



**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](mailto:woo131@gmail.com)



Kimberly Woo

**Research Supervisor:**

Dr. Lorena Llosa  
New York University  
[lorena.llosa@nyu.edu](mailto: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

- Aarts, R., Demir, S., & Vallen, T. (2011). Characteristics of academic language register occurring in caretaker-child interaction: Development and validation of a coding scheme. *Language Learning, 61*(4), 1173-1221.
- Abedi, J. (2007). English language proficiency assessment and accountability. In Abedi, J. (Ed.), *English language proficiency assessment in the nation: Current status and future practice* (pp. 3-10). Berkley, CA: The Regents of the University of California.
- Aud, S., Hussar, W., Planty, M., Snyder, T., Bianco, K., Fox, M. A., Frohlich, L., Kemp, J., & Drake, L. (2010). *The condition of education 2010 (NCES 2010-028)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Anderson, S., & Messick, S. (1974). Social competency in young children. *Developmental psychology, 10*(2), 282-293.
- Arjoudis, S., & O'Loughlin, S. (2004). Tensions between validity and outcomes: Teacher assessment of written work of recently arrived immigrant ESL students. *Language Testing, 21*(3), 284-304.
- Arnold, D. H., Kupersmidt, J. B., Voegler-Lee, M. E., & Marshall, N. A. (2012). The association between preschool children's social functioning and their emergent academic skills. *Early Childhood Research Quarterly, 27*, 376-386.
- Aukerman, M. (2007). A culpable CALP: Rethinking the conversational/academic language proficiency distinction in early literacy instruction. *The Reading Teacher, 60*(7), 626-636.
- Bachman, L. F., & Alderson, C. J. (2004). *Statistical analysis for language assessment*. Cambridge, UK: Cambridge University Press.
- Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford, UK: Oxford University Press.
- Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford, UK: Oxford University Press.
- Bailey, A. L. (Ed.). (2007). *The language demands of school: Putting academic English to the test*. New Haven, CT: Yale University Press.
- Bailey, A., & Butler, F. (2004). Ethical considerations in the assessment of the language and content knowledge of U.S. school-age English language learners. *Language Assessment Quarterly, 1*(2&3), 177-193.



- Bailey, A. L., & Heritage, M. (2008). *Formative assessment for literacy grades K-6: Building reading and academic language skills across the curriculum* (pp. 15-18). Thousand Oaks, CA: Corwin Press.
- Bailey, A. L., & Huang, B. H. (2011). Do current English language development/proficiency standards reflect the English needed for success in school? *Language Testing*, 28, 343-365.
- Baldwin, D., & Meyer, M. (2009). How inherently social is language? In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp.87-106). Malden, MA: Wiley-Blackwell.
- Bjorklund, D. F., & Pellegrini, A. D. (2002). *The origins of human nature: Evolutionary developmental psychology*. Washington, DC: American Psychological Association.
- Borsato, G. N. & Padilla, A. M. (2007). Educational assessment of English language learners. In L. A. Suzuki & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment* (3<sup>rd</sup> ed.) (pp. 471-489). San Francisco, CA: Jossey Bass Publishers.
- Brindley, G. (1998). Outcomes-based assessment and reporting in language learning programmes: A review of the issues. *Language Testing*, 15(1), 45–85.
- Bunch, G. C. (2009). Going up there: Challenges and opportunities for language minority students during a mainstream classroom speech event. *Linguistics and Education*, 20(2), 81-108.
- Butler, F. A. & Stevens, R. (2001). Standardized assessment of the content knowledge of English language learners K-12: Current trends and old dilemmas. *Language Testing*, 18(1), 409-427.
- Butler, F. A. & Stevens, R. (1997). Oral language assessment in the classroom. *Theory into Practice*, 36(4), 214-219.
- Cairns, H. S. (1996). *The acquisition of language* (2<sup>nd</sup> ed.). Austin, TX: Pro-Ed Inc.
- California Department of Education. (2010). *CELDT 2010–11 Edition Technical Report*. Sacramento, CA. Retrieved from <http://www.cde.ca.gov/ta/tg/el/techreport.asp>
- California Department of Education. (2011). *2011-2012 CELDT Information Guide*. Sacramento, CA. Retrieved from <http://www.cde.ca.gov/ta/tg/el/resources.asp>
- California Department of Education. (2012). *Released test questions*. Sacramento, CA. Retrieved from <http://www.cde.ca.gov/ta/tg/el/resources.asp>



