

Use Mainstream Language Tests For Language Minority Students To Screen Communication Disorders?

Ruixia Yan

Speech-Language Pathology Department
Misericordia University
301 Lake Street, Dallas, PA 18612
United States

Abstract

Test development and test use carry important social responsibilities. There is consensus that language testing and assessment should reflect sociological and ethical aspects of testing along with concerns for naturalistic language use in tests (Canale, 1988; McNamara & Roever, 2006; Yan, 2009, 2013). This requirement is more urgent for language testing and assessment involving English Language Learners (ELL) and individuals of cultural and linguistic diversity. In assessing students from such background whose native language is a minority language, and in determining whether they have communication disorders, tests written in the mainstream language are generally used. This leads to over-representation of language minority groups as communication disordered and their under-representation in classes for the gifted (Oller, 1997, 2000, 2001; Yan, 2013; Yan & Oller, 2007). This paper examines the difficulties that tests written in the mainstream language pose for language minority students. A three-stage assessment process is discussed to avoid the disproportionate representation of language minority students in classes for learning difficulties and for the gifted.

1. Background

Language assessment plays an essential role in screening communication disorders (Damico, 1991; Hegde & Pomaville, 2008; Marquardt, 1999; Owens, Metz & Haas, 2002; Yan, 2013). According to the regulations stipulated in the *Individuals with Disabilities Education Act* (IDEA, 1997, 2004), it is required that children with communication disorders be identified before entering school. Historically, standardized procedures have been used by speech-language pathologists for this purpose (Sturmer et al., 1994). Two types of errors are commonly observed in the screening process. On the one hand, the screening process is susceptible to Type 1 errors, where the test-taker has no language problems, yet fails the test. On the other hand, Type 2 errors results form the opposite scenario, in which a test-taker has some problems, but manages to pass the test. These errors account for either the over- or under-representation of culturally and linguistically diverse students in special education. Currently, however, the major problem continues to be the over-representation of these students as communication disordered and misinterpreting language problems of these students as disabilities (Baca, 1990; Figueroa, 1989; Oller, 2000, 2001; Yan, 2009; 2013). This paper, therefore, sets out to examine language tests used to screen students from linguistically and culturally diverse backgrounds, and discusses ways to deal with the problem of disproportionate representation of language minority students in classes for communication disorders and for the gifted.

2. The language proficiency factor in testing and screening

Based on an extensive review of the literature, Ortiz (1997) concluded that tests are the primary explanation for the disproportionate placement language minority students. Oller (1997) argues that language or dialects must be called on to make sense of the surface-forms of speech or other signs used by communities to share abstract meanings. The denial or unawareness of this fact is called monoglossosis. Therefore, there is a language proficiency factor in testing and screening procedures. However, tests in the mainstream language have been widely used to diagnose language communication. Jitendra and Rohena-Diaz (1996, p. 42) observed that “assessment procedures used to diagnose communication disorders” have “virtually ignored the linguistically diverse nature” of students from minority language/dialect backgrounds.

This inadvertently leads to the disproportionate representation of language minority users in classes for communication disorders (Artiles et al., 2010; Kong & Orosco, 2015; Oller, 1997; Oller et al., 2001).

3. Critique of mainstream language tests

“There could be no science as we know it without measurement” (Henning, 1987, p.1). Language testing is one form of measurement. Specifically, language tests are “procedures for eliciting language performances which allow us to make inferences about learners’ language abilities” (Douglas, 2004b, p.17). Douglas (2004a, 2004b) reminded us that since language knowledge is a cognitive attribute and is internal to a person, it cannot be measured or observed directly. That is, language is testing functions just as a mirror, a medium or a tool enabling us to indirectly infer a person’s internalized language ability. This mirror can reflect the image of the object, but not the object itself. Actually, a wall of separation, a gap, or at least a theoretical boundary exists between language testing and actual language use in real world (Badon et al., 2005). Different testing procedures have been designed to cross the walls and gaps. The idea is to make test scores or assessment results mirror as much as possible how an examinee addresses tasks beyond the testing situations. More particularly we need to enable “valid judgments about ‘what they know and can do’ in situations that matter” (Eisner, 1999, p.2).

Two widely used language tests are discrete point language tests and processing dependent measures written in a mainstream language. However, neither of these two approaches is appropriate to screen communication disorders in minority language groups.

