PROTOTYPING A CALL SYSTEM FOR STUDENTS OF JAPANESE USING DYNAMIC DIAGRAM GENERATION AND INTERACTIVE HINTS

Christopher Waple*, *Yasushi Tsubota*†, *Masatake Dantsuji*†, *Tatsuya Kawahara**†

*School of Informatics, Kyoto University,

†Academic Center for Computing and Multimedia Studies, Kyoto University,

Sakyo-ku, Kyoto 606-8501, Japan

`chris@ar.media.kyoto-u.ac.jp`

Abstract

In this paper we discuss the concept and design of a new CALL (Computer Assisted Language Learning) system being developed to aid students learning Japanese as a second language. The system is being designed to allow students to create and speak their own sentences based on visual prompts, before receiving feedback on their mistakes. The students may choose to receive guidance in order to complete each task, selecting the level of help that best suits their needs. Having described the concept of the system and its design, we discuss some tests recently carried out using a prototype of the system, summarize the results obtained, and give some thought as to the significance of these results on the future development of the system.

Index Terms: CALL, second language learning, Japanese, sentence generation, interactive help.

1. Introduction

There has been much interest in the development of CALL systems, and the various approaches represented to their development. Some of these systems focus on correcting pronunciation errors in the students speech[1]. Others concentrate on vocabulary, or grammar learning[2]. And of course, there are systems that look to combine these fields. The systems also differ in the level of interactivity that they provide the student. A study comparing the relative advantages and disadvantages of a system that allows a free-form of input compared to those which restrict the students answers (having them reading a given line, or choose from multiple choice answers) has been carried out previously [3]. The advantages and disadvantages of making use of speech recognition technology in CALL systems have also been considered [4].

In the design of this system, we have considered various options that would allow more freedom for the students with regards to how they study the language. Firstly, we have decided that the tasks for the students should involve constructing their own sentences, rather than being given complete sentences to read, or selecting from multiple options. Of course, the system must create the restraints and context within which the sentence is to be formed. These are expressed to the student in terms of a Concept Diagram, a picture dynamically generated from an underlying concept of a situation, along with the appropriate grammar rules and restraints for any given lesson.

Secondly, we believe that the students should be able to choose just how much guidance they would like to receive in order to solve a task. Thus, an interactive hint system has been designed, allowing the students to reveal sections of a target answer, whilst leaving other sections hidden. Also, the system has the facility to suggest the choice of next hint to the students. Both methods are accompanied by a scoring system, through which each hint used incurs a penalty or cost.

Finally, we believe that a student or teacher should have the option to configure the software to fit with their own priorities and/or needs. To this end we are looking at the option of customizing the hint penalties such that the cost system is inline with the student's own learning priorities. Lessons should also be customizable, allowing options such as just practicing verb forms in isolation, or embedding that challenge within the task of constructing sentences of varying complexity. This kind of interactive help and customization is only possible through a carefully designed CALL system, and can not be provided by textbooks for example.

In order to scope out the potential of such a system, and in particular the benefits of the features outlined above, we carried out a set of trials in which students undertook a number of challenges presented by a prototype of the system, and then gave their opinions via a questionnaire. Another aim of this experiment was to capture the kind of grammatical mistakes typically made by the students. We present analysis of the data obtained towards the implementation of the full system.

2. CallJ - System Overview

In this section we firstly cover the scope of the system, that is the proficiency level of Japanese at which the content of the software is aimed at providing. We will then present an overview of the processes involved in the system, before focusing on two specific areas. Firstly, we will describe the generation of the concept diagrams that are used to convey the situation to the student. Secondly, we will describe the interactive help system, focusing on "Hints".

2.1. System Scope

The system is aimed at beginner to intermediate level students of Japanese. Specifically, it contains contents from levels 4 and 3 of the Japanese Proficiency Language Test (JPLT) [5]. The material specified within these boundaries consists of approximately 1,500 words (of which around 200 are verbs), 300 kanji, and 95 grammatical points. We would be looking to introduce these grammatical points to the student over the course of approximately 30 lessons.

2.2. Process Overview

- Lesson Definition - the contents of each lesson, in terms of the target grammar, vocabulary, etc must be defined.
- Concept Generation - each question involves the student being asked to describe a situation which is dynamically generated within the confines of the lesson. Appropriate vocab-
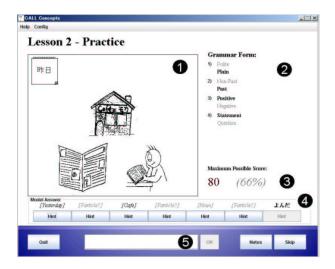
Figure 1: *Screenshot of practice screen; 1: Concept diagram, 2: Desired Form, 3: Score, 4: Hints, 5: Answer Entry*

ulary must be used depending on the verb, and to achieve this a case-frame selection would be used [6].

