

Measurement of Accuracy in Group Oral Interaction

Junko Negishi

Open Education Center, Waseda University

jnegishi@dream.com

Abstract

This study aimed to explore some of the characteristics relating to accuracy of second language learners' speech samples by means of error analysis and self-corrections. One hundred thirty-five students of junior/senior high schools and universities carried out oral interactions in a group of three and ten raters assessed the speakers utilizing the criteria in the Common European Framework of Reference (CEFR; Council of Europe, 2001). Among the 12 main error types, the following were the top five errors: verb, noun, article, pronoun, and preposition errors. The five types of errors were further examined by dividing them each into lexical choice errors, omission errors, and additional errors. A rank order correlation demonstrated that there was no strong relationship between neither the number of errors and the CEFR scores, nor between the number of self-corrections per total number of errors.

Keywords

Accuracy, Error analysis, Speaking, Group oral interaction

Introduction

According to Skehan and Foster (1999), accuracy is "the ability to avoid error in performance, possibly reflecting higher levels of control in the language, as well as a conservative orientation, that is, avoidance of challenging structures that might provoke error" (p. 96). Accordingly, novice learners' language contains a number of errors. Even returnees who can speak fluently with native-like pronunciations often make such errors. Second language learners, whose primary concern is accuracy, tend to have control over their internal elements and show a conservative attitude toward their language (Ellis & Barkhuizen, 2005).

Accuracy has been said to have trade-off effects with complexity and fluency. Relating to the acquisition order of these three aspects of performance, Swain and Lapkin (2001) reported that students in an immersion program acquired fluency first. By contrast, linguistic complexity kept

on developing in order to meet their academic goals with a cognitive load, while accuracy remained to be the last objective to attain. Such cognitive load has been evidenced to exert an impact on learners' performance, but interestingly were not always negative. For example, Wigglesworth (1997) found that when intermediate level students underwent a cognitively demanding test, they showed more accuracy in their language.

1 Measurement of Accuracy

1.1 Background

Ellis and Barkhuizen (2005) introduce several types of measurement: for example, the number of self-corrections per total number of errors, the percentage of target-like verbal morphology (the number of correct finite verb phrases divided by the total number of verb phrases multiplied by 100, used by Wigglesworth, [1997]) or the percentage of target-like use of plurals (the number of correctly used plurals divided by the number of obligatory occasions for plurals multiplied by 100, employed by Crookes, [1989]). Ellis and Barkhuizen point out a danger of investigating particular features because such an indicator "may not be representative of a learner's overall ability to use the L2 grammar" (p. 139) and "the learner's L1 may make a particular feature more or less easy" (p. 150). In other words, they point out the necessity to investigate overall ability of the learners' accuracy rather than pointing out a specific characteristic.

One of the widely used measurements of accuracy is error analysis; for example, the error-free clauses per total number of clauses have been employed in many studies (cf. Bygate, 1999; Foster & Skehan, 1996; Kormos & Denes, 2004). Although error-free clauses have widely been used in accuracy measurement, Ellis and Barkhuizen (2005) raised question about using a clause for the analysis when the data was obtained from oral interaction which contained a number of elisions. Ellis and Barkhuizen asserted that employing the errors per 100 words would eliminate the problem. Following their assertion, analysis of accuracy was carried out through calculating errors per 100 words

with self-corrected clauses as error-free. This measurement was employed by Mehnert (1998) who calculated the number of errors divided by the total number of words uttered divided by 100.

Owada (2005) conducted a 15 minute oral discourse test to 30 English learning Japanese by giving grammar error tag sets based on corpus analysis. The result demonstrated that the largest number of errors was omission (30.9%), followed by lexical choice (26.3%), tense (11.8%) and noun number (10.4%). In the case of omission and additional errors, the article errors were the largest and preposition in second. Checking all types of errors may give us more information about overall characteristics on the participants' accuracy.

1.2 Purpose of the Study

This study aimed first to find some features and developmental phenomena of the participants as second language learners of English at junior high school, senior high school and university level when they underwent group oral discussions in a group of three. Second, the relationship between the results obtained from the assessment given by raters utilizing CEF rating scales and the valuables obtained by the analysis of accuracy.

1.3 Present Methods of Analysis

Based on the previous studies described above, the following measures were taken to investigate the participants' accuracy. Overall errors were explored rather than analyzing particular grammatical features.

- 1) Types of errors made: following the error tag sets created by Izumi, Uchimoto, and Isahara (2004) specifically for analyzing spoken language by Japanese learners of English.
- 2) Ratio of self-corrected errors: the number of self-corrected errors divided by the total number of errors committed.

