

Assessing L2 Academic Speaking Ability: The Need for a Scenario-based Assessment Approach

Yuna Seong

Teachers College, Columbia University

In second language (L2) testing literature, from the skills-and-elements approach to the more recent models of communicative language ability, the conceptualization of L2 proficiency has evolved and broadened over the past few decades (Purpura, 2016). Consequently, the notion of L2 speaking ability has also gone through change, which has influenced L2 testers to constantly reevaluate *what* needs to be assessed and *how* L2 speaking assessment can adopt different designs and techniques accordingly.

The earliest views on speaking ability date back to Lado (1961) and Carroll (1961), who took a skills-and-elements approach and defined language ability in terms of a set of separate language elements (e.g., pronunciation, grammatical structure, lexicon), which are integrated in the skills of reading, writing, listening, and speaking. According to their views, speaking ability could be assessed by test items or tasks that target and measure different language elements discretely to make inferences on one's speaking ability. On the other hand, Clark (1975) and Jones (1985) put emphasis on communicative effectiveness and the role of performance. Clark (1975) defined speaking ability as one's "ability to communicate accurately and effectively in real-life language-use contexts" (p. 23), and this approach encouraged the use of performance tasks that replicate real-life situations.

However, the most dominant approach to viewing L2 speaking ability and its assessment has been influenced by the models of communicative competence (Canale, 1983; Canale & Swain, 1980) and communicative language ability (Bachman, 1990; Bachman & Palmer, 1996), which brought forth a multicomponential approach to understanding speaking ability in terms of various underlying and interrelated knowledge and competencies. For instance, adopting Bachman and Palmer's (1996) model of communicative language ability, Fulcher (2003) defined speaking ability as having five components: (1) *language competence* described as phonology, accuracy of syntax, vocabulary and cohesion, and fluency; (2) *textual knowledge* or the understanding of discourse structures such as turn-taking, adjacency pairs, and openings and closings; (3) *pragmatic knowledge* of appropriacy, implicature (doing things with words), and expressing being (defining status and role through speech); (4) *sociolinguistic knowledge* that is situational, topical, and cultural; and finally (5) *strategic capacity* that entails the speakers' use of achievement and avoidance strategies in order to overcome or avoid communication problems.

Due to the difficulty in defining such a complex construct as language ability, theoretical models and frameworks which empirical studies adhered to naturally varied in terms of the types of components they included. Yet, using analytic rubrics for scoring speaking test performance data elicited by tasks has been the most common practice in both large scale and classroom-based speaking assessment. Rubrics consist of scales that represent the different components of the speaking ability construct that is deemed suitable for the assessment context and needs.

While maintaining the importance of multiple underlying traits and knowledges, Chapelle (1998) expanded the notion of L2 proficiency by emphasizing the role and influence of *context* in language performance. In other words, she claimed that language performance is context-dependent; thus, in order to achieve the goal of successfully completing a given task, the L2 user is compelled to employ relevant cognitive or metacognitive strategies in order to activate and use

necessary knowledge or competencies. Therefore, her view is that underlying traits, context (i.e., test task), strategies, and their interaction should all be taken into account when designing L2 assessment and making inferences from test performance.

Purpura (2014, 2016) also proposes the importance of examining one's processing abilities as part of L2 proficiency in terms of "both the mind's cognitive architecture (e.g., attention, perception, memory) and its functions (e.g., processing, strategies)" (Purpura, 2016, p. 197). In addition, he points out that current L2 proficiency models and assessments pay little attention to conceptualizing and assessing *meaning* or content. He states L2 testers must also clearly specify and address the scope and type of content to be measured. This would especially be important in academic L2 assessment where conveyance of topical or disciplinary knowledge is considered an integral part of one's communicative ability. Thus, test users would be interested in the extent to which an L2 user's response is "content-responsible" (Purpura, 2017).

The need for integration of language competencies, content, and cognitive processes in L2 assessment becomes more imperative in the academic domain. As the number of non-native speakers of English continues to grow at English-medium colleges and universities across the globe, the importance of preparing these students to be successful in their academic studies increases. As a result, the number of English for Academic Purposes (EAP) courses is on the rise, and a large number of the matriculated students are required to take additional English courses to be better prepared for their degree studies. Accordingly, there is a call for assessments that can effectively measure the competencies and processes necessary in real-life academic demands.

