An Automated Method to Generate e-Learning Quizzes from Online Language Learner Writing

Brendan Flanagan, Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan

Chengjiu Yin, Research Institute for Information Technology, Kyushu University, Fukuoka, Japan

Sachio Hirokawa, Research Institute for Information Technology, Kyushu University, Fukuoka, Japan

Kiyota Hashimoto, School of Humanities and Social Sciences, Osaka Prefecture University, Sakai, Japan

Yoshiyuki Tabata, Research Institute for Information Technology, Kyushu University, Fukuoka, Japan

ABSTRACT

In this paper, the entries of Lang-8, which is a Social Networking Site (SNS) site for learning and practicing foreign languages, were analyzed and found to contain similar rates of errors for most error categories reported in previous research. These similarly rated errors were then processed using an algorithm to determine corrections suggested by a native speaker. Subject matter experts then evaluated the processed sentences to determine the quality in relation to use in tests or exams for language learners. The method describes the automatic generation of multiple choice and fill-in-the-blanks quizzes using the writings of language learners on public web based learning sites, in order to support learner reflection on corrections and practicing past errors to overcome problems.

Keywords: *Analyzation, Fill-in-the-Blanks Quiz, Language Learning, Reflection, Social Networking Site (SNS)*

DOI: 10.4018/ijdet.2013100105

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

INTRODUCTION

As we know, there are four skill of language learning: reading, writing, listening, and speaking. This paper focuses on writing skills. In research many have discussed the effectiveness of learning systems on the Web in learning (English as a Second Language / English as a Foreign Language) (Meurant, 2010; Yang & Chan, 2008; Fan, 2011). It is very useful that another person can point out your errors and correct your entries, much like teachers correcting homework for students in a class. Corrective feedback is a technique to help learners correct errors by providing them with some kind of prompting (Lo et al., 2008). It is useful for learners to write in a foreign language and ask another person to correct it for them so they can be aware of their errors. However, a problem for learners' is how to reflect on corrections and then master the errors.

Reflection is a way of helping participants better understand what they know and do as they develop their knowledge and skills through reconsidering what they have learned (Loughran, 2002). Hatton and Smith (1995) identified that we should consciously account for the wider historic, cultural, and political values or beliefs in framing practical problems to arrive at a solution.

Nowadays, many web-based learning services are provided in many different forms, such as: chat forums, newsgroups, blogs, virtual environments, search engines and social networks. Of these, one of the fastest growing areas is that of Social Networks Services (SNS). There are many SNSs for language learning, such as Lang-8 (http://lang-8.com/). Lang-8 is a place for learning and practicing foreign languages. When you write an entry in the language you are learning, a native speaker will correct your entries.

In this paper, we used the writings and corrections created by users in Lang-8 as a source of data. In this sense, we used Lang-8 not as an SNS service, but as a crowdsourcing data collection method.

Using data collected from large groups of disparate people, known as crowdsourcing, has become a more and more popular with increasing Internet usage. As opposed to traditional data collection methods that use known groups of people to collect data, crowdsourcing data collection utilizes data that has been generated by an unknown public group. For example: Coleman (2012) used twitter in crowdsourcing data to create a web application aimed at learners of foreign languages.

We collected the diaries from Lang-8 (lang-8.com/) in which language learners have written a diary in a foreign language and a native speaker has corrected it for them. The error patterns and frequencies of writings in English were analyzed using these diaries. A quiz system was then created to train learners, using the error patterns and frequencies to determine which questions the user should undertake.

Kroll (1990) and Weltig (2004) classified the error patterns of English learners writings. These error patterns targeted the English students who were studying in their respective universities. With the development of the Web, many learners are studying English online on web sites such as Lang-8. Based on the error patterns of Kroll (1990) and Weltig (2004), analysis has been performed on the error patterns of Lang-8 users. Then, based on the error patterns of Lang-8 users, a demonstration quiz system has been developed, which is proposed to analyze the error characterizers of every student. Students can train their English focus on their error characterizers, and then improve their English.

In this paper, in order to help facilitate learners' reflection on corrections, we analyzed the entries of Lang-8 where we found a clear correspondence between original entries and the corrections. We consider that this correspondence can be used to generate fill-in-the-blank questions automatically from user's erroneous writings. By answering these questions, learners can review the corrections once again and reaffirm whether they really know the answer themselves. A demonstration

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

system was built which can generate fill-in-the-blank quizzes according to different situations such as going traveling, hotel reservation and gaining employment. Learners can find specific problems to practice their errors and reflect on the correction. Figure 1 shows the learning flow of using the system. If a learner made an error, then s/he reflects on the correction and then uses the system to find quizzes about the same error (which has made by other learners as well) to practice.

Fan (2011) identified three benefits of Web Learning systems: firstly, it can provide a large number and variety learning materials, secondly, it can maintain the learners' autonomy, thirdly, it is possible to facilitate collaborative learning between learners and teachers. Lang-8's web-based online corrective feedback system uses the same approach as traditional paper-based correction. Compared with Lang-8, our system generates easy quizzes for autonomous learners.