- Grammar Generation - the target grammar for the speech recognition task must be generated, containing paths for all the acceptable sentences (valid answers), along with all predicted errors.
- Hint Generation - a set of hints is generated for each question, with an appropriate penalty assigned to each hint.
- Recognition - The student's answer would be spoken (with text entry perhaps being supported to provide confirmation).
- Error Processing - Any errors the students make need to be detected, and categorized.
- Feedback - Feedback would be given on any trends in errors (mistakes that are commonly detected by the system).

2.3. The Concept Diagram

2.3.1. Outline

For each question in each lesson, the student is tasked with forming a sentence based on some situation. This situation or concept is depicted graphically via the application. In our system the concept diagram is generated dynamically at run-time. This approach offers a number of advantages:

- An appropriate image is generated promptly in all cases
- Significantly reduced cost time-wise in creating the images
- Consistency in style across all generated images

We are only considering simple Japanese sentences created within the confines of JPLT levels 3 and 4, and as such the meaning of these sentences can be expressed via such a diagram.

2.3.2. Realization

The diagram is currently created by combining smaller sub-images that represent each component in the situation. For example, consider the example "Yesterday I read a newspaper at the coffee shop". In this situation, we have a number of key components, namely "Yesterday", "I", "Read", etc. The diagram created given this situation can be seen in Figure 1, which shows a screen capture of the lesson practice screen. To maintain a consistent feel

Table 1: *A table demonstrating the generation of hints*

Sentence	*Kinou*	[none]	*Ko-hi-*	*Wo*	*Nonda*
Type	Time	Part.	Object	Part.	Verb
Hint L1	[Time]	[P?]	[Object]	[P?]	[Verb]
Hint L2	Yesterday	-	Coffee	-	Drink
Hint L3	*Kinou*	[none]	*Ko-hi-*	*Wo*	*Nomu*
Hint L4	-	-	-	-	*Nonda*

throughout all generated images, all the sub-images were obtained from a single source [7].

2.4. Dynamic Hint Generation

2.4.1. Outline

One of the key features that distinguishes this software from more traditional methods of learning is the ability for the student to choose the level of help they receive in answering a question, something that is difficult to achieve in printed textbooks for example. In this system we allow the student to uncover the target sentence word by word. Each word is not simply revealed in one step, but incrementally with the word class (Time, Person, Object etc) being given first, then the English word, and so on until the final Japanese form is given. The process of generating these hints is given below in section 2.4.2..

Moreover, we assign a "cost" to the revealing of each hint. This cost would deduct from an overall score for that question, and act as a motivator to encourage the student to attempt solving the question themselves before resorting to guidance. Deciding the cost for each hint is an important factor, and will be discussed below. The idea of having a score for each question, and thus lesson, was introduced to add a more game-like feel to the software, to keep the students interested in progressing.

The method through which a student accesses these hints is also important, and a couple of alternatives have been considered. The word which is to be unveiled next may be either selected by the student, or alternatively suggested by the system. Both of these methods are described further in section 2.4.4..

2.4.2. Generation

The hints are generated based on breaking down one target sentence (one grammatically correct answer to the question, although others alternative (valid) answers would be accepted) into its constituent components, and then for each component creating an ordered set of hints. The number of hints or hint levels per component varies with the base type of that component. For example, with a verb, the dictionary form (the base Japanese form) of the word is given as one hint, with the final form appropriate for the given situation revealed as a separate hint. With nouns, this extra step is not necessary.

Table 1 shows an example of a sentence being broken down into a set of hints. Regarding the level of hints, L1 is the component type, L2 the English equivalent, L3 the Japanese word in base form, and L4 the final form (for verbs and adjectives).

Note that in some cases, such as after a time-related word, whether to insert a particle or not varies depending on the time expression used. For example, after a specific time phrase, such as "8 O'clock" (*Hachiji*), a particle (*ni*) is required. However, if as in the example above a relative time word is used, such as "Yesterday" (*Kinou*), no particle is inserted. This is why the potential presence of a particle is given by hint level 1 ([P?]), but the student is not actually informed as to whether a particle is inserted or not until they reveal further hints. It is a challenge for the student to

determine the presence of a particle, and this challenge should not be solved for them too quickly.

