Title of Project:
The role of voice in high-stakes second language writing assessment

Researcher:
Cecilia Guanfang Zhao
New York University
gz312@nyu.edu

Research Supervisor:
Dr. Lorena Llosa

Cecilia Guanfang Zhao

Abstract:

　　Although a construct commonly found in writing textbooks and rubrics, voice remains a concept that is only loosely defined in the literature and mystically assessed in practice. Few attempts have ever been made to formally investigate whether an authorial voice in a written text can be reliably measured, and how the strength of an authorial voice may affect the assessment of overall writing quality. Using a mixed-method approach, this study first developed and validated an analytic voice rubric (Phase 1), and then formally investigated the relationship between voice strength and writing quality (Phase 2) in the context of a high-stakes L2 writing assessment. Specifically, 400 TOEFL® iBT argumentative writing samples were used in the study with permission from Educational Testing Service, the copyright owner. Six voice raters, who had extensive experience in L1 and/or L2 writing instruction and assessment, rated the writing samples using a preliminary voice rubric developed on the basis of Hyland's (2008) interactional model of voice. Each writing sample was double rated and the average of the two ratings was used in the subsequent statistical analysis. Qualitative data collected from rater think-aloud protocols and interviews were also analyzed to provide additional evidence on rubric reliability, validity and applicability.

　　Results from the first phase of the study showed that three categories in the preliminary voice rubric—personal aside, reference to shared knowledge, and the use of rhetorical and audience directed questions—were rarely identified in the TOEFL® iBT writing samples. They were thus excluded from all the subsequent statistical analyses. Factor analysis of the remaining voice elements pointed to a three-dimensional conceptualization of voice. The use of hedges, boosters and attitude markers together formed one dimension, which was later interpreted as a writer's "manner of presentation." In other words, a writer's use of these linguistic devices indicates whether the author's ideas are presented assertively, mildly, confidently, tentatively, enthusiastically, or maybe indifferently. A second dimension, consisting of authorial self-mention and direct reader reference, was interpreted as "writer and reader presence." Apparently, authorial self-revelation and direct reference to readers are also able to contribute to the realization of voice in written discourse. Finally, central point articulation and the use of directives together were considered as yet another component—"presence and clarity of ideas in the content." In argumentative writing, particularly, the presence or absence of a central point also does make a difference in terms of getting a clear voice across.

Analyses of qualitative data, while confirming the three main dimensions identified in the quantitative analysis, revealed that the frequency-based preliminary voice rubric was limited in capturing the strength of an authorial voice. All raters agreed that while the presence and salience of those individual voice elements did influence their perception of voice strength in a writing sample, how these elements were used was probably more important in creating a strong authorial voice. Based on their think-aloud and interview data, therefore, a qualitative piece was added into the rubric. Thus, for each dimension, there were both frequency-based ratings for individual voice elements and a more qualitative evaluation of the general voice strength evoked by that dimension (e.g., voice strength evoked by the author's manner of presentation).

This revised voice rubric was then used in the second phase of the study which examined the relationship between voice strength and writing quality. A series of correlation and regression analyses showed that overall voice strength was a significant predictor of L2 writing quality, and it was able to explain approximately 1/4 of the variance in the TOEFL writing scores (R^2 = 0.246, F = 21.297, (df) = (3, 196), P < 0.0001). After controlling for the effect of other voice dimensions, however, only the content-related dimension (presence and clarity of ideas) was a statistically significant predictor of writing quality (B = 0.39, $t(196)$ = 4.42, p < 0.001). Moreover, results from this phase also showed that a writer's background had little, if any, influence on either the voice ratings or the relationship between voice and writing quality. Even after controlling for the effect of writer background variables, overall voice strength was still a significant and positive predictor of writing quality. Finally, moderation analysis showed that the relationship between voice and writing quality did not vary as a function of any particular writer background variables.

As the first formal attempt to develop and validate an analytic rubric that measures this well-known yet slippery construct of voice, this study offers an alternative three-dimensional conceptualization of voice. Unlike the few existing measures of voice that are mostly holistic and impression-based, this analytic rubric is more pedagogically useful in that it could help demystify the seemingly intangible concept of voice for both writing instructors and student writers. Using this rubric, teachers could make this notion of voice more accessible to their students, especially L2 students, and thus better help them to construct a strong and effective authorial voice with the proper use of those linguistic- and discourse-level language features.