Academic speaking is associated with complex and challenging tasks such as participating in academic discussions or giving oral presentations. In addition to linguistic competencies, students must comprehend and communicate disciplinary-specific content, making one's topical or disciplinary knowledge unequivocally inherent to L2 speaking ability in the academic domain; therefore, it is necessary to explicitly address and assess *meaning*, the content of what is being said (Douglas, 2000; Purpura, 2017). The content can be derived from the speaker's existing background knowledge, or it could be based on information that the speaker has collected and synthesized from other sources. Moreover, information is oftentimes obtained from written or oral academic sources and then integrated into spoken performance. Therefore, academic speaking ability is multi-modal and must be assessed in connection with the integration of other L2 skills such as listening or reading (Leki & Carlson, 1997) and not solely based on one's speaking performance. Finally, academic L2 ability involves not only the ability to use linguistic resources to understand and convey a range of disciplinary meanings via multiple modalities, but also the ability to perform academic tasks that require complex thinking skills (Chamot & O'Malley, 2004; Zwiers, 2008). As illustrated above, students must evaluate, select, and reorganize information, and reason through problems by means of thinking processes such as applying, analyzing, and synthesizing (Bloom, 1956). They must also employ metacognitive and cognitive strategies to complete the task in a way that is goal-oriented.

The complex nature of academic speaking ability is, in fact, accounted for in a model of academic speaking ability that Douglas (1997, 2000) proposed. His framework includes background knowledge explicitly as one part of the knowledge component along with language knowledge. Additionally, the model included a strategic component (i.e., metacognitive, language, and cognitive strategies) that plays a central role in activating and directing the use of the knowledge components. Douglas' model does not only address the importance of expanding the notion of academic speaking ability to include content and strategy use, but it also highlights

the need for test developers to create L2 tasks with contextual features that mirror the target language use domain.

Scenario-based assessment (SBA) is a current and innovative assessment practice that has recently been adopted and examined in mainstream education (e.g., Bennett, 2010) and in L1 literacy contexts (e.g., Sabatini & O'Reilly, 2013). SBA allows the assessment of an expanded L2 proficiency construct that accounts for linguistic as well as non-linguistic factors, such as background/topical knowledge or strategy use. Instead of testing L2 ability through a collection of unrelated tasks that get at different dimensions of L2 proficiency, SBA allows test takers to demonstrate their language proficiency competencies and processing abilities in a meaningful and goal-oriented context that simulates real world language use. The test takers perform thematically-related tasks sequenced in a way that ultimately leads to an overarching goal (e.g., completing a project). Through the process of working through the test, the test takers are able to use their knowledge on a certain topic, collect information across different texts or modalities (e.g., listening and speaking), process and utilize the information using metacognitive and cognitive strategies, and finally integrate it into a spoken or written response in order to accomplish the goal.

To illustrate an example of SBA used for assessing L2 academic speaking ability, an online scenario-based academic English speaking test developed for an ESL academic speaking course at an American university will be introduced here. The purpose of the test was to provide formative or summative information to the students regarding achievement in their academic speaking course. Therefore, the speaking tasks were designed to align with commonly adopted learning outcomes in academic English speaking courses such as being able to give oral summaries, verbalize opinion with supporting information, and discuss ideas. The test was developed around a scenario, where the test taker participates in a synchronous online class called *Intro to Journalism* that focuses on the topic of "the changes and future of journalism." In this test, the test taker is instructed to complete a final project, which is posting an oral response to an online discussion forum. The discussion forum prompt reads:

Discuss some of the major changes in journalism and pick one controversial issue associated with the new trends that you find interesting. Present your position or opinion on the issue. Finally, talk about what positive or negative directions you think journalism would take in the future. You must use the information from the listening materials to explain the issues and support your opinion.

In order to achieve the goal of successfully posting a response that incorporates topical content necessary to answer this very complex question about journalism, the test taker must refer to multiple sources of information by watching or listening to audio-visual materials (e.g., lectures, interviews) to obtain necessary information about the topic and eventually formulate a coherent spoken response to the prompt by summarizing, synthesizing, and reorganizing the information retrieved. In addition, while working toward the completion of the discussion forum project, the test taker must also complete a series of tasks designed to specifically elicit students' use of cognitive strategies (e.g., predicting the content, recalling key points, examining relationships between the texts). The test taker's spoken response is then scored in terms of grammatical, phonological, and organizational control, and most importantly, content control, which looks at the extent to which the test taker accurately includes relevant and key topical information in the response.

With the call for L2 assessment practices that keep pace with the current ways of viewing L2 proficiency, this example was presented to demonstrate the usefulness of SBA and the opportunities it affords in assessing L2 speaking ability. In this single assessment, the student is able to demonstrate his or her ability to comprehend and speak about a discipline-specific topic using the appropriate strategies necessary in executing an academic speaking task that may very likely exist in the real world. Such a comprehensive and multi-modal assessment allows L2 testers and educators to tackle a broadened construct of L2 academic speaking ability that is construed of different facets such as topical/background knowledge and strategy use in addition to language knowledge.

