Development of a Dictogloss System Oriented for Focus on Form

Makoto KONDO[a*], Ryo SANO[b], Asanori TASHIRO[a], Yasuhiro NOGUCHI[a], Satoru KOGURE[a], Tatsuhiro KONISHI[a] & Yukihiro ITOH[c]
[a]*Faculty of Informatics, Shizuoka University, Japan*
[b]*Graduate School of Informatics, Shizuoka University, Japan*
[c]*Shizuoka University, Japna*
*mkondo@inf.shizuoka.ac.jp

Abstract: In this paper, we discuss how to develop a dictogloss system oriented for focus-on-form (FonF) instruction. Dictogloss is a multiple skills collaborative activity in which learners listen to a text and try to reconstruct the text in collaboration with their partners. Focus on form is a pedagogical approach aiming at improving learners' ability to produce grammatically correct sentences. A dictogloss system incorporating FonF instruction would therefore be of great help for improving learners' correct understanding and production of their target languages. We discuss what functions are necessary for a dictogloss system and how to implement the functions. Preliminary evaluation of our system shows that the system can act as a dictogloss partner for second/foreign language learners.

Keywords: Dictogloss, focus on form, second/foreign language education

Introduction

Dictogloss is a multiple skills collaborative activity proposed by Wajnyrb [15]. In a dictogloss activity, a teacher reads a short text to learners and the learners try to reconstruct the contents of the text. Dictogloss is similar to a traditional dictation activity in that both activities require learners to reproduce what they have listened to. Dictogloss however differs from dictation in some important aspects. While dictation requires learners to replicate a dictated text word for word, dictogloss encourages learners to use their own linguistic and grammatical knowledge to produce a parallel text. In other words, learners can make use of different linguistic forms from the original text as far as the reproduced text is grammatical and has the same contents as the original one. Another important difference between those activities is that dictogloss requires learners to reconstruct the original contents with their peers. Dictogloss therefore promotes collaboration among learners, which contrasts sharply with a traditional dictation activity, in which each learner is required to complete the task by him/herself.

Although there are several ways to put dictogloss into practice, a typical dictogloss activity involves the following three stages according to Izumi [3].

(1) A teacher reads a short text to learners twice at a natural speed (or plays CD-recorded sound of the text). On the first listening, the learners are instructed to grasp the gist of the text but not allowed to take any notes. On the second listening, they are instructed to jot down key words/phrases in the text.
(2) The learners make small groups and reconstruct the original contents by sharing and referring to their notes. They are instructed to speak in the target language.

(3) The reconstructed text of each group is compared with the original text by focusing on grammatical and semantic aspects of the differences.

Let us refer to those stages as (1) dictation stage, (2) reconstruction stage, and (3) analysis and correction stage, respectively.

A number of studies support the use of dictogloss in second/foreign language education [8,9,11,12,13]. Dictogloss has at least two very important properties. One is that learners use all four language skills (listening, speaking, reading, and writing) in order to complete a dictogloss activity. In the dictation stage, learners listen to their teacher read a text. In the reconstruction stage, they speak to their groupmates in the target language, and write a reconstructed version of the text. In the analysis and correction stage, they must read the original text.

Another important property is that dictogloss provides learners with opportunity to reflect on their use of the target language. In order to complete a dictogloss task, learners must reconstruct the contents of the original text. In the reconstruction, they talk about the language of the text they are reconstructing. Kowal and Swain [6,7] have found that dictogloss activities have elicited metalinguistic talks from learners; that is, learners talk about form-meaning/form-function relations in their target language.

The second property is of particular importance from the viewpoint of a pedagogical approach called focus on form (FonF). FonF has attracted much attention because it could solve a potential problem of another pervasively adopted approach called communicative approach (CA) [1]. While FonF aims at improving learners' ability to produce grammatically correct sentences, the CA puts a higher priority on conveying a speaker's intention than on making grammatically correct utterances. The CA therefore has a risk that learners would acquire incorrect grammatical rules for their target languages. FonF-based dictogloss is an effective activity to promote learners' correct understanding and production of their target languages.