RELATED WORK

Computer Technologies and Language Learning

Language Learning is a life-long activity, and with the development of computer and Internet technologies, many researchers are focusing on computer assisted and web-based language learning.

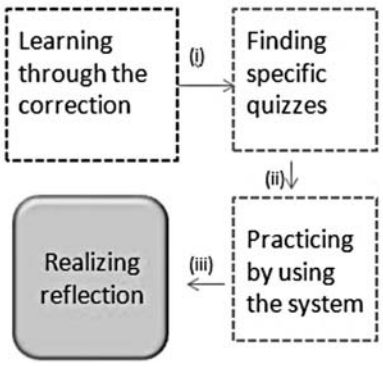
In the early 1990s, multimedia computers were used to enhance the four basic skills of listening, speaking, reading and writing, as it enabled text, images, sound and video to be combined in one device (Davies, 2011). Since the early 2000s, with the development of the Internet, web-based language learning has become popular worldwide and throughout generations.

There are many web-based applications that have been developed for language learning and a selection is listed as follows (Walker et al., 2011):

- Image storage and sharing;
- Social bookmarking;
- Discussion lists, blogs, wikis, social networking;
- Chat rooms, MUDs, MOOs and MUVEs (virtual worlds);
- Podcasting;
- Audio tools;
- Video sharing applications and screen capture tools;
- Animation tools - comic strips, movies, etc.;
- Mashups.

In this paper, it is proposed to develop a web-based language learning system that can automatically generate multiple choice and fill-in-the-blanks quizzes to support language-learning writing.

Figure 1. Learning flow



Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

Corpora and Concordancers

There are also some researchers who have focused on building corpora and concordancers. Johns (1991) proposed Data-driven learning (DDL). "DDL encouraged learners to work out their own rules about the meaning of words and their usage by using a concordancer to locate examples in a corpus of authentic texts. It is also possible for the teacher to use a concordancer to find examples of authentic usage to demonstrate a point of grammar or typical collocations, and to generate exercises based on the examples found." ("Computer assisted language learning", 2012).

We collected the entries of Lang-8, which is an SNS site for learning and practicing foreign languages. This data can be seen as a corpus of a language. By using these data, a system was developed to generate exercises automatically.

DATA ANALYSIS WITH RESPECT TO WORDS

Data Analysis of Entries of Lang-8

A profile choice was conducted on the Lang-8 data to select entries created by users studying English. As a result 15499 targets were selected and the files were extracted. These files were then filtered to find user written journals in English. This resulted in 1039 remaining entries that were used to create a data set on which analysis with respect to words was undertaken. These entries contained 7619 lines, of which 6165 lines were corrected by inserting or deleting words. There are 4019 words inserted and 1210 words deleted in total.

Format of Corrected Sentences

Each corrected sentence is displayed together with the original sentence. Some words may have been inserted or deleted by other users in the form of corrections, as can be seen in the following example:

Do you have the special event for the day?

Do you have the a special event for the on Thanksgiving day?

The corrected sentence is marked up in HTML and uses CSS based styles to indicate changes. The corrected segments are encapsulated by "span" tags, such as: for the, and Thanksgiving. The parameter class="sline" renders the enclosed words with a strikethrough indicating the deleted words and the parameter class="f_blue" renders the enclosed words in a blue color indicating the inserted words:

Do you have the a special event for the on Thanksgiving day?

It is proposed that the HTML structure can be used to extract the position of fill-in-the-blank-boxes and the answer that should be inserted or deleted.

Frequent Words in Corrections

Table 1 shows the top 20 frequent words that were found in the original and corrected sentence pairs. Under the *Corrections* group, the columns "ins" and "del" display the number of lines where the word was either inserted or deleted. Under the *Lines* group, the column "cor" displays the total number of lines where the word was the error and it was corrected by insertion, deletion or substitution. The column "Error" displays the number of lines that contain the word and error word(s) other than the word. The column "No Err" displays the number of lines that contain the word and have no corrections. The column "all" displays the total number of lines that contain the word. For example, the third row shows 3197 lines that

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

Table 1. Frequent words

		Corrections		Lines			
	Word	Ins	Del	Cor	Error	No Err	All
1	the	1010	281	1291	2104	457	3852
2	i	565	82	647	2564	531	3742
3	to	659	219	878	1925	394	3197
4	a	776	230	1006	1636	372	3014
5	in	402	134	536	1302	237	2075
6	of	430	165	595	1201	250	2046
7	and	308	81	389	1365	288	2042
8	is	401	107	508	1074	268	1850
9	my	241	50	291	965	245	1501
10	it	324	102	426	880	181	1487
11	was	223	69	292	703	165	1160
12	for	266	81	347	691	107	1145
13	that	256	88	344	667	133	1144
14	have	182	76	258	543	126	927
15	you	258	34	292	504	98	894
16	but	128	22	150	557	104	811
17	so	123	55	178	509	94	781
18	this	178	24	202	411	110	723
19	on	181	49	230	410	72	712
20	with	132	50	182	419	81	682

contain the word “to”, among which 394 lines have no corrections and 1925 lines contain corrections that do not contain the word. 659 lines have been corrected by inserting words and 219 lines have been corrected by deleting words, and this totals 878 lines that contain corrections of the word.

