Issues, Challenges, and Future Directions for Multilingual Assessment

Yangting Wang
University of Texas, San Antonio, U.S.A

Shikun Li
Nanjing Normal University, Nanjing, China

Abstract—Multilingual teaching and learning practices are often implemented in K-12 classrooms. However, issues related to multilingual assessment are rarely investigated. With the growing population of multilingual learners in the classroom, there is a great need to understand what multilingual assessment is and how to assess students who come from culturally and linguistically diverse backgrounds. The current study attempts to fill the research gap by reviewing the assessment literature over the past 15 years on multilingualism. We summarize and synthesize three main themes: 1) issues related to multilingualism, 2) difficulties, and challenges of multilingual assessment, and 3) approaches to assessment for multilingualism. We further divide the third theme into five subcategories: ideological shift, new ways of measuring English language proficiency, translanguaging-based assessment, dynamic assessment, as well as incorporating qualitative research. The study introduced the challenges of implementing multilingual assessment, offers an overview of the different approaches, and calls for more work to be conducted using the approaches.

Index Terms—multilingual assessment, multilingualism, sociocultural context

I. INTRODUCTION

With increasing globalization, mass migration, and growing mobility in modern society, multilingualism has become a widespread topic across many continents. Although multilingual teaching and learning practices are often implemented in classroom settings, issues involving assessment for multilingualism are rarely addressed. Assessment policies and practices have a history of lagging behind the advances in teaching (Mathew, 2008). As Shohamy (2011) points out “Although multilingual teaching and learning are currently promoted, encouraged, and practiced, especially in a European context, there are no voices that argue in favor of multilingual tests” (p. 421). The sociopolitical and economic reasoning behind this is to ensure that the dominant group maintains its power and the status quo of the society. The monolingual tests reflect “one language, one nation” collective ideology and perpetuate the traditional “native-like” norm. This deficit monolingual view overlooks the changing socio-cultural context and further marginalizes the minority groups.

Additionally, tests are powerful tools that have a tremendous influence on learning and teaching (Shohamy, 2011). Shohamy presents empirical data that monolingual tests such as national achievement tests, citizenship tests, and Program for International Assessment (PISA) have a negative influence on individual students, especially bilinguals. Second language learners and emergent bilinguals are often left behind because their content knowledge is often assessed in their second or third language (Solano-Flores, 2016). On the other hand, multilingual assessment can more accurately reflect test takers’ academic knowledge and provide more accurate, fair, and valid information about learner’s competence (Menken & Shohamy, 2015; Gorter, 2015).

Due to the sparse investigation on multilingual assessment and its importance in educational and social development, a need for an integrative review of the literature on assessment for multilingualism is imperative. The current issue highlights the trend of research studies on multilingual assessments over the past 15 years. Our guiding research question is: what are the central topics of multilingual assessment?

II. METHODOLOGY

We first searched for articles that contained all of the words “multilingual assessment” in the title published from 2003 to 2018 in Google Scholar. 116 articles appeared in the results. We then searched words consisting of “multilingual assessment” in article title/keywords/abstract in Linguistics and Language Behavior Abstracts (LLBA) and ERIC databases. Only peer-reviewed academic journals from LLBA and ERIC published from 2003 to 2018 were included. There are 137 articles from LLBA and 10 articles from ERIC. We then applied the following exclusion criterion: 1. Article duplicates. 2. Journal publications related to learning disabilities, applied psycholinguistics,
language pathology. 3. Articles published in languages other than English. A final result of 31 articles was used in the current literature review.

III. RESULTS

A review of literature on multilingual assessment from 2003 to 2018 yields three major research concentrations: Advantages of being bi/multilingual, challenges of multilingual assessment, and diverse approaches to multilingual assessment.

