S. L. Wong et al. (Eds.) (2010). Proceedings of the 18th International Conference on Computers in Education. Putrajaya, Malaysia: Asia-Pacific Society for Computers in Education.

Form-Wise Error Detection in a FonF-Based Language Education System

Makoto KONDO[a*], Yoko DAICHO[b], Ryo SANO[b], 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, Japan*
*mkondo@inf.shizuoka.ac.jp

Abstract: A language education system oriented for focus-on-form instruction should recognize which linguistic forms are focused in a particular lesson, and it should also evaluate how properly the focused forms are used in learners' utterances. This paper discusses how forms should be described in the system, how the form description helps the system recognize focused forms, and how the system evaluates learners' use of the focused forms. Preliminary evaluation of the system shows that the system correctly detects focused forms and that it successfully evaluates the use of the detected focused forms. Neither incorrect detection nor evaluation failure has been attested in its preliminary evaluation.

Keywords: Focus on form, foreign language education, dialog system

Introduction

In second language education, a pedagogical approach called focus on form (FonF) has attracted much attention because it could solve a potential problem of another pervasively adopted approach called communicative approach (CA) [2]. 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. If FonF is effectively incorporated into a CA-based education system, it should help overcome the problem.

In our previous studies, we developed a CA-based Japanese education system [10], and we have been trying to incorporate FonF into the system [9]. The system engages in role play with a learner under a given situation in which the learner must accomplish a given task. The system accepts ungrammatical input by referring to its situation knowledge. The situation knowledge is a set of semantic representations denoting what learners should convey in order to accomplish a given task. The system detects errors in learners' input by referring to the situation knowledge.

FonF instruction is performed focusing on particular linguistic forms (FonF forms). For example, FonF forms involve grammatical constructions like "verb-*te kudasai*" (denoting honorific request), "-*ga* (nominative case particle) + potential verb" (denoting ability or possibility), etc. A FonF-based language education system therefore should recognize which forms are focused in a particular lesson. In addition, the system should evaluate learners' use of the focused forms. Error detection in our previous system, however, is not form-wise; that is, the system detects every error irrespective of whether it is involved in a focused form. Since the error detection is performed by referring to the situation knowledge, storing information on focused forms in the situation knowledge would enable the system to recognize which forms are focused and to evaluate how properly the focused forms are used.

Section 1 describes the outline of the previous system and its extension for form-wise error detection. Section 2 deals with how to store information on focused forms in the situation knowledge. Section 3 explains how to evaluate learners' use of focused forms by referring to the situation knowledge. Section 4 provides the result of preliminary evaluation of the system. The final section provides a summary of this study.

1. Backgrounds and Goals

1.1 Outline of the Previous System

We are developing a Japanese education system by extending a Japanese dialog system developed by the Japanese Dialogue Tools (JDT) project [4,11]. The JDT system produces semantic representations from input sentences and accumulates the representations as context information. The system performs problem solving by referring to the context information and its problem-solving knowledge.

The JDT system is designed for Japanese native speakers, and our previous education system includes some additional modules. The outline of the previous education system is shown in Fig. 1, where the added modules are given in the gray shaded area.

