

## **“My Elbow is the Same As Your Elbow”: From Grammatical Structure to Communication in Classroom Assessment**

*Andrea Kalwara\**

Hawai‘i Pacific University

### **Abstract**

This paper explores the possibilities and challenges in testing grammar forms communicatively. Using an actual case of a beginning-level class of English as a Second Language, it outlines the steps of and considers issues in designing a grammar test on comparison constructions while attempting to maintain communicativeness and authenticity.

### **Introduction**

As a language teacher and language learner, I am well aware of the importance of communicative language teaching. When teaching grammar classes, I try to show students that grammar has relevance in everyday interactions by showing the usage of specific grammar structures in real-life situations. However, when it comes to testing grammar, I usually fall back on old-fashioned multiple-choice or gap-filling tests that are likely to be high in reliability and practicality. My situation is described by Purpura (2004), who observed that many teachers put a lot of effort into teaching grammar communicatively, focusing on form and meaning; however, for testing, “they rely exclusively on traditional multiple-choice or blank-completion tasks of grammatical form”(p. 21). From my own experience, this approach to testing grammar is inadequate, as students who usually perform very well on these traditional tests still fail to produce the tested form correctly when speaking in real-life communication. For example, the class average of a recent assessment I administered targeting present and past tense structures was 90%; however, when students communicated, they still produced sentences such as “She go,” “Yesterday I go,” and “I study last night.” Multiple-choice tests can be highly reliable, especially if they contain a sufficient number of items; yet, performance on the test might not be a valid measure of students’ ability to use the target structures proficiently in real-life communicative contexts (Anderson, Clapham, & Wall, 1996, p. 187). Therefore, I was excited by the opportunity to collaborate with my colleagues to design an assessment to measure students’ abilities to use target grammatical structures while performing an authentic speaking task. This way, I could determine not only whether students know the grammatical rules and structures but also whether



Kalwara, A. (2018). “My elbow is the same as your elbow”: From grammatical structure to communication in classroom assessment. *TESOL Working Paper Series, 16*, 97-103.

Website: Hawaii Pacific University <http://www.hpu.edu>

\* Email: [akalwara@my.hpu.edu](mailto:akalwara@my.hpu.edu). Address: MP 411, 1188 Fort Street Mall, Honolulu, HI 96816, USA.

students could proficiently use the target structures in authentic communicative contexts. Below, I will describe the assessment itself, discuss the challenges we faced, and reflect on how these challenges may be addressed in the future.

### **Assessment Objectives**

The purpose of this test was to find out whether students could accurately produce simple structures for comparisons using *the same as*, *similar to*, *different from*, *like*, and *alike*. These structures are from *Basic English Grammar* (Azar, 1996), sections 15-5 and 15-6 of chapter 15 (pp. 470-474).

The test was created for a 12-week High Beginner Grammar class that met for two hours each class day. The class used the *Basic English Grammar* (Azar, 1996) textbook, supplemented with some materials for communicative purposes. The students' proficiency level was approximately equivalent to the A2 level in the Common European Framework of Reference for Languages (CEFR). Students' goals were to become competent in English so that they could not only socialize and meet new people but also develop English skills that would help them in their careers. Based on the course syllabus, the class objectives focused on grammar study for spoken and written communication. However, the textbook's heavy focus on grammar, the school's curriculum, and the students' learning objectives, each was geared toward non-communicative pedagogy, made it challenging for students to develop communicative competence. As the teacher of this class, I tried to modify the grammar points into speaking activities using pair work and group work. Students practiced phrases and simple sentences using the target grammar structures and participated in a variety of pair/group discussions, surveys, and role-plays that were meant to encourage the development of their communicative competence.

The students were assessed biweekly to determine whether they had acquired the grammatical structures in the chapter. It would have been easy to assess only their knowledge of grammar, but I wanted to assess their communicative competence as well, as the major goal of language teaching is to enable students to use the target language as a means to accomplish their goals and successfully navigate their social environments. Therefore, my colleagues and I designed a communicative assessment that would measure the students' abilities to use the target grammatical structures in their speaking. The assessment itself is included in Appendix A.

