MEASURING ACCURACY AND COMPLEXITY OF AN L2 LEARNER'S ORAL PRODUCTION

Teguh Khaerudin
Center for Language Development,
Syarif Hidayatullah State Islamic University of Jakarta
t.khaerudin@uinjkt.ac.id

ABSTRACT

This paper aims at examining the influence of different tasks on the degree of task performance in a second language learner's oral production. The underlying assumption is that among the three aspects of language performance in L2, i.e. fluency, accuracy, and complexity, learners may prioritize only one of them (Ellis & Barkhuizen, 2005, p. 150) and that their decision to prioritize one particular area of language performance may be determined by the characteristics of the task given to the learners (Skehan & Foster, 1997). Having a written record of an oral production, the writer focuses this study on determining the degree of complexity and accuracy, and analyzing whether the different tasks change the level of learner's oral performance. The results show that learner's accuracy from both tasks remains in the same level. However, both task conditions, which do not allow speech plan, result in no improvement in accuracy level and a minor improvement in the complexity level.

Key Words: learner language, complexity, accuracy, oral production, task types

ABSTRAK

Artikel ini bertujuan untuk melihat bagaimana performa siswa dipengaruhi oleh penugasan berbahasa lisan yang berbeda. Penelitian ini ditulis berdasarkan asumsi bahwa siswa mungkin akan memilih untuk memprioritaskan salah satu dari tiga aspek kemampuan bahasa asingnya, yaitu kelancaran, ketepatan, dan kompleksitas (Ellis & Barkhuizen, 2005, p. 150) dan bahwa pemilihan prioritas yang dilakukan oleh siswa ditentukan oleh karakteristik dari penugasan yang diberikan (Skehan & Foster, 1997). Melalui catatan tertulis dari produksi lisan siswa, penulis memfokuskan penelitian ini untuk melihat tingkat ketepatan dan kompleksitas dari performa lisan siswa, dan menganalisa apakah penugasan berbahasa lisan yang berbeda menyebabkan perbedaan tingkat ketepatan dan kompleksitas dalam performa lisan siswa. Hasil penelitian ini menunjukkan bahwa siswa menunjukkan tingkat ketepatan yang sama meski dalam penugasan yang berbeda. Sementara itu, dua penugasan berbahasa lisan yang diberikan, keduanya adalah penugasan berbahasa secara spontan, tidak memperbaiki performa lisan siswa dalam hal tingkat ketepatan dan sedikit saja meningkatkan kompleksitas bahasa lisan mereka.

Kata kunci: bahasa siswa, kompleksitas, ketepatan, bahasa lisan, jenis penugasan

INTRODUCTION

There have been many different approaches used by linguists to examine the development of second language (L2) learning. Until the late 1960's, examining L2 proficiency, especially in criticizing its imperfection compared to the native speaker level, was primarily performed by contrasting L2 to the learners' first language (L1). L1 became a very important source of L2 learner from which L2 learners would compare the new rules of L2 to a more familiar L1 rules. Robert Lado was among the first to give foundation to the comparative method, which was later known as Contrastive Analysis Hypothesis, between second or target language and native language. In his most influential book, he expressed his belief that "in the comparison between native and foreign language lies the key to ease or difficulty in foreign language learning" (Lado, 1957, pp. 1-2). Furthermore, Contrastive Analysis Hypothesis would see L2 speech as a deviated version of the target language (Lightbown & Spada, 1999). Heavily influenced by structural linguistics and behavioral psychology, contrastive analysis claimed to be able to better prepare L2 learning materials by comparing the target language to the learners' L1. By the same procedure, this analysis also claimed to have the capability of predicting learners'

behavior and possible difficulties in learning L2. However, due to its incapability of explaining many aspects of L2 learners' language, such as its incapability of truly predicting L2 learners' errors, and its overwhelming emphasis on the negative transfer of L1 to L2 when explaining learner errors, this view shifted in the late 1960's.

As contrastive analysis hypothesis declined, another approach flourished in the late 1960s. The alternative theory, known as 'Error Analysis', had different assumption about L2 learners' language. It assumes that L2 learners' language is a rule-governed system of language and is predictable. Error analysis sees that learners' making faulty inferences about L2 rules has been a major cause of errors, as contrasted to negative transfer from L1. However, like its predecessor theory, error analysis was a subject to a great number of critics especially regarding its problematic methodology.