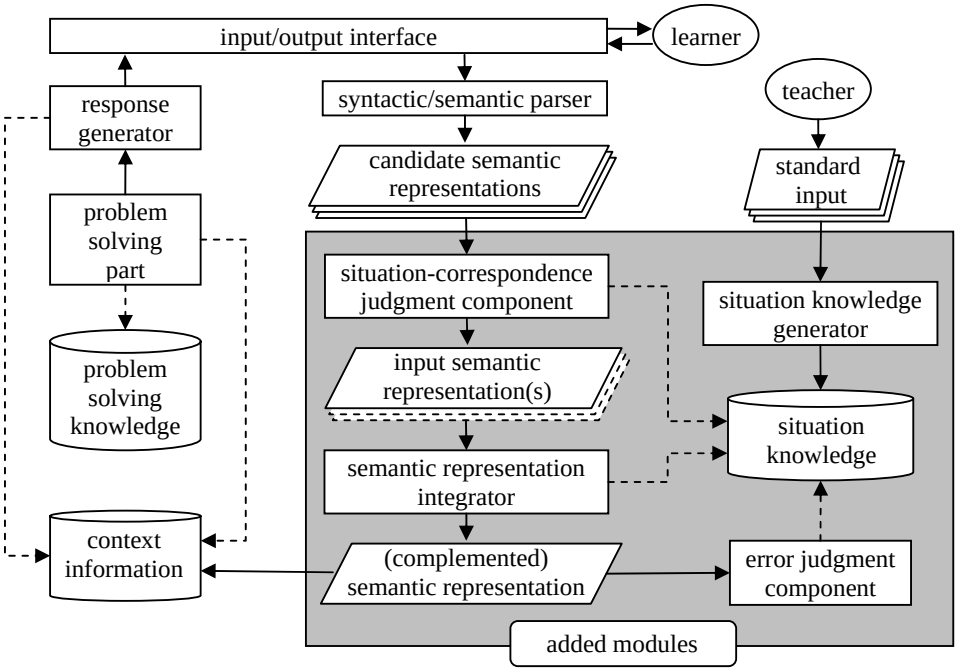

Fig. 1: Outline of the Previous System

A learner's input sentence goes through the input/output interface and is processed by the syntactic/semantic parser, which generates candidate semantic representations for the input. In the JDT semantic representation, the meanings of content words (verbs, nouns, etc.) are represented by concept frames containing attribute-value pairs, and the 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. 2 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.

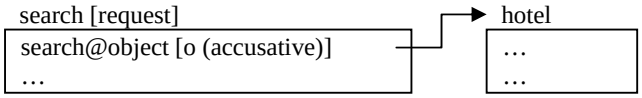

Fig. 2: Semantic Representation for Hoteru-o sagashi-te (Find [me] a hotel)

Accordingly, the system can compare the meanings of different sentences by comparing attribute-value pairs in their semantic representations. If an input involves "finding a hotel", its semantic representation has the attribute-value pair of "search@object" and "hotel" irrespective of its sentence style. As a result, the system absorbs the difference in sentence styles and accepts a wide variety of input sentences [4,11]

The candidate semantic representations go to the situation-correspondence judgment component, which compares each candidate with the situation knowledge and decides which one is the most plausible representation corresponding to the learner's intention.

The situation knowledge is generated by the situation knowledge generator. A teacher feeds the generator with the standard input, which is a set of sentences necessary for accomplishing a task in a given role-play situation. The generator produces the semantic representation of each standard input sentence. The generator then integrates the concept frames denoting the same concept into one frame. Fig. 3 shows the situation knowledge associated with two standard input sentences: *Tokyo-no hoteru-ni tomari-tai* (*[I] want to stay at a hotel in Tokyo*) and *Yasui hoteru-o sagashi-te* (*Find [me] a cheap hotel*). In Fig. 3, 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.

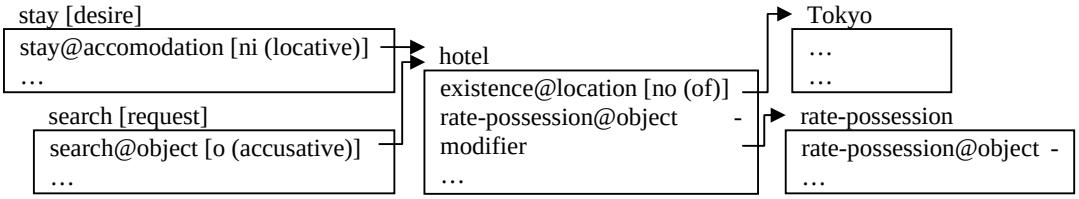

Fig: 3. Sample Situation Knowledge

The situation-correspondence judgment component correctly matches learners' input 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 situation knowledge as well as those in the standard input.