The lack of a close alignment between what is proposed in Hyland's (2008) theoretical model and what is observed in our empirical data, however, deserves careful interpretation. It should be noted that Hyland's model was developed based on his examination of voice-related language features in a large corpus of published academic articles, whereas data used in this study were writing samples produced by TOEFL® iBT test takers within a limited time and in response to a single prompt. The difference in the nature of the writing, the length of the written products, as well as the writers' level of language proficiency, thus, explains why certain more sophisticated voice elements (i.e., personal aside, reference to shared knowledge, and rhetorical or audience-directed questions) are rarely observed in our data.

What this suggests is two-fold. First, as much as we would want a general measure of voice that could be used across different assessment or educational settings, the evaluation of voice is probably not context-independent. Rather, we need to take into consideration the characteristics of the writing task, the genre and level of the writing, and the audience for whom the writing is intended. In different contexts, the realization of voice may be different; therefore, the criteria used to evaluate voice may vary. A second implication is that students' ability to employ voice-related features in their writing probably depends on their general language

proficiency. Hence, a potentially useful way to address the concept of voice in writing assessment and instruction is to take a developmental approach. In other words, at different L2 proficiency levels, students could be taught to use different voicing strategies appropriate at that level, moving from the simplest and most straightforward ways of voicing, such as the use of first-person pronouns, gradually to more complex and sophisticated ways, such as inserting personal comments or reflections in the midst of a statement (i.e., personal aside).

Results on the relationship between voice and writing quality could also inform L2 writing pedagogy and contribute to the development of L2 writing assessment theory and practice in the long run. As some researchers have also argued (e.g., Connor & Mbaye, 2002), in addition to the linguistic aspect of writing that is traditionally focused on when evaluating L2 writing, the sociolinguistic and communicative aspects of writing also deserve our attention when assessing L2 writing competence. And the close association between voice and L2 writing quality found in this study offers one more reason for us to start looking into the communicative aspect of writing competence in L2 writing assessment.

The observation that, after controlling for the effect of other voice dimensions, the content-related dimension was the only significant, strong, and positive predictor of L2 argumentative writing quality also bears some important implications for L2 writing instruction. Stapleton (2002) argues that L2 writing researchers and practitioners should pay less attention to voice and "return the spotlight to ideas" (p 177). What Stapleton implies here is that voice and ideas are two separate constructs that are largely unrelated to each other. The present dissertation study, however, not only identifies, in empirical data, this dimension of ideas and content as an integral part of the concept of voice, but it also shows that it is indeed this dimension of voice that is most critically related to writing quality. As one of the voice raters also shared in the interview session, whether the content "makes sense" and the ideas "stand out" is indeed the sine qua non of a strong voice in writing, especially argumentative writing. Consequently, writing instructors should not teach voice in isolation, nor treat it merely as a matter of expression; rather, they should explicitly link idea and content development to the concept of voice for their students in their writing instruction.

Results from this study also showed that key writer background variables such as age, gender, L1 background had very little impact on the observed voice strength in L2 writing samples. What this suggests is that regardless of age, gender, language and cultural background, it is still possible for writers from different backgrounds to achieve a similar level of voice strength in their writing. It is also worth noting that in this study, most of the L2 writers with a non-Indo-European L1 background were from East Asian countries, overwhelmingly China, Japan, and Korea. Since these cultures are often identified as collectively oriented and vastly different from the Western culture of individualism, researchers have argued that L2 writers from these collectively-oriented cultural backgrounds often have difficulty voicing themselves in writing (e.g., Ramanathan & Atkinson, 1999; Ramanathan & Kaplan, 1996). Nonetheless, results from this study about the lack of association between voice strength and writer background variables seem to counter this argument. It might be possible, therefore, that with proper instruction, L2 writers could also write with a very strong voice, just like their native English-speaking peers.

References

Aghbar, A. A., & Chitrapu, D. (1997). Teaching for voice in ESL writing from sources. *Maryland English Journal, 31*(2), 39-47.

Andrews, R. (2003). The end of the essay? *Teaching in Higher Education, 8*(1), 117-128.

