Running Head: NWRT and Vocabulary

Using Nonword Repetition in Vocabulary Assessment

Jill R. Hoover, Doctoral Student

Child Language Doctoral Program, University of Kansas

Holly L. Storkel, Associate Professor

Department of Speech-Language-Hearing Sciences and Disorders, University of Kansas

*This project was supported by NIH grant DC000052, DC004781, DC006545.

Key Words: vocabulary, assessment, nonword repetition, test bias, word learning

Address for Correspondence: Jill R. Hoover, M.A., Child Language Doctoral Program, Dole Center, 1000 Sunnyside Ave., Room 3031 Lawrence, KS 66045-7555, jrhoover@ku.edu, www.ku.edu/~wrdlrng.

Abstract

Standardized vocabulary tests have been criticized for their cultural/experiential biases and insensitivity to word learning differences. This review discusses the utility of supplementing the diagnostic process with a processing-based measure, such as a nonword repetition task. Nonword repetition tasks have been heralded as a more sensitive indicator of individual differences in vocabulary/word learning. Evidence on the relationship between vocabulary and nonword repetition is discussed along with a review of the currently available tests of nonword repetition. Suggestions for constructing a nonword repetition task specific to the needs of individual clinicians are offered.

Typically developing children acquire new words rapidly. The production of a few words around the age of 12-months rapidly takes off around 18-months, growing exponentially from that point forward. In fact, the number of words known by a child increases from approximately 3,500 in kindergarten to nearly 6,000 at the end of the second grade (Beimiller & Slonin, 2001). As typical language learners develop and expand their vocabulary with relative ease, children with language impairments generally do not. These children are typically late to acquire their first words and can continue to experience a slower rate of growth as they often have difficulties learning new words relative to age-matched peers (e.g., Leonard, 1998; Rice, Burh, & Nemeth, 1990). Consequently, one element to consider in the diagnosis of language impairment may be word learning ability. Given the contribution of vocabulary acquisition/oral language to later academic skills (NICHD Early Child Care Research Network, 2005), early assessment of vocabulary is critical in the early identification of language learning differences. However standardized vocabulary tests have been criticized for their experiential/cultural biases and insensitivity to language impairment (Campbell, Dollaghan, Needleman, & Janosky, 1997; Dollaghan & Campbell, 1998). These criticisms suggest that there is a need for including alternative measures of vocabulary that are sensitive to word learning skills, rather than experience or culture.

DRAWBACKS TO STANDARDIZED VOCABULARY TESTS

Vocabulary in children is most commonly assessed through the administration of standardized tests. Such tests provide clinicians with a way to compare a child's vocabulary to others of the same chronological age. A complete evaluation of vocabulary targets both the expressive and receptive domains through the administration of two separate tests. The two types of tests require children to either name a picture (i.e., expressive) or select one from a set of four

(i.e., receptive), allowing clinicians to obtain separate standardized scores for each domain. Furthermore, some standardized vocabulary tests provide a comparison of receptive and expressive standard scores which can be used in diagnosing word finding deficits. Specifically, a standard score on the receptive vocabulary test that is significantly higher than that on the expressive test can be taken as an indicator of a word finding deficit.

Although standardized vocabulary tests are important to the diagnostic process, they are not without criticism (e.g., Campbell et al., 1997; Gray, Plante, Vance, & Henrichsen, 1999; Rodekohr & Haynes, 2001). Campbell et al. (1997) argue that performance on these tests draws heavily upon knowledge and world experiences, rather than underlying processes responsible for acquiring new words. A specific item on a standardized vocabulary test that could be sensitive to experience might include one related to geographically-specific events (e.g., tornado). Experience with such events contributes to a child's general knowledge base of specific vocabulary items. Standardized vocabulary tests have also been criticized for reflecting cultural biases. Specifically, scores have been found to be less accurate for children from varying cultural backgrounds (Campbell, et al., 1997; Rodekohr & Haynes, 2001). Obtaining a norm-referenced score from a standardized test allows for the comparison between a target child and his or her peers. However, clinicians should bear in mind that performance on such a test heavily indexes experiences. Thus, in order to avoid misdiagnosis on the basis of experiential/cultural differences, clinicians ought to consider supplementing, but not replacing, knowledge-based measures with those that place greater demands on underlying linguistic processes.