3.1 Use discrete point language tests for language minorities?

The theoretical framework of the widely used discrete point language testing is developed from the structuralist school of linguistics and the behaviorist school of psychology (Jitendra & Rohena-Diaz, 1996; Oller & Damico, 1991). The discrete approach views language as an autonomous cognitive ability composed of numerous, discernible, and measurable linguistic elements such as sounds, stress, tone, intonations, morphemes, words, word order, phrases, and sentences (Jitendra & Rohena-Diaz, 1996). In this vein, the basic notion of the traditional language testing is that language consists of distinct and separable structures, and language proficiency involves knowledge of structures and rules that make up linguistic competence (Nelson, 1986). Specifically, this fragmentation or discrete point approach stresses assessing structural linguistic elements such as phonology, morphology, and syntax. Unfortunately, this approach neglects the dynamic, synergistic, integrative, interactive, context-dependent nature of language complexity, and functional language usage (Artiles et al., 2005; Canale, 1988; Damico, 1988, 1991; Nelson, 1986). Nelson (1986) indicated that functional language use calls for the complex ability to “integrate linguistic rules with higher order rule systems relating to interpersonal communication, sociolinguistic considerations of the situation, and other pragmatic features” (Nelson, 1986, p.88). Therefore, the discrete approach separates language assessments from communication contexts by attaching importance to form-oriented grammatical components of language but ignoring the pragmatic aspects of language use in different contexts. This is actually a narrow view of language and the fragmentation of language into parts, and it is insufficient to assuring validity and generalizability. In addition, there are the following limitations with the discrete approach:

1) Testing environment

In standardized discrete point language tests, language is treated as a collection of diverse but discrete behaviors. Since superficial aspects of language structures are emphasized, the result of such tests is a set of spurious behaviors with relatively little to do with authentic communication in real-life (Jitendra & Rohena-Diaz, 1996; Yan, 2013). Typically, discrete point’s tests are administered in school settings, and they do not attend to the individual’s competence in using language functionally in natural contexts. This underestimates the language proficiency and abilities of students who are linguistically diverse.

2) Instrument inadequacies

Poor technical adequacy (low reliability and validity) of many traditional tests is a serious concern. Hayes-Brown (1984) evaluated various discrete point language tests and found that different information about the same group of children was yielded by those different instruments. Merino & Spencer (1983) examined several language tests which contained English and Spanish versions and revealed that those tests failed to take into consideration the differences in the development of language skills in the two languages and dialectal differences in the second language. Although we know that a bilingual student’s languages influence each other, traditional discrete point language tests do not measure this influence (Valdes & Figueroa, 1994).

3) Score interpretation

For test items, only the simple binary analysis is given in widely used discrete point language tests: “correct” or “incorrect”, which in turn yields a set of numerical scores for interpreting the language level of a test-taker. Such scores may be able to show the extent to which a minority language speaker is better or worse than his/her peers on that particular task, but they tell us little about how well the he/she is able to perform in authentic contexts of communication. Furthermore, the scores achieved by language minority groups are always normed against those of the normative mainstream to evaluate the language performance of the language minority speakers. Establishing appropriate norms is an important issue (Genesee & Upshur, 1996), however, the validity of this norming is a concern without the differentiation of the purpose of a language test. This paper argues that if a language test tends to test an examinee’s performance in the mainstream language, this comparison with the normative mainstream makes sense. However, this comparison is not appropriate if a language test is designed to assess whether a language minority speaker is communication disordered. In the latter case, an appropriate norming group should encapsulate the individuals of the same ethnic, cultural, and linguistic background, and of the same, age and education level and type (Genesee & Upshur, 1996; Saenz & Huer, 2003; Yan, 209, 2013; Yan & Oller, 2007). Otherwise, results of such comparisons will be skewed favoring the mainstream, which will lead to the over-representation of language minority groups as communication disordered.

3.2 Use processing-dependent measures for language minorities?

Campbell and Dollaghan (1997, p. 520-25) introduced the processing-dependent language measures as contrasted with knowledge- or experience-dependent measures. They claimed the processing-dependent language measures could reduce the bias in language assessment attributable to prior knowledge or experience of an examinee. According to them, processing-dependent measures are intended to be “relatively more dependent on psycholinguistic processing and less dependent on language knowledge”. They further defined psycholinguistic processing as “comprising the mental operations required to manipulate linguistic units”. Based on this, they designed three kinds of test methods: Non word repetition task, competing language processing task, and revised token test.

