Test of Error Correction and Metalinguistic Knowledge

Dan Brown

Northern Arizona University

Abstract

This report summarizes the development, administration, and results of a language test that was

piloted in a ENG 105 course within the Program of Intensive English in the Fall 2012 semester.

The purpose of this research came out of course requirements for ENG 638: *Second Language

Assessment*, and the results of the piloting of the this test development project will ultimately be

used in conducting a research study on the relationship between writing students' metalinguistic

knowledge and their ability to make use of corrective feedback on grammar. This relationship is

has received scant attention in the second language writing literature, and could provide

theoretical grounds for pedagogical techniques in giving feedback on grammar in second

language writing. This summary will include the background for the project, a description of the

test, its design and purpose, along with a brief summary of the results and the relevance of this

research to the Program of Intensive English and to second language research.

Test of Error Correction and Metalinguistic Knowledge

Background

The motivation for designing the test that was piloted in the PIE this semester was to create an measurement instrument for an empirical research study that would investigate the role of second language (L2) learners' level of metalinguistic knowledge and their ability to utilize indirect (coded) feedback on their grammar in writing. Indirect feedback, in which L2 teachers use an abbreviated code to represent different error categories for feedback on grammar, is a commonly used approach for giving L2 feedback, although research to date has provided few clear guidelines for its effective use.

The opportunity to pilot a draft of this test in an upper level PIE course was a valuable opportunity to develop experience in analyzing test statistics with theoretical backing (Bachman & Palmer, 2010; Miller, Linn & Gronlund, 2010). It also helped paved the way for the development of a more reliable and valid measure of the constructs we are looking to compare in the eventual research study.

Research Questions

In order to apply theoretical backing to the quality of the test under development, we aimed to answer the following question as s result of the pilot testing: Using the concepts we learned in ENG 638, to what extent did our test measure the constructs we defined reliably and validly?

Methods

Participants in this study were students from a single section of the level 5 composition class (ENG 105). The design of the pilot test was based on the few examples we could find in our literature review of metalinguistic measures used in L2 research (Akakura, 2012; Morris, 2003). Following guidelines from the instruments developed for these studies, our test included 20 items: 10 items to test the sub-construct of procedural knowledge in error correction ability. In these items, test-takers were supplied 10 sentences, each containing a grammatical error. The focus was on five different types of errors: subject-verb agreement, articles, prepositions, pluralization, and sentence fragments, as these represent some of the most common categories that L2 teachers focus on in their feedback (Ferris, 2011), and we predicted these areas to be structures of focus in the eventual research project. These questions were scored dichotomously (correct or incorrect).

Each of the error correction questions were followed by a question representing the second sub-construct, declarative metalinguistic knowledge. For each of these 10 questions, test-takers were asked to write a short explanation of the grammar rule they used in making each correction. These items were scored on a scale using a rubric we developed for this test. The scoring was reported separately for each sub-construct and test-takers were supplied with a score report form that included a summary of each students' strengths and weaknesses in both error correction and declarative metalinguistic knowledge, across each error category investigated on the test.

The test was administered by the ENG 105 teacher who kindly agreed to use this test in her class, and it took between 15 and 30 minutes for students to complete. We consulted the Human Resource Protections Informed Consent Forms for the class and identified 13 students that provided consent to participate so that we could analyze their data.

Results

The results of the test were analyzed for descriptive statistics, including means, standard deviations, minimum and maximum scores, and Cronbach's Alpha for internal reliability of each sub-construct section. We also analyzed item discrimination and item difficulty for each individual item, which allowed us to identify four items that deserved attention in form of editing. Results of test piloting revealed that the internal reliability of the error correction test items were low (Cronbach's Alpha of 0.5), but with revision, we feel this section could improve to within an acceptable level to be a reliable measure for research purposes. The metalinguistic knowledge items resulted in dismal internal consistency (Cronbach's Alpha of 0.28), which can be partially attributable to the weaknesses in the errors we constructed. However, based on these results, we came to two conclusions about this section of the test. First, if we were to keep this structure for this portion of the test, we would like to pilot these type of questions again using a more advanced L2 sample population—possibly a group of L2 English teachers in training. This way, we could better understand the reliability of these items, as we expect some of the items to be quite difficult and that a large portion of the target sample population (intermediate to advanced L2 learners) might not know the answers. This would be acceptable for our research purposes, although a more proficient (or more metalinguistically knowledgeable) group for the pilot would be helpful. Alternatively, we could scrap this portion, based on our results, and re-design the metalinguistic portion using multiple choice items, which proves a useful format to measure basic declarative knowledge (Miller et al., 2010).

Relevance to PIE and Second Language Learning

Given that feedback on grammar in L2 writing is such a time intensive practice for many L2 teachers, a better understanding of how students interact and make use of this type of feedback would be beneficial for both pedagogical and theoretical purposes. With a valid and reliable instrument to measure these constructs, we hope to contribute to knowledge of this relationship, which could guide teachers in how to make indirect feedback on grammar in writing most effective for their students.

In the context of the Program of Intensive English, such a test could be beneficial for writing classes in which teachers give feedback on their students' grammar in their writing. Employing such a test could offer both students and their teachers awareness of strengths and weaknesses in their knowledge of grammatical rules, which could be applied towards improved performance in written grammatical accuracy. In the class that participated in this pilot study, we included a summary of analysis of strengths and weaknesses for each sub-construct for each student, in hope that this awareness would be beneficial, both for students' individual awareness, and for their teacher to discover what categories might deserve more or less attention if time is devoted to explicit grammar or awareness building on grammar in writing.

References

Akakura, M. (2012). Evaluating the effectiveness of explicit instruction on implicit and explicit

    L2 knowledge. *Language Teaching Research*, *16*(1), 9-37.

Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. New York NY: Oxford

    University Press.

Ferris , D. R . (2011). Treatment of Error in Second Language Student Writing. Ann

    Arbor, MI: University of Michigan Press.

Miller, M. D., Linn, R., & Gronlund, N. (2010). *Measurement and 'evaluation in teaching.* (10th

    Ed.). Upper Saddle River, NJ: Merrill, Prentice Hall.

Morris, L. (2003). Linguistic knowledge, metalinguistic knowledge and academic success in a

    language teacher education programme. *Language Awareness*, *12*(2), 109-123.