercises will affect further exercises they will learn.

5. Analysis & Result

Before using extracted quizzes in the system, we analyzed quality of the quizzes. There are two concerns, which are number of quizzes in each category and completeness of sentences. The number of quizzes is important, because the more quizzes we can extract, the more exercises the students can practice. The completeness shows how much the extracted sentence is understandable without its context appearing in the original corpus.

From the analysis details shown in Table 1, we chose the categories, which have enough number of quizzes, high unintelligibility, and completeness. We decided to construct the system on 6 error categories: normal preposition, collocation, adjective lexis, adverb lexis, verb forms, and Japanese English. For collocation and adverb lexis category, further adaptation is needed to have better completeness.

Table 1 A part of error categories with number of quizzes, unintelligibility of error, and completeness of sentence

Error Category	Count	Unintel ^[1]	Completeness ^[2]
Article	1837**	0.73	○
Normal preposition	475**	1.15*	◎
Verb lexis	404**	0.87	△
Noun lexis	302**	1.93*	×
Collocation	257**	1.15*	○
Subj-verb agreement	234**	0.64	○
Adjective lexis	216**	1.42*	○
Adverb lexis	144	2.4*	◎
Conjunction lexis	133	1.21*	△
Word order	73	1.44*	△
Verb form	55	2.03*	◎
Japanese English	36	1.27*	◎

[1] Proportion of frequency of the category indicated as unintelligible utterance compared to the frequency of the same category in all data (normalized per 1,000 words) (Izumi 2007)

[2] Completeness judged by proportion of context-independent understandable sentences in all test data (randomly picked)

* indicates unintelligibility that is higher than 1.

** indicates number of quizzes that is more than 200 per category.

Examples of erroneous sentences and correct sentences from 3 of the 6 categories selected are

shown in Table 2. Learners' task is to correct erroneous sentences into the correct ones.

Table 2 Examples of sentences of the 6 categories selected

Erroneous Sentence	Correct Sentence	Error Category
So she decided to buy foods <i>to</i> a discount supermarket.	So she decided to buy foods <i>at</i> a discount supermarket.	Preposition
And the cat was <i>like</i> killed.	And the cat was <i>almost</i> killed.	Adverb lexis
And Hanako was eating <i>soft cream</i> too.	And Hanako was eating <i>soft ice cream</i> too.	Japanese English

6. Conclusion & Future Work

We analyzed quizzes extracted from the corpus and chose error categories that produce enough number of quizzes, have high unintelligibility level, and complete without additional contexts. These will be used to construct the intelligent tutoring system in our future work. Evaluation will also be done by testing awareness of errors as stated in the goal of the system.

References

- Hendrickson, J. (1978). Error correction in foreign language teaching: Recent research and practice. *Modern Language Journal*, v.62, p.387-398.
- Izumi, E., Uchimoto, K. & Isahara, H. (2007). Vocabulary usage and intelligibility in learner language. In Davies, M., Rayson, P., Hunston, S. & Danielsson, P. (Ed.), *Proceedings of the Corpus Linguistics Conference*. Birmingham, UK: The University of Birmingham.
- Leeman, J. (2007). Feedback in L2 learning: Responding to errors during practice. In DeKeyser, R. (Ed.), *Practice in a second language. Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press, p.111-137.
- MacWhinney, B. (1987). The competition model. In MacWhinney, B. (Ed.), *Mechanisms of language acquisition*, New Jersey: Lawrence Erlbaum Associates, Inc., p.249-308.
- Shahin, N. (2011). Error treatment in TESOL classrooms. *J. Appl. Sci.: Humanities Series* 13 (1), p.207-226.
- Thisayakorn, P., Nishihara, A. (2011). The computer adaptive English tutoring system focusing on errors often made by Japanese learners, In *JSET 2011 Conference Proceedings*, Tokyo.
- Tono, Y., Izumi, E., Kaneko, E. (2005). The NICT JLE Corpus: the final report. In Bradford-Watts, K., Ikeguchi, C., Swanson, M. (Eds.), *JALT 2004 Conference Proceedings*, Tokyo, p.345-356.
- U.S. Department of Education, Office of Planning, Evaluation, and Policy Development (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*.