Correction Analysis

How does one discover typical words that are erroneously used and the typical situations in which these errors occur? It may be more likely that frequently used words will have a higher tendency to be erroneously used. However, the probability of erroneous usage with respect to the average of all words needs to be compared. Among the 7619 sentences, 6165 sentences by users were found to contain errors. Therefore there is an 81% average probability of a sentence containing an error. This score can be used as a benchmark of the difficulty of words. Another point that should be taken into consideration is that the word is not always the cause of the error in a sentence, even if the sentence has an error in some way. Another word may be the cause of the error. Taking this into account, two ratios were created: the error-ratio ErrorRatio(w), and the error-cause-ratio ErrorCauseRatio(w) of a word are defined as follows:

$$ErrorRatio(w) = Eline(w)/ Line$$
$$ErrorCauseRatio(w) = ECause(w)/Line(w)$$

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

Line(w) = #{s | w appears in s}
Eline(w) = #{s | w appears in s & s contains an error}
Ecause(w) = #{s | w is an error in s}

Here *s* is a sentence, *w* is a word and *Line* is the total number of the sentences.

Table 2 shows the top 10 words of high Error-Cause-Ratio and Figure 2 displays the relationship between the Error-Ratio and the Error-Cause-Ratio of words. In Figure 2, we can see typical words, such as "a" and "the", which are known to be difficult for Japanese learners to master, appear on the right hand side of the graph. Words such as: "about", "at", "for", "had", "or" and "on", which reside in the upper right region of the graph are not recognized as difficult words.

Using the two ratios, we can make a more detailed analysis of the difficulty of words. For example, both "the" and "and" appear in the same horizontal level in Figure 2, with respect to Error-Ratio. In fact, "the" and "and" have the Error-Ratio of 0.8216 and 0.8258 respectively. However, the Error-Cause-Ratio of "the" is 0.5860 compared to "and" which is 0.2733. This means that the chance of making an error using "the" is almost double that of "and" when comparing by Error-Cause-Ratio.

DATA ANALYSIS WITH RESPECT TO ERROR CATEGORIES

A subset of 507 pairs of sentences was selected for error pattern categorization. After removing invalid pairs, 482 pairs of sentences were manually categorized into 42 error types.

Linear regression analysis was used to establish whether a correlation exists between the frequency of errors in the common categories of previous studies (Kroll, 1990; Weltig, 2004) and that of the Lang-8 error analysis. The results of the analysis, seen in Figure 3, show that there is a significant correlation, showing a critical alpha level of $p < 0.05$, and $t = 4.3509$, 4.4179, and 3.8011 for Kroll Class, Kroll Home, and Weltig respectively, illustrated in Table 3.

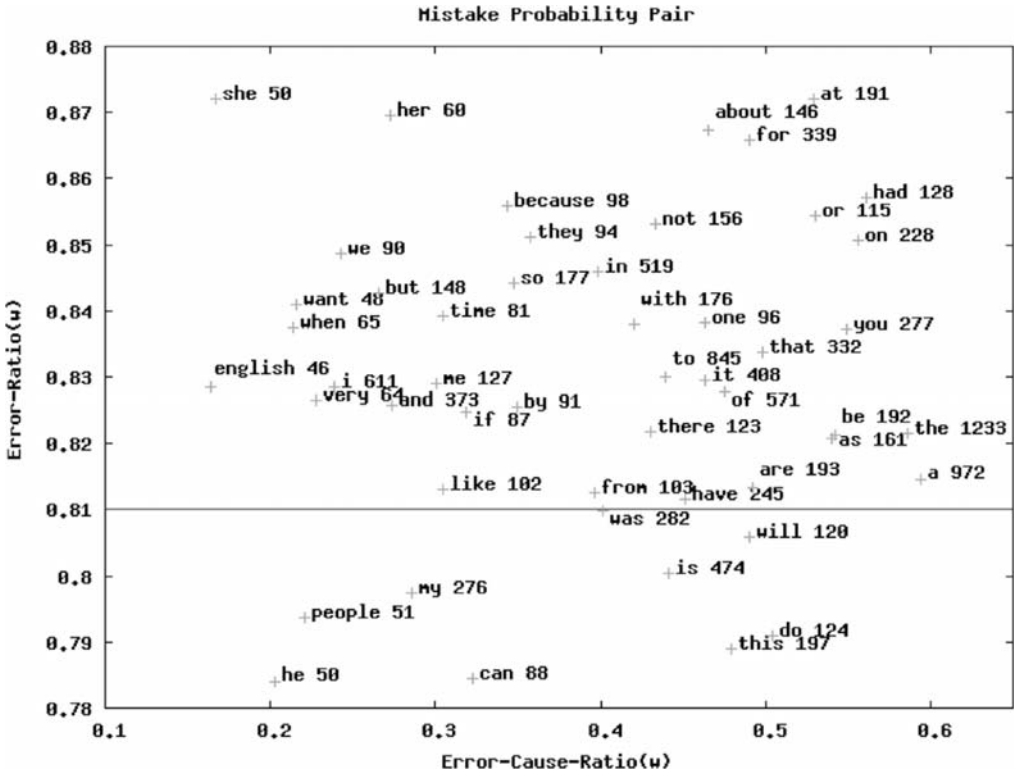
The feedback provided by native speakers often contained several different error pattern corrections within a single response. Taking this into consideration, sentences that contain more than one error type were categorized as having multiple error patterns accordingly. Some feedback contained comments about the correction and/or multiple suggestions for a single word or phrase that were mostly to do with lexical or phrase choices and categorized as such.

