CLD Corner: Non-Word Repetition Tasks as an Alternative Assessment Form for CLD Populations

By: Mary Bauman-Forkner, MS, CCC-SLP, CLD Committee Member

Regulations for the Individuals with Disabilities Education Act (IDEA) Part B outline best practices for service delivery to culturally and linguistically diverse (CLD) populations, including assessment and evaluation materials that should not be racially or culturally discriminatory. An addition to the final regulations on evaluation procedures (§300.304) requires that assessment and other evaluation materials are administered "in the form most likely to yield accurate information on what the child knows and can do academically, developmentally, and functionally." Professionals with training in assessment of individuals from culturally and linguistically diverse backgrounds understand that the "form" of assessment of CLD students will differ from students in mainstream culture. The language in this regulation allows alternative administration methods and assessment tasks that do not require standard testing procedures when not appropriate (ASHA, n.d.).

Norm-referenced tests often fail to accurately differentiate between typically developing bilingual children and bilingual children with specific language impairment (SLI), as these measures typically include norms based on a monolingual population. Even for tests in which bilingual individuals may be included in the normative sample, the test generally assesses only one of the languages known by the examinee. When the examinee's second language (L2) is assessed (e.g., assessing English of a non-native English speaker), lower scores may be interpreted as being a result of the second language not having yet been mastered, leading to a potential underdiagnosis. For some more widely spoken languages in the United States, such as Spanish, clinicians are able to assess a child's language skills using a standardized, norm-referenced test in the child's primary language, or first language (L1). However, a student who obtains below average scores may be exhibiting lower performance due to limited experience or formal education in the primary/first language. Particularly for a child in English-only instruction or an early-exit bilingual program, underdeveloped academic vocabulary in the L1 may lead to potential overdiagnosis as the child is being compared to monolingual peers or those with formal education in the child's L1; results may appear to show an impairment when, in reality, the child may be demonstrating evidence of language loss or lack of development in L1 (a false positive). A child with this profile and standard scores below the average range in L1 may be recommended for speech-language therapy but may, in fact, have typically developing (bilingual) language skills.

Evidence-based practice encourages clinicians to use alternative assessment tools to reduce the bias inherent in standardized, norm-referenced assessments. Alternative methods for non-biased (or reduced bias) assessment include the use of processing-dependent measures and dynamic assessment measures (Laing & Kamhi, 2003). Some alternative forms of assessment practices include dynamic assessment (the most widely known being the test-teach-retest model), narrative language analyses, conceptual scoring, and analysis of evaluation tools.

For years, there has been research among different cultural and linguistic groups worldwide providing evidence that another form of alternative assessment—non-word repetition (NWR) tasks— has the potential to inform differential diagnosis of children with specific language impairment from those with typically developing language. With non-word repetition tasks, children are asked to repeat a series of nonsense syllables, ranging in length from one to five syllables. This task then assesses their ability to perceive, store, recall, and reproduce phonological sequences. Because these phonological memory skills are also key in word and morpheme learning, difficulty in accurately performing this task would suggest an underlying deficit in the language-learning mechanism. Utilizing this form of assessment as one measure of language ability is also less biased

against children from minority backgrounds in comparison with many standardized, knowledge-dependent measures (Dollaghan & Campbell, 1998). Processing-dependent measures, such as non-word repetition tasks, better distinguish between children's low performance that is a result of an underlying language-processing deficit rather than due to a difference in overall language experiences.

Earlier research from a variety of sources has provided evidence that this form of alternative assessment can reduce the effect of limited experience and knowledge of language, such as those disparities or discrepancies that may exist between groups of low versus mid-high socioeconomic status (SES). Processing-dependent measures such as NWR offer reduced bias against children from minority backgrounds compared to knowledge-dependent measures. Children with language impairments demonstrated notable deficits in non-word repetition that could not be attributed to differences in their language knowledge (Dollaghan & Campbell, 1998; Laing & Kamhi, 2003); Chiat & Polišenská, 2016).

More recently, studies also have looked at the application of non-word repetition tasks to differentially diagnose children from non-mainstream English and more diverse linguistic backgrounds than the more common Spanish-English dichotomy. One study by H. J. Lee, Kim, and Yim (2013) compared typically developing Korean-English preschoolers to monolingual Korean preschoolers for both vocabulary tests and non-word repetition tasks. On vocabulary tests, the bilingual children exhibited overall scores lower than their monolingual peers; however, both groups showed similar performance on non-word repetition tasks in Korean. Results suggest that assessing the phonological memory of the children reduces the bias that is otherwise seen in tests measuring lexical knowledge.