The non word repetition task was designed to assess phonological working memory storage. Specifically, an examinee is asked to repeat phonotactically legal nonsense words such as *kiv*, *muk*, *bakum*. These are thought to be able to decrease the bias in the repetition performance by eliminating the difference of the subjects’ vocabulary knowledge, since all the subjects are unfamiliar with these non words. This is, however, just ideal or unpractical. Although the non word repetition task rules out the influences of vocabulary knowledge, yet another factor, phonetic knowledge plays an important role. Take the non words given by the authors ---- *kiv*, *muk*, *bakum* ---- for instance, it is easier to repeat these non words for a person who is very familiar with the English phonological system, but will put a person who has no knowledge of this phonological system at disadvantage. Take the following non words in Chinese for instance, “*diáo, shuī, dóng*”, which are very easy to repeat for a person with a Chinese phonological background, but it will present a similar problem for English speakers who do not know anything about Chinese phonology.

The competing language processing task was designed to test the storage and processing functions of working memory. An examinee is given certain group of sentences (the number of sentences is different according to difficulty levels) through audiotape, and in the end the examinee is asked to recall the last word of each sentence. For example, after listening to the sentences “I like this movie” and “Today is bright”, the examinee is supposed to recall these two words: “movie, bright”. The authors argue that this method can minimize the influence of vocabulary knowledge. Is this really so? However, it is the familiarity with the words that makes the task easy to finish. Just imagine hearing the following sentences in Chinese “*cǎofēichánlǜ*” (*The grass is very green*), “*jīn tiāntiānhěnlán*” (*Today the sky is very blue*), and “*wōhěnxīhuánzhçběnshū*” (*I like this book very much*). Without the knowledge of the Chinese vocabulary, can you recall exactly the three words “*lǜ, lán, shū*” after listening to the aforementioned three sentences as easily as you do in English? Without the vocabulary knowledge of Chinese together with the visual access to the linguistic signs of the Chinese sentences stated above, it may be difficult to define the boundary of each Chinese word, and it is logically more challenging to recall the last words of the given sentences. In addition, bilinguals process information more slowly in their less stronger language (Ascher, 1990), which implies that if their speed to finish the competing language processing task is slower, or even if they cannot finish this task at all, it cannot be concluded that they are communication or learning disordered. The condition of bilingualism may account for this matter.

The denial or ignorance of the dependence on language knowledge to carry out the competing language processing task is, in essence, monoglotosis. The result of monoglotosis is that proficiency in the language of the tests is mistaken for native intelligence. Oller (1997, 2001) also claimed that the confusion of language tests for measures of innate abilities is the main source of the disproportionate representation of language minority groups in disordered and gifted classes.

The last test method designed by the authors is the revised token test. An examinee is asked to perform actions based on the commands spoken by a tester. For instance, after the tester's instruction "*Touch the big square blue book*", the examinee is supposed to do so in response. But imagine carrying out the instructions "*qīngmōdàdē, fāngdē, bìngqiěshílánshèdē shū*" (Please touch the bigger, square, and blue book), can you do so without sufficient knowledge of the Chinese language?

Therefore, the so-called processing-dependent measures are still knowledge- or experience-based. In other words, an examinee's performance in such kind of tests may also be significantly affected by his/her prior knowledge or experience in addition to possible communication disorders. That is, both the validity and reliability of such measures to test language minority groups are questioned.

3.3 Use mainstream language tests as the primary measure?

In addition to the inherent limitations, traditional language tests also exhibit some external weaknesses originating from the use of the mainstream language testing as the primary measure to diagnose communication disorders.

As developed by Hymes (1971), communicative competence is used to signal an individual's ability to use language. Communicative competence covers at least the following two characteristics: The ability to analyze the context of communication, and the ability to utilize the linguistic knowledge with communication strategic appropriateness. In this sense, communicative competence is interactive and collaborative in nature. Since language use involves linguistic, social and cultural dimensions, assessment of language proficiency should take such dimensions into consideration and consist of language competence, strategic competence, and psycho physiological mechanisms (Bachman, 1990). The primary use of the mainstream language tests, however, fails to cover these components. This is likely to increase the possibility of misdiagnosis or misplacement of language minority students.

