Title of Project:
Oral Social Language in the English Language Proficiency Assessment of Young ELLs

Researcher:
Kimberly Woo
New York University (Doctoral Institution)
Teachers College, Columbia University (as of December 2015)
woo131@gmail.com

Research Supervisor:
Dr. Lorena Llosa
New York University
lorena.llosa@nyu.edu

Project Summary:

Background
With the growing population of English language learners (ELLs) in American public schools, standardized English language proficiency (ELP) tests have become a common response to federal mandates for accountability. Such exams hold additionally high stakes in that results are frequently used to inform decisions, which can have an appreciable impact on the instruction ELL students receive and the trajectory of their academic careers (Abedi, 2007; Schappe, 2005). Recognizing that school-age language learners in the earliest grades require special considerations for their developmental features and that they are particularly vulnerable to both how tests are administered and used, there have been calls for research considering the validity of standardized language tests and how their tasks operate with young learners (Bailey & Butler, 2004; Esquinca, Yaden, & Rueda, 2005; Genesee, Paradis, & Crago, 2004; McKay, 2006).

Among the issues complicating ELP assessment, especially for ELLs in the earliest grades, is the notion of a dichotomy between “social” and “academic” language. Cummins’ (1980) proposal that language proficiency falls along two dimensions, one academically-oriented and one socially-oriented, has had an enormous influence on how the sequence and nature of second and foreign language development are understood (Gu, 2014). As constructs, social and academic language are conventionally discussed as a binary, with each largely characterized in relation to the other. Over the last decade, and especially for ELLs, priority has been placed on defining and assessing academic language proficiency (e.g., Bunch, 2009; Lucero, 2012; Shiel, Cregan, McGough, & Archer, 2012), and, to this end, ELP testing in the United States purportedly targets students’ academic language (Abedi, 2007; Chaloub-Deville & Deville, 2008). However, in practice, it is not uncommon for the speaking section(s) of ELP tests to include items that aim to capture students’ ability to use English for social interaction.
Given that ELP tests are positioned as academic in nature, the presence of such items is unexpected and creates ambiguity in how the two constructs are delineated. Additionally, in the early-grade ELL setting, further overlap between the “academic” and “social” domains exists in that social interaction and the language it requires are frequently part of curricula and state and professional standards for language instruction. In this context, and given the relatively little work to date focusing on social language, it is unclear what social language items on ELP tests are measuring. Further, it must be considered whether the construct is defined consistently among ELP tests and with how it is taught and assessed in students’ classrooms.

This dissertation study sought to examine the construct of social language in the assessment of young ELLs’ oral English proficiency. It was guided by three research questions:

1. How is the construct “oral social language” defined and operationalized by tasks used on large-scale standardized tests of English language proficiency?
2. How do teachers of young ELLs define and assess oral social language in the classroom?
3. How do ratings of young ELL’s ability compare when they are assessed, respectively, according to tests’ and teachers’ criteria?

Methods

This study employed a mixed-methods approach, conducted in two phases over the 2012-13 academic year. To determine how tests defined and operationalized social language (i.e., Research Question 1), a content review was conducted, which included examination of publicly-available test documents for the Kindergarten through first grade (K-1) bands of three ELP tests: the California English Language Development Test (CELDT), the New York State English as a Second Language Achievement Test (NYSESLAT) and the ACCESS for ELLs, used by 33 states and territories under the WIDA consortium. Materials such as test samplers and technical guides were first reviewed for explicit descriptions of social language and then qualitatively analyzed using Bachman and Palmer’s (2010) test task characteristics framework to systematically compare and categorize test tasks, with attention to the characteristics of the input (i.e., prompt format) and expected response (i.e., content, linguistic and pragmatic requirements, etc.).