Arter, J. A., & Chappuis, J. (2007). *Creating and recognizing quality* rubrics. Upper Saddle River, NJ: Pearson Education, Inc.

Atkinson, D. (2001). Reflections and refractions on the jslw special issue on voice. *Journal of second language writing, 10*, 107-124.

Barton, E. L. (1995). Contrastive and non-contrastive connectives: Metadiscourse functions in argumentation. *Written Communication, 12*(2), 219-239.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Bentler, P. M., & Bonnet, D. G. (1980). Significant tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588-606.

Bentler, P. M., & Dijkstra, T. (1985). Efficient estimation via linearization in structural models. In P. R. Krishnaiah (Ed.), *Multivariate analysis VI* (pp. 9-24). Amsterdam: North-Holland.

Berkenkotter, C., & Huckin, T. (1995). *Genre knowledge in disciplinary communication: Cognition/culture/power*. Hillsdale, NJ: Lawrence Erlbaum.

Blair, L. (1991). Developing student voices with multicultural literature. *English Journal, 80*(8), 24.

Bowden, D. (1999). *The mythology of voice*. Portsmouth, NH: Boynton/Cook.

Breland, H., Lee, Y., Najaran, M., & Muraki, E. (2004). *An analysis of TOEFL CBT writing prompt difficulty and comparability for different gender groups*. TOEFL Research Report 76. Princeton, NJ: Educational Testing Service.

Brindley, G. (1998). Describing language development? Rating scales and second language acquisition. In L. F. Bachman & A. D. Cohen (Eds.), *Interfaces between SLA and language testing research* (pp. 112-114). Cambridge: Cambridge University Press.

Broughton, M. A., & Fairbanks, C. M. (2003). In the middle of the middle: Seventh-grade girls' literacy and identity development. *Journal of Adolescent & Adult Literacy, 46*(5), 426-435.

Cadman, K. (1997). Thesis writing for international students: A question of identity? *English for Specific Purposes, 16*, 3-14.

California State Board of Education. (1997). Retrieve December 15, 2007, from
http://www.cde.ca.gov/be/st/ss/enggrades11-12.asp

Canagarajah, S. (2001). The fortunate traveler: Shuttling between communities and literacies by
economy class. In D. Belcher & U. Connor (Eds.), *Reflections on multiliterate lives* (pp.
23-37). New York: Multilingual Matters.

Cappello, M. (2006). Under construction: Voice and identity development in writing workshop.
Language Arts, 83(6), 482-491.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research,
1*, 629-637.

Cheng, Y.-S. (2004). A measure of second language writing anxiety: Scale development and
preliminary validation. *Journal of Second Language Writing, 13*(4), 313-335.

Comfort, J. R. (2000). Becoming a writerly self: College writers engaging black feminism essays.
College Composition and Communication, 51(4), 540-559.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, New Jersey:
Erlbaum.

Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second language writing.*
Cambridge: Cambridge University Press.

Cumming, A., Grant, L., Mulcahy-Ernt, P., & Powers, D. E. (2005). *A teacher-verification study
of speaking and writing prototype tasks for a new TOEFL*. TOEFL Monograph Series 26.
Princeton, NJ: Educational Testing Service.

Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences
in written discourse in independent and integrated prototype tasks for next generation
TOEFL. *Assessing Writing, 10*, 5-43

Cumming, A., Kantor, R., Powers, D., Santos, T., & Taylor, C. (2000). *TOEFL 2000 writing
framework: A working paper*. TOEFL Monograph Series 18. Princeton, NJ: Educational
Testing Service.

Cummins, G. S. (1994). Coming to voice. In K. B. Yancey (ed.), *Voices on voice: Perspectives,
definitions, inquiry* (pp. 48-60). Urbana, IL: National Council of Teachers of English.

DeRemer, M. (1998). Writing assessment: Raters' elaboration of the rating task. *Assessing
Writing, 5*(1), 7-29.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks,
CA: Sage Publications.

Easton, L. B. (2005). Democracy in schools: Truly a matter of voice. *English Journal, 94*(5), 52.

Ede. L. (1992). *Work in progress: A guide to writing and revising* (2nd ed.). New York: St. Martin's Press.