Table 2. ErrorCauseRatio and ErrorRatio

Word	ELine(w)	Line(w)	ECause(w)	ECR(w)	ER(w)
a	972	1636	2008	0.5941	0.8147
the	1233	2104	2561	0.5860	0.8216
had	128	228	266	0.5614	0.8571
on	228	410	482	0.5561	0.8506
you	277	504	602	0.5496	0.8372
be	192	354	431	0.5424	0.8213
as	161	298	363	0.5403	0.8209
or	115	217	254	0.5300	0.8543
at	191	361	414	0.5291	0.8720
do	124	246	311	0.5041	0.7910

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

Figure 2. Distribution of error probability pairs



As the error categories of previous research by Kroll (1990) and Weltig (2004) utilize a different set of error number lists for their analysis, a merged error number list was created as seen in Table 5 in the Appendix.

These correlations were then used to identify possible outlier errors not residing within the 95% confidence interval. A total of 22 different error categories were found outside the 95% confidence interval, with 11 of these errors being common across all three regression analyses. These common outlier errors suggest a characteristic difference in the error frequency of writings and corrections on Lang-8 when compared to those from an academic setting, such as: Kroll and Weltig. This may be a result of the differences in influencing factors, such as: motivation, the subject of the writing, and personal factors (age, socioeconomic background, etc).

As seen in Table 4, seven error categories occur more frequently on Lang-8 when compared to results from Kroll and Weltig. Of these, the error categories "Word order", "Verb formation", "Preposition" and "Article errors" are considerably outside the 95% coincidence interval and may not be a good sample to use for content creation. As a result, the outlier error categories were removed as their occurrence rate is statistically different to that of previous research and therefore might not be a good representative sample.

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

Figure 3. Error correlation of Lang-8 vs Weltig

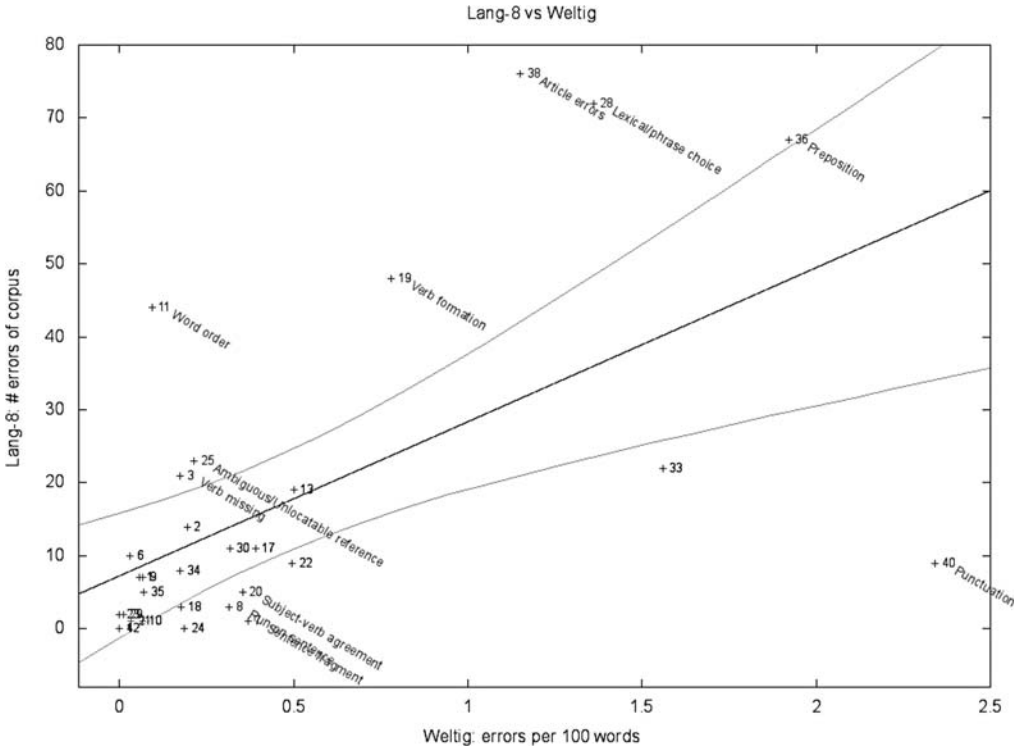


Table 3. Linear regression analysis results

	Kroll (Class)	Kroll (Home)	Weltig
r^2	0.6351	0.6409	0.5834
t	4.3509	4.4179	3.8011
p	0.0002	0.0001	0.0007
y	2.9376 + 4.2918x	4.9722 + 3.6384x	7.2613 + 21.1171x