Not only have standardized tests been criticized for their reliance on prior knowledge/cultural experiences, but also for their inability to correctly classify children with language impairment. Gray et al. (1999) evaluated the accuracy of diagnosing preschool

language impairments using standardized vocabulary tests in children between the ages of four- and five-years. While children with language impairments as a group scored significantly lower than their peers with typical language, their individual scores still fell within the normal range of development. Dollaghan and Campbell (1998) found similar results in older children between the ages of six- and nine-years. Although differences in scores between children with varying language ability are detected by standardized tests, such scores may not always be sensitive enough to identify the word learning deficits of children with language impairment. Thus, performance on standardized tests should be interpreted with caution if used as the primary indicator of word learning deficits during the screening process.

ALTERNATIVE DIAGNOSTIC TOOLS

One possible solution that has been offered to counteract the drawbacks of standardized vocabulary tests is to supplement them with measures that rely primarily upon linguistic processing rather than prior knowledge/experience (processing-dependent measures). Processing-dependent measures utilize items that are either novel (i.e., nonwords) or equally familiar to all children, thus relying more on linguistic processing than prior knowledge/experience (Campbell et al., 1997). One such task is the nonword repetition task (NWRT). In a NWRT, children hear a list of novel words that they are required to repeat one at a time. Novel words are constructed to resemble words that could be possible in the target language. Children's responses are phonetically transcribed and given two scores: whole-word score and proportion of correct sounds. The former is scored as correct only if the child repeats the novel word exactly as presented. The latter score allows the child to receive partial credit for recalling one or more sounds in the target word. Nonword repetition has been regarded as a processing-dependent measure because it requires children to recognize and produce unfamiliar

phoneme sequences as opposed to retrieving known words from their lexicon. Supplementing standardized vocabulary tests with a NWRT comes from a theory suggesting the presence of a link between phonological memory and vocabulary/word learning in preschool-age children (Gathercole & Baddeley, 1989). Baddeley (1986) suggests that phonological memory, one part of working memory, is responsible for recognizing, holding, and producing novel phonological information, such as the nonsense words in a NWRT. Similar to nonword repetition, one aspect of learning a word for the first time also requires a child to recognize, hold, and produce novel phonological information. Although nonword repetition does not involve a semantic (i.e., meaning) component, the phonological component of nonword repetition and learning a new word parallel one another. Nonword repetition has therefore been regarded as a supplemental task that may be used to evaluate the phonological aspects of word learning.

Using nonword repetition to supplement knowledge/based measures has also been supported by a correlation between performance on nonword repetition and vocabulary/word learning in children (Gathercole & Baddeley, 1989; Gathercole & Baddeley, 1990b; Gathercole, Hitch, Service, & Martin, 1997). Specifically, children with larger vocabularies are often better at repeating lists of novel words (Bowey, 2001; Gathercole, Willis, Emslie, & Baddeley, 1992; Roy & Chiat, 2004). As children's vocabularies increase, the relationship becomes even stronger suggesting that performance on nonword repetition is facilitated by a growing vocabulary with the strongest relationship demonstrated during the preschool and early school-age years (Gathercole & Baddeley, 1989; Gathercole & Adams, 1994). Finally, children with high, rather than low, scores on nonword repetition and standardized vocabulary tests are better able to learn and later recall novel names (Gathercole, & Baddeley, 1990b; Gathercole, et al., 1997). However, recent evidence suggests that additional complex working memory measures might be

more sensitive than nonword repetition in predicting later language ability (Gathercole, Tiffany, Briscoe, Thorn, & The ALSPAC team, 2005). Thus, using performance on a NWRT to predict language ability beyond the preschool years should be interpreted with caution.

A significant difference in performance on knowledge-, but not processing-dependent, measures in children from minority and majority groups in the United States has also been documented (Campbell et al., 1997). Since word learning deficits can be one aspect of language impairment, performance on nonword repetition might be more sensitive in identifying children with such characteristics than standardized tests (Dollaghan & Campbell, 1998; Edwards & Lahey, 1998; Gathercole & Baddeley, 1990a; but see Bishop, Adams, & Norbury, 2006; Ellis Weismer et al., 2000). Therefore, a NWRT might be used as a supplemental task that is free from cultural biases and more sensitive in detecting the phonological aspects of word learning deficits (Bishop, et al., 2006; Campbell et al., 1997; Edwards & Lahey, 1998; Ellis Weismer et al., 2000; Gathercole & Baddeley, 1990a; Rodekohr & Haynes, 2001).

ADMINISTERING A NWRT

Given the shortcomings of standardized vocabulary tests, supplementing the diagnostic session with a NWRT might be a promising option for clinicians. Administering a NWRT is straightforward and takes minimal time. Clinicians can use a NWRT that is either commercially available or has been published in a research report. Clinicians will need to consider their clinical needs when deciding which NWRT is most appropriate for their population.