When a learner's input is grammatical, the frames in the semantic representation constitute a single tree, and the representation matches with a part of the situation knowledge. When an input is ungrammatical, the situation-correspondence judgment component divides the input into sub-trees so that each of the sub-trees matches with a part of the situation knowledge. In other words, the situation-correspondence judgment component determines which part of the situation knowledge is uttered by the learner.

The semantic representation integrator receives the input semantic representation(s) from the situation-correspondence judgment component. If the input is grammatical, the semantic representation integrator receives a single tree, and the tree trivially passes through the integrator. If the input is ungrammatical, the integrator receives more than one sub-tree. Then the integrator integrates those sub-trees into a single tree by referring to the situation knowledge. The integrator complements the sub-trees with appropriate concept frames, and integrates them into one tree which matches with a part of the situation knowledge. In other words, the integrator reproduces the learner's intention from ungrammatical input.

The output from the semantic representation integrator is accumulated in the context information and is referred to by the problem solving part and the response generator in order for the system to reply to the learner's input. At the same time, the output of the integrator goes to the error judgment component, which determines whether the input involves any errors. The error judgment component is designed by taking account of actual error patterns made by Japanese learners [3,9].

1.2 Extension for Form-Wise Error Detection

Our previous system is capable of accepting (un)grammatical input and detecting errors in learners' input, but FonF instruction needs more than those capacities. In FonF instruction, learners receive instruction when they incorrectly use focused forms. On the contrary, no instruction is necessary when nonfocused forms are erroneously used. This is because too much instruction would discourage learners from using their target languages.

Accordingly, a FonF-based education system should be able to recognize which forms are focused in a particular lesson and to evaluate whether a learner correctly uses the focused forms. In order to realize form recognition, we construct a form dictionary, which stores information on every FonF form. In addition, we construct a form detector, which searches the situation knowledge for FonF forms by referring to the form dictionary. A teacher selects forms to be focused from the detected FonF forms and the situation knowledge stores the information on which part of the knowledge corresponds to the focused forms. The form dictionary is also used in form evaluation. We extend the situation-correspondence judgment component and the error judgment component so that the former should search the candidate semantic representations for the focused forms and the latter should evaluate whether the detected FonF forms are correctly used in the (complemented) semantic representation.

2. Form Recognition

2.1 FonF Forms and Form Dictionary

We looked through Japanese textbooks for beginners and their teacher's manuals [5,6,7,8] and found that 225 forms were involved there. We selected 159 FonF forms from the 225 forms. In selecting the FonF forms, we adopted the following criteria by referring to a FonF literature [12]: (1) multiple forms corresponding to a single form in another language, (2) forms rarely used in ordinary conversation, (3) forms bearing less importance for conveying intention, and (4) forms inducing typical errors. Some linguistic forms have multiple pragmatic functions. In counting the number of forms, we treated a form with multiple functions as separate forms each of which has a single function.

In CA-based second language education, instruction is designed based on pragmatic functions like order, request, etc. Accordingly, FonF forms should be described from two different viewpoints: (1) a surface pattern of each form (form pattern) and (2) its pragmatic function (form function). The form dictionary stores the form pattern and form function of every FonF form.

Form patterns are described by a combination of 6 elements: (a) parts of speech (noun, verb, etc.), (b) surface forms of morphemes (*-ga* (nominative case particle), *-kara* (*from*), etc.), (c) a particular inflection involved in a form (attributive form, continuative form, etc.), (d) inflection types (*i*-adjective, *na*-adjective, etc.), (e) conceptual classes (place, human, etc.), and (f) word types (potential verb, honorific verb, etc.). The system can recognize any combination of these elements. The elements (a)-(d) are recognized by referring to the surface form information attached to semantic representations. The element

(e) is recognized by referring to the concept hierarchy of the system. Although the word type recognition is necessary for a FonF-based education system, it is not required in an ordinary dialog system and the JDT system does not have any word type hierarchy. Therefore we have newly constructed a word type hierarchy. The system recognizes the element (f) by referring to this hierarchy. Accordingly, we can represent every form pattern in the form of the JDT semantic representation, and the form dictionary stores form patterns represented as semantic representations.