Actually, such mono-lingual test does not usually consider the language dominance of examinees. Jitendra and Rohna-Diaz (1996) defined language dominance as the language most often used by a person for communication and self-expression. For language minority language speakers, there are variability's of communicative and intercultural competences in the mainstream language. For example, they may not speak the mainstream language at all, or they may speak limited mainstream language, or fluent mainstream language. In a certain spatio-temporal point, any bilingual or student has a temporarily stronger language (Ascher, 1990; Hallman & Fradd, 1983). If the mainstream language happens to be the relatively unfamiliar or weaker language for a student, then the diagnosis or placement of this student based on this mainstream language assessment is biased against him/her, because bilinguals process information more slowly with the less skilled language (Ascher, 1990) and also, in such situations, bilinguals are more easily disturbed by the external environmental conditions of the test, which may negatively affect their scores on tests (Ascher, 1990; Figueroa, 1989; Oller, 2001). This mono-lingual perspective is also likely to neglect the phenomenon of language transfer. For example, Chinese is not inflectional, so Chinese learners of English may omit tense markers, plural markers, and gender markers in English (Cheng, 1987). It is often to find a language minority speaker whose native language is Chinese speaks or writes English like this: *I go there yesterday; He and she is good friend; He is my sister*. In short, it is necessary to consider the facts about bilingualism of the language minority groups in addition to measures by mainstream language tests.

Furthermore, the mono-cultural perspective of the mainstream language assessment must be expanded to embrace the possibility of diagnosing language differences as communication disorders due to cultural differences. For the language minority language speakers who are just entering unfamiliar social and cultural contexts, their native culture cannot be meshed immediately with the target culture (Damico & Damico, 1993; Yan, 2013). What is taken for granted or accepted in one culture may be weird or even offensive in another. The appropriate language codes or styles in one culture may be unintelligible, inappropriate or irrelevant in another. For example, *have you eaten?* Is a typical way of greeting somebody in China, but it seems strange in America? As a result, the cultural shock or difference may influence these speakers' target or mainstream language performance.

That is, the primary use of the mainstream language assessment may put language minority speakers at a disadvantage, which is against the ethics of assessment with culturally and linguistically diverse populations (Crowley, 2004).

Accordingly, the use of mono-lingual and cultural mainstream language tests as the primary measures is insufficient or even mistake-leading for language minority groups. The rapidly increasing diversity of population in the United States has called for the exploration of alternatives to traditional standardized screening procedures (Yan, 2013; Yan & Oller, 2007).

4. Strategies to reduce bias in language assessment

To avoid the disproportionate representation of language minority students in classes for communication disorders and for the gifted, a naturalistic and ethical language assessment should include the following three stages:

4.1 Language pre-assessment --- case history

Case history is the cornerstone of assessing a language minority student's linguistic skills. Since the backgrounds of language minority groups are so linguistically and culturally diverse, it is necessary to get the demographic information relating to a person's "history", his/her home language, as well as reports on the "new" language (mainstream language) performance, the time of arrival in this "new" country, and so forth (Anderson, 2004). Such information can serve as a general language pre-assessment, which is also the starting-point for further more detailed and more accurate assessment.

A case history can be gathered through a variety of means (Damico, 1991; Restrepo, 1998; Westby, 1990). For instance, self-report by a language minority student him/herself, interviews with parents and teachers with open-ended questions are helpful to explore a language minority student's language learning experiences. In addition, it is also productive to gather some sample language data of an examiner, in both his/her native language and the mainstream language. Case history can help to provide a rough picture of the examinee's overall language performance and skills, and language dominance.