In 1972, Larry Selinker coined the term *Interlanguage* to identify the learner's developing L2 knowledge (Selinker, cited in Lightbown & Spada, 1999). This interlanguage has been a subject of many researches in the study of Second Language Acquisition (SLA) ever since. SLA itself is a relatively new discipline as a sub-field of Applied Linguistics. It was born in early 1970s

resulted from the views concerning the 'systematicity of learner's language' (Larsen-Freeman, 2000). Unlike Applied Linguistics which focuses on language as its epicenter, Second Language Acquisition studies the process in which people acquire a second language and its problems. SLA, as Larsen-Freeman (2000, p. 165) puts it, "draws on multidisciplinary theoretical and empirical perspectives to address the specific issue of how people acquire a second language and the specific problem of why everyone does not do so successfully". In other words, the main goals of SLA are determining the level of interlanguage and its development (Lakshmanan & Selinker, 2001).

The analysis of a learner language development in Second Language Acquisition studies focuses on determining learner's language level of proficiency systematically and describing how aspects of the learner's L2 competence change over time (Ellis & Barkhuizen, 2005). L2 learners' language is generally analyzed in terms of its:

1. features: accuracy, complexity, and fluency;

2. forms: past tense –ed, articles, conventional indirectness; and

3. function: negation, temporality, aspect, request.

These three aspects of language is examined both of oral and written production (Ellis & Barkhuizen, 2005). This paper, however, will only focus on assessing L2 learner's language in terms of its accuracy and complexity in L2 learner's oral production. Accuracy is defined as "how well the target language is produced in relation to the rule system of the target language" (Skehan, cited in Ellis & Barkhuizen, 2005, p. 139). Furthermore, Ellis & Barkhuizen (2005, p. 139) consider that "*complexity* is the extent to which learners produce elaborated language while *fluency* is the production of language in real time without undue pausing or hesitation."

Different measures have been developed to analyze learner's oral production. Generally, the measurement tools can be used to asses both oral and written productions in terms of their complexity and accuracy, but with different procedure for fluency (Ellis & Barkhuizen, 2005). In analyzing L2 oral production, researcher can use a transcribed speech data to be segmented into unit for further analysis (Foster, Tonkyn, & Wigglesworth, 2000).

METHOD

Research design

This study reported in this paper examines the level of accuracy and complexity in an L2 learner's language. Two transcriptions of an L2 learner's oral production in responding to two different tasks are analyzed. The two different tasks are: 1) free conversation, and 2) movie retelling. The accuracy level is determined by dividing the number of error-free clauses by total number of AS-unit multiplied by 100% (Pauline Foster & Skehan, 1996). An AS-unit is defined as "a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with it" (P. Foster et al., 2000, p. 365). Then, the complexity level is determined by calculating the number of different verb forms used (Yuan & Ellis, 2003).

Participant

The participant of this study is an Italian native speaker, for the sake of privacy named Vito, who migrates to an English speaking country. With all four language skills assessed as poor and very poor, Vito represents those who survive living in foreign country with very little interaction with the target language. Relatively poor performance in his oral production can

be traced back as a result of lack of input and interaction between Vito and the target language. This paper studies two excerpts of Vito's oral production taken in December 6th, 1984, that is four years after his arrival in the target country.

Measures

Two excerpts of learner's oral production, LIEVI33A (free conversation) and LIEVI33I (movie retelling), are segmented into AS-unit. The first excerpt is a transcription from one side of a cassette while the other is from two sides of a cassette. Therefore, the writer uses only the first part of LIEVI33I (movie retelling) excerpt intending to have similar length of speech duration between two excerpts.

FINDINGS AND DISCUSSION

In assessing spoken language performance of the participant, the writer segments the text into unit, namely AS-unit. In the application of this analysis of speech unit (AS-unit), Foster et al. (2000) prescribe three levels of inclusiveness for different purposes of research and different types of spoken language data. As suggested, this study applies the Level Two considering its appropriateness to cope with interactional data. This level of inclusiveness excludes one word minor

utterance and verbatim echo responses (P. Foster et al., 2000).

Using two excerpts of an L2 learner's oral production, the writer analyzed the learner's utterances by calculating the number of AS-unit the learner can produce in each utterance. The calculation was divided into two categories based on the two different speaking tasks that the learner responded to. The calculation results in the following number of AS-unit, as described in table 1.

Table 1. AS-unit

EXCERPT	TASK TYPE		AS-UNIT
LIEVI33A.1TR	1.	conversation, anecdote about buying shoes	27
LIEVI33I.1TR	2.	broken off film i.e. Modern Times	34

The data shows a remarkable difference of the number of AS-unit in learner's production between two tasks. Task type 2 results in more oral production by the learner in relatively same length of time. This can be understood as the movie provides the learner more information to report than an impromptu conversation does. These figures, however, are not the main concern in this paper for the reason that the writer focuses more on accuracy and complexity. In the subsequent analysis, the writer examines the learner's L2 proficiency by assessing the speech data in terms of its accuracy and complexity.