A. Advantages of Being Bi/Multilingual

Studies suggest that bi/multilinguals significantly outperform monolinguals on both verbal and nonverbal intelligence tests (Peal & Lambert, 1962; Andrade et al., 1989). According to Thomas and Collier (1998), multilinguals not only exhibit advantages over monolinguals in linguistic domains such as larger vocabulary size, better proficiency and accuracy, and effective reading and grammar strategies, they are also better in some cognitive, social, and personal areas. Multilinguals have better working memory in terms of storing and processing information (Bialystok et al., 2004). Learning a second language was shown to have beneficial effects on the main elements of executive functioning (Diamond, 2006). Additionally, bi/multilinguals have fewer problems with concentration (Wodniecka, 2007). They are more effective in multitasking and better at resolving conflicting information (Costa et al., 2008). Learning another language was also found to help improve cultural awareness and intercultural competence (Paradowski, 2011). Moreover, there are social and employment benefits of being able to speak multiple languages. Lastly, bilingualism has long-term health benefits in developing the brain (NNELL, 1996), reducing the risk of attention-deficit/hyperactivity disorder (Toppelberg et al., 2002), and even delay dementia (Bialystok, Craik & Freedman, 2007). Ample studies demonstrated considerable benefits of learning a foreign language and being bi/multilingual. The concepts of multicompetence, multi-dialectal, multicultural, and multimodal were called for in the field of second language acquisition (Block, 2013). Therefore, assessment for multilingualism can significantly contribute to the development of multicompetent individuals in society.

B. Difficulties and Challenges of Multilingual Assessment

Multilingual assessments are often used in classrooms in the form of formative and classroom-based assessments. However, cases where multilingual assessments are administered in standardized and large-scale testing situations, are very rare (Garcia & Li, 2014). Since multilingual dimensions include complex features and it regards language as a fluid dynamic system (Canagarajah, 2011; Garcia & Li, 2014), developing tests and rubrics that are based on the multilingual constructs of language is a challenge (Shohamy, 2011). McNamara (2015) indicates that the challenges of conducting multilingual assessments include the lack of teachers’ multilingual competence, resources in developing, administering, and scoring dynamic assessments, as well as overcoming resistance from the dominant monolingual ideology. There has been a lack of bilingual certified teachers in the United States. A significant number of pre-service or in-service teachers did not have adequate training on English language learners. Teachers’ multilingual competence posed a serious problem in implementing a multilingual approach in schools.

Another important issue is the definition and assessment of English language proficiency (ELP). ELP is closely related to English language assessment, as indicated by Spolsky (1989), “one cannot develop sound language tests without a method of defining what it means to know a language, for until you have decided what you are measuring, you cannot claim to have measured it” (p.138). The question of what counts as the norm of English is extremely problematic. Language assessments have long been regarded the native speaker’s English as the scoring model. The traditional monolithic view of English as a Lingua Franca (ELF) is highly criticized as it represents imperialism and hegemonic power.

Seidhlofer (2005) and Murray (2016) discussed the view of “World English” and “Englishes.” They agreed with Rajagopalan who posited that “Word English (WE) belongs to everyone who speaks it, but it is nobody’s mother tongue” (2004, p.111). The widely used communicative competence model proposed by Canale and Swain (1980) and the CEFR framework took into account the cultural context in assessing language. However, the multilingual environment and the problem of what counts as correct and appropriate use of a language in specific settings are not addressed. The definition of language needs to change in the context of globalization and multilingualism in order for tests to better assess learners and achieve construct validity (Shohamy, 2011).

C. Approaches to Multilingual Assessment

To overcome these challenges and promote social equity and justice, various language conferences, organizations, and projects focused on promoting multilingual assessment. Ideological shifts are needed to reconceptualize assessment practices. Educational policies such as the Every Student Success Act passed in 2015 in the United States and the Every Child Matters legislation launched in 2003 in the UK were announced. Language assessment scales such as CEFR, ACTFL, and WIDA have moved to can-do statements.