4.2 Assessment in bilingual settings ---- pragmatic and dynamic

Assessment used for minority language users should include all languages with which the child has significant experience. The following types of assessment approaches can be utilized:

1) Pragmatic approach

Since the widely used traditional language tests are divorced from the dynamic communicative contexts, the result of which is to reduce the highly dynamic, synergistic, integrative, and context-bound process of communication to some numerical scores, which are stripped of the substantial content of human communicative behavior. A richer, deeper, and more detailed descriptive approach to record and assess the language use is called for. Oller (2000) suggested that language testing should adhere to three pragmatic criteria: Data should be meaning-based, contextually-embedded, and temporally-constrained (also in Yan, 2013; Yan & Oller, 2007). The pragmatic approach views language as a dynamic, synergistic, and integrative construct including both intrinsic cognitive factors and extrinsic contextual features, and it stresses the performance of an examinee in rich and naturalistic communication contexts. The procedure is not limited to the collection of language fragmentations; however, language data to be collected should be meaning-based and integrative. That is, the focus of this approach is on functional aspects of communication in terms of the effectiveness, fluency, and the appropriateness of meaning transmission (Damico, 1991; Damico, Oller, & Tetnowski, 1999).

Ethnography is a method of studying social actions and people to reveal the underlying rules, but it is often used now in the study of education and the learning process (Cheng, 1987). In a descriptive language approach, an examiner may take the role as an ethnographer, who can employ various qualitative research techniques. An effective technique is systematic observation of the examinee's communicative interaction. The technique involves observing the targeted behaviors of the examinee in multiple settings, matching the observed behaviors with the items of observational format, and analyzing the collected data after observation (for details, see Damico, Oller, & Tetnowski, 1999). In order to enhance reliability of this approach, this paper proposes that observations be given in bilingual settings, because it is often the case that an examinee is more active or talkative in a more familiar language circumstance, but relatively quiet or even silent if he/she is less skilled in the language used to communicate.

Therefore, observations in bilingual settings may balance the language performance of an examinee's in both less familiar and more skilled languages, which is consequently helpful for assessors to get a more reliable language assessment.

2) Dynamic approach

Dynamic assessment is based on Vygotskian social interactional view of development. According to Vygotsky (1978, 1986), the transformation from lower, natural mental function to higher, cultural mental function has its origins in social interaction. Vygotsky believed that behaviour exists socially before it is turned into a part of the internal behaviour of an individual. Such external social activities are internalized in a person to restructure his/her mental processes, during which the language or mediation of other more competent people played a pivotal role. This transformation process was called semiotic mediation by Vygotsky. The mediation, or guided social interaction, moves a person from other-regulated or external to self-regulated or internal functioning. Vygotsky also believed there was a difference between a person's actual development level and potential development level, which he called the zone of proximal development (ZPD). Put in another way, ZPD is the zone in which a learner cannot complete certain tasks without the help of a more capable person who provides scaffolding. In this vein, the focus of this approach is on how a person responds to scaffolding in acquiring new skills.

This direction of assessment incorporates a learning component into the testing situation (Benjamin & Lomofsky, 2002; Lidz, 1987). In terms of language minority students' tests, the Vygotskian theory implies that the scaffolding technique tailored to meet the needs of a student should be meshed into the language testing processes. After this learning stage for an examiner, a post-test will be administered. High responsiveness to scaffolding and a great improvement on post-testing indicate that the student is not communication or learning disordered (Baek & Kyoung, 2003; Bain, 1994; Barrera, 2003; Bountrogianni, 1990; Elliott, 2000, 2003; Lidz, 1987; Yan, 2009, 2013; Yan & Oller, 2007).

4.3 Follow-up re-assessment

In order to avoid either the type 1 or the type 2 error, it is necessary to have follow-up re-assessment as complementary. Re-evaluation can act as a safety net for those who have been misdiagnosed or misplaced. The follow-up stage can be more reflective of student growth than just one-time snapshot standardized assessment (Jalomo, 2000). Portfolio assessment involves collection of performances of a student over time and provides a more holistic assessment, which can function as a follow-up re-assessment to reduce the bias or errors of language assessment, and to record the development of the client's language abilities (Cook-Benjamin, 2001; Fouriea & van Niekerk, 2001; Hanson & Gilkerson, 1999; Hoepfl, 1993; Palardy, 1994; Yan, 2013). Portfolio assessment puts emphasis on continuous and critical reflection (Fouriea & van Niekerk, 2001). To be more specific, the first step in planning a portfolio is to set the purpose for conducting this assessment. Will the results be used for further assessment of a person's language ability, or for the instruction decisions, or for some other purposes? The second step is to determine the specific focus (for instance, oral or written language skills). Then, assessment measures should be identified, followed by the determination of the criteria to interpret the portfolio contents. The data collection for portfolio assessment will be more comprehensive if students, teachers, parents are involved in the evaluation process (Pierce & O'Malley, 1992).