- Canale, M. & McSwain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47.
- Carpenter, K., Fuji, N., & Kataoka, H. (1995). An oral interview procedure for assessing second language abilities in children. *Language Testing*, 12(2), 157-81.
- Chafouleas, S. M., Kilgus, S. P., & Hernandez, P. (2009). Using Direct Behavior Rating (DBR) to screen for school social risk: A preliminary comparison of methods in a kindergarten sample. *Assessment for Effective Intervention*, 34(4), 214-233.
- Chaloub-Deville, M. & Deville, C. (2008). Nationally mandated testing for accountability: English language learners in the US. In B. Spolsky & F. M. Hult (Eds.), *Handbook of Educational Linguistics* (pp. 510-522). Malden, MA. Blackwell.
- Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, 21(3), 360-389.
- Clark, B. (2000). Proceedings from the Lilian Katz Symposium: *First- and second-language acquisition in early childhood. Issues in Early Childhood Education: Curriculum, Teacher Education & Dissemination of Information*. Champaign, IL. (ERIC Document Reproduction Service No. ED470889).
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Coltrane, B. (2003). *Working with young language learners: Some considerations*. Washington DC: Center for Applied Linguistics. (ERIC Document Reproduction Service No. ED481690).
- Crain, S., & Fodor, J. D. (1993). Competence and performance in child language. In E. Dromi (Ed.), *Language and cognition: A developmental perspective* (Vol. 5) (pp. 141-171). Norwood, NJ: Ablex Publishing Corporation.
- Cumming, A., Grant, L., Mulcahy-Ernt, P., Powers, D. E., & MacDonald, S. L. (2004). A teacher-verification study of speaking and writing prototype tasks for a new TOEFL. *Language Testing*, 21(1), 107-145.
- Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, 14(2), 175-187.
- Cummins, J. (1991). Conversational and academic language proficiency in bilingual contexts. *AILA Review*, 8, 75-89.





- Dalton, S. (1979). Validation of the Language Assessment Scales. *Educational and Psychological Measurement*, 39, 1001-1003.
- Davison, C. (2004). The contradictory culture of teacher-based assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21(3), 305-334.
- Edelenbos, P. & Kubanek-German, A. (2004). Teacher assessment: The concept of 'diagnostic competence'. *Language Testing*, 21(3), 259-283.
- Educational Data Systems. (2013). *CELDT Fundamentals 1: CELDT Overview Transcript*. Retrieved from:  
[http://celdt.org/training/fundamentals/video.aspx?title=CELDT\\_Fundamentals1](http://celdt.org/training/fundamentals/video.aspx?title=CELDT_Fundamentals1)
- Eisenberg, N. (1984). Social competence: A developmental perspective. *School psychology review*, 13, 267-277.
- Elder, C., Iwashita, N. & McNamara, T. (2002). Estimating the difficulty of oral proficiency tasks: what does the test-taker have to offer? *Language Testing*, 19(4), 347-368.
- Espinosa, L. M. (2005). Curriculum and assessment considerations for young children from culturally, linguistically, and economically diverse backgrounds. *Psychology in the Schools*, 42(8), 837-853.
- Esquinca, A., Yaden, D., & Rueda, R. (2005). Current language proficiency tests and their implications for preschool English language learners. In J. Cohen, K. McAlister, K. Rolstad, & J. MacSwan (Eds.), *ISB4: Proceedings of the 4th International Symposium on Bilingualism* (pp. 674-680). Somerville, MA: Cascadilla Press.
- Fang, Z., Schleppegrell, M. J., & Cox, B. E. (2006). Understanding the language demands of schooling: Nouns in academic registers. *Journal of Literacy Research*, 38(3), 247-273.
- Feinberg, A. B. & Shapiro, E. S. (2003). Accuracy of teacher judgments in predicting oral reading fluency. *School Psychology Quarterly*, 18(1), 52-65.
- Fu, D. (1993). *An island of English: Teaching ESL in Chinatown*. New York, NY: Heinemann.
- Fu, D. (1998). Unlock their lonely hearts. *Voices from the middle*, 6(1), 3-10.
- Fulcher, G. (1996). Testing tasks: Issues in group design and the group oral. *Language Testing*, 13(1), 23-51.
- Fulcher, G. (1997). An English language placement test: issues in reliability and validity. *Language Testing*, 14(2), 113-139.