In many of the studies, when tests were administered looking first at language ability (e.g., one-word picture vocabulary tests, receptive vocabulary tests), researchers did observe significant differences in vocabulary performance between monolingual and bilingual children and also between groups of children from mid-high SES and low SES groups. We would expect to see this difference as the effects of limited experience and language knowledge for bilingual groups in their L2 and lower SES groups come into play due to bias inherent with the test instrument and items. Due to limited vocabulary and language experiences, standard scores on language assessments of these groups of bilingual and lower SES children were lower, suggesting that their prior experiences rather than language-learning ability were actually assessed. However, in these same studies, when researchers administered the non-word repetition tasks to both monolingual and bilingual groups, no group differences were found when the tasks were controlled for various factors like child age, length and complexity of non-words, phonotactic probability (the frequency with which a phonological segment or sequence of segments can occur in a given position in a word), etc. (Chiat & Polišenská, 2016; Santos & Ferré, 2018).

Considerations for Selecting and Administering an Effective NWR Task

Phonotactic Constraints of Language/'Word-Like' Quality of Stimuli

Sorenson Duncan and Johanne Paradis (2016) investigated the performance of English language learners (ELLs) with limited English exposure on an English non-word repetition task from the *Comprehensive Test of Phonological Processing (CTOPP)*. Due to differences in syllable structure for different L1s, children from two main language groups performed differently on the NWR task. In their study, children in the South Asian (Hindi, Punjabi, or Urdu) group had higher performance than children in the Chinese (Cantonese or Mandarin) group on the English NWR task. Differences in syllable structure and coda consonants (i.e., syllable final consonants) were significant because the vast majority of words in Chinese languages consist of one open syllable. For Mandarin, the only consonants that occur in the coda position are [n] and [ŋ]. In Cantonese, only unreleased voiceless stops are permitted in coda positions: [p̚] [t̚] [k̚]. In contrast, South Asian languages such as Hindi allow up to 33 consonants in coda position in addition to consonant clusters. Languages with more permissible consonants in the coda (or final) position more closely

align with the syllabic structure of English; therefore, ELLs with these L1s will likely perform better on English NWR tasks. In other words, non-words that are not word-like in the L1 are more difficult for ELL children who are relatively new to English (i.e., within the first three years of their English exposure).

Similar findings were revealed in other languages as well; bilingual children perform better on non-words that conform to the phonotactics of their L1 (the language with greater exposure) than those constructed on the basis of L2 phonotactics (the language with less exposure) (Chiat and Polišenská, 2016; Summers, et. al, 2010; Windsor, et. al, 2010).

To develop a non-word repetition task for use with native Vietnamese speakers, Pham, et. al. (2018) created stimuli with syllables that were phonologically possible in all Vietnamese dialects in addition to being highly word-like. Their stimuli included phonemes with high phonotactic probability—a high probability of those syllable segments occurring in the child's language. Phonemes chosen where those consistently found in the phonemic inventories of young Vietnamese-speaking children and children who speak different dialects of Vietnamese. Because of differences in productions across dialects of the language, the NWR task they developed included the four consonants that would not be influenced by vowel contexts—/pˋ, kˋ, m, ŋ/—and therefore fit the constraints regardless of dialect spoken. As Vietnamese is a tonal language, the tones selected for their NWR task included three of the six tones of Vietnamese (rising, sắc; level, ngang; and falling, huyền), as each of those three are found across dialects.

Additionally, the stimuli rated as being highly word-like were more easily and more accurately repeated by children across age groups compared with non-words with low phonotactic probability. Although word-like, the stimuli were carefully developed to exclude real words in Vietnamese in order to reduce the effects of lexical knowledge on performance.

Clinical Implications
Administering a non-word repetition task in English to an ELL with an L1 that is highly dissimilar to the phonological system of English could result in reduced performance and resulting overdiagnosis of the child. A natural tendency for ELLs is to apply phonological rules or constraints from their L1 to their L2 as they are learning the language. Therefore, considerations must be made of the similarities and differences between the primary language(s) used by the child and the phonological structure of the syllables used in the NWR task. Non-word repetition stimuli must adhere to the phonotactic constraints of a language in order for a child to demonstrate his or her true ability to process the targets. In other words, stimuli can contain only admissible phonemes, phoneme sequences, and suprasegmental patterns for the target language.

Non-word repetition tasks also may contain more bias when used with ELLs with limited English exposure compared with monolingual children from diverse backgrounds or simultaneous bilingual language learners. Therefore, English NWR tasks may be more appropriate as a reduced bias form of assessment for children from diverse SES status, speakers of non-mainstream dialects of English, or sequential (or near sequential) bilingual language learners with a longer exposure to the English language. To avoid overidentification of language impairment for sequential ELLs with more limited exposure to English, clinicians may want to locate and administer NWR tasks in the child's L1 if possible. Another strategy could be to develop ELL-specific strategies for interpreting NWR performance, such as developing ELL norms for NWR (Paradis et al, 2013) or alternative scoring that takes into account an ELL's L1 background.

