GenieTutor: a computer assisted second-language learning system based on semantic and grammar correctness evaluations

Oh-Woog Kwon[1], Kiyoung Lee[2], Young-Kil Kim[3], and Yunkeun Lee[4]

Abstract. This paper introduces a Dialog-Based Computer-Assisted second-Language Learning (DB-CALL) system using semantic and grammar correctness evaluations and the results of its experiment. While the system dialogues with English learners about a given topic, it automatically evaluates the grammar and content properness of their English utterances, then gives corrective feedback on grammar and semantics. The system consists of a non-native optimized speech recognition module and a semantic/grammar correctness evaluation based tutoring module. The tutoring module decides to continue the dialogue or asks learners to try again by evaluating semantic correctness of their utterances, and also gives them turn-by-turn semantic and grammatical corrective feedback. The semantic correctness evaluation consists of a 2-classes classifier for the 'pass or try again' and a 6-classes classifier for semantic corrective feedback, using the domain knowledge and language model. The grammatical correctness is evaluated by a hybrid grammatical error correction system composed of four approaches: a rule-based, a machine learning-based, an n-gram based, and an edit distance based approach. In the experiments, in which 30 subjects in a real environment took part, we acknowledged that the 'pass or try again' evaluation has a success rate of 97.5%, the semantic feedback classification has a success rate of 87.8%, and the precision and recall for grammar error correction are 79.2% and 60.9%, respectively.

1. Electronics and Telecommunications Research Institute, Korea; ohwoog@etri.re.kr
2. Electronics and Telecommunications Research Institute, Korea; leeky@etri.re.kr
3. Electronics and Telecommunications Research Institute, Korea; kimyk@etri.re.kr
4. Electronics and Telecommunications Research Institute, Korea; yklee@etri.re.kr

How to cite this article: Kwon, O.-W., Lee, K., Kim, Y.-K., & Lee, Y. (2015). GenieTutor: a computer assisted second-language learning system based on semantic and grammar correctness evaluations. In F. Helm, L. Bradley, M. Guarda, & S. Thouësny (Eds), *Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (pp. 330-335). Dublin: Research-publishing.net. http://dx.doi.org/10.14705/rpnet.2015.000354

Keywords: computer-assisted second-language learning system, grammar error correction, semantic correctness evaluation, non-native-optimized speech recognition.

1. Introduction

For second language learning, the use of language learning software is recently considered as natural. The software is very important in blended learning environments and individual learning for second language learning. Most second language learning systems mainly focus on vocabulary memorization, pronunciation practice, grammar acquisition, and simple repetition of given conversation. These one-way teaching-learning and simple repetition learning methods cannot attract voluntary participation from learners. To overcome the shortcomings, we have been investigating an interactive system which plays the role of the language tutor and native friends, using spoken dialog processing and natural language processing technologies. We developed a DB-CALL system, called GenieTutor[5], which has a conversation with learners and gives them semantic and grammar corrective feedback. GenieTutor does not currently provide language learning and practices for free conversation which is suitable for learners with high proficiency levels. However, GenieTutor allows learners to freely speak whatever learners think within the fixed scenario and provides corrective feedback. GenieTutor has been developed with the purpose of assisting learners with low and middle levels to achieve higher proficiency levels. Although GenieTutor was developed for English learning, we plan to extend GenieTutor for Korean learning and also improve the technologies of GenieTutor for free conversation.

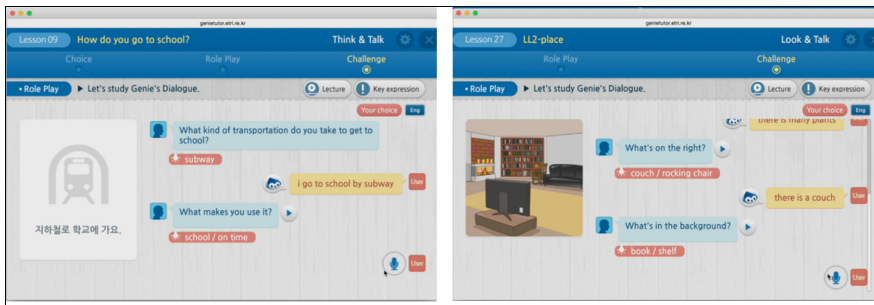
2. GenieTutor

GenieTutor is a role-play dialogue system for second language learners that uses a spoken dialog understanding technology. GenieTutor promotes dialogue with learners on through two types of English learning stages, called Think & Talk and Look & Talk. In Think & Talk, each topic consists of several fixed role-play dialogues. Learners first select a topic, and then select their favorite scenario among the several scenarios available. GenieTutor and learners have a communication based on the selected role-play scenario. In Look & Talk, learners select a picture and then GenieTutor asks the learner to describe the picture. Figure 1 shows the dialogue exercises in the Think & Talk and the Look & Talk. Although GenieTutor runs according to a fixed scenario of a given topic, it allows learners to freely speak

5. Information about GenieTutor can be found at http://genietutor.etri.re.kr/index.asp

with diverse responses to each utterance of GenieTutor, then evaluates semantically and grammatically the response and provides feedback on semantic and grammar correctness.