### **Dilemmas and Future Solutions in Assessment Design**

#### **Form vs. Function**

In attempting to design an authentic assessment which would measure students' production of language forms to achieve communicative functions, we quickly encountered a paradox. Language form includes morphology and syntax, and during assessments, we are mainly concerned with determining the accuracy of the students' production of these language forms (Purpura, 2004, p. 59). In contrast, "language function" refers to how people use the forms in communication and interaction with others when expressing their ideas (Rutherford-Quach, Hsieh, Zerkel, Skarin, & Wiles, 2015). Integrating these two language aspects—form and function—in a single test seems like an ideal combination for testing students' language abilities. However, this ideal combination requires careful and thorough analysis of how language forms are used in the employment of language functions in real-life communication. Further, when it

comes to production in communication, students may choose among many forms to express a given function, which means that in order to elicit a sample of the target grammatical forms, we need to restrict students' freedom in their choice of forms. As we discovered during the test administration, by forcing students to use the assigned language form, we sacrificed the communicative part of the task as well as the task's authenticity.

Because of the constraints of a grammar-driven syllabus, our group felt obliged to design and restrict the test to particular grammar forms that the students were learning that week. Students' task for the assessment was to compare mundane objects using these exact 5 structures: *the same as*, *similar to*, *different from*, *like* and *alike*. In our attempt to achieve some authenticity through relevance of content, we asked students to bring their own objects, which included pens, cups, wallets, shoes, bags, phones, and books. To ensure that students would produce the target forms, my colleagues and I wrote the target grammatical forms and the names of the objects on the whiteboard, instructing students to only use the forms they see written (Appendix B). Although the students performed very well (90% of the students scored above the passing grade), I felt that the assessment design gave students ample time to think ahead and check the forms on the whiteboard before producing their sentences. As Carroll (1961) stated, "if testers limit the test only to one structure at a time, the students have more time to think than they would have in normal communication" (as cited in Bachman, 1990, p. 300). Thus, by giving students the extra time to consider the target forms, the assessment lost its authenticity, as the conditions for natural and communicative contexts were not achieved. Although our students had five grammatical structures to use, it was still a very controlled task that allowed students to read the grammar structures first and then produce them. Students were too dependent on these particular structures and used the whiteboard as a crutch when they forgot the structures. To avoid this dependency in my future assessments, one should avoid writing such structures on the board. This way one could also assess whether the students can use the structures in conditions which simulate real-life interaction. However, without a reminder of the target forms, students may not use them in their oral production, and the assessment would fail to collect a sample of the target structures. Another solution is to study how these target structures are actually used in real life and then recreate those situations in the assessment. Some situations that may be likely to elicit the target forms are two friends choosing clothing items to buy, speed dating, wherein people learn about each other to find similarities and differences, or evaluating cities in another country to determine whether they are good candidates for being a sister city (see further discussion below). Even then, there is no guarantee that students will actually use the target forms unless they are required to do so, and this is due to the nature of language: the same function can be expressed through the use of more than one form.

Not writing the target forms on the board could also resolve another problem which arose during the implementation of the assessment. The communicative aspect of the assessment as well as the interaction between students also suffered because of our form restriction and the lack of a natural purpose for the interaction. Students produced forced language as required, and their communication was not as interactive as we would have liked. Before the test, students were divided into groups, and in their groups, they had to take turns producing the given structures and comparing their objects. Students did not have a conversation with a meaningful purpose;

rather, they waited patiently for each other's turn to end before initiating their turn. Most sentences that the students produced were predictable ("My book is the same as your book," "My phone is similar to your phone," "My pen is like your pen"), which is not always the case in a real conversation. Interactiveness that contributes to the accomplishment of a task is one aspect of a good speaking assessment (Bachman & Palmer, 2009, p. 25). I noticed the lack of interaction during the administration of the assessment, as some students kept their eyes on the whiteboard to make sure they were using all the assigned forms instead of interacting with their group members.