Shohamy (2011) believed that a multilingual approach can lead to a more construct valid assessment that enables further knowledge of bilingual students’ skills and abilities. Gorter and Cenoz (2016) stated three trends of a holistic
multilingual approach. First, is assessing comprehension. Ensuring that students are assessed in their first language (L1) or both their first and second language (L2) in content-based tests. Second, is regarding multilingual scoring which takes into account scores from both L1 and L2 exams and computes a composite score. Third, is the translanguaging approach in assessment. That is, legitimizing translanguaging in language tests and allowing tests to reflect actual language practices and design bilingual rubrics. Solano-Flores (2016) proposes to use a 'confusion matrix' and ‘linguagrams’ to assess multilinguals. Confusion matrix refers to the way where multilinguals were assessed twice using comparable tests. Based on the consistency of two test results, educators can eliminate false positive or negative English proficiency classifications of students. Another suggestion was the use of linguagrams which measure students’ four subskills in both languages. By using these two methods, educators can obtain a more comprehensive evaluation of students’ learning abilities (Solano-Flores, 2016; Wang, 2017).

We now turn to the five key topics to multilingual assessment which include an ideological shift in assessing languages, new ways to measure English language proficiency, dynamic assessment, translanguaging-based assessment, and qualitative approaches.

1. Ideological Shift.

An ideological shift from a monolingual to a multilingual perspective was presented by leading scholars over the past 50 years. Gorter and Ceno (2016) proposed an ideological shift from traditional views of defining and assessing languages separately, to an innovative emergent view of teaching and assessing multilinguals as a whole. That is, rethinking multilingualism in a situated, interactional, dynamic, and changing context (Bloemmmert, Collins, & Slembrouck, 2005). Multilingual learners should not be treated as a homogenous group (De Backer et al, 2016) and emergent bilinguals should not be viewed as two different monolingual individuals (Solano-Flores, 2016). Thus, moving from an ideology of language as problems, to language as resources (Ruiz, 1984). Key assessment practices include: assessment for learning rather than assessment of learning (Leung, 2004; Rea-Dickins 2006); culturally responsive assessment practices (Laher & Cockcroft, 2017); Nelson-Flores’ (2018) ecological approach and “water cycle-based” assessment; Shohamy’s (2001) democratic approach; De Backer et al.’s (2016) Functional Multilingual Assessment for Learning (FMA4L); as well as ethical issues and social impacts on language assessment (McNamara & Roeber, 2006).

Additionally, other reframing and reconceptualization, including ELF and CEFR framework, were put forth to better understand and assess multilinguals. House (2003) called for a change of conceptualization on English as a lingua franca (ELF). He believed that ELF involves code-switching and negotiation of meaning during interactions using a variety of languages. He further discusses several terms associated with ELF such as interlanguage, multicompetence, and “World Englishes”. He stated that ELF should not be viewed as a threat and discusses three research projects which demonstrate how local discourse norms are not influenced by deficit views of ELF. According to House (2003), ELF speakers’ interactions are unique and consist of a hybrid of different discourses. McNamara (2018) also saw ELF as a positive phenomenon and used ELF as a term similar to multilinguals.

Moreover, Jones and Saville (2009) called for the use of CEFR as a multilingual assessment framework for language learning and teaching. The CEFR framework is created as a standard for all languages, however, the one-sided adoption of the CEFR framework focuses primarily on English language assessments. This caused people to criticize the framework and its impact on learning. They believed that CEFR should be implemented flexibly in a sociocultural context, not restricted in its own prescriptive standards. Three initiatives that were proposed include Asset Languages which is a multilingual assessment framework created for pupils, European Language Portfolios, and European Survey on Language Competences (Jones & Saville, 2009). These initiatives were aimed at supporting learning, teaching in a multilingual context, and achieving comparable measures across different languages.

2. New Approaches in Measuring ELP.

Since the 1990s, new approaches were put forth to define and measure English language proficiency. Martinez-Flor et al. (2006) agreed with the multimodality dimension of language and believed that non-verbal behaviors, such as body language, facial expressions, eye contact, should be considered during a spoken interaction. Following the multimodal dimension of language and interaction patterns, Kramsch (2005) argued to include the notion of intercultural communicative competence in the construct of communicative competence. Martinez-Flor et al. (2006) suggested the inclusion of language functions, including cohesion and coherence, and turn-taking mechanisms in interactional competence. In CEFR (2001), turn-taking, cooperation, and asking for clarification were deemed as three illustrative acts for measuring interaction strategies.