Since dictogloss is a collaborative activity, it requires the presence of a partner. If a language education system plays a role of a dictogloss partner, it would be of great help to second/foreign language learners. In sections 1, 2 and 3, we discuss what functions are necessary in each of the three stages of dictogloss activities, respectively. We also discuss how we have implemented each of the functions. Section 4 shows the result of preliminary evaluations and Section 5 gives the summary of what has been achieved and some remaining issues for the future work.

1. Functions for Dictation Stage

A dictogloss system must be able to play the sound of a dictogloss text in the dictation stage. There are several ways to implement this function. The simplest way is to have sound files of dictogloss texts recorded by native speakers of target languages. A technically more interesting/challenging way is to automatically generate sound from dictogloss texts. The latter approach is beyond the scope of this paper and we have chosen the former approach.

In the dictation stage, learners listen to a dictogloss text a designated number of times; that is, they should not be able to listen to the text as many times as they like. At the same time, rewind and fast-forwarding functions should not be included in the system. Accordingly, our system has dictogloss texts and corresponding sound files; the GUI for the dictation stage only has a "play" button that can be used only a designated number of times.

One can add a note-taking function to the system; however, effectiveness of this function heavily depends on learners' typing skills. We have therefore omitted the note-taking function from our system and decided to let learners freely take notes on a sheet of paper.

2. Functions for Reconstruction Stage

In the reconstruction stage, the system must play a role of a dictogloss partner. At least two functions are necessary for the system: a reconstruction function and a dialog function. The reconstruction function generates a dictogloss partner's answer to a given dictogloss task. Let us call it a system answer. The system answer is shared by learners to complete the task. The dialog function allows the learners to collaborate with the system in reconstructing the contents of the original text. Section 2.1 and section 2.2 describes how we have designed the reconstruction function and the dialog function, respectively.

2.1 Reconstruction Function

Ideally, we want learners to notice errors in their answers by themselves through the reconstruction process with the system. In order to help them notice their errors, system answers should be generated in such a way that learners easily notice their errors by comparing their answers with the system answers. The system therefore should be capable of generating different system answers according to which part of learners' answers involves errors.

In order for the system to behave differently in accordance with learners' answers, the system must be capable of detecting errors in learners' answers. Notice that a simple matching between learners' answers and the original text does not suffice for this purpose. This is because natural language generally has more than one way to express a single semantic content. Learners' answers may be perfectly grammatical and semantically well-formed even if surface forms in the answers are quite different from the ones in the original text. In other words, the system must be able to perform semantic comparison between learners' answers and the original text. Section 2.1.1 explains how to detect errors based on semantic comparison. Section 2.1.2 describes how to generate different system answers in accordance with learners' errors.

2.1.1 Error Detection based on Semantic Comparison

We have employed JDT semantic representation [2,14], which enables the system to conduct semantic comparison. In the JDT semantic representation, meanings of content words (verbs, nouns, etc.) are represented by concept frames containing attribute-value pairs, and meanings of function words (case particles, auxiliary verbs, etc.) are represented as attributes or markers attached to frames. Dependency relations between content words are represented by pointers which link attribute values to the concept frames denoting the values. Fig. 1 shows an example semantic representation for *Hoteru-o sagashi-te* (*Find [me] a hotel*), where markers are given in square brackets. (Unlike English, Japanese allows phonetically null subjects/objects. We put them into square brackets in English translation.) For the sake of simple illustration, we omit irrelevant details throughout this paper.

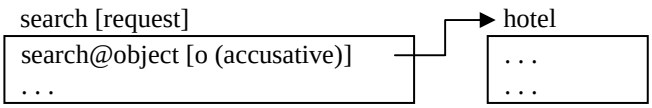


Fig. 1: Semantic Representation of *Hoteru-o sagashi-te* (*Find [me] a hotel*)

The JDT semantic representation enables semantic comparison based on attribute-value pairs. If two sentences have the same semantic content, they have the same set of attribute-value pairs in their JDT semantic representations. Consequently the JDT semantic representation enables the system to perform semantic comparison between two

sentences with different syntactic structure by comparing the attribute-value pairs contained in each of the semantic representations.