Educational Testing Service. (2007). *About TOEFL® iBT*. Retrieved January 15, 2009, from http://www.ets.org/portal/site/ets/menuitem.1488512ecfd5b8849a77b13bc3921509/?vgnext oid=f138af5e44df4010VgnVCM10000022f95190RCRD&vgnextchannel=b5f5197a484f40 10VgnVCM10000022f95190RCRD

Educational Testing Service. (2008a). *TOEFL iBT Dataset – Writing Form 1*. Princeton, NJ: Author.

Educational Testing Service. (2008b). *TOEFL iBT Dataset – Writing Form 2*. Princeton, NJ: Author.

Elbow, P. (1968). A method for teaching writing. *College English, 30*(2), 115-125.

Elbow, P. (1981). *Writing with power*. New York: Oxford University Press.

Elbow, P. (1994). What do we mean when we talk about voice in texts? In K. B. Yancey (ed.), *Voices on voice: Perspectives, definitions, inquiry* (pp. 1-35). Urbana, IL: National Council of Teachers of English.

Elbow, P. (1999). Individualism and the teaching of writing: Response to Vai Ramanathan and Dwright Atkinson. *Journal of Second Language Writing, 8*(3), 327-338.

Elbow, P. (2007). Reconsiderations: Voice in writing again: Embracing contraries. *College English, 70*(2), 168-188.

Farmer, F. (1995). Voice reprised: Three etudes for a dialogic understanding. *Rhetoric Review, 13*(2), 301-320.

Fecho, B. (2002). Madaz publication: Polyphonic identity and existential literacy transactions. *Harvard Educational Review, 72*(1), 93-119.

Ford, B. W. (1991). Essays of the act of the mind: Authentic voices in student writing. *English Journal, 80*(8), 30.

Fox, H. (1994). *Listening to the world: Cultural issues in academic writing*. Urbana, IL: National Council of Teachers of English.

Fulcher, G. (1996). Does thick description lead to smart tests? A data-based approach to rating scale construction. *Language Testing, 13*(2), 208-238.

Gale, X. L. (1994). Conversing across cultural boundaries: Rewriting "self". *Journal of Advanced Composition, 14*, 455-462.

Gass, S. M., & Selinker, L. (2001). *Second language acquisition: An introductory course.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Gillis, C. (2002). Multiple voices, multiple genres: Fiction for young adults. *English Journal, 92*(2), 52.

Graves, D. H. (1983). *Writing: Teachers & children at work*. Portsmouth, NH: Heinemann.

Graves, D. (1994). *A fresh look at writing*. Portsmouth, NH: Heinemann.

Halliday, M. A. K. (1985). *Introduction to functional grammar*. London: Arnold.

Hamalian, L. (1970). The visible voice: An approach to writing. *English Journal, 59*(2), 227-230.

Harris, J. (1989). Constructing and reconstructing the self in the writing class. *Journal of Teaching Writing, 8*(1), 21-29.

Hashimoto, I. (1987). Voice as juice: Some reservations about evangelic composition. *College Composition and Communication, 38*(1), 70-80.

Helms-Park, R., & Stapleton, P. (2003). Questioning the importance of individualized voice in undergraduate L2 argumentative writing: An empirical study with pedagogical implications. *Journal of Second Language Writing, 12*(3), 245-265.

Hickey, D. J. (1993). *Developing a written voice*. London: Mayfield.

Hirvela, A., & Belcher, D. (2001). Coming back to voice: The multiple voices and identities of mature multilingual writers. *Journal of Second Language Writing, 10* (1-2), 83-106.

Holding, M. (2005). Liberating the student's voice: A teacher's story of the college essay. *English Journal, 94*(4), 76-82.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure anlaysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1-55.

Hyland, K. (1994). Hedging in academic writing and EAP textbooks. *English for Specific Purposes, 13*(3), 239-256.

Hyland, K. (1998). *Hedging in scientific research articles*. Amsterdam: John Benjamins.

Hyland, K. (2000a). Hedges, boosters and lexical invisibility: Noticing modifiers in academic texts. *Language Awareness, 9*(4), 179-301.

Hyland, K. (2000b). 'It might be suggested that…': Academic hedging and student writing. *Australian Review of Applied Linguistics, 16*, 83-97.