- Fulcher, G. & Reiter, R. M. (2003). Task difficulty in speaking tests. *Language Testing*, 20(3), 321-344.
- Gee, J. P. (2011). *Social linguistics and literacies*. New York, NY: Routledge.
- Genesee, F., Paradis, J., & Crago, M. B. (2004). *Dual language development and disorders: A handbook on bilingualism and second language learning*. Baltimore, MD: Paul H. Brookes Publishing.
- Geva, E. (2000). Issues in the assessment of reading disabilities in L2 children – Beliefs and research evidence. *Dyslexia*, 6, 13-28.
- Gottlieb, M. Carnuccio, L., Ernst-Slavit, G., & Katz, A. (2006). *PreK-12 English language proficiency standards*. Alexandria, VA: Teachers of English to Speakers of Other Languages.
- Gresham, F. M, Elliott, S. N., Vance, M. J., & Cook, C. R. (2011). Comparability of the Social Skills Rating System to the Social Skills Improvement System: Content and psychometric comparisons across elementary and secondary age levels. *School Psychology Quarterly*, 26(1), 27-44.
- Halle, T. G., Whittaker, J. V., Zepeda, M., Rothenberg, L., Anderson, R., Daneri, P., . . . Buyesse, V. M. (2014). The social-emotional development of dual-language learners: Looking back at existing research. *Early Childhood Research Quarterly*, 29(4), 734-749. <http://dx.dpo.org/10.1016/j.ecresq.2013.12.002>.
- Halle, T., Hair, E., Wandner, L., McNamara, M., & Chien, N. (2012). Predictors and outcomes of early versus later English language proficiency among English language learners. *Early Childhood Research Quarterly*, 27, 1-20.
- Hardesty, D.M. & Bearden, W.O. (2004). The use of expert judges in scale development: Implications for improving face validity of unobservable constructs. *Journal of Business Research*, 57, 98-107.
- Hasselgreen, A. (2005). Assessing the language of young learners. *Language Testing*, 22(3), 337-354.
- Hasselgren, A. (2000). The assessment of English ability of young learners in Norwegian schools: an innovative approach. *Language Testing*, 17(2), 261-77.
- Hawkins, M.R. (2005). ESL in elementary education. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 25-43). New York, NY: Routledge.



- Hoge, R. D. & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research*, 59(3), 297-313.
- Honig, A.S. (2007). Oral language development. *Early Child Development and Care*, 177(6&7), 581-613.
- Hsieh, H. & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288.
- Hymes, D. (1972). On communicative competence. In A. Duranti (Ed.), *Linguistic anthropology: A reader* (pp.53-73). Malden, MA: Blackwell Publishers Inc.
- Jerome, E. M., Hamre, B. K. & Pianta, R. C. (2009). Teacher-child relationships from kindergarten to sixth grade: Early childhood predictors of teacher-perceived conflict and closeness. *Social Development*, 18(4), 915-945.
- Jeynes, W. H. (2006). Standardized tests and Froebel's original kindergarten model. *Teacher's College Record*, 108(10), 1937-1959.
- Johnstone, R. (2000). Context-sensitive assessment of modern languages in primary (elementary) and secondary education: Scotland and the European experience. *Language Testing*, 17(2), 123-143.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement, Fourth edition* (pp. 17-67). Westport, CT: Praeger Publishers.
- Kane, M.T. (2008). Terminology, emphasis, and utility in validation. *Educational Researcher*, 37(2), 76-82.
- Kim, Y. (2008). The effects of integrated language-based instruction in elementary ESL learning. *The Modern Language Journal*, 92(3), 431-451.
- Kondo-Brown, K. (2004). Investigating interviewer-candidate interactions during oral interviews for child L2 learners. *Foreign Language Annals*, 37(4), 602-615.
- Kwong, P. (1979). *Chinatown, N.Y.: Labor and politics, 1930-1950*. New York, NY: Monthly Review Press.
- Kwong, P. (1987). *The new Chinatown*. New York, NY: The Noonday Press.
- Kwong, P., & Mišćević, D. (2005). *Chinese America: The untold story of America's oldest new community*. New York, NY: The New Press.