Figure 1. Examples of dialogue exercises in Think & Talk and Look & Talk



(a) A dialogue exercise of the lesson 9 'How do you go to school?' in the stage Think & Talk

(b) A dialogue exercise of the lesson 27 'Place' in the stage Look & Talk

The schematic diagram of GenieTutor consists of Automatic Speech Recognition (ASR) and tutoring modules. We optimized the ASR module to recognize the English utterances of Korean learners as well as native speakers' utterances (Chung, Lee, & Lee, 2014; Lee, Kang, Chung, & Lee, 2014), and also to recognize grammatically wrong sentences uttered by learners (Kwon et al., 2015). To minimize the effects of ASR errors, GenieTutor forces learners to confirm their utterances recognized by the ASR module. If they are not correct, learners speak again or edit the wrong sentences to correct themselves.

The tutoring module consists of semantic and grammar correctness evaluation, turn-by-turn feedback generation, and overall feedback generation. Semantic and grammar correctness evaluation evaluates the semantic properness of learners' responses which is appropriate to previous utterances of GenieTutor, and detects their grammar errors and finds the corrections. The semantic correctness evaluation classifies learners' responses into 6 classes ("perfect", "too few modifiers", "inflection error", "subject-verb error", "keyword error", and "illegal expression"), using the domain knowledge and language model. The 6 classes are defined as follows:

- "Perfect" class: the utterance is semantically perfect in the dialogue context.
- "Too few modifiers" class: the utterance is semantically good, but has the modifier mistakes.

- "Inflection error" class: the utterance is semantically good, but has the inflection mistakes.

- "Subject-verb error" class: its subject and main verb aren't semantically appropriate as the response of the previous utterance.

- "Keyword error" class: learners omit to speak some important words which are necessary for the response to the previous utterance of GenieTutor.

- "Illegal expression" class: the utterance has some wrong English expressions.

The grammar correctness evaluation is performed by a hybrid grammatical error correction system composed of four approaches: a rule-based, a machine learning-based, an n-gram based, and an edit distance based approach. Because of false alarms in which correct words are detected as errors is critical in second language learning, the hybrid system was aimed to decrease the false alarms by filtering implausible correction candidates using votes of several different methods (Lee, Kwon, Kim, & Lee, 2015).

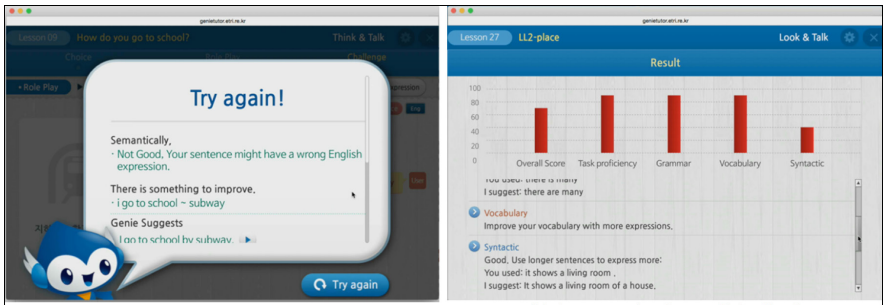
The turn-by-turn feedback generation generates the corrective feedback in a step-by-step and sequential manner. It firstly shows pass or fail feedback to the learner using the results of semantic correctness evaluation. If the result is "perfect", "too few modifiers", or "inflection error", pass feedback is generated, and otherwise, fail feedback is generated. If the learners' utterance is "perfect", it doesn't generate any feedback, but otherwise, it shows the result of semantic correctness evaluation and grammar error words detected by grammar error correction. The last part is the corrective feedback which consists of some recommendation sentences as semantic corrective feedback and the corrections of grammatical errors described in the second part. Figure 2 (a) shows an example of turn-by-turn corrective feedback when a learner replied "I go to school subway." to the system question "How do you go to school?" in the stage Think & Talk.

Once the dialogue is over, the module provides an overall feedback consisting of 4 assessments for "task proficiency", "grammar accuracy", "vocabulary diversity", and "syntactic complexity" to show which part the learner should focus more on (Kwon et al., 2015). Figure 2 (b) shows an example of overall feedback after a dialogue ended in the lesson 27 'Place'.

GenieTutor provides an authoring tool that enables English teachers and course designers to construct new topics and role-play scenarios for the Think & Talk

and the Look & Talk. The authoring tool also provides the customizing semantic and grammar correctness evaluation into the new scenarios. The customization is simply performed by training the domain knowledge and language model from the new scenarios and human annotated keywords for each utterance.