Kondo et al. [5] have extended the JDT semantic representation to develop what they call situation knowledge. Situation knowledge is a set of JDT semantic representations in which concept frames denoting the same concept are integrated into one frame. Fig. 2 shows the situation knowledge associated with two sentences: Tokyo-no hoteru-ni tomari-tai ([I] want to stay at a hotel in Tokyo) and Yasui hoteru-o sagashite (Find [me] a cheap hotel). In Fig. 2, the meaning of yasui (cheap) is represented by the rate-possession frame, and the value of the rate-possession@object attribute, "- (minus)", is transferred to the value of the same attribute in the hotel frame based on the fact that they are the same attribute.

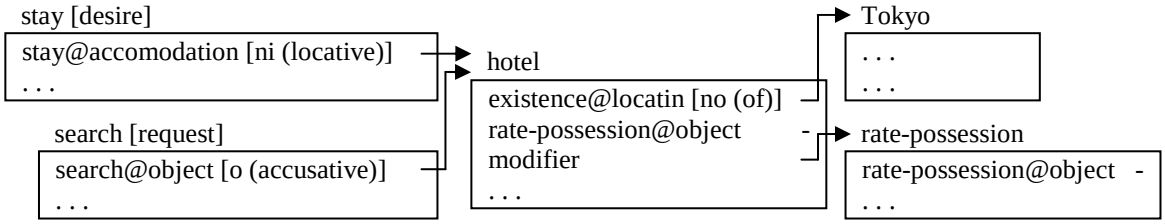


Fig. 2: Sample Situation Knowledge

We have adopted this extension of JDT semantic representation in order to represent the semantic contents of the original dictogloss text. By storing the original semantic contents in this manner, the system correctly matches learners' answers like *Yasui hoteru-ni tomari-tai* (*[I] want to stay at a cheap hotel*) and *Tokyo-no hoteru-o sagashi-te* (*Find [me] a hotel in Tokyo*) with the original semantic contents as well as those sentences in the original text.

In order to implement the reconstruction function, the system must be able to detect errors in learners' answers since the system should behave differently according to errors in learners' answers. Kondo et al. [5] and Sano et al. [10] have developed error judgment technique based on the JDT semantic representation and its extension discussed above. We have employed the technique in our system and the system can detects the following four types of errors.

(1) Erroneous omission: Learners incorrectly omit necessary linguistic forms (e.g., omission of a case particle).
(2) Erroneous addition: Learners incorrectly add unnecessary linguistic forms (e.g., addition of the past verbal suffix "*-ta*" when simple present tense is appropriate).
(3) Confusion of different linguistic forms: Learners mistake a linguistic form for another form (e.g., confusion of a progressive "verb-*teiru*" form and a perfective "verb-*tearu*" form).
(4) Incorrect word order

Basic idea behind the technique is that errors in learners' answers would result in difference between their semantic representations and the semantic representations of the original text. If a learner fails to refer to some entity, event, or property of an entity/event, the semantic representation of the learner's answer does not have the corresponding concept frame. If a learner fails to reconstruct the meaning denoted by a function word, it results in the semantic representation in which the corresponding marker is absent. If a learner's answer involves erroneous addition, the semantic representation of the learner's answer contains the corresponding concept frames/markers, which is absent in the semantic representation of the original text does not. Confusion errors are detected by combination of erroneous omission and addition. If erroneous omission and addition are detected simultaneously in the same position, the sentence should involve confusion of the omitted form and the added form. Incorrect word order is also detected by the combination of

erroneous omission and addition because this type of error results in the semantic representation in which erroneous omission of a form is detected in one position and erroneous addition of the same form is simultaneously detected in another position.

