WHAT DO WRITING ERRORS TELL US?
A CASE STUDY OF JAPANESE EFL STUDENTS

Yuji Shuhama
Center for Fundamental Education, Dept. of Creative Engineering
National Institute of Technology, Tsuruoka College
Tsuruoka, Japan
shuhama@tsuruoka-nct.ac.jp

Abstract— **This article explores learner language of Japanese EFL students by analyzing linguistic errors in their essays and investigates the significant difference of error types by their levels of writing proficiency. The collected errors were categorized by error types to compile frequent error types characterizing different proficiency levels. The results show that the most common errors for all the students are related to wrong words and sentence structures. They also show that the elementary-level students frequently make spelling errors, while the intermediate-level students often make verb-related errors. The case study presented here contributes to a better pedagogical approach to promote Japanese EFL students' accurate writing skills evidenced by error analysis from a comparative linguistic perspective.**

Keyword— **Error Analysis, L2 English Writing, Japanese L1 Transfer, Learner Language.**

I. INTRODUCTION

The purpose of this article is to analyze learner language of Japanese high school–level EFL (English as Foreign Languages) students in the elementary to intermediate proficiency level through error analysis of their essays. The article also aims to contribute to developing more effective teaching methods of EFL writing by presenting our analysis.

Learning how to write seems quite difficult for many EFL students in Japan. According to a recent survey, the English ability of most Japanese high school students aged 17-18 ranges between A1 to A2 by Common European Framework of Reference (CEFR) standards. Of all the four essential skills, speaking and writing skills are remarkably low: more than 80% of the students are in A1 level and a lot of students were unable to answer almost anything [1].

Although writing is hard for Japanese EFL students, their teachers teach how to write, expecting that once they understand the basic rules for combining words into a sentence, they will be able to build a simple sentence. However, this is in fact not so easily realized due to various factors such as motivation, lack of correct knowledge, and L1 transfer. Of these factors, this article focuses on linguistic factors reflecting *learner language* [2].

Before going on to the next sections, let us present two research questions to be answered in this article: (1) Are the characteristics of writing errors different according to English proficiency levels? (2) What is the learner language of elementary and intermediate EFL students like?

II. BACKGROUND

Learner language (or *interlanguage*, [3]) refers to learners' developing knowledge of the target language. It is not merely an imitation of the target language but rather a systematic knowledge evolving as long as learners continue learning the target language. Since they are always in the process of perfecting their knowledge of the target language, learners keep updating their systematic knowledge, making linguistic mistakes and errors.

Mistakes and *errors* must be distinguished in terms of error analysis. A mistake is defined as a failure to utilize a known system correctly [4]. It is verbally realized as a random guess or a slip of the tongue, and it can be noticed and corrected by native speakers.

An error, on the other hand, comes from a learner's lack of correct knowledge of the target language, and. importantly, it cannot easily be corrected by the learners themselves. Such errors play an important role in error analysis, because by analyzing errors, what learners (mis)understand about the target language can be revealed.

Not only errors themselves but also the sources of errors are important. There are various potential sources of errors, but two types should be noted here: *interlingual* and *intralingual transfer*.

Interlingual transfer refers to positive/negative interference from the learners' first language [4]. For example, a lot of Japanese EFL learners tend not to match verb endings with subjects (e.g., *Yuta play the drums*; '*' means "grammatically incorrect"). From an interlanguage viewpoint, this mismatch can be explained by L1 Japanese interference, since no overt subject-verb agreement appears in Japanese. It should also be noted that the early stages of second language learning are often characterized by interlingual transfer from a learner's native language.

Another kind of transfer, intralingual transfer, does not result from a learner's first language characteristics but from their generalization based on the incomplete second language knowledge. One example of such a transfer might be *Does Yuta can play the drums?*, told by a learner who learns that *do* can alter a statement into a yes/no question.