AUTOMATIC GENERATION OF QUIZZES

Automatic Generation of Fill-in-the-Blanks Questions from Corrected Sentences

The segments that have been corrected are encapsulated by tags in the form of in HTML files. The corrected sentence can be transformed into a quiz question by replacing the corrected part(s) of the sentences in HTML format with a fill-in-the-blanks textbox. This is then used to create another HTML file that contains “input” tags within a form to automate the quiz generation process. Instances of “span” tags that contain the attribute class=”sline” are transformed into <input type=”text” name=xi_j>, where “xi_j” represents a dynamically generated variable. Other instances that contain the attribute class=”f_*” are converted into <input

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

Table 4. Outlier error categories and relation to Lang-8 error frequency

More Frequent in Lang – 8		Less Frequent in Lang - 8	
#	Error Cat.	#	Error Cat.
3	Verb missing	7	Sentence fragment
11	Word order	8	Run-on sentence
19	Verb formation	20	Subject-verb agreement
25	Ambiguous/Unlocatable reference	40	Punctuation
28	Lexical/phrase choice		
36	Preposition		
38	Article errors		

type=”hidden” name=xi_j value=*>. Figure 4 is an example of original HTML source code before transformation. Figure 5 shows the example corrected sentence’s HTML source code output that has been transformed into a fill-in-the-blank question.

The button at the end of the form calls the javascript function called “check” which verifies if the value of the input text field named “x1_1” is equal to that of the value of the hidden field “x1_2”, and also if the value of the input text field named “x2_1” is equal to that of the value of the hidden field “x2_2” as shown in Figure 6. If they are both identical then an alert box displays the message “OK” to the user when the check button is clicked. If either of them is not identical then an alert box displays the message “WRONG”.

When the original sentence contains an error word, the word is shown in the input text

Figure 4. Corrected sentence HTML source code

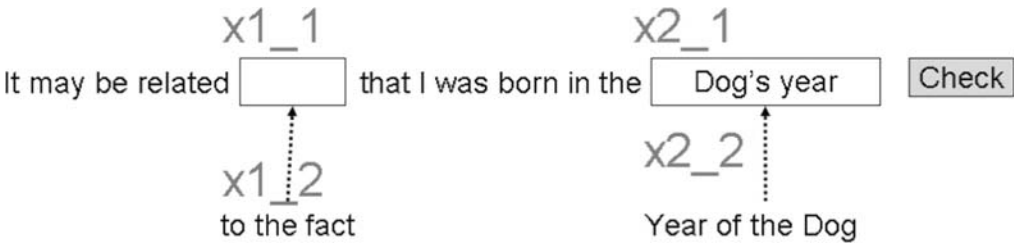
```
It may be related <span class="f_blue">to the fact </span>
that I was born in the
 <span class="sline">Dog's year</span>
<span class="f_blue"> Year of the Dog</span>.
```

Figure 5. Fill-in-the-blanks question HTML source code

```
<form name="quiz">
 It may be related
 <input type="text" name="x1_1" value="" size="11" />
 <input type="hidden" name="x1_2" value="to the  fact" />
 that I was born in the
 <input type="text" name="x2_1" value="Dog's year" size="10">
 <input type="hidden" name="x2_2" value="Year of the Dog">.
 <input type="button" onClick="check(this.form)" value="check">
</form>
```

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

Figure 6. Fill-in-the-blanks question



field named "x1_1". If a word was deleted in the corrected sentence, then the hidden text field named "x1_2" is blank and the user should just delete the contents of the text field in the response. If a word was inserted into the corrected sentence, then the hidden text field named "x1_2" contains the correction word and the user should input the same word into the text field named "x1_1". The javascript function "check" verifies if the contents of "x1_1", "x2_1" and their respective hidden fields are identical or not. The HTML file containing the form is generated automatically from the original corrected sentence pair.

Automatic Quiz Generation Algorithm

When automatically creating a quiz from the Lang-8 data, it is not necessary to use the correction tags in the feedback sentence. These tags often do not represent the actual corrections that have been applied to the sentence, and even in some cases indicate changes or corrections where no corrections exist. This occurs because the use of tags to indicate correction is not strictly enforced in the Lang-8 user experience. There are no governing rules or checks on the use of tags when submitting corrections, and as such diminish the usefulness.

To overcome this problem it is proposed that all the tags in the corrected sentence should be removed, and leaving the words inside the tags in the sentence untagged. Words contained within the sline class tags should be deleted along with the tags as it is rendered as string with a strikethrough or crossed out, and represents the deletion of the contain words.