Hyland, K. (2001). Humble servants of the discipline? Self mention in research articles. *English for Specific Purposes, 20*, 207-226.

Hyland, K. (2002a). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics, 34*, 1091-1112.

Hyland, K. (2002b). Directives: Argument and engagement in academic writing. *Applied Linguistics, 23*(2), 215-239.

Hyland, K. (2002c). Options of identity in academic writing. *ELT Journal, 56*(4), 351-358.

Hyland, K. (2002d). What do they mean? Questions in academic writing. *TEXT, 22*(4), 529-557.

Hyland, K. (2005). *Metadiscourse*. London: Continuum.

Hyland, K. (2008a). Disciplinary voices: Interactions in research writing. *English Text Construction, 1* (1), 5–22.

Hyland, K. (2008b). 'Small bits of textual material': A discourse analysis of Swales' writing. *English for Specific Purposes, 27*(2), 143-160.

Hyland, K., & Milton, J. (1997). Hedging in L1 and L2 student writing. *Journal of Second Language Writing, 6*(2), 183-296.

Iida, A. (2008). Poetry writing as expressive pedagogy in an EFL context: Identifying possible assessment tools for haiku poetry in EFL freshman college writing. *Assessing Writing, 13*(3), 171-179.

Ivanič, R., & Camps, D. (2001). I am how I sound: Voice as self-representation in L2 writing. *Journal of Second Language Writing, 10*(1-2), 3-33.

Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfiel, V., & Hughey, J. (1981). *Testing ESL composition: A practical approach.* Rowley, MA: Newbury House.

Jeffery, J. (2009). Constructs of writing proficiency in US state and national writing assessments: Exploring variability. *Assessing Writing, 14*(1), 3-24.

Jeffery, J. (2007, November). *Discourses of writing in high-stakes direct writing assessments*. Paper presented at the National Reading Conference, Austin, TX.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151.

Kamani, G. (2000). Code switching. In M. Nana-Ana Danquah (Ed.), *Becoming American* (pp. 95-104). New York: Hyperion.

Kaplan, R. B. (1966). Cultural thought patterns in inter-cultural education. *Language Learning,16*, 1-20.

Kaplan, R. B. (1972). *The anatomy of rhetoric: Prolegomena to a functional theory of rhetoric*. Philadelphia: Center for Curriculum Development.

Knoch, U. (2008). The assessment of academic style in EAP writing: The case of the rating scale. *Melbourne Papers in Language Testing, 13*(1), 34-67.

Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing, 26*(2), 275-304.

Lam, Wan Shun Eva. (2000). L2 literacy and the design of the self: A case study of a teenager writing on the internet. *TESOL Quarterly, 34*(3), 457-482.

Lee, Y., Breland, H., & Muraki, E. (2004). *Comparability of TOEFL CBT writing prompts for different native language groups*. TOEFL Research Report 77. Princeton, NJ: Educational Testing Service.

Leggo, C. (1991). Questions I need to ask before I advise my students to write in their own voices. *Rhetoric Review, 10*(1), 143-152.

Leonard, E. A. (1997). Assignment #9. A text which engages the socially constructed identity of its writer. *College Composition and Communication, 48*(2), 215-230.

Lensmire, T. J. (1994). *When children write: Critical revisions of the writing workshop*. New York: Teachers College Press.

Lensmire, T. J. (1998). Rewriting student voice. *Journal of Curriculum Studies, 30*(3), 261-291.

Li, X. (1996). *"Good writing" in cross-cultural context*. Albany, NY: State University of New York Press.

Llosa, L. (2005). Assessing English learners' language proficiency: A qualitative investigation of teachers' interpretations of the California ELD standards. *The CATESOL Journal, 18*(1), 7-18.

Llosa, L., Beck, S. B., & Zhao, C. G. (2009, March). *Defining the construct of academic writing to inform the development of a diagnostic assessment*. Paper presented at the 31st LTRC conference, Denver, CO.

Matsuda, P. K. (2001). Voice in Japanese written discourse: Implications for second language writing. *Journal of Second Language Writing, 10*(1-2), 35-53.