One of several recent studies of Japanese EFL students' writing errors relevant to our research is Kato's 2006 report [5]. She analyzed the errors in 148 essays written by Japanese high school students. Her data shows that while the most common errors made by the first year students are lexical errors (related to word usage; 24% out of total errors), the most common errors made by the second- and third-year students are syntactic errors (related to sentence structures; 35% and 31%, respectively). Kato's research findings suggests that the type of errors Japanese EFL students commonly make gradually changes from lexical errors to syntactic ones as their learner language develops.

III. METHODS

The participants were 60 Japanese EFL students in a technical college in Japan. They were chosen from 158 students in total after all of them took the same English proficiency test discussed in more detail, below. They were all enrolled in the first year of the college and their ages ranged from 15 to 16. All of them speak Japanese as their first language. They have been learning English as a foreign language at least 3 years since they started learning English in junior high school.

In December 2014, all the 158 first year students took an English proficiency test titled 'GTEC for Students,' a standardized test of English proficiency produced by Benesse Corporation in Japan and especially designed for Japanese high school students. The test consists of reading, listening, and writing sections and gives the examinees feedback in the form of sectional scores, total scores, and grades corresponding with their English proficiency level.

In the writing section, the students were asked to write freely for 20 minutes about the following topic: "Write freely about what you want to achieve in one year." During the writing section, students were not allowed to use dictionaries or any devices to look up words. All the answer sheets were sent overseas and graded by several qualified native speakers of English. According to their score report, 5 students are ranked as Grade 1 (Preparatory Level; lowest), 17 as Grade 2 (Introductory Level), 112 as Grade 3 (Primary Level), and 25 as Grade 4 (Intermediate Level).

For this research, 60 essays were chosen from 158 essays and divided into two groups based on the writers' proficiency. One group is labeled as "Upper," a collection of 30 essays including 25 Grade-4 essays plus 5 highest-scored essays ranked as Grade 3. The Upper students' proficiency level is A2 in CEFR standards, judging from the average of their total scores. Another group is labeled as "Lower," a collection of 30 essays including 17 Grade-2 essays plus 13 lowest-scored essays ranked as Grade 3. The Lower students' proficiency level is A1 in CEFR standards. 5 Grade-1 essays were disregarded for consideration, because almost nothing that could be analyzed was written in these essays.

ERROR CATEGORIES

Error Types	Error Type Description
Verb errors (V)	All errors in verb tense or form, including relevant subject-verb agreement errors
Noun ending errors (N)	Plural or possessive ending incorrect, omitted, or unnecessary; includes relevant subject-verb agreement errors
Article errors (Art)	Article or other determiner incorrect, omitted, or unnecessary
Wrong word errors (W)	All specific lexical errors in word choice or word form, including preposition and pronoun errors. Spelling errors only included if the (apparent) misspelling resulted in an actual English word.
Sentence structure errors (SS)	Errors in sentence/clause boundaries (run-ons, fragments, comma splices), word order, omitted words or phrases, unnecessary words or phrases, other unidiomatic sentence construction
Spelling errors (SP)	Errors in spelling (other than those already classified as word choice)
Other errors (O)	Errors that do not fit into previous categories (may include capitalization, punctuation not already included in above types, and so on)

After dividing the essays into two groups, errors in the essays were identified and classified basically following the error categories proposed by Ferris and Hedgcock (2005) [6], which are listed in Table 1. Then, the frequency of each error type was calculated for Lower and Upper groups, and the frequent types of errors were compared between the two groups. The sources of errors were also considered for better writing instructions for Japanese EFL students.

IV. RESULTS

The results are summarized in Table 2, showing total numbers of errors found in 30 Lower essays and in 30 Upper essays per error category, frequency of each error category, and the rank orders according to frequency.