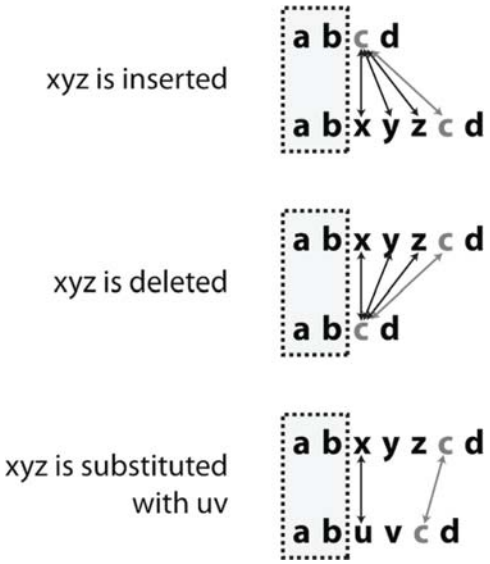
To identify the corrections that have been applied to the sentence it is proposed that the comparison of the words in the original and correction sentence strings be undertaken in a similar manner to that of the sequence alignment problem. The words of the original and correction sentences are compared in order as shown in Figure 7. If a word is found in the corrected sentence, but not in the original sentence then it is considered as an insertion. Conversely, if a word is found in the original sentence, but not in the corrected sentence then it is considered as a deletion. When a deletion and insertion occur at the same position in the original and corrected sentences then it is considered as a substitution. The corrected sentence is then retagged based on the analysis preformed using this algorithm. By using this algorithm to analyze the changes that have been made in the corrected sentence it is possible to determine the actual changes that have occurred. This does not relying on the HTML tags that are used to markup the sentence and therefore negates the possibility of incorrect error tagging, therefore improving the accuracy of the corrected sentence.

EVALUATION BY EXPERTS

We developed a demonstration system in order to evaluate the quality of the algorithm's output by language teaching experts. Using

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

Figure 7. Sequence alignment used for detection of inserted, deleted and substituted words



data sourced from Lang-8, a sample of 400 quizzes were created and then reviewed by two specialists: a language-learning researcher, and an experienced university English teacher (male, 30 years history of teaching English, has written textbooks and the like). They evaluated the quality and viability of the questions when compared to those that have been handmade by English teachers for the purposes of education and testing.

Standard Guidelines for Evaluating English Quizzes

On consultation with the English teacher, it was determined that the following three criteria are thought of as a practical example of standard practice for teachers in regards to linguistic fill-in-the-blanks questions:

1. Is the question appropriate for the student's academic level? (Are questions that are too hard or easy appropriately eliminated?)
2. Are the blanks suitably placed with regard to the question contents?
3. Is the answer valid and unique?

It was pointed out that since the questions are automatically generated from corrections from Lang-8, the system in its current state couldn't provide questions corresponding to the academic level of a user. The evaluator determined that in reality "On one hand many questions contain basic elementary spelling errors, but on the other hand there are problems that cannot be answered easily even by advanced learners, and at this rate it is probably difficult to use in applications that include actual learning". It was suggested that analyzing the typical error patterns could alleviate these problems, as there is a "relatively clear level of difference". This analysis could then be used to adapt the quiz to the user's academic level.

Characteristic Errors in Lang-8

The following are the errors that most often appear in the analyzed samples of writings from Lang-8:

1. Spelling errors;
2. Preposition errors;
3. Articles etc., determiner errors;

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

4. Tense errors;
5. Lexical selection;
6. Sentence structure (word order).

It was observed that among the evaluated example questions that relate to a, b, c, d were on the whole often of elementary level. According to the evaluator, the first error A (spelling errors) is considered to be a common error for beginner to intermediate learners and occurs in vocabulary that learners use in writing without a relatively high level of vocabulary. The errors B (preposition errors) and C (articles etc., determiner errors) generally occur in a number of different patterns of varying difficulty, however the evaluator observed that the difficulty of these errors in the sample questions were predominately at elementary level. Considering this the evaluator suggested that the system be used to "actually create questions to focus on the patterns that are commonly used in practice".

In contrast, it was determined that it would be difficult to correctly answer automatically generated questions that contain errors E (Lexical selection) and F (Sentence structure, word order). The evaluator commented that vocabulary selection problems are difficult because "even if one understands up to the point of the error, it is generally hard to answer correctly especially if the context of the question has been stripped from the sentence". It was suggested that the use of multiple-choice fill-in-the-blank questions for vocabulary selection problems could be generated from other corrections sourced from Lang-8 data and this would enhance the system. It was determined that it was hard to correctly answer questions containing sentence structure errors as it involved rewriting the whole sentence, and that this also leads to another issue.

It was observed that the corrections of sentence structure errors on Lang-8 fall under two different categories: making minimal corrections that barely fix the sentence, and making substantial corrections with the aim of ensuring the highest quality of English is achieved. The evaluator commented, "on the whole, a lot of writings on Lang-8 are of Japanese High School level". In order to achieve a suitable level of English, most sentences would require substantial rewriting, however it was observed that most were only minimally corrected.

Comments Made by the Evaluators

There are three main groups of positive responses that we received from the evaluators as following:

- Sentence structure problems;
- Usefulness for teachers;
- Adaptive teaching based on statistical analysis of errors.