- Ladd, G.W., Herald, S. L., & Kochel, K. P. (2006) School readiness: Are there social prerequisites? *Early Education & Development*, 17(1),115-150.
- Ledbetter, P. J. & Dent, C. H. (1988). Young children's sensitivity to direct and indirect request structure. *First Language*, 8, 227-246.
- Leung, C. (2005). Classroom teacher assessment of second language development: Construct as practice. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 25-43). New York, NY: Routledge.
- Limbos, M. M. & Geva, G. (2001). Accuracy of teacher assessments of second-language students at risk for reading disability. *Journal of Learning Disabilities*, 34, 136-151.
- Llosa, L. (2005). Assessing English learners' language proficiency: A qualitative investigation of teachers' interpretations of the California ELD standards. *The CATESOL Journal*, 17(1), 7-18.
- Llosa, L. (2007). Validating a standards-based classroom assessment of English proficiency: A multitrait-multimethod approach. *Language Testing*, 24(1), 489-515.
- Llosa, L. (2008). Building and supporting a validity argument for standards-based classroom assessment of English proficiency based on teacher judgments. *Educational Measurement: Issues and Practice*, 27(3), 32-42.
- Llosa, L., Beck, S. W., Zhao, C. G. (2009, April). *Developing a diagnostic assessment of academic writing for high school students: The construct definition phase*. Paper presented at the AERA annual meeting, San Diego, CA.
- Louie, V. (2004). *Compelled to excel: Immigration, education, and opportunity among Chinese Americans*. Stanford, CA: Stanford University Press.
- Lucero, A. (2012). Demands and opportunities: Analyzing academic language in a first grade dual language program. *Linguistics and Education*, 23, 277-288.
- Lumley, T., & O'Sullivan, B. (2005). The effect of test-taker gender, Audience, and topic on task performance in tape-mediated assessment of speaking. *Language Assessment*, 22, 415-437.
- Mashburn, A. J., & Pianta, R. C. (2006) Social relationships and school readiness. *Early Education & Development*, 17(1), 151-176,
- MacSwan, J., Rolstad, K., & Glass, G.V. (2002). Do some school-age children have no language? Some problems of construct validity in the Pre-LAS Español. *Bilingual Research Journal*, 26(2), 395-420.



- Maxwell, L. A. (2013). Common core ratchets up language demands for language-learners. *Education Week*, 33(10), 14-16.
- McClelland, M. M., & Morrison, F. J. (2003). The emergence of learning-related social skills in preschool children. *Early Childhood Research Quarterly*, 18, 206-224.
- McClelland, M. M., Morrison, F. J., & Holmes, D. L. (2000). Children at risk for early academic problems: The role of learning related social skills. *Early Childhood Research Quarterly*, 15(3), 307-329.
- McGinnis, M. (1983). Social language: Towards fluency and flexibility. *The Volta Review*, 85(5), 101-115.
- McKay, P. (2005). Research into the assessment of school-age language learners. *Annual Review of Applied Linguistics*, 25, 243-263.
- McKay, P. (2006). *Assessing young language learners*. New York, NY: Cambridge University Press.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3<sup>rd</sup> ed.) (pp. 13-103). New York, NY: Macmillan.
- Nazari, A. (2007). EFL teachers' perception of the concept of communicative competence. *EFL Journal*, 61(3), 202-210.
- New York City Department of Education. (n.d) *Identification, eligibility and programs for English language learners*. Retrieved August 11, 2011 from <http://schools.nyc.gov/Academics/ELL/FamilyResources/ELL+Programs.htm>.
- New York State Education Department. (2004). *Learning standards for English as a Second Language: Early Childhood Grades PreK-1*. Retrieved from <http://www.p12.nysed.gov/biling/resource/ESL/standards.html>
- New York State Education Department. (2007). *New York State Testing Program NYSESAT Sampler Grades K-1*. Retrieved From: <http://www.p12.nysed.gov/assessment/nyseslat/samplers/2007/>
- New York State Education Department. (2009). *Learning standards for English as a second language: Early childhood grades PK-1*. Retrieved November 15, 2010, from <http://www.p12.nysed.gov/biling/resource/ESL/standards.html>.
- New York State Education Department. (2011). *Common core standards FAQ*. Retrieved May 26, 2011, from [http://www.p12.nysed.gov/ciai/common\\_core\\_standards/faq.html](http://www.p12.nysed.gov/ciai/common_core_standards/faq.html)