In total, 172 errors were found in 30 Lower essays and 199 in 30 Upper essays. It should be noted here that there was a remarkable gap in the total number of written sentences and the rate of error production between Lower and Upper groups. The total number of sentences written in Lower essays was 134, and 1.28 errors appeared per sentence in Lower essays. In contrast, the total number of sentences written in Upper essays was 334 and only 0.60 errors were made per sentence. Thus, the data shows that much fewer errors were found in Upper essays than in Lower essays.

As for frequent error categories, it is shown in Table 2 that a word-related error type was the most frequent error type for

the Lower group, while an error type related to sentence structure for Upper group was the most frequent. These two error types are the most common for both groups.

The most different error types between Lower and Upper groups are spelling and verb-related error types. The spelling error rate of the Lower group is 18.6%, which is almost doubled compared to the spelling error rate of the Upper group. Conversely, the verb-related error rate of the Upper group is 16.1%, and it is almost doubled compared to the Lower group. Article errors were also found in more than 10% frequency in both groups.

TABLE II. NUMBERS OF ERRORS AND THEIR FREQUENCY

Rank Order	Lower			Upper		
	Error	No. of errors	Percentage	Error	No. of errors	Percentage
1	W	47	27.3	SS	55	27.6
2	SS	39	22.7	W	49	24.6
3	SP	32	18.6	V	32	16.1
4	Art	22	12.8	Art	20	10.1
5	V	14	8.1	SP	19	9.5
6	N	12	7.0	N	16	8.0
7	O	6	3.5	O	8	4.0
	Total	172		Total	199	
	No. of sentences		134 (1.28/s)	No. of sentences		334 (0.60/s)

V. DISCUSSIONS

A. Word-related Errors

Now, let us discuss the results by looking at the actual data in detail. First, let us begin with errors related to wrong word usage, which are the most common error type in Lower essays and the second most common in Upper essays.

> Lo1: I am going to win to the tennis ^
> "I am going to win the tennis match"
> Lo2: Because I am helped ^ my mather
> "… because I am helped by my mother"
> Lo3: I want to ^ the Live on stage
> "I want to sing/perform in the live on stage"
> Lo4: I want ^ go to Kosen handball ship
> "I want to go to Kosen handball championship"
> Lo5: I'm not speak English
> "I don't/can't speak English"
> Lo6: I was play handball in juni high school
> "I played handball in junior high school"

One common feature of Lower essays is that prepositions are often misused: they can be unnecessarily inserted between verbs and their object (Lo1) or omitted before nouns inappropriately (Lo2). Also, as shown in Lo3 and 4, a simple phrasal verb *want to + verb* is used incorrectly. In addition, as in Lo5 and 6, *be* is wrongly used with lexical verbs. Both *want to + verb* and the distinction between lexical verbs and *be* must be learned in the students' early learning stage, but the above data suggest that these structures have not been mastered by the students in the Lower group.

Next, let us turn to several samples of word-related errors found in Upper essays for comparison. Wrong word choice observed in Upper essays seems to be caused by L1 Japanese interference. Adjectives like *strict* in Up1 and *tired* in Up2 do not match with the subject in meaning, but their Japanese counterparts (*kibishii* 'strict/competitive' and *tsukareru*, *akiru* 'tired/tiring') can be used with either human or event subject. Thus, there is a possibility that the students in Upper group tend to use English words wrongly associated with Japanese meanings.

> Up1: It is very strict
> "It (=the contest) is very competitive"
> Up2: It is tired, but very important
> "It (=practice swings) is tiring but important"
> Up3: I can't sing parfect yet
> "I can't sing perfectly yet"
> Up4: I will practice my weekness point.
> "I will overcome my weak point/weakness"
> Up5: I have to practice hard to be more better
> "I have to practice hard to be better"
> Up6: So I study English more harder
> "So I study English harder"

Choosing the wrong part of speech also seems to influence word-related errors in the Upper group. In Up3 and 4, *parfect* and *weekness* should be *perfect* and *weak*, respectively. The students in the Upper group have learned more English words than the students in the Lower group, but the above data imply that they will need to be able to choose a word in appropriate parts of speech from its derivatives and related words (e.g., *danger*, *dangerous*, *dangerously*, etc.).