Sentence Structure Problems

Both evaluators identified problems that arise from the use of language learner writings. The language-learning researcher reported this as a demerit of the system, highlighting that as a result of using actual language learner writings that "sometimes the output questions are too long or short as the original sentence, from which it was generated, had the same long sentence length". This also affects the output of the automatically generated quizzes as the evaluator identified that "this also can sometimes impact on the relationship between the selections of generated multiple-choice questions and the contents of the sentence, and result in a weakening of this relationship". The English teacher also commented on the same criticism saying, "If the original English sample in the first place is not appropriate, then it is not going to be appropriate for use as a question". The evaluator had in fact pointed out that this was a prior apprehension. The English teacher also identified other types of problem sentences as listed below:

- In general, the questions become unnecessarily difficult when a character, word, or phrase of blanks need to have unity, and should be isolated when generating questions;

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

- Sentences consisting only of a long noun phrase and extremely short sentences should be excluded from candidate sentences;
- For actual practice, there needs to be some configuration that enable focus on topics such as grammar, but it is desirable to have separate functions of the parts of speech, such as the case of words that go into blanks.

These are challenges that we plan to investigate in the future to find possible solutions, for example: if a sentence contains an unusually large amount of n-grams when compared to an n-gram database of general corpus texts, it is therefore isn't appropriate for use as a sample. This provides a method of checking sentences before using them as samples for automatic question generation. We plan to investigate filtering candidate sample sentences in the future to improve the quizzes generated by the system.

Usefulness for Teachers

The language-learning researcher commented that teachers use a lot of time and effort in order to create test questions for practicing and testing of language knowledge and that using this system can enable teachers to produce these with efficiency and ease, for example in the following comment: "I think by using this system teachers can reduce their workload and burden". The output of the system can also be used in different ways depending on the teachers teaching method, such as an example that was provided by the language learning researcher: "Teachers may elect to use the output questions of the system as is, or use it as a base from which they can customize to meet specific needs".

The evaluator also commented on using the system to evaluate the linguistic level, for example: "As the difficulty of the question can be determined, it can be considered as a method to test a student's abilities". This would also help students realize their language ability and enable them to focus on weak points to improve their overall performance. The evaluator also made the following comments: "As the questions produced by the system are based on real writings of students, the aim of the questions is clear". In closing the evaluator commented "We are considering applying the system to actual use in the future". From this we can deduce that the evaluator found the system to be useful and they have an interest in using the system in actual learning environments.

Adaptive Teaching Based on Statistical Analysis of Errors

The English teacher commented that general when creating problems in an attempt to practice various patterns, frequent and non-frequent errors are often not distinguish between. Although many areas of improvement for the system to be of practical use were identified, the evaluator remarked, "The system in this paper provides the ability to automatically extract errors that a lot of learners are prone to, and then generate natural questions. On this point the system receives a positive evaluation". This indicates that the system has the potential for actual use, once the future work is carried out to improve the quality of the automatically generated quizzes. The evaluator also expressed that upon implementation of the aforementioned recommendations they would like to once again evaluate the system.

CONCLUSION AND FUTURE WORK

It is useful for learners to write in a foreign language and ask another person to correct it for them so they can know their errors. However, a problem for learners' is how to reflect on the correction of their problems and then master the error. Corrective feedback is an area that is of concern to teachers, researchers, and instructional designers alike. Although it is generally agreed that students expect teachers to correct written errors, and that teachers are willing to provide them, the immediate concern of many

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

teachers "is not so much to correct or not to correct", but rather when and how to respond to what students write (Lee, 2003).

This paper provides an alternative way to respond to students and provides them with a tool that can facilitate reflection, checking and mastering problems that caused errors in the past. Utilizing user created data in the form of journals on Lang-8, a SNS web site for language learning, analysis was undertaken to identify possible weaknesses in the usage of typical words and which words were typically misused in sentences. It was proposed that the use of two ratios, ErrorRatio and ErrorCauseRatio, provides a better measure of which words are used incorrectly within a sentence. In addition to the word-based analysis, error category based analysis was undertaken in order to compare the rates of errors that occur in second language writings on Lang-8 to that of previous English language learning writings. This resulted in the error categories that were of similar rates to past research being selected as candidate samples to reduce possible anomalies. Then a demonstration system for generating fill-in-the-blank quizzes using erroneous original user created sentences and the sentences corrected on Lang-8 to facilitate questions and answers.

Using the word analysis and error category analysis along with the generated fill-in-the-blank quizzes, it is proposed that users will be able to reflect on not only their own past errors, but also those of other students who have made same errors in the past, thus providing a greater capacity to enable students to master their past writing errors.

A Language Learning researcher and an English teacher undertook evaluations of the demonstration system with the outcome identifying both positive and negative characteristics. Possible uses of the system were stated; along with recommendations for future improvement as system development moves forward.

In the future, we plan to improve the system based on the comments and recommendations provided in the evaluations, and then by quantifying how useful it is in facilitating students reflection on their past errors. It is also planned to further evaluate the system using interaction design methodologies to investigate the effectiveness of its functions and user perceived satisfaction.

ACKNOWLEDGMENT

We are grateful to Lang-8 Inc., Japan for providing us with the data. This work was supported by JSPS KAKENHI Grant Number 24500176.

REFERENCES

Coleman, G. W., & Hine, N. A. (2012, October). Twasebook: A crowdsourced phrasebook for language learners using Twitter. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design* (pp. 805-806). ACM.