2.1.2 Generation of the System Answers

While the system should generate different answers depending on which part of learners' answers involves errors, it should also behave differently according to types of forms involving errors. Since FonF instruction focuses on a few linguistic forms in a lesson, the system should behave differently according to whether errors are involved in focused forms or not. In addition to distinction between focused forms and non-focused forms, we divide non-focused forms into three categories: key words/phrases, FonF forms, and other forms. Accordingly, we divide forms in a given dictogloss text into four categories: (1) forms focused in a lesson employing the dictogloss text (focused forms), (2) key words/phrases in the text, (3) forms suitable for FonF instruction (FonF forms), and (4) other forms. Since the purpose of a FonF-based dictogloss is to improve learners' grammatical correctness in using focused forms, these are the forms which should be given the highest priority. The second priority should be given to key words/phrases in a text. Those are important in understanding semantic contents of the text. Since the very first proposal of dictogloss [15], it has been assumed that learners should be prepared for the vocabulary in a dictogloss text. Kondo et al. [4] have selected 159 FonF forms for FonF instruction. Since these forms are suitable for FonF instruction, they receive the third priority.

We have implemented the following rules for generation of system answers according to whether a learner's answer involves errors of each category mentioned above.

(1a) If a learner correctly uses a focused form, the system generates an answer involving erroneous omission of the focused form.
(1b) If a learner incorrectly uses a focused form, the system generates an answer involving a confusion error of the focused form.
(2a) If a learner correctly uses a key word/phrase, the system generates an answer involving erroneous omission of the key word/phrase.
(2b) If a learner incorrectly uses a key word/phrase, the system generates an answer involving the correct use of the key word/phrase.
(3a) If a learner correctly uses a FonF form, the system generates an answer involving the correct use of the FonF form.
(3b) If a learner incorrectly uses a FonF form, the system generates an answer involving a confusion error of the FonF form.
(4a) If a learner correctly uses one of other forms, the system generates an answer involving the correct use of the form.
(4b) If a learner incorrectly uses one of other forms, the system generates an answer involving the correct use of the form.

In (1a,b), the system always generates an erroneous answer in order to induce as much interaction between the system and the learner as possible. This is because focused forms have the highest priority in a given lesson and it is desirable to induce as much discussion on them as possible. In (2a,b), the system generates an erroneous answer if and only if the learner' answer is correct; otherwise, it generates a correct answer. This is because an erroneous system answer would induce interaction on the key word/phrase if the learner correctly uses it. At the same time, a learner's erroneous use of a key word/phrase suggests that the learner needs some help for their correct use; hence, the system should generate a correct answer. In (3a,b), the system generates an erroneous answer if and only if the learner's answer involves an error, because discussion on non-focused FonF forms is unnecessary if the learner correctly uses them. In (4a, b), the system always generates a

correct answer because forms in this group have the lowest priority and discussion on them should be made as little as possible.

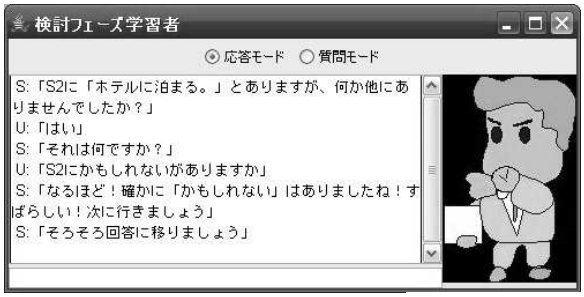
2.2 Dialog Function

In the reconstruction stage, the system and the learner collaborate to make the final answer. A learner engages in a dialog with the system through keyboard input. A leaner makes his/her question to the system or answer to the system's question using predetermined templates. Currently, the system has one question template and two answer templates. The question template is "Does S*n* have *a form*?", where "S*n*" stands for "sentence *n*" in the answer (each sentence in the answer is given a unique sentence number), and "*a form*" is replaced by any linguistic form of the learner's choice. As for answer templates, a learner can use a simple "Yes/No" answer.