5. Practicality of the proposed strategies to screen language minority groups

The issue of practicality covers a range of aspects, such as the cost of development and maintenance, test length, ease of marking, time required to administer the test, ease of administration and equipment required (Davies et al., 2002). *No Child Left behind Act* is an education reform to build the mind and character of every child, from every background, in every part of America. For testing, the purpose of state assessments required under *No Child Left* is to provide an independent insight into each child's progress, as well as each school's. This information is essential for parents, schools, districts and states in their efforts to ensure that no child--regardless of race, ethnic group, gender or family income -- is trapped in a consistently low-performing school. (E.D., 2005). Along this line, the responsibility of test developers and test users should suppress the concern with money and time. In addition, *No Child Left behind Act* has called for more than \$675 million – a 51 percent increase since 2001 to help ensure that students with limited English proficiency learn English and meet high academic standards (E.D., 2005). This money, if used appropriately, can to a high degree promote the practicality of the suggested assessment approaches.

6. Conclusion

This paper argues that different language testing approaches should be utilized for language minority groups according to different testing purposes: Be it to assess an examinee's performance in the mainstream language, or to test whether he/she is communication disordered. This paper illustrates the latter case. Based on the inherent limitations of the widely used mainstream language tests, and the mono-lingual and cultural weaknesses related to them as a primary use, this paper argues that it is insufficient or even error-leading to assess language minority groups by primarily using mainstream language tests. In order to overcome these weaknesses and improve the validity and reliability of the language assessments for language minority people, this paper suggests the following strategies: Conducting pre-assessment, assessing in bilingual settings, and doing follow-up re-assessment.

References

- Anderson, R. (2004). First language loss in Spanish-speaking children: Patterns of loss and implications for clinical practice. In B. Goldstein (Ed.), *Bilingual language development and disorders in Spanish-English speakers*. Baltimore, MD: Brookes.
- Artiles, A. J., Klingner, J., Sullivan, A., Fierros, E. (2010). Shifting landscapes of professional practices: English learner special education placement in English-only states. In P. Gándara & M. Hopkins (Eds.), *Forbidden language* (pp. 102-117). NY: Teachers College Press.
- Artiles, A. J., Rueda, R., Salazar, J. J., & Higareda, I. (2005). Within-group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children*, 71(3), 283-3000.
- Ascher, C. (1990). Assessing bilingual students for placement and instruction. ERIC/CUE Digest No. 65.
- Baca, L. M. (1990). Theory and practice in bilingual/cross cultural special education: major issues and implications for research, practice, and policy. ERIC Database. ED341267.
- Bachman, L. (1990). Fundamental considerations in language testing. Oxford: Oxford University Press.
- Badon, L. C., Oller, S. D., Yan, R., & Oller, J. W. (2005). Gating walls and bridging gaps: Validity in language teaching, learning, and assessment. Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics, 5(1). <http://www.tc.columbia.edu/academic/tesol/Webjournal/>
- Baek, S-G & Kyoung, J.K. (2003). The effect of dynamic assessment based instruction on children's learning. *Asia Pacific Education Review*, 4(2), 189-198.
- Bain, B.A. (1994). A framework for dynamic assessment in phonology: Stimulability revisited. *Clinical Communication Disorders*, 4(1), 12-22.
- Barrera, M. (2003). Curriculum-based dynamic assessment for new- or second-language learners with learning disabilities in secondary education settings. *Assessment for Effective Intervention*, 29(1), 69-84.
- Benjamin, L. & Lomofsky, L. (2002). Effects of the observation of dynamic and static assessment on teachers' perceptions of the learning potential of less academic learners. *Journal of Cognitive Education and Psychology*, 2(2), 102-123.
- Bountrogianni, M. (1990). Dynamic assessment: Implications for classroom consultation, peer tutoring and parent education. In E. Cole & J.A. Siegel (Eds.). *Effective consultation in school psychology* (pp. 129-140). Kirkland, WA: Hogrefe & Huber.
- Campbell, T., & Dollaghan, C. (1997). Reducing bias in language assessment: processing-dependent measures. *Journal of Speech, Language & hearing Research*, 40(3), 519-526.
- Canale, M. (1988). The measurement of communicative competence. *Annual Review of Applied Linguistics*, 8, 67-84.
- Cheng, L. L. (1987). Assessing Asian language performance: Guidelines for evaluating limited-English-proficient students. Aspen Publishers, Inc. 125-151.
- Cook-Benjamin, L. (2001). Portfolio assessment: Benefits, issues of implementation, and reflections on its use. *Assessment Update*, 13(4), 6-7.
- Crowley, C. J. (2004). The ethics of assessment with culturally and linguistically diverse populations. *The ASHA Leader*, 9(5), 6-7.
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., & McNamara, T. (2002). *Dictionary of language testing*. Cambridge University Press.