Matsuda, P. K., & Tardy, C. M. (2007). Voice in academic writing: The rhetorical construction of author identity in blind manuscript review. *English for Specific Purposes, 26*, 235-249.

Matsuda, P. K., & Tardy, C. M. (2008). Continuing the conversation on voice in academic writing. *English for Specific Purposes, 27*, 100-105.

McAlexander, P. J., & Marston, J. (1994). Writing and suppression of self: A case study of a basic writer. *Research and Teaching in Developmental Education, 11*(1), 75-83.

McHaney, P. A. (2004). Let every voice be heard: Focus essays create democratic classrooms. *English Journal, 93*(5), 72.

Murphy, E. (1991). Whomp! Real voices in college admission essays. *English Journal, 80*(8), 34.

Murray, D. M. (1978). Write before writing. *College Composition and Communication, 29*(4), 375-381.

Murray, D. M. (1991). *The craft of revision*. Fort Worth: Holt, Rinehart, and Winston.

Muthén, B. O., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 38*, 171-89.

New York State Education Department. (2007a). *The university of the state of New York, regents high school examination: Comprehensive examination in English*. Retrieved April 15, 2007 from http://www.nysedregents.org/testing/engre/1-107.pdf

New York State Education Department. (2007b). *Comprehensive examination in English: Scoring key and rating guide*. Retrieved April 15, 2007, from http://www.nysedregents.org/testing/engre/1rg-107.pdf

North, B. (1995). The development of a common framework scale of descriptors of language proficiency based on a theory of measurement. *System, 23*(4), 445-465.

O'Leary, M. E. (1993). *A voice of one's own: Born, achieved, or thrust upon one?* San Diego, CA: Retrieved October 10, 2006, from ERIC database.

Palacas, A. (1989). Parentheticals and personal voice. *Written Communication, 6*(4), 506-527.

Peterson's. (2007). *A brief introduction to the TOEFL*. Retrieved October 1, 2007, from http://www.petersons.com/common/article.asp?id=1147&path=ug.pft.advice&sponsor=1

Prior, P. (2001). Voice in text, mind, and society: Sociohistoric accounts of discourse acquisition and use. *Journal of Second Language Writing, 10*(1-2), 55-81.

Raimes, A., & Zamel, V. (1997). Response to Ramanathan and Kaplan. *Journal of Second Language Writing, 6*(1), 79-81.

Ramanathan, V., & Atkinson, D. (1999). Individualism, academic writing, and ESL writers. *Journal of Second Language Writing, 8*(1), 45-75.

Ramanathan, V., & Kaplan, R. B. (1996). Audience and voice in current L1 composition texts: Some implications for ESL student writers. *Journal of Second Language Writing, 5*(1), 21-34.

Ramanathan, V., & Kaplan, R. B. (1997). Response to Raimes and Zamel. *Journal of Second Language Writing, 6*(1), 83-87.

Reid, J. (2001). Advanced EAP writing and curriculum design: What do we need to know? In T. Silva & P. K. Matsuda (Eds.), *On second language writing* (pp. 143-160). Mahwah, NJ: Lawrence Erlbaum.

Ritchie, J. S. (1989). Beginning writers: Diverse voices and individual identity. *College Composition and Communication, 40*(2), 152-174.

Romano, T. (2004). *Crafting authentic voice*. Portsmouth, NH: Heinemann.

Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *Proceedings of the American Statistical Association*, 308-313.

Sawaki, Y. (2007). Construct validation of analytic rating scales in a speaking assessment: Reporting a score profile and a composite. *Language Testing, 24*(3), 355-390.

Schaub, M. (1995). *Cross-cultural dialogics: Bakhtinian theory and second language audience*. U.S.; Indiana: Retrieved October 10, 2006, from ERIC database.

Schwartz, T. A. (2003). Urban Appalachian girls and writing: Institutional and 'other/ed' identities. *Pedagogy, Culture and Society, 11*(1), 69-87.

Shen, F. (1989). The classroom and the wider culture: Identity as a key to learning English composition. *College Composition and Communication, 40* (4), 459-466.

Shohamy, E. (1996). Competence and performance in language testing. In G. Brown, K. Malmknaer, & J. Williams (Eds.), *Peformance and competence in second language acquisition* (pp. 138-151). Cambridge: Cambridge University Press.