- New York State Education Department. (2011). *New York State P-12 Common core learning standards for English language arts & literacy*. Retrieved May 25, 2011, from [http://www.p12.nysed.gov/ciai/common\\_core\\_standards/pdfdocs/p12\\_common\\_core\\_learning\\_standards\\_ela\\_final.pdf](http://www.p12.nysed.gov/ciai/common_core_standards/pdfdocs/p12_common_core_learning_standards_ela_final.pdf)
- New York State Education Department. (2012). *2012 New York State English as a Second Language Achievement Test (NYSESLAT) school administrator's manual*. Retrieved from <http://www.p12.nysed.gov/assessment/nyseslat/archive.html>
- New York State Education Department. (2013). *New York State Testing Program NYSESLAT: Guide to the 2013 NYSESLAT*. Retrieved from: <http://www.p12.nysed.gov/assessment/nyseslat/>
- O'Connor, E. & McCartney, K. (2007). Examining teacher-child relationships and achievement as part of an ecological model of development. *American Educational Research Journal*, 44(2), 340-369.
- O'Loughlin, K. (2002). The impact of gender in oral proficiency testing. *Language Testing*, 19, 169-192.
- O'Sullivan, B. (2000). Exploring gender and oral proficiency interview performance. *System*, 28, 373-386.
- O'Sullivan, B., Weir, C.J., & Saville, N. (2002). Using observation checklists to validate speaking tests. *Language Testing*, 19(1) 33-56.
- Papageorgiou, S. (2010). Investigating the decision-making process of standard setting participants. *Language Testing*, 27(2), 261-282.
- Paradis, J. (2009). Second language acquisition in childhood. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 387-405). Malden, MA: Wiley-Blackwell.
- Pearson. (2009). *New York State testing program English as a Second Language Achievement Test (NYSESLAT) 2008 administration technical manual*. Retrieved from: <http://www.p12.nysed.gov/assessment/reports/archive3.html>
- Pellegrini, A. (1984). A speech analysis of preschoolers' dyadic interaction. *Communication & Cognition*, 17(4), 425-436.
- Pellegrini, A. D. (1998). Play and the assessment of young children. In O. N. Saracho & B. Spodek (Eds.), *Multiple perspectives on play in early childhood education* (pp. 220-239). Albany, NY: State University of New York Press.



- Pinter, A. (2006). *Teaching young language learners*. Oxford, UK: Oxford University Press.
- Porath, M. (2003). Social understanding in the first years of school. *Early Childhood Research Quarterly, 18*, 468-484.
- Porter, S. G. & Vega, J. (2007). Overview of existing English language proficiency tests. In Abedi, J. (Ed.). *English language proficiency assessment in the nation: Current status and future practice* (pp. 93-104). Berkley, CA: The Regents of the University of California.
- Pray, L. (2005). How well do commonly used language instruments measure English oral-language proficiency? *Bilingual Research Journal, 29*, 387-409.
- Raver, C. C. & Zigler, E. F. (1997). Social competence: An untapped dimension in evaluating Head Start's success. *Early Childhood Research Quarterly, 12*, 363-385.
- Rea-Dickins, P. (2000). Assessment in early years language learning contexts. *Language Testing, 17*(2) 115-122.
- Rea-Dickins, P. (2000). Current research and professional practice: Reports of work in progress in the assessment of young language learners. *Language Testing, 17*(2), 245-249.
- Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing, 18*(4), 429-462.
- Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing, 21*(3), 249-258.
- Rea-Dickins, P. & Gardner, S. (2000). Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing, 17*(2), 215-243.
- Rice, M.L. & Kemper, S. (1984). *Child language and cognition: Contemporary issues*. Baltimore, MD: University Park Press.
- Roach, A. T., McGrath, D., Wixson, C., & Talapatra, D. (2010). Aligning an early childhood assessment to state kindergarten content standards: Application of a nationally recognized alignment framework. *Educational Measurement: Issues and Practice, 29*(1), 25-37.
- Russell, R. L. & Grizzle, K. L. (2008). Assessing child and adolescent pragmatic competencies: Towards evidence-based assessment. *Clinical Child and Family Psychology Review, 11*, 59-73.