### **Test Structure vs. Student Creativity**

During the test, the students compared their own and their partners' objects, which allowed them to relate to the task on a personal level and went some way in encouraging interaction. Their actual performance on the test pointed to a contrast between the assessment's designed structure and students' creativity. To my pleasant surprise, some students ventured outside the assigned objects, making the task more interesting, meaningful, and interactive for them. Instead of sticking to books, pens, wallets or other assigned objects, students creatively exploited other items as well: *my mind is different from your mind; my elbow is the same as your elbow; my finger is similar to your finger; my thumb is like Dan's; my hairstyle is like JJ's hairstyle; or my face and Oguri Shun's face are alike* (Oguri Shun is a famous Japanese actor). As Bachman & Palmer (2009) asserted, a task that requires test takers to relate the topical content of the test input to their own topical knowledge is likely to be more interactive than one that does not (p. 25). In our case, the topical knowledge included the familiarity with the objects that the test takers compared. However, natural interactions normally include two or more people talking to each other about subjects that they think are mutually interesting and relevant in the situation (Luoma, 2004, p. 20). Although familiarity with the assigned objects which students were asked to compare introduced a measure of personalization, the missing communicative function (or purpose) of the task prevented the students' production from resembling natural conversations. To compensate for this problem, some students creatively invented their own purpose (humor or social relationships with classmates) by going outside the list of the assigned objects. Students and assessments benefit when this kind of spontaneity is introduced. For example, one might ask students to look around and make comparisons between them and their classmates based on the objects they brought to class, the objects they see, or imaginary objects. One might challenge students to think of as many humorous comparisons as possible, making this test potentially more interactive. Alternatively, the goal of this activity could be to find similarities and differences among members of the class in order to conduct a class survey.

### **Topic Authenticity vs. Task Authenticity**

Along with designing a communicative task to test language forms, our group also strived to introduce authenticity so that not only could we measure their communicative competence in a real-life context, but also encourage students to see the relevance of the test to their everyday language use. Bachman (1990) defined authenticity as the real life (RL) approach, i.e., "the extent to which test performance replicates a non-test language performance," mirroring the "reality of

non-test language use” (p. 301). Likewise, Bachman and Palmer (2009) stated that authenticity and the relevance to target language use of the test content and task may help promote test takers’ positive and affective response to the task, which in turn may lead students to perform their best (p. 24). The authors maintained that learners’ performance on the language test need to be relatable to learners’ language use in a “specific domain other than the test itself” (p. 23). Bachman (1990) added that authenticity includes “the interaction between the test taker and the task,” and it becomes identical with “communicative language use or the negotiation of meaning” (p. 317). The real-life component in our task was meant to be the authentic objects that students brought in from their daily lives and compared, e.g., their cell phones, books, wallets and other objects. However, when I think back on the task, asking students to compare their objects does not meet Bachman and Palmer’s criteria of the relevance of the target language produced or even Bachman’s requirement of an authentic test, including negotiation of meaning. In our real-life conversations and interactions, we do not have many instances in which we compare objects just for the sake of comparing them. In addition, if we do compare objects in our environment, we usually do not use the forms *the same as*, *similar to*, *different from*, *like* and *alike*, all together. To increase the authenticity of such test tasks in the future, it is vital to strive to connect the target structure with real life situations. One promising route involves using the Corpus of Contemporary American English (COCA), where one can identify situations in which these grammar structures are used and adapt this usage to the language level of the test takers. This way, students might have the opportunity to experience how the target expressions are used outside of the classroom. To find out the most common use of comparison expressions, I looked at the SOAP and GlobeWE’s corpora. The contexts usually involved situations wherein people compared personalities or the past with the present, as in the following: *Ryan was different from your father; we’re so much alike; my income is the same as it was ten years ago; and her father’s house was almost the same as years ago* (GlobeWe). There was also a trend of comparing music and TV shows that ran through different websites, together with comparisons of people and life situations, as was found in COCA and SOAP: *single songs for a music library is [sic] the same as Twitter feeds; the show was similar to Oprah’s show; the boy is similar to my dad; and a broken promise is not the same as a lie*. Thus, for a test, one might design a situation wherein students survey their classmates’ tastes of music, TV shows, and/or their classmates’ personalities. Another option could also be a comparison of their homes in their countries and in the United States. A more complex proposal would involve setting up four stations with different tasks. One station would have pictures of typical homes from the students’ home country and typical homes from the U.S. Students would have conversations comparing these different homes. Other stations could include pictures of movies, TV shows, food, or people’s personalities. Students could work in groups of three or four. This way, students would have their eyes directed at the pictures or their classmates, rather than turning their heads toward the whiteboard. As this test would be for a high-beginner class, I would expect short responses. For instance:

- A: My home in Japan is *different from* my home in Hawaii. In Japan I had two bedrooms and in Hawaii, I only have one.
- B: Really? My home in Japan *is similar to* my home in Hawaii. I have two bedrooms and one kitchen.

- C: My home in Japan and Hawaii are *alike*. They both have a backyard and a patio.  
D: In my home, the kitchen is *the same* size as my kitchen in Japan.