- Damico, J. S. & Damico, S. K. (1993). Clinical forum: Language and social skills in the school-age population. *Language, Speech, and Hearing Services in Schools*, 24, 236-243.
- Damico, J. S. (1988). The lack of efficacy in language therapy: A case study. *Language, Speech, and Hearing Services in Schools*, 19, 51-67.
- Damico, J. S. (1991). Descriptive assessment of communicative ability in limited English proficient students. In E. V. Hamayan & J. S. Damico (eds.) *Limiting bias in the assessment of bilingual students*. Austin, TX: PRO-ED. 157-217.
- Damico, J. S., Oller, J. W., Jr., & Tetnowski J. A. (1999). Investigating the inter observer of a direct observational language assessment technique. *Advances in Speech Language Pathology*, 1, 77-94.
- Department of Education. (2005). NCLB Executive Summary. Retrieved on November 9, 2014, from <http://www.ed.gov/nclb/overview/intro/execsumm.html>
- Douglas, D. (2004a). Assessing the language of international civil aviation: issues of validity and impact. In Proceedings of the International Professional Communication Conference, IEEE Professional Communication Society. Minneapolis: IEEE. 248-252.
- Douglas, D. (2004b). English language testing in the context of aviation English. *ICAO Journal*, 59(3), 17-18, 25-26.
- Eisner, E. (1999). The uses and limits of performance assessment. *Phi Delta Kappan*, 80, 658-660.
- Elliott, J. (2003). Dynamic assessment in educational settings: Realising potential. *Educational Review*, 55(1), 15-32.
- Elliott, J.G. (2000). Dynamic assessment in educational contexts: Purpose and promise. In C.S. Lidz & J.G. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications* (713-740). Amsterdam: JAI/Elsevier Science.
- Figueroa, R. A. (1989). Psychological testing of linguistic-minority students: Knowledge gaps and regulations. *Exceptional Children*, 56 (2), 145-153.
- Fouriea, I. & van Niekerk, D. (2001). Follow-up on the use of portfolio assessment for a module in research information skills: An analysis of its value. *Education for Information*, 19(2), 107-126.
- Genesee, F., & Upshur, J. (1996). Classroom-based evaluation in second language education. New York: Cambridge University Press.
- Hallman, C. L. & Fradd, S. H. (1983). Implications of psychological research for assessment and instruction of the culturally different bilingual multicultural education training project for school psychologists and guidance counsellors. ERIC/ED269537.
- Hanson, M. F. & Gilkerson, D. (1999). Portfolio assessment: More than ABCs and 123s. *Early Childhood Education Journal*, 27(2), 81-86.
- Hayes-Brown, Z. (1984). Linguistic and communicative assessment of bilingual children. In C. Rivera (Ed.), *Placement procedures in bilingual education: Educational policy and issues* (p. 40-105). Great Britain: Oxford University.
- Hedge, M.N., & Pomaville, F. (2008). Assessment of communication disorders in children. Plural Publishing.
- Henning, G. (1987). A guide to language testing: development, evaluation and research. Heinle & Heinle Publishers, a division of Wadsworth, Inc.
- Hoepfl, M. (1993). Portfolio assessment. *Technology Teacher*, 53(2), 28-29.
- Hymes, D. (1971). On communicative competence. Philadelphia: University of Pennsylvania Press.
- Individuals with Disabilities Education Act Amendments (IDEA). (1997). Retrieved on September 18, 2015, from <http://www.ed.gov/offices/OSERS/Policy/IDEA/index.html>
- Individuals with Disabilities Education Act Amendments (IDEA). (2004). Retrieved on September 23, 2015, from <http://www.ed.gov/policy/speced/guid/idea/idea2004.html>
- Jalomo, R., Jr. (2000). Assessing minority student performance. *New Directions for Community Colleges*, 112.
- Jitendra, A. K., & Rohen-Diaz, E. (1996). Language assessment of students who are linguistically diverse: Why a discrete approach is not the answer. *School Psychology Review*, 25(1), 40-57.
- Lidz, C. (Ed.). (1987). *Dynamic assessment*. New York: Guilford Press.
- Kong, J.E., & Orosco, M.J. (2015). Word-problem-solving strategy for minority students at risk for math difficulties. *Learning Disability Quarterly*, September. doi: 10.1177/0731948715607347.
- Marquardt, T. P. (1999). Assessment in communication disorders: Some observations on current issues. *Language Testing*, 16(3), 249-269.