The second research question, investigating teachers’ definitions and assessment of social language, was addressed through quantitative analysis of teachers’ responses to a survey examining their beliefs, perceptions, and instructional practices regarding academic and social language. Analysis of survey responses looked for commonalities in how teachers characterized social and academic language and their classroom assessment methods. Follow-up interviews were conducted with a subset of participating teachers to provide further insight into teachers’ definitions and assessment in practice. Interviews were transcribed and inductively analyzed and compared against survey data.

Finally, in response to the third research question, students’ use of oral social language was assessed by two measures, a “test task measure” (TTM) reflecting social language speaking tasks found on the three target ELP tests, and a summative report-card- like “classroom rating scale” (CRS) completed by participating students’ teachers. Ratings of students’ ability on both measures
were compared using descriptive and inferential statistical procedures to determine whether and to what extent a relationship between assessments existed.

**Population**
Participating teachers and students were drawn from public elementary schools located in and around New York City’s Chinatown, a neighborhood selected for its historical significance as a cultural and linguistic immigrant enclave community, as well as its high proportion of early-childhood ELLs. In the year leading up to the study (2011-12), just over 40% of all K-1 students enrolled in Chinatown public schools had been identified as ELLs. Five out of the eight Chinatown schools (62.5%) agreed to serve as sites for this study. From these schools, 30 Kindergarten, first grade, and early-grade ESL teachers participated in the survey that was part of the first phase of the study. Four surveyed teachers continued into the second phase and, from their classrooms, 53 K-1 ELLs participated in student assessments.

**Key Findings**

**Research Question 1**
Content analysis of documents for the three tests revealed that, although materials often spoke to social language in their linked standards, they rarely included explicit discussion of the construct. Tests’ definition(s) of social language could only be inferred through how they were reflected in the tasks’ characteristics. Across the three tests, eleven speaking tasks were identified as targeting social language proficiency. Qualitative analysis using Bachman and Palmer’s test task characteristics framework (2010) found that although tasks shared surface characteristics (e.g., use of directive statements and questions to elicit student speech), they varied considerably in both their input (e.g., how they positioned the relationship between the student and the adult administering the test, the degree and nature of contextual support) and the kind of language they expected students to produce (i.e., single-word answers to extended narratives). Surprisingly, expected language production was fairly limited across all tasks; despite commonly framing of social language in terms of conversation or interaction, test tasks rarely engaged students for more than one turn of speech. Of the tests reviewed, the NYSESLAT and CELDT were found to have relatively narrow, specific operational definitions (e.g., initiating a conversation or responding appropriately to politeness phrases). In contrast, the variety of tasks on the ACCESS for ELLs pointed to a more global definition in which social and academic language may not necessarily be mutually exclusive.

**Research Question 2**
Compared to the tests, teachers’ approach to defining social language was found to be broader, and their assessments, when occurring, more greatly emphasized students’ ability to engage with others and to be understood. When asked to describe social language, teachers tended not to define the construct in terms of specific linguistic features or objectives, but instead spoke to more general aspects of language use, with particular emphasis on the participants (e.g., “with friends”) and setting (e.g., “outside of school,” “the playground”). Frequently, teachers defined social language in terms of broad descriptors (e.g., “conversational,” “interactive,” “natural,” “everyday”). Fitting with traditional discourse surrounding social language, teachers’ definitions
were often couched in a dichotomous contrast to academic language (e.g., “social language is non-academic”) or circular descriptions (e.g., “social language is language used in social situations”). However, fitting with recent literature that questions the academic/social dichotomy (e.g., Aukerman, 2007; Bailey & Huang, 2011; Lucero, 2012), teachers also acknowledged areas of “overlap” in which social language could be used in the school setting. No specific content was deemed exclusively indicative of social language, and many common topics, especially those that occur in classroom settings, were viewed by teachers to be both social and academic in nature.