However, according to Leung’s analysis of the CEFR’s three components of competences and their can-do descriptors, resided solely in the individual’s communicative language ability and only considered advanced level learners using English in a sociocultural conventionalized way. The definition was still limited and restricted as it disregarded the social dynamics of the interaction. Therefore, Leung (2014) posited that participatory involvement must be taken into account in considering a language learner’s proficiency. Similarly, Davis, Norris, Malone, McKay, & Son (2018) have incorporated turn-taking, collaboration, engagement, and appropriateness in looking at a learner’s English interactional competence.

Furthermore, scholars suggested the importance of the inclusion of context and environment while examining multilinguals. Blommaert et al. (2005) proposed the use of scale and spatial analysis to assess multilinguals and
introduced the term “truncated multilingualism.” Truncated multilingualism argues against the deficit view and sees multilingualism as being fully competent in various languages. Truncated multilingualism also refers to linguistic competences organized on the basis of topics and specific activities. For example, indexicality and positional accuracy are important to understand the situatedness of people’s linguistic competence. Blommaert et al. (2005) believed that the value of what is deemed appropriate in a given environment is related to the notion of negotiation and repair that occurs during conversation exchanges. Multilingual speakers scale their expectations and change the variety of their speech depending on their environment and space (Blommaert et al., 2005). Therefore, when assessing a person’s linguistic competence, it is important to include the context in which the conversation is based upon.

3. Translanguaging-based Assessment.

According to De Backer et al. (2016), translanguaging practices need to be acknowledged and accommodations should be implemented in large-scale testing. Additionally, “interlanguage” needs to be considered in the context of assessment (Mathew, 2008). Heugh et al. (2016) showcased the possibility of using translanguaging and code-switching in large-scale assessment. Given that a nation-wide multilingual assessment practice was difficult to implement in both Europe and North America, Heugh et al. (2016) presented a reverse case in South Africa where for centuries the majority of the population has been instructed and assessed in multiple languages. The authors implemented system-wide education policies and school exit examinations in South Africa and argued that large-scale assessment can include the use of students’ multilingual capabilities and repertoires. Students were given exam papers in two languages, one language on the left, another on the right side. What was found to be surprising was that even English-speaking students found the translation helpful. The system-wide assessment provided support for each linguistic background. Heugh et al., (2016) indicated that systematic support of bilingual education was missing for low socioeconomic status African students. However, such support is vital for overall student achievement. Solano-Flores (2015) also proposed the use of a translanguaging-based testing approach to address the tension between the following two: general accessibility of test materials for all students and addressing the language heterogeneity of bilingual students. Several specific strategies were provided to address this tension. One of them is to develop an alternative framework of standardized exams that measure the relationship between language and the test constructs.

4. Dynamic Assessment

Activities such as using digital videos, classroom-based assessment, and self-assessment are ways for teachers to understand and assess learners’ abilities and potential in a multilingual environment. Gajek (2011) illustrated the use of digital video projects in classrooms to raise students’ awareness of nonverbal communication and develop visual literacy. Through the use of dynamic assessment, the project facilitated cultural awareness and improved intercultural, media, and cognitive competences. The FMA4L model provided by De Backer et al. (2016) showed the potential for schools to move from a narrow monolingual assessment to a broader multilingual assessment context. The paper presented classroom-based assessment as a possible way to implement a learning-oriented assessment approach where social interaction and evaluation, in the form of teacher feedback, were used in the classroom. Students’ well-being and self-efficacy were improved as a result. Thompson (2015) provided an alternative way to assess the language background of participants through the use of self-assessment in SLA research. Using the CEFR can-do statements as a way to measure learners’ linguistic data, she presented empirical data where self-assessment scores were collected from seventy-nine participants enrolled in an English language program. It was found that there was a moderate, significant correlation between the CEFR can-do statements and Michigan State University English Language Test (MSUELT). CEFR self-assessment materials were translated into participants’ L1. In situations where participants with different languages were involved, designing and administering tests in multiple languages were costly and impractical. She argued that the self-assessment technique could provide information about the diverse linguistic background of students and build multilingual language profiles of the learner. This is similar to projects illustrated by Jones and Saville (2009) that a European Language Portfolio was created based on CEFR, in order to gain information about the culturally and linguistically diverse population across Europe.

