

Difficulties of Scoring Accuracy in Students' Writing

Sato Rintaro (佐藤 臨太郎)

In an English Writing Class we often ask students to write an essay. However, evaluating students' writing is not an easy task at all. Their writing can be scored holistically or analytically, and can be analyzed from the points of view of accuracy, fluency or complexity. In this essay, I will introduce some difficulties I found in scoring students' essays, specifically focusing on accuracy. The essays were written by my second-year high school students on the topic of "My Summer Vacation."

In the students' essays, there were sentences with grammatical errors, which although conveyed an intelligible idea, were considered incorrect because I decided, in evaluating the essays, to regard only grammatical sentences as accurate. In the following examples, (1) is grammatical and (2) is just acceptable.

(1) Last week I went to the beach with my friend.

(2) *Last week I go to the beach with my friend.

(2) is acceptable enough for communication but due to its ungrammaticality it was regarded as inaccurate, and I assumed that this measurement was valid. In some instances, sentences which seemed to be impressive were judged as inaccurate for their ungrammaticality, and monotonous sentences without grammatical errors were judged as accurate. The following are examples:

(3) *During the summer vacation I *would be*

more happy, if I had been with my father who had gone to Sapporo to work very hard for us, our family.

(4) I was happy. My father came home.

Students who tried to write grammatically complicated sentences so that they would express what they really meant were more likely to create inaccurate sentences than those who made simple sentences. It is clear that the former students are more proficient than the latter, which unfortunately was not reflected in the accuracy scores. The following are the impressive but inaccurate sentences written by the students:

(5) *On the first day of Minato Matsuri I really enjoyed the Firework Festival from the top of Mt. Hakodate when *the sky is very clean*.

(6) *I didn't like practicing hard in the gym for *much times* because I became very thirsty after *play* hard in the games.

(7) *Every day I came to school to take classes *taached* by my teachers because my dream is to be a high school teacher who *teach* a lot of important things to students such as music and English.

The fact that the samples introduced here (3, 5, 6, 7) were treated as erroneous, inaccurate sentences while (4) was considered as accurate, that is to say, the fact that the learners' avoidance¹ cannot be well dealt with, makes me wonder about the difficulty of measuring

accuracy in writing.

Ellis and Barkhuizen (2005) mention a problem arising with linguistic forms of “possible-but-not-preferred” (p.59) sentences. They insist on the importance for researchers to decide whether or not they should treat possible-but-not-preferred sentences, or sentences with “dispreferred forms” (p.59) as erroneous. They are in favor of not counting that type of sentence as inaccurate due to the subjective judgment of teachers to decide whether a sentence is dispreferred or not. In evaluating students’ essays, taking their position, I restricted the analysis to absolute errors as opposed to dispreferred forms.

(8) ? This summer I had a plan *of studying* very hard.

(9) ? *When I was playing basketball, I became very tired.*

(10) ? We went to the restaurant and ate very *fast.*

These are the samples of dispreferred forms. A preferred sentence, for (8), is “This summer I had a plan to study very hard”, (9) must be “After I played basketball, I became very tired,” and in (10) “quickly” should be used instead of “fast.” However, as these sentences were communicative enough, I was not able to stand that I should regard them as absolutely erroneous, and ended up deciding to treat them as sentences without errors. The following are sentences which were written concerning the same event.

(11) ? There was a soccer training camp.

(12) *During the summer vacation there was a soccer training camp for five days in this school because *to win the games.*

(11) is a dispreferred sentence but was judged as accurate, while (12) was regarded as

erroneous though the sentence is lexically and grammatically more dense. We could assume that the student who wrote (12) is more proficient in writing than the student who wrote (11). How we should deal with the issue of dispreferred forms remains to be solved to improve accuracy scoring.

In scoring accuracy, I was concerned only with absolute lexical and grammatical errors. Apparently this measure is valid, but even among absolute errors I found something to think about. The following sentences were all dismissed as erroneous:

(13) *If I did not busy, I finished homework.

(14) *If I had time, I finished my homework.

(15) *If I had had enough time, I could have finish my homework.

We could easily guess what the writers meant to say, but due to grammatical errors the sentences were not counted as accurate. This measure must be valid, but we have to notice that distances or gaps between each sentence and the reconstructed accurate one (“If I had had enough time, I could have finished my homework”) are different. It is clear that (15) shows, even though incorrect, the greatest degree of acquisition of the subjunctive mood, with (13) being the least. Further analysis of some erroneous samples which were intended to use the same grammatical forms implied a different degree of acquisition of the forms. That is to say, we can see the different degree of how their productions deviate from the target language norm. Ideally this difference, if it is definite, should be reflected in the accuracy score, which seems to be very difficult, if not impossible, because of its reliance on subjective judgment.

The term “error,” which broadly means wrong productions, has been used in this essay. James (1998) distinguishes errors from mistakes

in learners' wrong productions. He defines mistakes as "...self-correctable only with the benefit of feedback," adding that "...if self-correction is still impossible despite the provision of simple factual feedback without specification of the nature of deviance it is defined as an error" (p.238). When we detect wrong productions in their writing we had better classify them into either mistakes or errors, and it might be better to count not mistakes but errors for accuracy scoring. The following are samples from the students' writings.

(16) *First, we enjoyed *to watch* TV and next we enjoyed *to playing* soccer.

(17) *We went to Sapporo to enjoy shopping in the morning and go to Asahikawa in the afternoon.

Student (16) never showed the correct use of gerunds after "enjoy" while student (17) made a mistake of tense only in the sentence, which makes me assume (16) is an absolute error while (17) is a mistake. However, this classification has not yet reached beyond just my assumptions. Student (16) may show the correct use of gerunds after "enjoy" when he is given unspecified implicit feedback. How errors and mistakes can be distinguished is still a big issue. To have an individual interview with the learner to examine whether he/she is able to self-correct the deviant form can be one of the possible solutions to the problem. However, this is not practical considering the number of students (Ellis and Barkhuizen, 2005).

It is definitely important for students to be provided with opportunities to write essays to improve their English writing and beyond that their thinking. However, unless we establish a more valid and reliable measurement of their writing to score students' essays, they may be discouraged from writing. We should strive

for the solution.

Note

1. "Avoidance" refers to a phenomenon in which students avoid using a difficult word or a structure, and instead use an easier word or structure. Corder (1978) states avoidance strategies will not foster acquisition.

References

- Corder, S.P. (1978). Language-learner language. in J. Richards (Ed.), *Understanding Second and Foreign Language Learning*. Rowley, MA: Newbury House.
- Ellis, R. and Barkhuizen, G.P (2005). *Analyzing Learner Language*. Oxford: Oxford University Press.
- James, C. (1998). *Errors in Language learning and use: exploring error analysis*. London: Longman

(市立函館高等学校教諭)