

On Empiricism and Objectivity

David Wiese

Teachers College, Columbia University

Even under the best of circumstances, placement testing is subjective. The ACTFL Proficiency Guidelines (ACTFL, 1985) and the updated ACTFL Writing (Breiner-Sanders, Swender, & Terry, 2001) and Speaking (ACTFL, 1999) Proficiency Guidelines are yet another example. The core of the problem is the use of purely descriptive guidelines in lieu of measurement guidelines that incorporate empirical evidence into the scales.

The four language skill proficiency guidelines are divided into five proficiency levels, which are then further subdivided into sublevels, for a total of ten distinct categories. Each of these categories is elaborated in rich detail, usually in a paragraph-long description. Many very discrete examples of what learners can do at each stage are given. For example, the speaking guidelines at the Intermediate-Mid level state that learners can “talk simply about self and family members” (ACTFL, 1999, p. 4). Certainly this is specific; however, it is also subjective. What does it mean, for example, to talk simply? In my class, beginning students talk simply about their family. What is the correlation, then, between simplicity and correctness? What about the quantity of language that is produced? If Learner A produces *I have two brothers*, and Learner B, albeit erroneously, produces **I have one old brother, and one sister more younger than me*, which learner has performed better given the existing criteria?

However detailed they may be, the ACTFL Guidelines are full of these vagaries. At the Reading-Intermediate-High level, structural complexity may interfere with comprehension (ACTFL, 1985). Does this mean that it may, or that it may not? Is this also to say that for advanced students, structural complexity does not interfere with comprehension? No matter what, the placement tester is forced to make an interpretation. In many cases, a tester could rate a learner within a range of three or four categories (from the total of ten) and still legitimately defend his or her choice. The abundance of examples makes no difference—the central issue of subjectivity still remains.

What is needed is a set of guidelines for operationalizing abstract concepts such as “talking simply” or “structural complexity” into quantifiable, and thereby more objective, measures. Possible measures of complexity include the average number of clauses per sentence, average number of words per sentence, or what percentage of words used is not in the 2,000 most common word families in the target language. Complexity can also be cross-referenced with accuracy by comparing the average number of clauses or words per sentence with the average number of syntactic errors per sentence. If the purpose of developing new rating scales in second language (L2) assessment is to reduce subjectivity, the current ACTFL Guidelines fail in this pursuit; they simply give the rater a wider range of criteria from which to make subjective judgments. Why not equip raters with an objective means of measurement as well? I recognize that rating scales are designed only to reduce subjectivity, and not to eliminate it altogether. To truly reduce subjectivity, raters should be equipped with an objective means of measurement as well. For example, why not say Speaking-Intermediate-Mid learners can “talk simply about self and family members”, etc., and typically average X clauses per sentence, with Y errors per sentence?

Certainly, the ACTFL guidelines serve a broad audience, parts of which may not have the desire or wherewithal to employ such empirical guidelines. Therefore, the empirical guidelines must be presented to raters simply as an additional means of assessment. Some raters might choose not to utilize them; however, those who do would be making more informed decisions.

ACTFL Proficiency Guidelines used for placement purposes should serve not only testers, but teachers and students as well. Testers and teachers automatically make subjective judgments about learner proficiency, with or without guidelines, and students have very subjective ideas of their own proficiency, as well as that of their classmates. Any proper L2 assessment will be based on empirical evidence in some way, so why wait until the last minute when the learner is assigned to a level between one and ten?

REFERENCES

- American Council on the Teaching of Foreign Languages. (1985). *ACTFL proficiency guidelines*. Hastings-on-Hudson, NY: ACTFL Materials Center.
- American Council on the Teaching of Foreign Languages. (1999). *Speaking guidelines revised*. Retrieved November 1, 2007, from <http://www.actfl.org/files/public/Guidelinespeak.pdf>
- Breiner-Sanders, K. E., Swender, E., & Terry, R. M. (2001). *Preliminary proficiency guidelines – writing: Revised 2001*. Hastings-on-Hudson, NY: ACTFL Materials Center.

David Wiese is pursuing an MA in TESOL at Teachers College, Columbia University, and is an instructor at New York University's American Language Institute, and at the New School. His previous experience includes instructor positions at Brooklyn College and overseas in Barcelona, Spain.