Exploration into teachers’ classroom assessment practices found that teachers more frequently asserted that they regularly assessed students’ overall English language development (82.1%) or academic language (79.3%), than social interactions (50%). Although at least half of teachers reported that they assessed students’ oral and social language on a regular basis, teacher interviews indicated that assessment of any form of oral language, much less that focusing on social interaction, was infrequent, incidental, and typically informal in nature (e.g., observations). Interviewed teachers stressed that although they viewed oral and social language to be important for early-grade ELLs, growing academic demands placed constraints on their time, which limited the extent to which these could be addressed. It was also found that some teachers counted evaluation of students’ writing as part of their assessments of students’ oral language ability, which could be interpreted as additional prioritization of academic “literate” language (Schleppegrell, 2004).

**Research Question 3**

Student assessment data from the Test Task Measure (TTM) and Classroom Rating Scale (CRS) revealed patterns in how students’ social language proficiency was respectively rated by test and teacher criteria. It was found that test tasks assumed a specific set of criteria for proficient performance applied across the entire K-1 band, where teachers employed a more developmental orientation in which expectations were adjusted to account for students’ grade level. Comparison of assessment data found a correspondence in test and teacher ratings, which offered some evidence of criterion validity for using these tasks, particularly at the narrow-but-crucial point of determining whether or not students met expectations. It had been initially hypothesized that if tests and test tasks were valid indicators of students’ social language proficiency, then their ratings should mirror classroom evaluation of students’ proficiency as assigned by their teachers. This hypothesized pattern was broadly found to hold true; on average, higher mean ratings on test tasks coincided with higher teacher ratings.

**Implications and Discussion**

Beyond finding evidence speaking to the validity of social language tasks on ELP tests, (i.e., commonalities between what tests and teachers look for in students’ social language use, tests’ apparent ability to adequately distinguish social language ability by grade level, and concordance with how teachers rate students’ ability at the crucial point of decision-making), this study also constituted broader examination of the construct of social language in the assessment of young language learners. Two recurring questions emerged as an underlying theme: assuming an academic/social dichotomy, can social language actually be assessed? And, if not, what are
teachers and tests assessing when they claim to do so? Referring to Bailey and Heritage’s (2008) schema for academic and social language, tests and teachers both appeared to be not assessing social language so much as a different construct, a category of language Bailey and Heritage (2008) refer to as ‘School Navigational Language’ (SNL), ‘the language needed to communicate with teachers and peers in the school setting in a very broad sense’ (p.15). School Navigational Language is offered as a middle ground, capturing features of both social and academic language, while broadly categorized as a variant of ‘academic’ language due to its use in supporting in-school learning.

Areas in which tests’ and teachers’ definitions and assessment were found to have less alignment may, in part, speak to a crucial difference in orientation towards SNL. In the testing field, SNL is occasionally explicitly acknowledged as a construct and is implicitly recognized as part of academic language through the very inclusion of “social language” tasks on academic ELP tests. Among early grade teachers, while they acknowledged areas of overlap in the social/academic binary, it was unclear the extent to which they recognized SNL as a category of language. Further, in efforts to work within the dichotomous framework, it seemed teachers subsumed SNL into the social domain and downplayed it in light of what they perceived as more pressing academic demands.

These findings point to an important inconsistency between assessment and instruction, which has implications for policy and practice. Social language, or SNL as the case may be, is a priority by virtue of its presence on high stakes ELP tests. Although purportedly valued by teachers, SNL is not emphasized in classrooms due to teachers’ perceptions that it is not academically relevant or prioritized by the larger education system. Beyond consequences for the validity of test tasks, de-emphasizing SNL in instruction and assessment means that students are not receiving necessary language support. Extensive literature points to the importance of social language in students’ linguistic and general academic development. For young students, the classroom represents a major social setting, and for ELLs, it may be one of the few in which they may be required to use English. Contrary to common assumptions that students naturally acquire social language, it has been argued that young language learners require instruction to aid in its development (Bailey & Heritage, 2008; Gu, 2014; Gresham, Elliott, Vance, & Cook, 2011). This work suggests that, to help bridge the disconnection, there is need for both tests and teachers to acknowledge and address areas of overlap in the academic/social binary, such as SNL.
References