- Sanchez, L. (2006). Bilingualism/second-language research and the assessment of oral proficiency in minority bilingual children. *Language Assessment Quarterly*, 3(2), 117-149.
- Sasaki, M. (2000). Effects of cultural schemata on students' test-taking processes for cloze tests: a multiple data source approach. *Language Testing*, 17(1), 85-114.
- Schaefer, B. A., Schur, K. F., Macri-Summers, M., & MacDonald, S. L. (2004). Preschool children's learning behaviors, concept attainment, social skills, and problem behaviors: Validity evidence for preschool learning behaviors scale scores. *Journal of Psychoeducational Assessment*, 22, 15-32.
- Schappe, J. F. (2005). Early childhood assessment: A correlational study of the relationship among student performance, student feelings, and teacher perceptions. *Early Childhood Education Journal*, 33(3), 187-193.
- Schleppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education*, 12(4), 431-459.
- Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. New York, NY: Routledge.
- Shiel, G., Cregan, Á., McGough, A., & Archer, P. (2012). *Oral language in early childhood and primary education (3-8 years)*. Research Report No 14. Dublin: National Council for Curriculum and Assessment.
- Shohamy, E. (1994). The validity of direct versus semi-direct oral tests. *Language Testing*, 11(2), 99-123.
- Siegal, M., & Surian, L. (2009). Conversational understanding in young children. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp.304-323). Malden, MA: Wiley-Blackwell.
- Stanton-Chapman, T. L., & Snell, M. E. (2011). Promoting turn-taking skills in preschool children with disabilities: The effects of peer-based social communication intervention. *Early Childhood Research Quarterly*, 26(3), 303-319.
- Teasdale, A., & Leung, C. (2000). Teacher assessment and psychometric theory: A case of paradigm crossing? *Language Testing*, 17(2), 163-184.
- Tienda, M., & Haskins, R. (2011). Immigrant children: Introducing the issue. *The Future of Children*, 21(1), 3-18.





- University of Texas in Austin. (2007). *Response rates*. Retrieved October 25, 2011, from <http://www.utexas.edu/academic/ctl/assessment/iar/teaching/gather/method/survey-Response.php>
- Upshur, J. (1967). Testing foreign-language function in children. *TESOL Quarterly*, 4, 31-34.
- U.S. Department of Education. (n.d.) *No child left behind act of 2001*. Retrieved May 19, 2011, from <http://www2.ed.gov/policy/elsec/leg/esea02/index.html>.
- Vanderplank, R. (2008). The significance of first language development in five to nine year old children for second and foreign language learning. *Applied Linguistics*, 29(4), 717-722.
- Weir, C. (2005). *Language testing and validation: An evidence-based approach*. Basingstoke, UK: Palgrave MacMillan.
- Wells, G., & Montgomery, M. (1980). Adult-child interaction at home and at school. In P. French & M. Maclure (Eds.), *Adult-child conversation* (pp. 210-243). New York, NY: St. Martins Press.
- West, J., Denton, K., & Germino-Hausken, E. (2000). *America's kindergarteners*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- WIDA Consortium. (2008). *ACCESS for ELLs listening, speaking, writing, and reading sample items 2008: Grade Kindergarten*. Wisconsin: Board of Regents of the University of Wisconsin System.
- WIDA Consortium. (2008). *WIDA MODEL Measure of Developing English Language student response booklet: Grade K* (version 1.0). Wisconsin: Board of Regents of the University of Wisconsin System.
- WIDA Consortium. (2008). *WIDA MODEL Measure of Developing English Language test administration manual: Grade K* (version 1.0). Madison, Wisconsin: Board of Regents of the University of Wisconsin System.
- WIDA Consortium. (2008). *WIDA MODEL Measure of Developing English Language test administrator script: Grade K* (version 1.0). Madison, Wisconsin: Board of Regents of the University of Wisconsin System.
- Willett, J. (1995). Becoming first graders in an L2: An ethnographic study of L2 socialization. *TESOL Quarterly*, 29(3), 473-503.



- William, D., & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22(5), 537-548.
- Wood, H. (1983). Questioning the pre-school child. *Educational review*, 35(2), 149-62.
- Wu, W. M., & Stansfield, C. W. (2001). Towards authenticity of task in test development. *Language Testing*, 18(2), 187-206.
- Yin, M. (2010). Understanding classroom language assessment through teacher thinking research. *Language Assessment Quarterly*, 7(2), 175-194.
- Zangl, R. (2000). Monitoring language skills in Austrian primary (elementary) schools: a case study. *Language Testing*, 17(2), 250-60.
- Zehr, M. (2004). Tests of youngest English-learners spark controversy. *Education Week*, 24(12), 1.
- Zehr, M., (2009). Oral-language skills for English-learners focus of researchers. *Education Week*, 29(8), 8.