Before analyzing, a native speaker of English checked the participants' transcript and errors corrected. The correction was restricted to minimum and short in order to make the learners' utterances intelligible, not to modify the utterances to be native-like, proficient conversation. Then all the corrected errors were classified into 12 categories based on the error tagset (Izumi et al., 2004). Each error type was further divided into three features, that is, lexical choice error, omission, and addition, partly following Owada's (2005) classification. The twelve error types were as follows: noun, verb, auxiliary, adjective, adverb, preposition, article, pronoun, conjunction, relative, interrogative, and others. The others include Japanese-English, word coinages, lexical choice errors in a set phrase or collocation, word order

errors, and unintelligible utterances.

2 Procedure

2.1 Participants

The participants in the study were 135 students. They were divided into a total of 45 groups, each containing three students. The groups comprised fifteen junior high school student groups, fifteen senior high school student groups, and fifteen university student groups.

2.2 Data Collection

The speaking data were collected from each educational institution through the following process: (1) The students were randomly allocated into groups of three; (2) Each group drew a card on which one of the seven interaction topics—School, Family, Friends, Hobbies, English, Dream, and Culture (the last being only for university students)—was written down, and they were asked to speak on the topic; (3) Five minutes were allotted to each member of the group to plan his/her speech without speaking to the other members of the group; (4) Each member of the group introduced themselves for about half a minute as a warm-up activity; (5) Finally, the three students interacted orally as a group for five minutes on the selected topic. The interaction was videotaped to make DVDs for rating purposes.

2.3 Ratings

Ten Japanese teachers of English rated the participants. Before rating, they received training using a DVD produced by the Council of Europe. They rated the students by applying both a holistic rating scale and analytic rating criteria of the CEFR. The latter consists of five subcategories; Range, Accuracy, Fluency, Interaction, and Coherence. The raters assessed the students while watching a DVD by 7 scales: Below A1, A1, A2, B1, B2, C1, and C2. This paper solely reports the results of analysis concerned Accuracy.

3 Results and Discussion

3.1 Error Analysis

3.1.1 Overall Characteristics of Errors

The analysis of Accuracy was carried out by giving an error tag-set consisting of 12 main error types and each of the error types fell into the following three conditions—lexical choice error, omission, and addition. Along with describing the types of errors, the number of errors per 100 words will be explained.

Table 1 shows the total number of errors made by all the participants sorted into 12 main error

types. Out of the total of 1,318 errors, the largest number of errors, 365, is in the verb errors, followed by the noun errors' 230, the article errors' 165, the pronoun errors' 131, and the preposition errors' 123. The errors in other categories are less than 100 in each category.

Table 1: Total Number of Errors Made by All the Participants

Error Type	1	2	3	4	5	6	7	8	9	10	11	12	T O T A L
Number	230	365	4	46	70	123	165	131	75	8	23	78	1,318

The average numbers of errors for each educational institution were 5.8 for the junior high school, 10.6 for the senior high school, and 12.9 for the university. These values seemed to increase in line with their educational levels; however, when calculating the number of errors per 100 words, there was not much difference among them, that is, 13.2 for the junior high school, 15.4 for the senior high school, and 13.2 for the university.

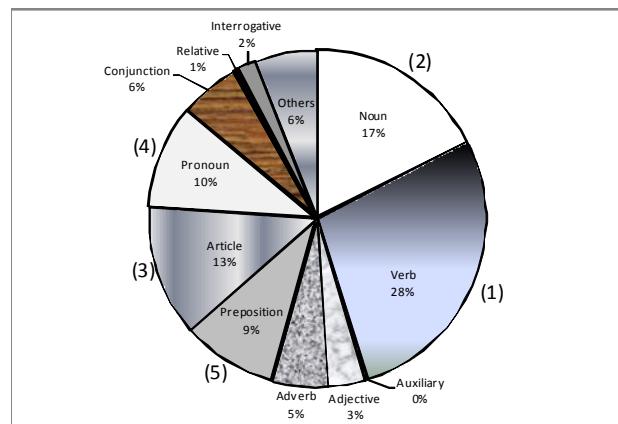


Figure 1: Proportion of each error type

Note: The numbers in parenthesis indicate rank orders

Figure 1 depicts the proportion of each error type within the total number of errors. The most common error is the verb errors (28%), followed by the noun errors (17%), the article errors (13%), the pronoun errors (10%), and the preposition errors (9%). The following sections look closely into these top five errors in detail.

3.1.2 Verb Errors

More than one-fourth of errors, that is 28%, were verb-related. For example, these were person and number disagreement, tense and aspect errors, finite/infinite and gerund errors, present/past participle errors, negative and interrogative errors, and lexical choice errors.