Each FonF form has its form function in addition to its form pattern. Form functions can be divided into two classes according to whether (i) a form as a whole has a single function or rather (ii) a form function is a union of functions which come from the elements constituting the form. For example, the form *-te itadake-mase-n-ka* is made up of (inflected) forms of *-te itadaku*, *-masu*, *-n* and *-ka*, whose functions are honorific receiving of an action, politeness, negation and interrogation, respectively. On the other hand, the form as a whole has the function of polite request. Notice that the form as a whole does not have the negation function and the interrogation function, and that the request function is not denoted by any of the form-constituting elements. (The same holds true for the English sentence *Why don't you ...?* It is used as a suggestion instead of a question asking the reason.) The form dictionary explicitly stores form functions of type (i) in addition to form patterns. Form-constituting elements in a form pattern have their own functions and these functions are represented in form patterns. Form functions of type (ii) are synthesized from these functions represented in form patterns.

We fed the system with input sentences containing the selected 159 FonF forms, and found that the system generated 151 correct semantic representations. Among the 8 failures, 3 cases were due to incorrect morphological analysis and 5 cases were due to the JDT framework; the JDT framework has not established the way to represent these 5 natural language expressions. (All the 5 cases involve parallel arrangement of phrases.) Since the reasons for the 8 failures are irrelevant to what is being proposed in this paper, let us put the 8 cases aside. In what follows, the discussion will be concerned only with the 151 FonF forms whose semantic representations are correctly generated by the system.

2.2 Form Detector

The error judgment component in our previous system compares the (complemented) semantic representation with the situation knowledge, and detects the difference between them. Therefore, if we extend the situation knowledge so that it should contain information on focused FonF forms, then the system can recognize which forms are focused in a particular lesson. Although the JDT semantic representation holds information on surface forms, the situation knowledge in our previous system deletes the information. This is because semantic equivalence is enough for CA-based instruction. Since FonF instruction needs information on surface forms, we have extended the situation knowledge and the knowledge now holds surface form information.

The situation knowledge is generated by using the standard input from a teacher. The standard input is a set of sentences necessary for accomplishing a task in a given role play situation. The form detector searches the situation knowledge for FonF forms by referring to the form dictionary. The teacher then selects FonF forms to be focused in his/her lesson and the situation knowledge stores information on which part of the knowledge corresponds to the focused forms.

Since the JDT semantic representation has a tree structure, each form pattern in the form dictionary also has a tree structure headed by a concept frame. Accordingly, the detector performs detection of each form pattern in the following manner. (1) The detector picks up the head of a form pattern and detects corresponding frames in the situation

knowledge. (2) The detector compares the head with each of the detected frames with respect to conceptual classes of the frames, markers attached to the frames, attributes in the frames, and pointers connecting the corresponding attributes and their values. (3) The detector recursively compares the corresponding value frames. In comparing a form pattern with the situation knowledge, the detector also checks whether their surface forms match with each other. Finally, if the form pattern matches with a part of the situation knowledge, the detector judges that the form is used in the corresponding part of the situation knowledge. Fig. 4 describes how the form *-te kudasai* (denoting honorific request) matches with the semantic representation of *Hoteru-o yoyaku-shi-te-kudasai* (*Would you reserve a hotel?*).