5. Qualitative Research

Drawing from the sociocultural perspective, a few qualitative studies were carried out to investigate multilingual assessment. Nelson-Flores (2018) conducted field notes, interviews, and collected classroom artifacts from ethnographic classroom case studies in the United States. He discussed the New York State Common Core State Standards (CCSS) policy and interpreted that the No Child Left Behind exams are “monoglossic”. He proposed a “water cycle-based” assessment, and interpreted CCSS as a way to promote dynamic assessment by using bilingual scaffolds. The ecological approach to assessment includes the use of ethnographic methods in linking assessments and classroom practices.

Laher and Cockcroft (2017) investigated multilingual assessment practices in the context of South Africa. They pointed out that the majority of the assessments in South Africa were delivered in English and were psychological assessments which cause problems in social equality. They argued that psychological tests were one-sided and were in conflict with the culturally, ethnically, socioeconomically, and linguistically diverse South African population. They believed that while psychological assessments were necessary, as it provided applicability in a large context, qualitative and flexible approaches should also be supplemented to obtain a comprehensive picture of the background and behavior of the diverse multicultural population.
IV. CONCLUSION

Assessment development needs to catch up with the fast-paced and changing multi-cultural, multi-dialectical, and multi-ethnic society. Ideological shift, multiple measures, and new approaches to assessing language competence should be advocated, promoted, and implemented in international contexts. Assessment instruments such as portfolio, self-assessment, peer-assessment, can-do statements, and reflective essays need to be incorporated. A review of the literature indicated that although studies on multilingual assessment are emerging and various conceptualization and approaches are proposed, it has remained largely unexplored. Studies with concrete examples of rubrics, test developments, implementations of various learning-oriented assessment in classrooms, and standardized contexts are greatly needed. Efforts to close the gaps between teaching, learning, and assessment need to be made. Additionally, most of the investigations centered around emergent bilinguals and K-12 multilinguals. Multilingual assessment in the higher education context is greatly needed. Despite the plethora of difficulties and challenges ahead, it is evident that multilingual assessment will improve language learning and teaching, support social equality, and enhance assessment practices.