Figure 2 shows the number of three types of verb errors per 100 words which are sorted by the

three educational institutions. In terms of the lexical choice errors, the senior high school students make the largest number of errors and the university students the least. Most of the verb errors are related to the lexical choice errors. As for the verb omission, the number of errors decreases along with the participants' educational level. The number of omission errors is a little more than half of those of the lexical choice. By contrast, the number of additional errors is much fewer, counting less than one-tenth of the lexical choice errors. Another difference concerning the additional errors is that the higher the education level, the more errors are apparent. In other words, it may be more difficult for lower level students to make additional errors.

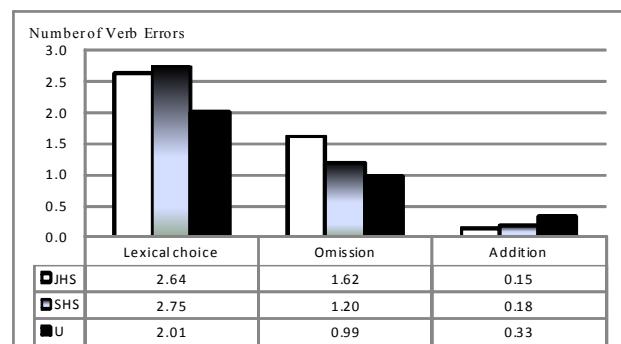


Figure 2: Number of three types of verb errors per 100 words sorted by educational institutions

3.1.3 Noun Errors

The noun errors account 17% of the entire errors made. They were singular/plural errors, noun case errors, countable/uncountable noun errors, lexical choice errors, and so on.

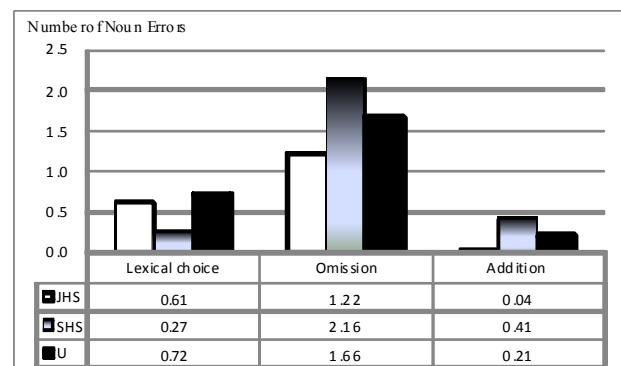


Figure 3: Number of three types of noun errors per 100 words sorted by educational institutions

Figure 3 demonstrates the number of three types of noun errors per 100 words. The number of lexical choice errors are less than one-third to one-tenth compared with the verb errors. The number of lexical choice errors made by the senior high school students are much less than the other two. On the contrary, the high school students make more omission and additional errors, followed by

the university students, and the junior high school students make the least. In the case of verb errors, the number of lexical choice errors is the largest but the noun errors are caused mostly by omission errors.

3.1.4 Article Errors

The third is the article errors which occurred 13% of the time, which is shown in Figure 4.

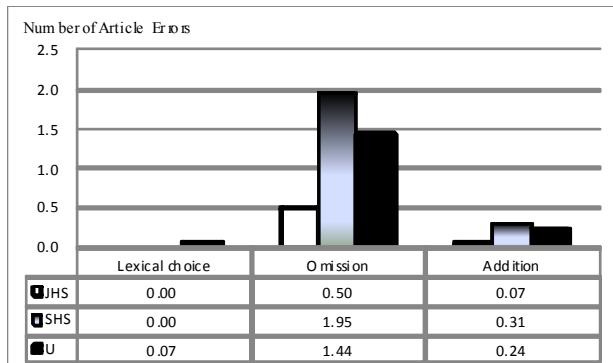


Figure 4: Number of three types of article errors per 100 words sorted by educational institutions

There are very few numbers of lexical choice errors in the use of articles. Few university students use definite article instead of indefinite article. No lexical choice errors concerning articles are found among the junior/senior high school students. The addition of articles is also few so that most of the errors are from omission of articles. Interestingly, the junior high school speakers make the least number of errors whereas the senior high school speakers make the most. It can be assumed that the junior high school speakers tend to use simple formulaic sequences repetitively, which may refrain them from article errors. The senior high school and university students do not repeat formulaic sequences so often; instead, they try to express themselves without being afraid of making mistakes. Regarding the fact that there is less number of errors made by the university students, it can be hypothesized that this is a process of their development.

With respect to articles, definite articles are said to be acquired first and later comes the indefinite articles. An analysis was conducted to find out some characteristics of definite and indefinite articles.