The collaboration process varies depending on which participant makes an error. We must therefore consider the following four cases: (1) both a learner's answer and the system answer are correct, (2) neither a learner's answer nor the system answer is correct, (3) a learner's answer is correct whereas the system answer is not, and (4) a learner's answer is incorrect while the system answer is correct. The first case does not require any further collaboration. The forth case does not require further collaboration, either. This is because the system shows a correct answer and we can expect that the learner would notice his/her error by comparing his/her answer with the system answer. Accordingly, let us examine the remaining two cases.

When neither a learner's answer nor the system answer is correct, the system ask the user whether the sentence should have any other form: "Your Sn is '. . . (the learner's answer)'. Does it have anything else?" If the learner correctly points out the correct form by using the question template ("Does Sn have the correct form?"), the system thanks the learner and proceeds to the next sentence. If the learner's answer is "No", the system proceeds to the next sentence after expressing a doubt about the learner's answer: "Something might be wrong with the answer, but let's go on to the next sentence" or "There might be something else, but let's go on to the next sentence".

When a learner's answer is correct and the system answer is not, we can expect that the learner would point out the error in the system answer. If the learner points out the correct form by using the question template, the system thanks the learner and proceeds to the next sentence. If a learner does not initiate a dialog, the system asks the learner whether the correct form should be included in the answer. If the learner's answer to that question is positive, the system thanks the learner and proceeds to the next sentence; otherwise, the system expresses a doubt about the answer and go on to the next sentence. This strategy is implemented for focused forms alone. This case occurs if and only if the form in question is a focused form or a key word/phrase. In the latter case, a learner knows that the key word/phrase must be used somewhere in the answer. Since the system's incorrect answer in this case involves erroneous omission of the key word/phrase, we can reasonably expect that the learner would initiate a dialog. Fig. 3 shows a dialog example. In Fig. 3, the learner's answer does not involve the focused form "kamoshirenai (might)". The system detects the omission error since the marker associated with the form is absent in the semantic representation of the learner's answer. The system then produces the question "Does it have anything else?" in the first line of the example. Since the learner correctly points out, in his/her second utterance, that the answer should involve the focused form, the system thanks the learner and proceeds to the next sentence.



[Translation] (S: system, U: user/learner)
S: Your S2 is "*Hoteru-ni tomaru* (I will stay at the hotel)". Does it have anything else?
U: Yes.
S: What is it?
U: Does S2 have *kamoshirenai* (might)?
S: Yes! It certainly has *kamoshirenai*. Great! Let's proceed to the next.
S: Let's submit the final answer, shall we?

Fig. 3: Example Dialog

3. Functions for Analysis and Correction Stage

In the analysis and correction stage, the system shows the learner's final answer and the correct answer. Focused forms are highlighted in the correct answer so that the learner can confirm his/her use of the focused forms. If a learner makes an error with regard to focused forms, the system shows grammatical explanations about the forms. Four types of errors discussed in section 2.1.1 are also highlighted by giving them different font colors, and the system shows what type of error is involved in each of the sentences.

4. Preliminary Evaluation

We have implemented the functions for dictogloss activities discussed in sections 1, 2 and 3, and conducted two preliminary evaluations: an evaluation for generation of the system answers and an evaluation for the dialog function.

In evaluating the answer generation, we manually made dictogloss texts which include focused forms, key words/phrases, FonF forms and other forms. In addition, we manually made test sentences that involve erroneous omission, erroneous addition and confusion errors. Incorrect word order is not examined in this evaluation because the JDT system fails in morphological analysis of many sentences involving incorrect word order. The target of evaluation is characterized by whether a test sentence involves an error and what type of form involves an error. Since there are four types of forms, we examined eight cases: correct/incorrect use of focused forms, correct/incorrect use of key words/phrases, correct/incorrect use of FonF forms and correct/incorrect use of other forms. We examined each of the eight cases and confirmed that the system correctly generates correct/incorrect answers in accordance with the generation rules discussed in section 2.1.2.

In evaluating the dialog function, we fed the system with the above eight types of test sentences, and confirmed that the system generated appropriate questions to the learner in accordance with the dialog strategy discussed in section 2.2. The system also responded to the learner's answer to the system's question as desired. We also confirmed that the system replied to a question originally made by the learner.