2.4.3. Cost Assignment

The usage of hints should come with a penalty, and the cost of each hint needs to be considered. The rationale behind this is that simple guidance, such as revealing the English meaning of a word, or identifying where in the sentence the verb is expected, should not come with as high penalty as revealing the correct form of the verb for example. In deciding the costs for each component at each hint level, we consider the following three properties:

1. Perceived Difficulty
2. Communication Value
3. Relevance to Main Lesson Point

"Perceived Difficulty" expresses the likelihood of a mistake being made at the point otherwise revealed by the hint, and would be calculated by analyzing answers given by students in trial runs of the system. It may also be desirable for this value to be automatically updated as a student runs the software, to tune the likelihoods to their own performance. "Communication Value" is a gauge of how much impact a mistake at this point would have on a listener's understanding of the sentence, and would be estimated by having Japanese natives evaluate answers given by students. "Relevance to Main Lesson Point" represents how the hint relates to the main point of the lesson. For example, if the lesson was about learning a specific verb form, the hint that reveals the verb form would score very highly in this category.

The total cost is calculated by weighting each of the above three factors, and combining them. The tuning of these weights is another point that demands careful consideration, and is of course dependent on what areas are seen as more important with regards to language learning (Grammatical correctness vs Communication etc). It is also possible that the teachers and/or students will have different views on what is important, and what is not, and thus it may be desirable to allow them to alter these weights depending on their own priorities.

Determining the preferences of the student regarding hints and their associated costs was one of the main goals of an experiment carried out using a prototype version of the software, as detailed in section 3. For this experiment, the costs were defined by a-priori knowledge.

2.4.4. Accessing the Hints

We considered two different ways in which the students may access the hints whilst attempting to answer a question. The first would be to show the student the framework of the sentence, and allow them to click on any word they are unsure of, to get the next hint for that component. This method would allow the students to get to the help that they need quickly, and without the need of revealing information that they already know.

A second method would be to have just one hint button, and the system would then reveal the hint which was seen as being the most appropriate at that point. By appropriate, we mean that the available hints are ordered in terms of their score, with lowest cost hints (the least significant) being revealed first, and the most expensive hint revealed last. This option removes the problem that a student might not know which word to unveil, but does mean the student will likely have to unveil numerous words to get to the information that they desire.

Both options were included in the prototype software used in a trial run, in order that we could obtain feedback from the students as to which method was more satisfying.

3. Experiment

3.1. Overview

A trial was conducted using a prototype of the proposed system, with a number of students who have either studied, or are currently studying Japanese as a second language, running through a set of lessons, and giving their feedback on the system. Log files produced by the software, along with a questionnaire filled out by the students were used in our analysis.

3.2. Aim

The main goals of this experiment were to:

- Evaluate the benefits of the hint and cost system
- Determine the preferred method of hint access
- Judge the comprehensibility of the concept diagrams
- Capture the typical errors made by the students
- Gain insight into the students' language priorities
- Gauge the students' overall opinion on the system

3.3. Software and Configuration

The experiment was carried out using a prototype version of the system. This system was designed to appear much as the final system would, in terms of the interface, the concept diagrams etc, whilst not including dynamic question generation, nor speech recognition. Using speech recognition would inevitably have introduced speech recognition errors into the process, thus making analysis of the students' actual mistakes problematic. For this trial, the students submitted their answers for each question via text entry. To evaluate the different proposed hint methods, the software was designed to run in a different configuration for each lesson. There were three configuration options that we were looking at evaluation. First was whether the hints were selectable by the student, or given in a fixed order as suggested by the system; second was whether the score was displayed; third was whether any initial hints were displayed for "free" (such as the component type of each word). These options, with two settings each, led to running the trial over 8 lessons ($2^3 = 8$). There were five questions per lesson.

3.4. Results

3.4.1. Hint and Cost system

From the questionnaire (Table 2, section 1), we can see that the students found the hints to be useful in terms of solving the problem and that they preferred to be able to select which hint to see as opposed to having a fixed order. This preference could be due to a dissatisfaction in the ordering of the hints in the fixed-order case. In the version of the software used, the hints were ordered such that the easier hints were unveiled first, whilst the students generally stated they would prefer the order to be reversed. One potential solution that we will consider is to dynamically tune the order via the students own hint usage and mistakes, with the aim being that the component judged most likely to be causing the student problems should be revealed first. The cost system was seen as a positive feature by all the students in terms of encouraging self-motivation.

Regarding hint usage, we were able to determine that the amount and category of hints used varied both with the student, and with the lesson. We were also able to see that the hint usage tended to decrease throughout the running of a lesson.