Figure 5 depicts the number and proportion of the three types of article errors per 100 words. As for the omission of indefinite articles, both the senior high school and the university students omit about one article per 100 words, and the numbers are three times as many as those of the junior high school students. The senior high school students omit definite articles the most, more than two times

as many compared with the university students and four times as many than the junior high school students, and these are one of the reasons that the senior high school students demonstrate the largest number of article errors. The addition of indefinite articles is few and no examples are observed among the senior high school students. Although the addition of definite articles is few as a whole, the senior high school students count the most. In summary, the article errors are predominantly omissions in indefinite articles more than that of definite articles. According to Biber, Johansson, Leech, Conrad, and Finegan (1999), the ratio of indefinite articles to definite articles in native speakers' conversation is approximately 13 versus 17. Meanwhile, the ratio of the indefinite article errors to the definite article errors of all the participants is approximately 13 versus 10. In this study, the participants likely to have used or made errors in the usage of the indefinite articles more than the definite articles.

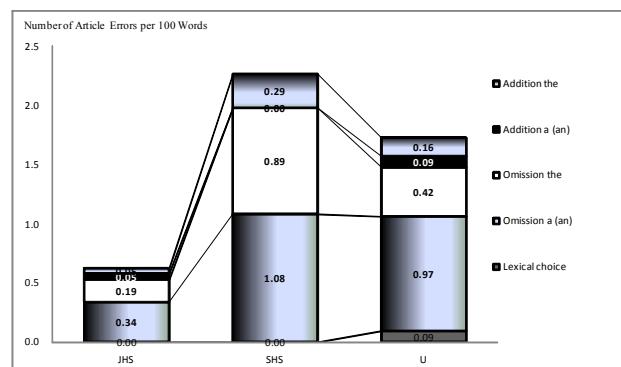


Figure 5: Number of three types of article errors per 100 words sorted by definite and indefinite articles and by educational institutions

3.1.5 Pronoun Errors

Throughout the entire errors made, pronoun errors occur approximately one-tenth of the time. Common errors are numbers and sex disagreements, case errors, and lexical choice errors.

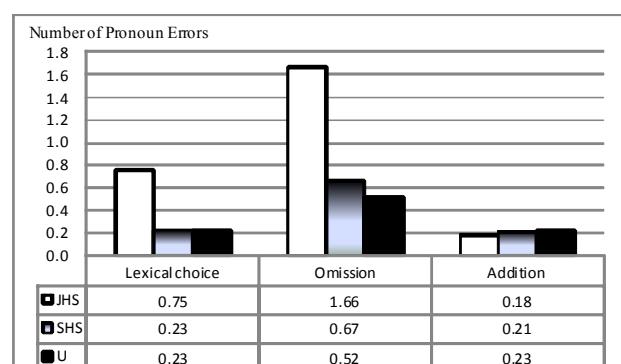


Figure 6: Number of three types of pronoun errors per 100 words sorted by educational institutions

Figure 6 shows the number of three types of pronoun errors per 100 words. In terms of the pronoun addition, there are few errors made across the three educational institutions. By contrast, the junior high school students make errors on lexical choice and omission. It can be inferred that the junior high school speakers have not yet acquired the usage of pronouns and it is difficult for them to use different types of morphemes according to context, which subsequently leads to errors. They often omit pronouns before nouns.

3.1.6 Preposition Errors

Lastly, preposition errors that occur nearly one-tenth of all the errors will be introduced.

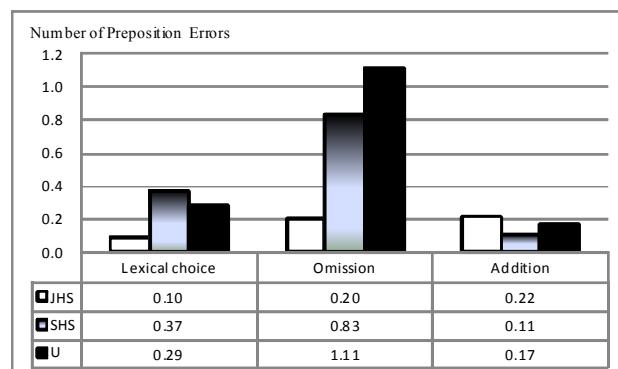


Figure 7: Number of three types of preposition errors per 100 words sorted by educational institutions

Figure 7 demonstrates the number of three types of preposition errors per 100 words. The characteristic in the preposition error is that the number of omission increases in reverse proportion to the speakers' educational level. It is difficult to infer what causes this phenomenon. Presumably, more proficient speakers may try to express themselves while taking some risks.