Accuracy

Some alternatives of measures have been used by researchers for determining accuracy. Ellis & Barkhuizen (2005) summarize them in the following table 2.

Table 2. Measures of Accuracy (Ellis & Barkhuizen, 2005, p. 150)

MEASURES	DEFINITION	STUDY
Number of self-correction	The number of self-corrections as a percentage of the total number of errors committed.	Wigglesworth (1997)
Percentage of error-free clauses	The number of error-free clauses divided by the total number of independent clauses, sub-clausal units and subordinate clauses multiplied by 100.	Foster and Skehan (1996)
Errors per 100 words	The number of errors divided by the total number of words produced divided by 100.	Mehnert (1998)
Percentage of target-like verbal morphology	The number of correct finite verb phrases divided by the total number of verb phrases multiplied by 100.	Wigglesworth (1997)
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.	Crookes (1989)
Target-like use of vocabulary	The number of lexical errors divided by the total number of words in the text (excluding dysfluencies).	Skehan and Foster (1997)

Among others, percentage of error-free clauses along with errors per 100 words are frequently used (Ellis & Barkhuizen, 2005). Using the framework from Foster and Skehan (1996), i.e. percentage of error-free clauses, the writer gets the following figures in table 3.

Table 3. Accuracy

EXCERPT	ERROR-FREE CLAUSE	PERCENTAGE OF ERROR-FREE CLAUSE
LIEVI33A.1TR	3	3/27 * 100% = 11%
LIEVI33I.1TR	4	4/34 * 100% = 11.7%

Table 3 shows that free conversation task (excerpt LIEVI33A.1TR) yielded very slightly lower accuracy (11%) than the movie retelling task (excerpt LIEVI33I.1TR), which could yield 11.7% accuracy. In movie retelling task, learner produce more extra error-free clause than the first task, however, this does not contribute much to the percentage as more error clauses are produced as well. This, however, shows that while the learner tried to produce more utterances, his low level of L2 vocabulary and grammar hindered him from producing accurate utterances. There seems to be a trade- off phenomenon between the learner's complexity and accuracy in performing this task.

Complexity

Assessing the level of L2 complexity can be performed by several methods. Furthermore, determining which method to be used would need to be based on what aspect of language a study intends to measure. Ellis and Barkhuizen (2005) summarize measures that have been used byresearcher in analyzing complexity.

Table 4 Measures of Complexity (Ellis & Barkhuizen, 2005, p. 154)

MEASURE	DEFINITION	STUDY
Interactional		
1. Number of turns	The total number of turns performed by each speaker is counted. This can be then be expressed as a proportion of the total turns in the interaction. Alternatively, the average number of words for each speaker can be calculated.	Duff (1986)
2. Mean turn length	The total number of words (or pruned words) produced by a single speaker divided by this speaker's total number of turns.	
Propositional		
3. Number of idea units	The total number of (a) major and	Zaki and Ellis (1999)

MEASURE	DEFINITION	STUDY
encoded	(b) minor idea units **in** the text is counted. Major and minor ideas are established with reference to a baseline performance of the message (e.g. by a native speaker).	
Functional		
4. Frequency of some specific language function (e.g. hypothesizing)	The total number of times a specific language function is performed by a learner is counted. This measure can be expressed as a raw frequency or relationally (e.g. in terms of total AS-unit)	Brown (1991)
Grammatical		
5. Amount of subordination	The total number of separate clauses divided by the total number of c- (or AS) units.	Foster and Skehan (1996)
6. Use of some specific linguistic feature (e.g. different verb forms)	The number of different verb forms used.	Yuan and Ellis (2003)
7. Mean number of verb arguments	The total number of verb arguments (subject, direct objects, indirect objects, adjectival complements, prepositional phrases) divided by the total number of finite verbs.	Bygate (1999)
Lexical		
8. Type-token ratio	The total number of different words used (types) divided by the total number of words in the text (tokens).	Robinson (1995)

In this study, the writer tried to measure the L2 learner's complexity by calculating the number of subordination he could produce in his utterances. However, measuring complexity by calculating the number of subordination in this excerpt would give too little information as the writer found only one subordination was produced in excerpt LIEVI33I.1TR and none in excerpt LIEVI33A.1TR. Therefore, in this paper, the writer uses the grammatical measure proposed by Yuan and Ellis (2003) by counting the syntactical variety in learner's oral production. The number of verbs occupied by the learner is measured as follow.