Slayton, J., & Llosa, L. (2005). The use of qualitative methods in large-scale evaluation: Improving the quality of the evaluation and the meaningfulness of the findings. *Teachers College Record, 107*(12), 2543-2565.

Stapleton, P. (2002). Critiquing voice as a viable pedagogical tool in L2 writing: Returning spotlight to ideas. *Journal of Second Language Writing, 11* (3), 177-190.

Stapleton, P., & Helms-Park, R. (2008). A response to Matsuda and Tardy's "Voice in academic writing: The rhetorical construction of author identity in blind manuscript review". *English for Specific Purposes, 27*, 94-99.

Steinman, L. (2003). Culture collisions in L2 academic writing. *TESL Canada Journal, 20*(2), 80-91.

Stewart, D. C. (1969). Prose with integrity: A primary objective. *College Composition and Communication, 20*(3), 223-227.

Stewart, D. C. (1972). *The authentic voice: A pre-writing approach to student writing*. Dubuque, IA: Brown.

Stewart, D. C. (1992). Cognitive Psychologists, Social Constructionists, and Three Nineteenth-Century Advocates of Authentic Voice. *Journal of Advanced Composition, 12*(2), 279-290.

Strange, C., & Alston, L. (1998). Voicing differences: Encouraging multicultural learning. *Journal of College Student Development, 39*(1), 87-99.

Tang, R. (2006). Addressing self-representation in academic writing in a beginners' EAP classroom. *Journal of Language and Learning, 5*(2), 76-85.

Tang, R., & John, S. (1999). The 'I' in identity: Exploring writer identity in student academic writing through the first person pronoun. *English for Specific Purposes, 18*, S23-S39.

Turner, C. E., & Upshur, J. A. (1996). Developing rating scales for the assessment of second language performance. In G. Wigglesworth & C. Elder (Eds.), *The language testing cycle: From inceptions to washback* (Australian Review of Applied Linguistics Series S, No. 13, pp. 55-79). Melbourne: Australian Review of Applied Linguistics.

Turner, C. E., & Upshur, J. A. (2002). Rating scales derived from student samples: Effects of the scale maker and the student sample on scale content and student scores. *TESOL Quarterly, 36*(1), 49-70.

Upshur, J. A., & Turner, C. E. (1995). Constructing rating scales for second language tests. *English Language Teaching Journal, 49*, 3-12.

Upshur, J. A., & Turner, C. E. (1999). Systematic effects in the rating of second language speaking ability: Test method and learner discourse. *Language Testing, 16*(1), 82-111.

Van Sluys, K. (2003). Writing and identity construction: A young author's life in transition. *Language Arts, 80*(3), 176-184.

Wolfe, E. W., Bolton, S., Feltovich, B., & Niday, D. M. (1996). The influence of student experience with word processors on the quality of essays writing for a direct writing assessment. *Assessing Writing, 3*(2), 123-147.

Wu, S. Y., & Rubin, D. L. (2000). Evaluating the impact of collectivism and individualism on argumentative writing by Chinese and North American college students. *Research in the Teaching of English, 35*(2), 148-178.

Wyoming State Board of Education. (2003). *Wyoming Language Arts Content and Performance Standards*. Retrieved December 15, 2007, from http://www.k12.wy.us/SAA/standards/lang.pdf

Yancey, K. B. (1994). Introduction: Definition, intersection, and difference—mapping the landscape of voice. In K. B. Yancey (ed.), *Voices on voice: Perspectives, definitions, inquiry* (pp. vii-xxiv). Urbana, IL: National Council of Teachers of English.

Yeh, S. S. (1998). Validation of a scheme for assessing argumentative writing of middle school students. *Assessing Writing, 5*(1), 123-150.

Young, M. (1996). Narratives of identity: Theorizing the writer and the nation. *Journal of Basic Writing, 15*(2), 50-75.

Zhao, C. G. (2008, April). *Issues of voice in ESL writing*. Paper presented at the 42nd annual TESOL convention, New York, NY.

Zhao, C. G., & Llosa, L. (2008). Voice in high-stakes L1 academic writing assessment: Implications for L2 writing instruction. *Assessing Writing, 13*(3), 153-170.