Syllable Complexity of Non-Words
Two non-word tests that have been used in many studies investigating the diagnosis of specific language impairment (SLI) and working memory in English-speaking children are the Children's Nonword Repetition Test (CNRep) developed by Gathercole and Baddeley (1996) and the Non-Word Repetition Test (NRT) developed by Dollaghan and Campbell (1998). The design of the CNRep is comprised of 40 total words, some of which contain consonant clusters; none of the 16

words on the NRT include consonant clusters. In a study created by Santos & Ferré (2018), the authors developed and analyzed performance for a language-independent list of non-words with phonemes and syllable structures that occurred in a large number of world languages. By isolating syllable length as a factor to reduce the effects of working memory limits, they discovered that both monolingual and bilingual children with SLI differed from both groups of typically developing children when two consonant clusters are present within non-words. Their results suggested that phonological structure of the non-words is more important than length for impairment detection.

Clinical Implications
When administering a non-word repetition task, a clinician might consider not only the phonological constraints of the language but also the complexity of the syllables or segments in the target words. For bilingual students whose L1 is characterized by more frequent multisyllabic words, they may perform better on repetition tasks with longer three- and four-syllable non-words; however, the differentiating factor between typically developing bilingual children and those with SLI may lie in the ability to correctly repeat syllables with higher phonological complexity. Children with an underlying impairment in their language-learning mechanism would likely struggle with the productions of non-words that include consonant clusters or sequences within some of the syllables (again, as long as they do not violate the phonological constraints of the child's language).

Length of Non-Words and Connection with Languages Assessed
Across tests of non-word repetition, repeated results show that as word length increases, accuracy tends to decrease. Dollaghan and Campbell (1998) argued that children with SLI repeated non-words less accurately than typically developing children especially when the non-words have three and more syllables. Pham and colleagues (2018) cited several studies (Dispaldro et al., 2013; Dollaghan & Campbell, 1998; Girbau & Schwartz, 2007; Munson et al., 2005) that also found word length affects the performance of children with language impairment more than their typically developing peers.

Generally, children with language impairment tend to show deficits beginning with longer three- and four-syllable words. However, for languages that have a greater number of multisyllabic words (e.g., Spanish), some bilingual (Spanish-English) children with language impairment are better able to produce a few of the longer three-syllable words on English NWR tasks, reducing the sensitivity of the task to identify an underlying impairment. Gutiérrez-Clellen & Simon-Cereijido (2010) investigated the use of an English Non-Word Repetition Task (ENWRT) and a Spanish Non-Word Repetition Task (SNWRT) for both English- and Spanish-dominant bilingual children with and without language impairment. They discovered that administering the task in just one language was not sufficient to rule out LI because of the varying levels of language skills the children demonstrated across the two languages. When tasks were administered in both languages, their results showed a greater accuracy of classification of language impairment.

Clinical Implications
Clinicians must keep in mind performance differences of linguistically diverse groups based on the length of the words repeated, the characteristics of the language spoken by the child, and the language(s) assessed with the NWR task. When assessing a child in his or her dominant language, also take note of results from related resources or local norms collected. To aid in differentiation for bilingual or multilingual groups, comparing performance to local norms can help the clinician determine at which syllable length typically developing children may or may not experience difficulty. Clinicians also must consider that a child's performance on tasks in each language may differ and should attempt administration in both or all languages spoken by the child.

Scoring Method (percentage of phonemes/consonants correct vs. word-level scoring)
Several of the studies concluded that scoring of an entire non-word production as correct or incorrect was better at distinguishing the groups with impairment compared with scoring based on consonants or vowels correct (Guiberson & Rodriguez, 2013; Le Clercq, et. al, 2017; Pham, et. al, 2018). In clinical practice, this type of scoring is beneficial to speech-language pathologists as not only is the

scoring much faster but it also can be completed online after each of the student's productions (rather than scoring from an audio recording of the child's productions). One disadvantage to scoring the entire word rather than calculating percentages of phonemes correct is that word-level analysis limits the ability to determine the types of errors made and in which phonetic contexts they occurred.

Clinical Implications
At times, we may utilize the assistance of an interpreter to administer a non-word repetition task in order to elicit non-word productions that fit within the constraints of the child's L1. As an interpreter would typically not have much experience with transcribing or analyzing word productions, a simpler determination of whether or not the overall word was correctly repeated will assist the interpreter and assessor in scoring.