Another frequent error type is adjectives and adverbs in comparative forms. As *more better* in Up5 and *more harder* in Up6 show, *more* is often misused with adjectives and adverbs inflected with a regular comparative suffix—*er*. This error seems to be an intralingual error caused by L2 overgeneralization: the students in Upper group tend to follow a wrong rule that adds *more* to any forms of adjectives and adverbs changes them into comparative forms.

B. Sentence Structure Errors

Second, let us observe errors related to sentence structures. These were the most common error types in Upper essays and the second most common in Lower essays. The data shown in the following box are three typical examples of sentence structure errors found in Lower essays. The students in the Lower group tend to form a sentence like Lo7, where topics appear in the sentence-initial position like subject but real subject is omitted. The reason for this tendency seems to come from a negative transfer of their L1 Japanese. In Japanese utterances, topics are often put in the beginning of sentences

and the subject is frequently omitted, so a sentence like *Winter, cannot use a tennis court* is possible in Japanese. It is known that some leaners heavily rely on the topic-comment structure of their L1 [7], so Japanese characteristics of topics and subject seem to negatively interfere in forming sentences for the students in Lower group.

> Lo7: <u>Winter is not use a court</u>
> "In winter we don't/can't use a (tennis) court"
> Lo8: I <u>hard</u> practice basketball
> "I practice basketball hard"
> Lo9: <u>Because</u> my dream is ^ singer
> "…because my dream is to be a singer"

Another typical error type is adverb placement. In Lo8, for example, an adverb *hard* is put right before a verb phrase *practice basketball*. Similar data to Lo8 is like **I have to <u>more</u> practice*, which should be *I have to practice (something) more*. Such adverb-placement errors can be seen as additional evidence of a negative transfer of the students' L1, because adverbs appear before predicates in Japanese.

The third frequent error related to sentence structure is the wrong use of *because*. As in Lo9, the students tend to begin a sentence with *because* but the clause is completely separated from a main sentence. In many examples where *because* was used, the students' sentences look like sentence fragments lacking main parts of statements.

To look closer at these kinds of sentence fragment errors, let us turn to similar and related errors found in Upper essays. Three typical error examples in Upper essays are shown in the following box.

> Up7: <u>Because</u> ^ very important subject and need in the life
> "…because they are very important subjects and necessary in our life"
> Up8: I think <u>teammate to practice</u>
> "I think my teammates will practice"
> Up9: My friends ^ very good player
> "My friends are very good players"

The sentence fragments introduced with *because* as in Up7 are frequently found in the Upper essays, too. Similar errors related to sentence fragments are found in other chunks introduced by *I think* as in Up8 and even in a sentence itself as in Up9. To use *because* and *I think* correctly, it is necessary for both Lower and Upper students to be aware that clauses consist of subject and tensed verbs and where clauses should be used. Being unaware of this is the reason why sentence fragment errors and the wrong usage of *because* are frequently observed in many students' essays.

C. Spelling Errors

As we have seen in Section 4, spelling errors are found in Lower essays more frequently than in Upper essays. Some typical spelling errors found in Lower essays are listed in the following box.

> Type 1: m<u>a</u>ther (mother), co<u>n</u>pan<u>ii</u> (company), t<u>o</u>rnament (tournament), cons<u>a</u>rt (concert)
> Type 2: chemist<u>l</u>y (chemistry), p<u>r</u>astic<u>k</u> (plastic), te<u>c</u>nic (technic), practi<u>s</u> (practice)

There are roughly two types of spelling errors: errors related to vowels (Type 1) and the ones related to consonants (Type 2). Most of the spelling errors found in Lower essays are related to vowel parts.