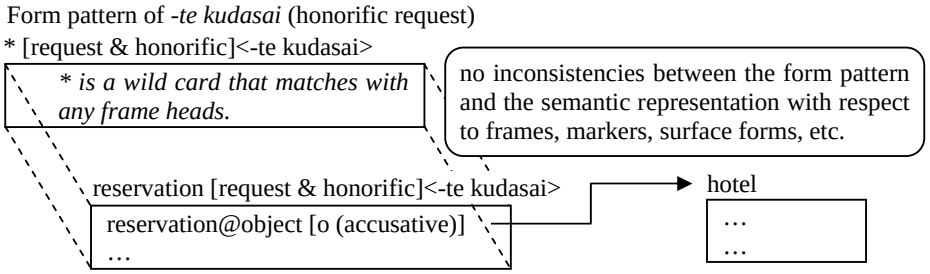

Fig. 4: Example Form Detection

3. Form Evaluation

In our previous system, the situation-correspondence judgment component compares the candidate semantic representations with the situation knowledge and judges which part of the situation knowledge each candidate corresponds to. In comparing the candidate representations with the situation knowledge, the component does not compare their surface forms. This is because semantic equivalence is enough for CA-based instruction. We have therefore extended the situation-correspondence judgment component so that it should perform surface level comparison in addition to the original semantic level comparison. The extension enables the component to detect whether each of the focused forms or part of it is uttered in learners' input.

The situation-correspondence judgment component detects which part of the input corresponds to which part of the focused forms and the semantic representation integrator differentiates uttered part of the focused forms from unuttered part of the focused forms. Accordingly, the error judgment component performs form-wise error detection based on the results of the situation-correspondence judgment and the semantic representation integration. The error judgment component looks through the (complemented) semantic representation and detects errors involved in the focused forms by referring to the situation knowledge and the form dictionary.

The error detection is performed based on the following two viewpoints: (1) whether the input involves a structure which is equivalent to the structure of the form pattern of a focused form, and (2) whether the surface form of the input matches with the surface form of the focused from. Consequently, the result of the error detection is classified into four types: (i) the input matches with a focused form with respect to its structure and the surface form, (ii) the input matches with a focused form with respect to its structure but not to the surface form, (iii) the input does not match with a focused from with respect to its structure nor to the surface form, but the input involves every concept in the form pattern of the focused form, and (iv) the input falls into none of (i)-(iii).

The classifications (i)-(iii) roughly corresponds to the distinction among TL, IL, and NTL encoding, where TL is completely targetlike encoding, IL is interlanguage encoding, and NTL is nontargetlike encoding [1]. In other words, TL or (i) indicates that the learner correctly understand the target form. IL or (ii) shows that the learner understands the function of the target form but encodes it in a nontargetlike way. NTL or (iii) shows that the learner's understanding of the target form is insufficient with respect to both the form pattern and the form function. The system judges the case (iv) as an utterance irrelevant to the target form. In addition, there are cases in which a focused form is used when it should not be used. The error judgment component also detects this type of error and judges it as OVERUSE.

4. Preliminary Evaluation

In order to evaluate the system, we picked up 41 FonF forms. Recall that form patterns are represented as semantic representations (section 2.1) and that the semantic representation is composed of (a) frames, (b) elements constituting frames, and (c) pointers (section 1.1). We therefore analyzed each FonF form from the viewpoint of which of (a)-(c) constitutes the form. The result is that 1 form is composed of (a) alone, 113 forms are made up of (a) and (b), 24 forms are constituted of (a) and (c), and 13 forms are made up of all the three types of elements. Accordingly, we selected 27 (a)-(b) forms, 12 (a)-(c) forms, and 2 (a)-(b)-(c) forms as the test set.

The evaluation of form recognition took the following steps. We first made 41 pieces of the situation knowledge containing the 41 forms. We then examined whether the form detector correctly detected the 41 forms from the situation knowledge. The result is shown in Table 1. The form detector successfully detected all the 41 forms and no erroneous detection was attested.