- McNamara, T., & Roever, C. (2006). Language testing: The social dimension. Wiley-Blackwell.
- Merino, B. J. & Spencer, M. (1983). The comparability of English and Spanish versions of oral language proficiency instruments. *NABE: The Journal of the National Association for Bilingual Education*, 8, 1-4.
- Nelson, S. H. (1986). The assessment of English language proficiency: Standards for determining participation in transitional language programs. *Journal of Law and Education*, 15(1), 83-103.
- Oller, J. W., Jr. (1997). Monoglotosis: What's wrong with the idea of the IQ meritocracy and its racy cousins? *Applied Linguistics*, 18(4), 467-507.
- Oller, J. W., Jr. (2000). Testing verbal (language) and non-verbal abilities in language minorities: A socio-educational problem in historical perspective. *Language Testing*, 17(3), 341-360.
- Oller, J. W., Jr. (2001). Testing relations between language (verbal) and nonverbal abilities in children and adults acquiring a nonprimary language. *Language Testing*, 18(1), 33-54.
- Oller, J. W., Jr., & Damico, J. S. (1991). Theoretical considerations for the assessment of LEP Students. In E. V. Hamayan & J. S. Damico (eds.) *Limiting Bias in the assessment of bilingual students*. Austin, TX: PRO-ED. 77-110.
- Ortiz, A.A. (1997). Learning disabilities occurring concomitantly with linguistic differences. *Journal of Learning Disabilities*, 30, 321-32.
- Owens, R. E., Jr.; Meta, D. E., & Haas, A. (2002). *Communication disorders: A life span perspective*. Pearson Education, Inc.
- Palardy, J. M. (1994). Theoretical underpinning of portfolio assessment. *Journal of Instructional Psychology*, 21(1), 90-93.
- Pierce, L. V., & O'Malley, J. M. (1992). Performance and portfolio assessment for language minority students. NCBE Program Guide Series. N0.9. Spring.
- Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research*, 41, 1398-1411.
- Saena, T. I., & Huer, M. B. (2003). Testing strategies involving least biased language assessment of bilingual children. *Communication Disorders Quarterly*, 24(4). 184-193.
- Sturner, R., Layton, L., Evans, A., Heller, J., Funk, S., & Machon, M. (1994). Preschool speech and language screening: A review of currently available tests. *American Journal of Speech-Language Pathology*, 3(1), 25-36.
- Valdes, G., & Figueroa, R. A. (1994). *Bilingualism and testing: A special case of bias*. Norwood, NJ: Ablex.
- Vygotsky, L. S. (1986). *Thought and language*. Cambridge, Massachusetts: The MIT Press.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, Massachusetts: Harvard University Press.
- Westby, C. (1990). Ethnographic interviewing: Asking the right questions to the right people in the right ways. *Journal of Childhood Communication Disorders*, 13, 101-111.
- Yan, R. (2009). Assessing English Language Proficiency in International Aviation: Issues of Reliability, Validity, and Aviation Safety. VDM Publishing House Ltd.
- Yan, R. (2013). Assessing the English language proficiency of international aviation staff. In A.J. Kunnan (Eds.), *The companion to language assessment (CLA)*. Wiley-Blackwell. pp. 484-496.
- Yan, R., & Oller, J.W.Jr., (2007). Processing-dependent measures as a failed solution to the assessment of individuals from language and dialect minorities, *Communicative Disorders Review*, 1(3-4), 201-213.