The students in Lower group tend to spell English words as similarly as they write and pronounce in *katakana*, which is one of the Japanese syllabary systems used for foreign loanwords. For example, 'mother' in katakana is written as マザー /maza:/, and 'company' as カンパニー /kanpani:/. Long vowels such as /o:/ and /a:/ are likely to be misspelled as *or* and *ar*, respectively, which seems to be influenced by corresponding katakana words (e.g. トーナメント /to:namento/ 'tournament' and コンサート /konsa:to/ 'concert').

The spelling errors in Type 2 are due to the difference of consonants and how to spell them in English and Japanese. English has a phonological distinction between the sounds of /r/ and /l/ (e.g., *right* and *light*), but Japanese has only one sound corresponding to these sounds. This fact seems to make it difficult for the students to distinguish spelling *r* and *l*, as shown in **chemistly* and **prastick* in Type 2 above. Likewise, the /k/ sound can be represented by *c, ch,* or *k* in spelling and /s/ by *s* or *c* in English. This kind of irregular sound-spelling correspondence is not observed in Japanese, so this influences the students' misspelling such as **tecnic* in Type 2 above.

These facts of spelling errors made by the students in the Lower group lead us to assure that misspelling of elementary-level Japanese EFL students is not just a mistake. Rather, their spelling errors should be regarded as L1 negative transfer influenced by the gap of sound-spelling correspondence between the two languages.

D. Some Common Mistakes

Finally, we will look at verb-related errors, which are typical of Upper essays.

> Up10: We ^ have a brass band concule next year
> "We will have a brass band contest next year"
> Up11: I ^ play very hard in winter vacation
> "I will play it very hard in winter vacation"
> Up12: I start_ play_ the sax last month
> "I started playing the sax last month"
> Up13: I must <u>thinking</u> for teammates
> "I must think (more) for teammates"
> Up14: I will <u>running</u> after school
> "I will run after school"
> Up15: I <u>practicing</u> table tennis hard every day
> "I practice table tennis hard every day"

The most common errors of this type are the use of bare verb forms without any tense, as shown in Up10, Up11, and Up12. *Next week* in Up10 and *in (coming) winter vacation* in Up11 both express future time and *last month* in Up12 past time, but neither *will* nor past tense forms are used. The students in Upper group tend not to match verb tense and time expressed in the sentences where the verbs are used.

The misuse of *–ing* forms are also observed frequently. As shown in Up13 and Up14, *thinking* and *running* are incorrectly used following modal verbs. In Up15, *practicing* must be replaced by *practice* since regularly scheduled routines can be expressed by the present tense. The *be + –ing* form has been learned in the early stages of the students' education as a form for expressing actions in progress, but the above data imply that the students in Upper group misunderstand that the *–ing* form can be used with modal verbs or even as a main tensed verb in a sentence.

In this survey, there were few errors of subject-verb agreement. However, this does not assure us that the students in Upper group have successfully mastered agreement. This is because in the essays they wrote, there were not many sentences using third person for the subject. As the above examples show, first person subject is quite frequently used.

E. Learner Language

So far, we have discussed four types of errors in detail, which include the most frequent and typical types of Lower and Upper groups. Now, let us characterize learner language of the two groups by summarizing the above discussion.

As we have seen in 5.1 and 5.2, word- and sentence structure-related errors are the most frequent types for both groups, but as Kato's (2006) report suggests ([5]; see Section 2), there is a tendency that the Lower group makes more lexical errors while the Upper group makes more syntactic errors. The Lower group's word-related errors are from a lack of correct knowledge of basic word usage: where to put prepositions or not and how to make a distinction between *be* and lexical verbs. On the other hand, the Upper group's word-related errors result from more complex factors such as L1 Japanese interference of word meaning and L2 overgeneralization of comparative forms of adjectives and adverbs.

As for sentence structure-related errors, the Lower group's errors are influenced by L1 Japanese grammatical structure. The topic-comment structure and adverb-predicate word order in Japanese are often relied on when the students in the Lower group make sentences. In addition, both the Lower and Upper groups are likely to be unaware of clauses being made up of subject and tensed verbs, which cause them to make sentence fragments frequently.