Table 1: Preliminary Evaluation of Form Detector

Forms to Detect	Detected Forms	Erroneous Detection
41	41	41

Since the test set includes every possible combination of (a)-(c) in FonF forms, the result shows that the system, in principle, has the ability to correctly detect every FonF form.
We examined the accuracy of the system's form evaluation in the following manner. (1) We made the situation knowledge containing the 41 forms. (2) We set the detected 41 forms as focused forms. (3) We made 4 types of input (TL, NTL, OVERUSE and irrelevant input) for each piece of the situation knowledge. As for IL, we made the input for 32 pieces of the situation knowledge but could not make IL input for the other 9 pieces of the situation knowledge. For example, forms like *-sugi masu* (denoting excess) and *-te shimai-mashi-ta* (denoting regret) do not have any synonymous expressions with the same structure of the semantic representation. Accordingly, any input with the same semantic structure as these expressions necessarily becomes a TL example. If we change the structure, it is not an IL example by definition. As a result, we fed the system with the 41 TL examples, 32 IL examples, 41 NTL examples, 41 OVERUSE examples, and 41 irrelevant examples. The result is shown in Table 2. The result confirmed that the error judgment component correctly detected all the 5 types of examples. No erroneous detection was attested.

Table 2: Preliminary Evaluation of Form-Wise Error Detection

Form Type	TL	IL	NTL	OVERUSE	Irrelevant
Number of Input	41	32	41	41	41
Success	41	32	41	41	41
Failure	0	0	0	0	0

5. Concluding Remarks

In this study, we have constructed the form dictionary and the form detector in order to enable our FonF-based education system to recognize which forms are focused in a particular lesson. The form detector searches the situation knowledge for FonF forms by referring to the form dictionary. The situation knowledge stores information on which forms are focused. We have also extended the situation- correspondence judgment component and the error judgment component in the previous system. The extended components determine which of the focused forms are uttered in learners' input, and evaluate the learners' use of the focused forms. The preliminary evaluation of the system has confirmed that the system successfully detects FonF forms in the situation knowledge, and that the system correctly performs form-wise error detection. Neither erroneous form detection nor erroneous form evaluation has been attested. Important topics of future research are (a) full evaluation of form detection and form-wise error detection and (b) evaluation of the system's usability from the viewpoints of both learners and teachers.

Acknowledgements

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

References

[1] Doughty, C.,& Varela, E. (1998). Communicative focus on form. in C. Doughty, & J. Williams (Eds.) (1988). 114-138.
[2] Doughty, C., & Williams, J. (Eds.) (1998). *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press.
[3] Ichikawa, Y. (1997). *Nihongo Goyo Reibun Shojiten [A Dictionary of Japanese Language Learners' Errors]*. Tokyo: Bonjinsha.
[4] 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.
[5] Kaigai Gijutsusha Kenshu Kyokai (Ed.) (1990) *Shin Nihongo no Kiso I*. Tokyo: 3A Corporation.
[6] Kaigai Gijutsusha Kenshu Kyokai (Ed.) (1992) *Shin Nihongo no Kiso I Kyoshiyo Shidosho*. Tokyo: 3A Corporation.
[7] Kaigai Gijutsusha Kenshu Kyokai (Ed.) (1993) *Shin Nihongo no Kiso II*. Tokyo: 3A Corporation.
[8] Kaigai Gijutsusha Kenshu Kyokai (Ed.) (1994) *Shin Nihongo no Kiso II Kyoshiyo Shidosho*. Tokyo: 3A Corporation.
[9] Kondo, M., Kure, U., Daicho, Y., Kogure, S., Konishi, T., Y Itoh, Y. (2009). Error judgment in a language education system oriented for focus on form. *Proceedings of ICCE 2009*, 43-50.
[10] Shiratori, T., Minai, N., Itoh, T., Konishi, T., Kondo, M., & Itoh, Y. (2004). Error detection and response-manner selection in a Japanese education system for nonnative speakers. *Proceedings of ICCE 2004*, 1383-1389.
[11] 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.
[12] Williams, J. & Evans, J. (1998) What kind of focus on which forms? in C. Doughty, & J. Williams (Eds.) (1988).139-155.