5. Concluding Remarks

We have developed a dictogloss system oriented for FonF instruction. The system plays the sound of a dictogloss text in the dictation stage. A learner can listen to the sound only a designated number of times. In the reconstruction stage, the system generates its own answer in order to collaborate with the learner. The system answers vary depending on whether the learner's answer involves an error and what type of form is erroneous. The

system also engages in a dialog with the learner in the reconstruction stage. The system makes questions on the answer if necessary. The learner can also make his/her own question to the system using the question template. The system changes its response to the learner's question according to whether the learner reaches the correct answer. In the analysis and correction stage, the system shows the correct answer and highlights errors in their final answer.

Needless to say, there are a lot of tasks to be completed in the future work. Since the evaluation discussed in this paper is a preliminary small-scale evaluation, we need a larger-scale evaluation with respect to both the size and variety of test sentences. One of the most important tasks to be completed is improvement of the dialog function. The current system accepts questions from a learner by using the question template. If the system takes full advantage of the JDT dialog system and accepts free input from a learner, it would surely improve the usability of the system.

Acknowledgements

This work was supported by Grant-Aid for Scientific Research (C) (21520398).

References

[1] Doughty, C., & Williams, J. (Eds.). (1998). *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press.

[2] Ikegaya, Y., Noguchi, Y., Kogure, S., Konishi, T., Kondo, M., Asoh, H., Takagi, A., & Itoh, Y. (2005). Integration of dependency analysis with semantic analysis referring to the context. *Proceedings of PACLIC 19*, 107-118.

[3] Izumi, S. (2009). *"Focus on Form"-o Toriireta Atarashii Eigo Kyoiku*. Tokyo: Taishukan.

[4] Kondo, M., Daicho, Y., Sano, R., Noguchi, Y., Kogure, S., Konishi, T., & Itoh, Y. (2010). Form-wise error detection in a FonF-based language education system. *Proceedings of ICCE 2010*, 9-16.

[5] Kondo, M., Kure, U., Daicho, Y., Kogure, S., Konishi, T., & Itoh, Y. (2009). Error judgment in a language education system oriented for focus on form. *Proceedings of ICCE 2009*, 43-50.

[6] Kowal, M. and Swain, M. (1994). Using collaborative language production tasks to promote students' awareness. *Language Awareness. 3*, 73-93.

[7] Kowal, M. and Swain, M. (1997). From semantic to syntactic processing: How can we promote metalinguistic awareness in the French immersion classroom? In R. K. Johnson, & M. Swain (Eds.), *Immersion Education: International Perspectives* (pp. 284-309). Cambridge: Cambridge University Press.

[8] Lim, W. L., & Jacobs, G. M. (2001). *An Analysis of Students' Dyadic Interaction on a Dictogloss Task.* ERIC Document Reproduction Service No. ED 456 649.

[9] Nabei, T. (1996). Dictogloss: Is it an effective language learning task? *Working Papers in Educational Linguistics, 12*(1), 59-74.

[10] Sano, R., Noguchi, Y., Kogure, S., Konishi, T., Kondo, M., & Itoh, Y. (20109. Extension of error judgment component in educational dialog system oriented for focus on form. *Proceedings of the 35th Annual Conference of JSiSE*, 28-G1-2, 485-486.

[11] Storch, N. (1998). A classroom-based study: Insights from a collaborative reconstruction task. *ELT Journal*, *52*(4), 291-300.

[12] Swain, M., & Miccoli, L. S. (1994). Learning in a content-based, collaboratively structured course: The experience of an adult ESL learner. *TESL Canada Journal*. *12*(1), 15-28.

[13] Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, *82*, 320-337.

[14] Takagi, A., Asoh, H., Itoh, Y., Kondo, M., & Kobayashi, I. (2006). Semantic representation for understanding meaning based on correspondence between meanings. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *10*, 876-912.

[15] Wajnryb, R. (1990). *Grammar Dictation*. Oxford: Oxford University Press.