Benefits and Limitations of Using Non-Word Repetition Tasks

Because of the multiple factors involved, we don't yet know how well non-word repetition tasks accurately identify children with language impairment. Despite this limitation, non-word repetition tasks are unique as part of the overall assessment picture in that they focus more on linguistic processing than on accumulated linguistic knowledge. These tasks isolate the language processing skills of a child apart from his or her semantic knowledge and experience. Being able to recognize deficits in one element of a child's language-processing system may be an important clinical marker that aids in reaching appropriate conclusions about a child's language ability and may offer converging or diverging evidence in the diagnostic decision-making process.

References

American Speech, Language, Hearing Association. (n.d.) IDEA Part B Issue Brief: Culturally and Linguistically Diverse Students. Retrieved from https://www.asha.org/Advocacy/federal/idea/IDEA-Part-B-Issue-Brief-Culturally-and-Linguistically-Diverse-Students/

Chiat, S. & Polišenská, K. (2016). A framework for crosslinguistic nonword repetition tests: Effects of bilingualism and socioeconomic status on children's performance. *Journal of Speech, Language, and Hearing Research, 59*, 1179-1189. https://doi.org/10.1044/2016_jslhr-l-15-0293

Dollaghan, C. & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1136-1146. https://doi.org/10.1044/jslhr.4105.1136

Gathercole S. E., & Baddeley A. D. (1996). *The Children's Test of Nonword Repetition*. Pearson.

Guiberson, M. & Rodriguez, B. L. (2013). Classification accuracy of nonword repetition when used with preschool-age Spanish-speaking children. *Language, Speech, and Hearing Services in Schools, 44,* 121-132. https://doi.org/10.1044/0161-1461(2012/12-0009)

Gutiérrez-Clellen, V. F. & Simon-Cereijido, G. (2010). Using nonword repetition tasks for the identification of language impairment in Spanish-English-speaking children: Does the language of assessment matter? *Learning Disabilities Research & Practice, 25*(1), 48-58. https://doi.org/10.1111/j.1540-5826.2009.00300.x

Laing, S. P. & Kamhi, H. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech, and Hearing Services in Schools, 34,* 44-55. https://doi.org/10.1044/0161-1461(2003/005)

Le Clercq, C. M. P., van der Schroeff, M. P., Rispens, J. E., Ruytjens, L., Goedegebure, A., van Ingen, G., & Frankena, M. C. (2017). Shortened nonword repetition task (NWR-S): A simple, quick, and less expensive outcome to identify children with combined specific language and reading impairment. *Journal of Speech, Language, and Hearing Research, 60*, 2241-2248. https://doi.org/10.1044/2017_jslhr-l-16-0060

Lee, H. J., Kim, Y. T., & Yim, D. (2013). Non-word repetition performance in Korean-English bilingual children. *International Journal of Speech-Language Pathology, 15*, 375-382. https://doi.org/10.3109/17549507.2012.752866

Paradis, J., Schneider, P., & Sorenson Duncan, T. (2013). Discriminating children with language impairment among English-language learners from diverse first-language backgrounds. *Journal of Speech, Language, and Hearing Research, 56*, 971–981. https://doi.org/10.1044/1092-4388(2012/12-0050)

Pham, G., Ebert, K. D., Dinh, K. T., & Dam, Q. (2018). Nonword repetition stimuli for Vietnamese-speaking children. *Behavior Research Methods, 50*, 1311-1326. https://doi.org/10.3758/s13428-018-1049-0

Santos, C. D. & Ferré, S. (2018). A nonword repetition task to assess bilingual children's phonology. *Language Acquisition, 25*(1), 58-71. https://doi.org/10.1080/10489223.2016.1243692

Sorenson Duncan, T. & Paradis, J. (2016). English language learners' nonword repetition performance: The influence of age, L2 vocabulary size, length of L2 exposure, and L1 Phonology. *Journal of Speech, Language, and Hearing Research, 59*, 39-48. https://doi.org/10.1044/2015_jslhr-l-14-0020

Summers, C., Bohman, T. M., Gillam, R. B., Peña, E. D., & Bedore, L. M. (2010). Bilingual performance on nonword repetition in Spanish and English. *International Journal of Language & Communication Disorders, 45* (4), 480-493. https://doi.org/10.3109/13682820903198058

Windsor, J., Kohnert, K., Lobitz, K. F., Pham, G. T. (2010). Cross-language nonword repetition by bilingual and monolingual children. *American Journal of Speech-Language Pathology, 19*, 298-310. https://doi.org/10.1044/1058-0360(2010/09-0064)