Table 5. Complexity

EXCERPT	VERBS USED	NUMBER OF VERBS	TYPE-TOKEN RATIO
LIEVI33A.1TR	Know, is, want, speak, think, do, work, working, look, pick, like, say, try, match, keep, come back	15	15/36 = 0.42

EXCERPT	VERBS USED	NUMBER OF VERBS	TYPE-TOKEN RATIO
LIEVI33I.1TR	think, have, is, walking, walk, look, come back, keep, crush, run, tell, pick, took, speak, work, buy, did, say, do, come in	20	20/55 = 0.36

The Type-token ratio analysis shows little complexity in vocabulary use as shown by the table above. The table shows that movie retelling task (excerpt LIEVI33I.1TR) could facilitate the learner to produce more verbs than the free conversation task (excerpt LIEVI33A.ITR). Within an approximately same length of time, 55 verbs were produced in movie retelling task compared to 36 verbs in free conversation task. This figure shows that the complexity in movie retelling task was somewhat improved in terms of the number of verbs. However, when the data was further analyzed using the measure of type-token ratio, the complexity level in movie retelling task is lower than that in the free conversation task.

Movie retelling task could motivate the learner to produce more utterances even though he had very limited vocabulary. However, because of the limited number of verbs produced, the type-token ratio of movie retelling task was lower than in the free conversation task. In the free conversation, on the contrary, L2 production was somewhat limited, but somehow it helped increase the type-token ratio.

CONCLUSION AND SUGGESTIONS

In this part of paper, the writer sums up the findings of the analysis in the following table 6.

Table 6. Accuracy and Complexity

TAS TYPES	ACCURACY	COMPLEXITY	
		Vocabulary	Type-token
Free conversation	11 %	15	0.42
Movie retelling	11.7 %	20	0.36

The table shows that the level of accuracy yielded from the two tasks are very slightly different and insignificant, between 11% and 11.7%. The movie retelling task results in slightly more accurate oral production than the free conversation task does. Provided that the percentage is separated only by .7%, it can be perceived that free conversation task and movie retelling task facilitate the same level of accuracy. Whereas in the level of complexity, the writer identifies a slightly more complexity resulted from movie retelling task which is shown by the more number of verbs produced from the task. The writer does not take into account the type-token ratio, which

shows the opposite result, as it is perceived merely caused by learner's lack of vocabulary.

Given the condition of both tasks as spontaneous speech or not allowing time for speech plan, the writer concludes that complexity is improved in movie retelling task but not accuracy even though the degree of difference is not so much significant. No extra time for planning the speech might be the case of this limited improvement even though learner has more information to share in different task. This finding corresponds to that of Mehnert (2000) and correlates to that of Yuan and Ellis (2003).

Further research on how different tasks could facilitate L2 learners in terms of accuracy, complexity, and fluency will still be a valuable endeavor in the study of Second Language Acquisition. The learner's low level of L2 proficiency has supposedly prevented this study from getting more accurate data of how movie retelling and free conversation tasks could facilitate L2 accuracy and complexity. For that reason, assessing more learners with more varied level of L2 proficiency will be more conclusive. Furthermore, more extensive data as more proficient learners tend to produce could be more preferable in providing more convincing analysis.

REFERENCES

Ellis, R., & Barkhuizen, G. (2005). Analysing accuracy, complexity, and fluency *Analysing learner language* (pp.139-164): Oxford University Press.

Foster, P., & Skehan, P. (1996). The Influence of Planning and Task Type on Second Language Performance. [143 Reports: Research; 080 Journal Articles]. *Studies in Second Language Acquisition, 18*(3), 299-323.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: a unit for all reasons. *Applied Linguistics, 21*, 354-375.

Lado, R. (1957). Linguistics across Culture: Ann Arbor: University of Michigan Press.

Lakshmanan, U., & Selinker, L. (2001). Analysing interlanguage: how do we know what learners know? *Second Language Research, 17*(4), 393-420. doi: 10.1177/026765830-101700406

Larsen-Freeman, D. (2000). Second Language Acquisition and Applied Linguistics. *Annual Review of Applied Linguistics, 20*(-1), 165-181. doi: 10.1017/ S026719050020010X.

Lightbown, P. M., & Spada, N. (1999). Learner language *How languages are learned* (2nd ed., pp. 71 - 90): Oxford University Press.

Mehnert, U. (2000). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition, 20*(01), 83-108. doi:10.1017/ S0272263198001041

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research, 1*(3), 185-211. doi: 10.1177/136216889700100302

Yuan, F., & Ellis, R. (2003). The Effects of Pre-Task Planning and On-Line Planning on Fluency, Complexity and Accuracy in L2 Monologic Oral Production. *Applied Linguistics, 24,* 1-27.