Figure 2. The educational feedback of GenieTutor



(a) An example of turn-by-turn corrective feedback (b) An example of overall feedback

3. Experiments

To evaluate the semantic and grammar correctness evaluation component of GenieTutor, we made a semantic evaluation set consisting of 3,024 utterances and a grammar error correction evaluation set consisting of 858 sentences which were randomly selected from dialogues produced by about 50 Korean learners using GenieTutor over two months. The students were college students or college graduates.

In the experiments for semantic evaluation, the pass or try again evaluation has a success rate of 94.1% and the semantic feedback (6 categories) classification has a success rate of 85.5%. In the experiments for grammar error correction evaluation, we achieved a precision of 91.3% with a recall of 45.1%.

Experiments in a real environment were also conducted, and 30 subjects were recruited (15 subjects had TOEIC scores lower than 500, while 15 subjects had TOEIC scores between 500 and 900). Each subject had a dialogue with GenieTutor on 30 learning topics of Think & Talk and Look & Talk. Contrary to our expectations, the results were very similar across the two groups. The experiments showed that the pass or try again evaluation has a success rate of 97.5%, the semantic feedback classification had a success rate of 87.8%, and the precision and recall for grammar error correction are 79.2% and 60.9%, respectively.

4. Conclusion

This paper described GenieTutor - a DB-CALL system based on semantic and grammar correctness evaluations and learner performance. GenieTutor has a fixed list dialogue flow given a lesson (topic), so it doesn't generate diverse utterances to respond to the learner. Despite this it allows learners to freely respond to the fixed questions on the given topic. GenieTutor evaluates semantically and grammatically the learners' freely spoken utterances, then decides to continue the conversation or requests the learner to respond again after providing feedback such as semantic evaluation results, grammatical error correction results, and some recommendation sentences. Through the experiments of semantic and grammar correctness evaluation, the evaluation showed good performance to provide educational feedback to second-language learners. However, we did not explore the extent to which the feedback provided by GenieTutor is useful to learn a second language, and whether second language learning using GenieTutor can improve learners' language skills. In the near future we plan to evaluate the effectiveness of English learning using GenieTutor.

5. Acknowledgements

This work was supported by the ICT R&D program of MSIP/IITP [10035252, Development of dialog-based spontaneous speech interface technology on mobile platform].

References

Chung, H., Lee, S. J., & Lee, Y. K. (2014). Weighted finite state transducer-based endpoint detection using probabilistic decision logic. *ETRI Journal, 36*(5), 714-720. doi:10.4218/etrij.14.2214.0030

Kwon, O. W., Lee, K., Roh, Y.-H., Huang, J.-X., Choi, S.-K., Kim, Y.-K., Jeon, H. B., Oh, Y. R., Lee, Y.-K., Kang, B. O., Chung, E., Park, J. G., & Lee, Y. (2015). GenieTutor: a computer assisted second-language learning system based on spoken language understanding. *IWSDS, 2015*. Retrieved from http://www.uni-ulm.de/fileadmin/website_uni_ulm/allgemein/2015_iwsds/iwsds2015_submission_29.pdf

Lee, S. J., Kang, B. O., Chung, H., & Lee, Y. (2014). Intra- and inter-frame features for automatic speech recognition. *ETRI Journal, 36*(3), 514-517. doi:10.4218/etrij.14.0213.0181

Lee, K, Kwon, O.-W., Kim, Y.-K., & Lee, Y. (2015). A hybrid approach for correcting grammatical errors. In F. Helm, L. Bradley, M. Guarda, & S. Thouësny (Eds), *Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (pp. 362-367). Dublin Ireland: Research-publishing.net. doi:10.14705/rpnet.2015.000359

Published by Research-publishing.net, not-for-profit association
Dublin, Ireland; info@research-publishing.net

Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy
Edited by Francesca Helm, Linda Bradley, Marta Guarda, and Sylvie Thouësny

Typeset by Research-publishing.net
Fonts used are licensed under a SIL Open Font License

ISBN13: 978-1-908416-28-5 (Paperback - Print on demand, black and white)
Print on demand technology is a high-quality, innovative and ecological printing method; with which the book is never 'out of stock' or 'out of print'.

ISBN13: 978-1-908416-29-2 (Ebook, PDF, colour)
ISBN13: 978-1-908416-30-8 (Ebook, EPUB, colour)

Legal deposit, Ireland: The National Library of Ireland, The Library of Trinity College, The Library of the University of Limerick, The Library of Dublin City University, The Library of NUI Cork, The Library of NUI Maynooth, The Library of University College Dublin, The Library of NUI Galway.

Legal deposit, United Kingdom: The British Library.
British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.

Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: décembre 2015.