We have seen spelling errors in 5.3, which are more frequently made by Lower group than Upper group. These are not just mistakes but really errors influenced by the L1 syllabary system and the way of matching sounds and spellings.

In 5.4, we have observed verb-related errors, which are more frequently made by the Upper group than the Lower group. The most frequent error is that bare verbs are used and do not match with time expressed in a sentence. Also, the *–ing* form can be misused following modal verbs or as a main tensed verb in a sentence, which results from imperfect knowledge of how to use *be* + the *–ing* form.

Based on what we have found from the above error types, we can simply characterize the Lower and Upper groups' learner language as shown in the following box.

> Lower group: Lack of L2 knowledge of (i) basic word usage and (ii) clause structures; L1 negative transfer of (i) sentence structures and (ii) spelling
>
> Upper group: L1/L2 influence of word choice; Lack of L2 knowledge of (i) clause structures, (ii) matching verb tense, and time and (iii) correct usage of *be + –ing*

VI. CONCLUSIONS

In this article, error analysis has been made of the essay samples written by Japanese high school–level EFL students, and based on the analysis, learner language of the two different proficiency groups has been characterized. Finally, the conclusion of this article is presented as follows in a form of the replies to our two research questions.

(1) Are the characteristics of writing errors different according to English proficiency levels?

–Yes. The elementary-level students tend to make more word-related and spelling errors than the intermediate-level students. They also tend to be negatively influenced by L1 Japanese in terms of basic sentence structures and sound-spelling correspondence. However, the intermediate-level students tend to make more sentence structure errors than the elementary-level students, and their word choice is influenced by L1 Japanese and L2 English.

(2) What is learner language of elementary and intermediate EFL students like?

–Learner language of the elementary-level students lacks L2 knowledge of basic words and clause structures and also has L1 transfer of sentence structures, which often causes them

to produce incomprehensible English sentences. On the other hand, it is not so often the case that the intermediate-level students produce a lot of incomprehensible sentences, because such L2 knowledge of basic words and sentence structures has already been acquired in their learner language. But still, some local errors are made because of L1/L2 influence of appropriate word choice from their vocabularies and their lack of L2 knowledge of specific grammatical features, such as verb tense and the *be + -ing* form.

Our future task is to carry out the writing instruction based on the research findings of this article in an EFL classroom context. The keys of the instruction will be how to promote EFL students' in noticing and correcting their errors in the form of pre-writing activities or various kinds of feedback. The classroom writing instruction will be designed and its effectiveness will be presented in our future works.

REFERENCES

[1]. MEXT (The Japan Ministry of Education, Culture, Sports, Science and Technology), "A report on the survey of English abilities for the betterment of English education" [in Japanese], 2014, Downloadable as of Aug. 30, 2015 at: http://www.mext.go.jp/component/a_menu/education/detail/__icsFiles/afieldfile/2015/07/03/1358071_01.pdf

[2]. P. M. Lightbown and N. Spada, How Languages Are Learned, Oxford: Oxford University Press, 2006.

[3]. L. Selinker, "Interlanguage", International Review of Applied Linguistics, vol. 10, pp. 209–241, 1972.

[4]. H. H. Brown, Principles of Language Learning and Teaching, 4th ed., New York: Longman.

[5]. A. Kato, "Error Analysis of high school student essays", *Accents Asia* [Online], vol. 1, no. 2, pp. 1-13. 2006. Downloadable as of Aug. 30, 2015, at: http://www.accentsasia.org/1-2/kato.pdf

[6]. D. R. Ferris and J. S. Hedgcock, Teaching ESL composition: Purpose, process and practice, 2nd ed., Mahwah, NJ: Lawrence Erlbaum Associates, 2006.

[7]. T. Odlin, Language Transfer: Cross-linguistic Influence in Language Learning, Cambridge: Cambridge University Press, 1989.