As for the junior high school speakers, the additional errors are seen the most, followed by the omission, and the lexical choice errors the least. By contrast, the senior high school and university speakers tend to omit prepositions followed by the lexical choice errors. Considering that the junior high school students are inclined to use fixed, simple sentences, the characteristics of the senior high school and university students may be more predictable.

3.1.7 Correlation Between Errors and Raters' Scores

To explore whether or not there are any relationships between the error types and the CEFR scores, a rank order correlation coefficient was calculated. Table 2 shows the correlation coefficients (Kendall's tau) between the top five

errors—verb, noun, article, pronoun, preposition—the total number of errors, and the CEFR scores for accuracy. The highest correlation stays $.235 p < .01$ in the article errors so the result indicates that there are no high correlation coefficients observed. Considering the feature of error analysis, the correlation coefficients need to show negative values, that is, the more errors the participants make, the less CEFR scores they get. It seems difficult to estimate the CEFR scores by means of error analysis. More accuracy-focused analysis should be conducted.

Table 2: Correlation Coefficients (Kendall's tau) Between Types of Errors and CEFR Scores for Accuracy

	Verb	Noun	Article	Pronoun	Preposition	Total error	CEFR
Verb	1	.117	-.027	.284**	.153*	.541**	-.070
Noun		1	.106	.180**	.130*	.378**	.147*
Article			1	-.107	.166*	.192**	.235**
Pronoun				1	.111	.386**	-.116
Preposition					1	.285**	.216**
Total error						1	.008
CEFR							1

Note: * $p < .05$, ** $p < .01$

3.2 Self-Corrections

The number of self-corrections per total number of errors is presented in this section. First, the number of self-corrections was counted and the total number of errors was calculated by the total of all errors obtained earlier. Then, the number of self-corrections was divided by the total number of errors.

The result demonstrated that the senior high school students corrected 15.5 % of the errors they made, which showed the largest percentage. The university speakers self-corrected about 12.1 %, while the junior high school students could self-correct the least among the three educational institutions, which was only 8.5%. It can be inferred that the phenomenon stems from the following reasons. The junior high school speakers have not acquired enough grammatical knowledge to correct errors or the knowledge to notice of their errors. The senior high school speakers tend to focus the most on grammar because their future entrance examination for universities weighs heavily on grammar, while the university speakers have already been released from entrance examination-based grammar. The correlation coefficient between the self-correction values and the CEFR scores was $.209$ at $p < .01$, a significant but not a strong result.

4 Conclusion

There were no items that showed a strong correlation with the CEFR scores. Although each type of errors demonstrated some features, not very many of them showed the participants' developments. Some of such few examples were the omission of verbs and pronouns. The junior high school students omitted verbs and pronouns the most, followed by the senior high school students and then the university students; that is, the more proficient the students became, the less omission of verbs and pronouns occurred. To be more specific, the junior high school students' pronoun omissions were more than three times, a higher number than those of the other two groups. This implies that the novice learners have difficulties in the practical use of pronouns. It is inevitable to encourage them use pronouns in everyday class until they become used to them.

Observing the three types of errors as a whole, omission errors were likely to occur the most, which subdivided into the four types of errors—noun, article, pronoun, and preposition; meanwhile, only the verb errors were seen as lexical errors the most. Interestingly, the additional errors were made by more proficient speakers, except the preposition errors. Additional errors seem to be more difficult as they may occur by overgeneralization which requires speakers to have a basic knowledge about grammar and rules.

References

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Pearson.

Bygate, M. (1999). Quality of language and purpose of task: Patterns of learners' language on two oral communication tasks. *Language Teaching Research*, 3, 185–214.

Council of Europe (2001). *Common European framework of reference for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.

Crookes, G. (1989). Planning and interlanguage variability. *Studies in Second Language Acquisition*, 11, 367–383.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299–323.

Izumi, E., Uchimoto, K., & Isahara, H. (2004). *Nihonjin 1200 nin no eigo supihking kohpas*. Dokuritsu Gyosei Hojin: Joho Tsushin Kenkyu Kiko. [English Speaking corpus produced by 1,200 Japanese]. National Institute of Information and Communications Technology (NiCT). Tokyo: Alc.

Kormos, J., & Denes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145–164.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20, 83–108.

Owada, K. (2005). Gakushusha kohpas [Learners' corpus]. In M. Nakano (Ed.), *Eigo kyoiku gurohbal dezain* [English education global design] (pp. 44–55). Tokyo: Gakubunsha.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language learning*, 49(1), 93–120.

Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks* (pp. 99–118). Harlow, England: Longman.

Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing*, 14, 85–106.