This review attempted to briefly discuss the evolution of L2 proficiency and speaking ability, and to suggest SBA as a promising approach to designing language assessments that can provide a more comprehensive interpretation of one's L2 speaking ability. An example of SBA use in the assessment of L2 academic speaking ability was also introduced to illustratively demonstrate the broad range of knowledge, skills and abilities SBA can address. Although SBA is a fairly new form of assessment that needs to be further examined in terms of its usefulness and validity, it is hoped that this review is able to bring attention to SBA's efficacy and potential in designing L2 assessment that aligns with the modern and broadened views of L2 proficiency.

REFERENCES

- Bachman, L. F. (1990). Communicative language ability. In *Fundamental considerations in language testing* (pp. 81-109). Oxford: Oxford University Press.
- Bachman, L. F. & Palmer, A. S. (1996). Describing language ability: Language used in tests. In *Language testing in practice* (pp. 61-79). Oxford: Oxford University Press.
- Bennett, R. (2010). Cognitively based assessment of, for, and as learning (CBAL): A preliminary theory of action for summative and formative assessment. *Measurement, Interdisciplinary Research and Perspectives*, 8, 70–92.
- Bloom, B. S. (1956). *Taxonomy of educational objectives, Handbook I: The cognitive domain*. New York: David McKay Co Inc.
- Canale, M. (1983). On some dimensions of language proficiency. In J. Oller (Ed.), *Issues in Language Testing Research* (pp. 333-342). Rowley, MA: Newbury House.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47.
- Carroll, J. B. (1961). Fundamental considerations in testing for English language proficiency of foreign students. In *Testing the English proficiency of foreign students* (pp. 30-40). Washington, DC: Center for Applied Linguistics.
- Chamot, A.U., & O'Malley, J.M. (1994). *The Calla handbook: Implementing the cognitive academic language learning approach*. White Plains, NY: Addison Wesley Longman.
- Chapelle, C. (1998). Construct definition and validity inquiry in SLA research in L. Bachman and A. Cohen (Eds.), *Interfaces between second language acquisition and language testing research*. (pp. 32-70) Cambridge University Press, Cambridge.
- Clark, J. L. D. (1975). Theoretical and technical considerations in oral proficiency testing. In R. L. Jones & B. Spolsky (Eds.), *Testing language proficiency* (pp. 10-28). Arlington, VA: Center for Applied Linguistics.
- Douglas, D. (1997). *Testing speaking ability in academic contexts: Theoretical considerations*. TOEFL Monograph Series, Number 8. Princeton: Educational Testing Service.

- Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge, UK: Cambridge University Press.
- Fulcher, G. (2003). *Testing second language speaking*. London: Longman.
- Jones, R. L. (1985). Second language performance testing: An overview. In P. C. Hauptman, R. LeBlanc & M. B. Wesche (Eds.), *Second language performance testing* (pp. 15-24). Ottawa: University of Ottawa Press.
- Lado, R. (1961). *Language testing: The construction and use of foreign language tests*. London: Longman.
- Leki, I., & Carson, J. (1997). "Completely different worlds:" EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly*, 31(1), 39–69.
- Purpura, J. (2014). Cognition and language assessment. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1452-1476). Oxford, UK: Wiley-Blackwell.
- Purpura, J. (2016). Second and foreign language assessment. *The Modern Language Journal*, 100, 190-208.
- Purpura, J. (2017). Assessing meaning. In E. Shohamy, L. G. Or, & S. May (Eds.), *Language testing and assessment, encyclopedia of language and education* (3rd ed.) (pp. 33-61). Springer International Publishing AG.
- Sabatini, J., & O'Reilly, T. (2013). Rationale for a new generation of reading comprehension assessments. In B. Miller, L. Cutting, & P. McCardle (Eds.), *Unraveling reading comprehension: Behavioral, neurobiological, and genetic components* (pp. 100– 111). Baltimore, MD: Brookes Publishing.
- Zwiers, J. (2008). *Building academic language: Essential practices for content classrooms*. San Francisco: Jossey-Bass.

Yuna Seong is a doctoral student in Applied Linguistics at Teachers College, Columbia University. Her research interests include language assessment, teaching and measuring second language speaking ability, scenario-based assessment, and metacognitive and cognitive processes in second language test performance. She is currently teaching ESL at the American Language Institute, New York University as a full-time language lecturer. Correspondence should be sent to Yuna Seong at: ys2458@tc.columbia.edu