Computer assisted language learning. (n.d.) In *Wikipedia*. Retrieved December 27, 2012, from http://en.wikipedia.org/wiki/Computer-assisted_language_learning

Davies, G. (2011). Introduction to multimedia CALL. Module 2.2. In Davies, G. (Ed.), *Information and communications technology for language teachers (ICT4LT)*. Slough, Thames Valley University. Retrieved December 27, 2012, from http://www.ict4lt.org/en/en_mod2-2.htm

Fan, J. (2011). Constructing web-based learning environment for college English teaching. In *Computing and Intelligent Systems* (pp. 515–521). Springer Berlin Heidelberg. doi:10.1007/978-3-642-24010-2_69

Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, *11*(1), 33–49. doi:10.1016/0742-051X(94)00012-U

Johns, T. (1991). From printout to handout: grammar and vocabulary teaching in the context of data driven learning. In Johns, T., & King, P. (Ed.), Classroom concordancing (pp. 27-45). Special issue of ELR Journal 4, University of Birmingham, Centre for English Language Studies.

Kroll, B. (1990). What does time buy? ESL student performance on home versus class compositions. In B. Kroll (Ed.), *Second language writing: Re- search insights for the classroom* (pp. 140–154). Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9781139524551.014

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

Lee, I. (2003). L2 writing teachers' perspectives, practices and problems regarding error feedback. *Assessing Writing*, *8*(3), 216–237. doi:10.1016/j.asw.2003.08.002

Levy, M. (1997). *CALL: Context and conceptualisation*. Oxford, UK: Oxford University Press.

Lo, J. J., Wang, Y. C., & Yeh, S. W. (2008). WRITE: Writing revision instrument for teaching English. In Technologies for e-learning and digital entertainment (pp. 1-8). Springer Berlin Heidelberg.

Loughran, J. J. (2002). Effective reflective practice in search of meaning in learning about teaching. *Journal of Teacher Education*, *53*(1), 33–43. doi:10.1177/0022487102053001004

Meurant, R. C. (2010). Dematerialization and deformalization of the EFL/ESL textbook-literature review and relevant issues. In *Signal processing and multimedia* (pp. 235–244). Springer Berlin Heidelberg. doi:10.1007/978-3-642-17641-8_28

Walker, R., Davies, G., & Hewer, S. (2011). Introduction to the internet. Module 1.5 in Davies G. (Ed.), *Information and communications technology for language teachers (ICT4LT)*. Slough, Thames Valley University. Retrieved December 27, 2012, from http://www.ict4lt.org/en/en_mod1-5.htm

Weltig, M. S. (2004). Effects of language errors and importance attributed to language on language and rhetorical-level essay scoring. *Spaan Fellow Working Papers in Second or Foreign Language Assessment, 2*(1001), 53.

Yang, Y. T. C., & Chan, C. Y. (2008). Comprehensive evaluation criteria for English learning websites using expert validity surveys. *Computers & Education*, *51*(1), 403–422. doi:10.1016/j.compedu.2007.05.011

Brendan J. Flanagan received a BS in Information Technology (Computing Studies) from RMIT University in 2010. Since 2013, has been a graduate student in the advanced information technology of Kyushu University. Research interests include: text mining, search engines, language learning, and CSCL (Computer Supported Collaborative Learning).

Chengjiu Yin received his Ph. D degrees from the Department of Information Science and Intelligent Systems, Tokushima University, Japan, in 2008. He is an Assistant Professor in the Research Institute for Information Technology, Kyushu University. Currently he is committing himself in mobile learning, ubiquitous computing, language learning, text mining and SNS. He received the best paper award from ICIE in 2009. Dr. Yin is a member of JSiSE, JSET, and APSCE.

Sachio Hirokawa is professor of Research Institute for Information Technology, Kyushu University, Japan. He studied mathematics and computer science at Kyushu University. He was appointed to a research assistant at Shizuoka University in 1979, moved to Kyushu University in 1988 as Associate Professor and Professor in 1996. He received PhD in 1992. He has been involved in research and teaching in the area of mathematical logic and computer science. Since late 90s, His research focuses on search engine and text mining, where frequency analysis and visualization are the key features. He conducted 3 years project on search engine and became founder of start-up company Lafla (http://www.lafla.co.jp) to realize his technologies for commercial services.

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

Kiyota Hashimoto: B.A. and M.A. in Linguistics from Kyoto University, and Doctor of Engineering from Nara Institute of Science and Technology. He worked at Seiwa College and Osaka Women's University, and currently is an associate professor at College of Knowledge & Information Systems, Osaka Prefecture University, Japan. Some of his interests are natural language processing, educational engineering for language learning/teaching, tourism informatics, and soft computing in general.

Yoshiyuki Tabata is a Professor in the Research Institute for Information Technology, Kyushu University. He received his B.A. and M.A. degrees from Tokyo University of Foreign Studies, Japan, in 1983 and 1986. He is a member of JGG, JDV, GDDJ, JAECS and JEI. His current interests are in language learning environments supported by ICT.

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.