REFERENCES

[1] Andrade, C., Kretschmer Jr, R. R., & Kretschmer, L. W. (1989). Two languages for all children: Expanding to low achievers and the handicapped. In Languages in elementary schools. International Education Series. American Forum, 177-203.
[2] Bialystok, E. & Senman, L. (2004) Executive Processes in Appearance-reality Tasks: The Role of Inhibition of Attention and Symbolic Representation. Child Development 75.2, 562-579.
[3] Bialystok, E., Craik, F.I.M. & Freedman, M. (2007) Bilingualism as a Protection against the Onset of Symptoms of Dementia. Neuropsychologia 45.2, 459-464.
[4] Blommaert, J., Collins, J., & Selsmexbouch, S. (2005). Spaces of multilingualism. Language & Communication 25.3, 197-216.
[5] Block, D. (2013). Moving beyond “lingualism”: multilingual embodiment and multimodality in SLA. The multilingual turn. London, UK: Routledge, 64-87.
[6] Canagarajah, S. (2011). Codemeshing in academic writing: Identifying teachable strategies of translanguaging. The Modern Language Journal 95.3, 401-417.
[7] Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. Applied linguistics 1.1, 1-47.
[8] De Backer, E., Van Avermaet, P., & Selsmexbouch, S. (2016). Schools as laboratories for exploring multilingual assessment policies and practices. Language and Education 31.3, 217-230.
[9] Diamond, A. (2006). The early development of executive functions. In Lifespan cognition: Mechanisms of change. Oxford, UK: Oxford University Press, 70-95.
[10] Flores, G. S. (2016). Assessing English language learners: theory and practice. London, UK: Routledge.
[11] Garcia, O. & L. Wei. (2014). Translanguaging: Language, bilingualism and education. New York: Palgrave Macmillan.
[12] Gajek, E. (2011). Learner-oriented teaching and assessment through digital video projects. Issues in promoting multilingualism teaching-learning-assessment. Foundation for the Development of the Education System, 288-307.
[13] Gorter, D. (2015). Multilingual interaction and minority languages: Proficiency and language practices in education and society. Language Teaching 48.1, 82-98.
[14] Gorter, D., & Cenoz, J. (2016). Language education policy and multilingual assessment. Language and Education 31.3, 231-248.
[15] Heugh, K., Prinsloo, C., Makgamatha, M., Diedericks, G., & Winnraa, L. (2016). Multilingualism(s) and system-wide assessment: a southern perspective. Language and Education 31.3, 197-216.
[16] House, J. (2003). English as a lingua franca: A threat to multilingualism? Journal of sociolinguistics 7.4, 556-578.
[17] Jones, N., & Saville, N. (2009). European Language Policy: Assessment, Learning, and the CEFR. Annual Review of Applied Linguistics 29, 51-63.
[18] Kramsch, C. (2005). Post 9/11: Foreign languages between knowledge and power. Applied Linguistics 26.4, 545-567.
[19] Laher, S., & Cockcroft, K. (2017). Moving from culturally biased to culturally responsive assessment practices in low-resource, multicultural settings. Professional Psychology: Research and Practice 48.2, 115-121. doi:10.1037/pro0000102
[20] Leung, C. (2004). Developing formative teacher assessment: Knowledge, practice, and change. Language Assessment Quarterly: An International Journal 1.1, 19-41.
[21] Martínez-Flor, A., Usó-Juan, E., & Soler, E. A. (2006). Towards acquiring communicative competence through speaking. In Current trends in the development and teaching of the four language skills. Berlin, Germany: Mouton de Gruyter, 139-157.
[22] Mathew, R. (2008). Assessment in multilingual societies. In Encyclopedia of language and education. Boston, MA: Springer, 2164-2181.
[23] McNamara, T., & Roeve, C. (2006). Language testing: The social dimension. Malden, MA, and Oxford, England: Blackwell.
[24] McNamara, T. (2015). Negotiating the complexities of multilingual assessment. Discussant for the Invited colloquium on AAAL Conference 2014.
[25] McNamara, T. (2018). A Challenge for Language Testing: The Assessment of English as a Lingua Franca. Keynote Speaker at LTRC Conference 2018, Ames, Iowa.
[26] Menken, K., & Shohamy, E. (2015). Invited colloquium on negotiating the complexities of multilingual assessment, AAAL Conference 2014. Language Teaching 48.3, 421-425.
[27] Murray, N. (2016). Standards of English in higher education: Issues, challenges and strategies. Cambridge, UK: Cambridge University Press.

© 2020 ACADEMY PUBLICATION
Yangting Wang earned her Ph.D. degree in Cultural, Literacy, and Language from the University of Texas at San Antonio (UTSA), Texas, United States. She currently serves as a research consultant for UTSA and she will be working as a Research Associate at Louisiana State University in September, 2020. She has published articles in peer-reviewed language journals such as International Journal of Bilingualism, CALICO journal, and Applied Linguistics.

Shikun Li earned her Ph.D. in cultural, literacy, and language from the University of Texas at San Antonio. Her research interests are in teacher cognition, individual differences, and curriculum innovation. Her recent research projects include learner need oriented differentiated instruction design, and the EFL learners’ emotion responses towards English acquisition during COVID 19 pandemic. Dr. Shikun Li is a lecturer in the School of Foreign Languages and Cultures at the Nanjing Normal University.