### Conclusion

In sum, the class syllabus' constraint to adhere to testing specific grammar structures took away from the interactivensness and authenticity associated with communicative language testing. Although we attempted to add authenticity to the task by introducing comparisons of students' own objects, the lack of purpose and interaction in the students' groups significantly decreased the authenticity of the task. This assessment experience shows that it is not a simple task to establish a natural relationship between testing grammar forms and testing communication in which functions are the primary focus. Having a clear definition of what to test and deciding which is more important to test, isolated grammar points or communicative functions, is a first step in developing an authentic grammar-based communicative test. I propose two kinds of tests: one form-driven (less communicative) and the other one function-driven (more freedom in forms). First, one could administer to students the standard writing test that is mandated by the school, after which students could use the forms in a speaking test that would reflect everyday language use of those forms. Teachers could view form-driven and function-driven tests as a continuum rather than a dichotomy. This is mainly because while it is possible to have a completely form-driven or a completely function-driven test, a combination of elements from both ends of the spectrum may be pedagogically necessary to inform teachers and learners of the learners' strengths and weaknesses in language skills and knowledge.

### References

- Alderson, J., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge, UK: Cambridge University Press.
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford, UK: Oxford University Press.
- Bachman, L., & Palmer, A. (2009). *Language testing in practice: Designing and developing useful language tests*. Oxford, UK: Oxford University Press.
- Carroll, J.B. (1961a). Fundamental considerations in testing for English language proficiency of foreign students' in *Testing the English Proficiency of Foreign Students*. Washington, DC: Center for Applied Linguistics: 30-40.
- Davies, Mark. (2008-). *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. Available online at <https://corpus.byu.edu/coca/>.
- Davies, Mark. (2013). *Corpus of Global Web-Based English: 1.9 billion words from speakers in 20 countries (GloWbE)*. Available online at <https://corpus.byu.edu/glowbe/>.
- Luoma, S. (2004). *Assessing speaking*. Cambridge, UK: Cambridge University Press.
- Purpura, J. (2004). *Assessing grammar*. Cambridge, U.K.: Cambridge University Press.
- Rutherford-Quach, S., Hsieh, H., Zerkel, L., Skarin, R. & Wiles, S. (2015). Module Two: Task analysis. Stanford University. Retrieved from the ELPA21 Consortium Website.

## Appendix A

### Test Instructions

Instructions on March 6<sup>th</sup>, 2018: On the test day, bring to class your pen, a cup, a wallet, a bag, a phone, and a book.

Instructions on the test date:

- Put your objects in front of you
- Compare objects using constructions: *the same as*, *similar to*, *different from*, and *alike*.
- Use the various constructions independently and produce 2 unique, grammatically correct sentences comparing your and your classmates' objects
- Listen to each other's sentences to avoid saying the same comparisons your classmate

### Speaking Test Grading Rubric

	Fluency	Grammar	Vocabulary	Size
<b>1</b>	Constant pauses, very slow speech, nervous hesitation, and/or inability to complete sentence.	Constant mistakes that take away meaning.	Minimal range of vocabulary. Constant errors that interfere with meaning.	Not able to produce two sentences using the learned structures
<b>2</b>	Frequent pauses, slow speech, and/or hesitation. Able to complete sentence.	Frequent mistakes that take away meaning.	Limited range of vocabulary. Frequent errors that regularly interfere with meaning.	Able to produce only half a sentence
<b>3</b>	Irregular Pauses, and/or minimal hesitation.  Complete sentences.	Regular mistakes that do not overly interfere with meaning.	Regular errors that do not overly interfere with meaning.	Able to produce one sentence with help from the Interlocutor
<b>4</b>	Minimal Pauses.  Complete sentences.	Some mistakes that do not interfere with meaning.	Mostly accurate with only a few errors that interfere with meaning.	Able to produce two sentences with minimal help from the Interlocutor
<b>5</b>	Natural pattern of speech for the level.	Very accurate, few mistakes.	A wide range of vocabulary used accurately.	Able to produce two sentences independently

## Appendix B

Test instructions written on the whiteboard for the students

1. Look at all the objects in front of you.
2. Choose one and compare it to your object
3. Use two different structures (*the same as, similar to, different from, like, and alike* )

---

### **About the Author:**

*Andrea Kalwara* (MA TESOL candidate, Hawai'i Pacific University) is an ESL Instructor at Central Pacific College in Honolulu. Her research and teaching interests include formative classroom assessment, teacher-student rapport, and teacher talk.