Table 2: *Questionnaire answers*

1	Selectable hints are useful	4.0
	Fixed-order: hints are useful	3.2
	Fixed-order: order makes sense	2.6
	Fixed-order: hardest should be given first	3.0
	Fixed-order: easiest should be given first	1.8
	Would like more free hints	4.0
	Would like less free hints	2.0
	Hint costs encourage motivation and effort	4.0
2	Easy to access relevant information	4.6
	Lesson goals are always clear	4.2
	It is always clear what to do	3.2
3	System could help improve skills	4.0
	Would potentially use system	4.0
4	Difficulty: Grammar	3.8
	Difficulty: Vocabulary	3.8
	Difficulty: Pronunciation	2.8
	Difficulty: Reading	4.4
	Difficulty: Writing	4.6
	Difficulty: Listening	3.0
5	Priority: Grammar	0%
	Priority: Communication	40%
	Priority: Both	60%
6	Priority: Reading/Writing	20%
	Priority: Conversation	40%
	Priority: Both	40%

Table 3: *Summary of the most common student errors recorded*

Component	Occurrences	Errors	ER(%)
Counter	30	10	33.3
Verb	206	40	19.4
Noun	320	33	10.3
Particle	328	29	8.8
Adjective	5	1	5.0
Adverb	5	1	5.0
Sentence	200	9	4.5

3.4.2. Concept Diagram Comprehensibility

Through the questionnaire (Table 2, section 2), the students expressed a general satisfaction that, through the concept diagram, the situation which they were being asked to describe was clear. There were a number of comments suggesting further improvements to the diagram, such as the need for consistency regarding the characters seen to be carrying out the action with regards to the subject of the sentence. The changes required to fulfill these suggestions are all seen as feasible within the current framework.

3.4.3. Student Errors

Student errors captured during the trials were categorized by the component type where the error occurred (noun, verb, particle). The categorization was not dissimilar from that seen in other works[8]. Table 3 summarizes the most common errors recorded, showing the error rate per component type. It should be noted that "Sentence" refers to sentence-level errors, such as word ordering etc. Also, within each component category of errors, further analysis was carried out into the nature of the mistake.

The following is a typical example of a student's response to the diagram given in Figure 1, containing mistakes:

Correct Answer: "*Kinou kissaten de shinbun wo yonda.*"
Student Answer: "*Kinou Ko-hi de shinbun wo yomimashita*"

In this case there were two errors. The first was using the word for "Coffee" instead of "Coffee Shop" (semantically related noun substitution). The second mistake was on the verb form, using the polite form where the plain form was requested.

3.4.4. Difficulties and Priorities in Language Learning

Through the questionnaire (Table 2, sections 4-6) it was noted that Reading/Writing were seen as the most challenging aspects of Japanese. Grammar and vocabulary were also seen as challenging,

whilst listening and pronunciation not so. Of the students questioned, 40% felt that Communication was more important than Grammatical correctness whereas 60% believed both areas were of equal importance. Regarding Reading/Writing and Conversational skills, there was no clear bias in priority either way.

3.4.5. Overall Opinions on System

The students were generally enthusiastic towards the system's potential. They expressed a belief that such a system could be useful in learning a foreign language, and would be interested in using such a system (Table 2, section 3).

4. Conclusions

We have presented the groundwork for a new interactive CALL system, and have successfully carried out tests using a prototype version of the system. We have been able to confirm that the hint and cost systems proposed are seen as useful by current students, and that the dynamic generation of concept diagrams is feasible. We were also able to capture some information on the typical errors that students make, which is critical for calculating the relative costs that hints should be assigned, as well being very useful for the creation of error paths in recognition grammars. The feedback from these trials was generally very positive, and useful for the development of the system.

5. References

[1] Tsubota,Y., Kawahara,T., and Dantsuji,M., "Practical Use of English Pronunciation System for Japanese Students in the CALL Classroom", ICSLP, 2004

[2] Nagata,N., "Japanese Courseware for Distance Learning", AILA, 2000

[3] Yang, J.C., and Kanji, A., "An Evaluation of Japanese CALL Systems on the WWW Comparing a Freely Input Approach with Multiple Selection", CALL Journal, Vol.12 No.1, pp.59-79

[4] Ehsani, F., and Knodt, E., "Speech Technology in Computer-Aided Language Learning: Strengths and Limitations of a New CALL Paradigm", LLT Journal, Vol.2, No.1, July 1998 pp.45-60

[5] Japanese Language Proficiency Test (JPLT), Japan Educational Exchanges and Services, http://www.jees.or.jp/jlpt/en/

[6] Kawahara, D., and Kurohashi, S., "Japanese Case-Frame Construction by Coupling the Verb and its Closest Case Component", In Proceedings of the HLT Conference, pp.204-210, 2001

[7] Hatasa, K., "Royalty-Free Clip Art Collection for Foreign/Second Language Instruction", http://tell.fll.purdue.edu/JapanProj/FLClipart/default.html

[8] Yang, J.C., and Kanji, A., "Error Analysis in Japanese writing, and it's implementation in a Computer-Assisted Language Learning System on the World Wide Web", CALICO Journal, Vol.15 No.1-3, pp.47-66