The Children's Test of Nonword Repetition (CNRep) is a standardized test of nonword repetition for children between the ages of four- and eight-years (Gathercole & Baddeley, 1996). This test consists of 40 nonwords varying in length from two to five syllables, takes approximately 4 minutes to administer, and uses the whole-word scoring method. All nonwords

have been pre-recorded by a female speaker of British English. The Comprehensive Test of Phonological Processing (CTOPP) is a standardized test that includes nonword repetition as one of many subtests (Wagner, Torgesen & Roshotte, 1999). This NWRT includes fewer items than the CNRep (18 compared to 40) but can be administered to a wider age-range of individuals (5- to 24-years as opposed to 4- to 8-years). Nonwords vary in length from 3 to 15 sounds and are spoken by a speaker of American English. The scoring method on this task is the same as the CNRep (i.e., whole-word score).

If the clinical population of interest does not match the age-range appropriate for standardized tests of nonword repetition, clinicians can refer to tasks that have been used in research where nonword lists and normative data have been published. Typically, published research reports provide normative data on both scoring methods (i.e., whole-word and proportion of correct sounds). Mean scores and standard deviations are usually available for various ages and language abilities and can be used to benchmark a child's performance relative to others his or her age. Clinicians interested in administering a NWRT to preschool children could refer to Roy and Chiat (2004), whereas those interested in the school-age population should refer to Dollaghan and Campbell (1998) or Ellis Weismer et al. (2000). While the normative data in Roy and Chiat (2004) pertains only to typically developing children, that in Dollaghan and Campbell (1998) pertains to children with and without language impairments. Ellis Weismer et al. (2000) present normative data for groups of children differing in nonverbal cognition and language ability. Clinicians interested in administering the NWRT to students in the secondary grades and older can refer to Gupta (2003) where normative data are provided for individuals with normal language between the ages of 18- and 26-years.

Clinicians can calculate a standardized score, namely a z-score, to determine where a child's score falls relative to others his or her age by using means and standard deviations from published research reports. A z-score is used to determine how far an individual score lies from the mean of a larger sample. Z-scores are calculated by subtracting the child's score from the mean of the larger published sample and dividing that value by the standard deviation of the larger sample. Z-scores can be positive or negative and reflect how far above or below the mean a score deviates. For example, a child with a z-score of +1.00 has a score that is one standard deviation above the mean, whereas a z-score of -1.00 reflects a score that is one standard deviation below the mean. Typically, scores between –1.00 and +1.00 are considered to be age-appropriate.

CONSTRUCTING A NWRT

Not all population characteristics (e.g., ages, clinical populations, regional dialect) are represented in the currently available tests of nonword repetition. When mean performance scores are not available for a specific age-range or population (e.g., adults with aphasia, children with delayed articulation, speakers of a specific dialect), clinicians might consider constructing a list of nonwords and collecting their own data in order to develop a set of norms. Using means and standard deviations from the sample of interest, z-scores can then be calculated for a given individual. This self-constructed task can then be used to supplement standardized tests in the diagnostic process.

Typically, nonwords that match the sound patterns of the target language are selected for a NWRT. In fact, it has been suggested that *word-likeness*, or how similar a novel word is to a real word, affects repetition (Dollaghan, Biber, & Campbell, 1995; Munson, Kurtz, & Windsor, 2005). Nonwords that are more word-like are repeated more easily than those that are less word-

like (Dollaghan et al., 1995; Munson et al., 2005) and therefore clinicians should consider this factor when selecting nonwords. Typically, NWRTs use less word-like nonwords.

Clinicians should also consider the age of acquisition of the sounds in the nonwords. This might be particularly useful in designing a NWRT for very young children or those with articulation errors. By constructing novel word stimuli using early-acquired sounds, the risk of poor performance due to articulation constraints is minimized. One disadvantage of standardized tests of nonword repetition is that several of the nonwords are composed of late-acquired sounds (e.g., "frescovent", Gathercole & Baddeley, 1996), making it difficult to determine the source of poor performance.

A final issue to consider when constructing a NWRT is the regional dialect of the individual to be tested. Gerken (1979) suggests that when a speaker's dialect is judged to differ from an adult listener, accuracy on language production tasks can be affected. The CNRep (Gathercole & Baddeley, 1996) features a speaker of British English, while American English is used on the CTOPP (Wagner et al., 1999). A unique feature of the CNRep is that with permission from the publishing company, clinicians may request to have a speaker of a given regional dialect prerecord nonwords from the test for use with speakers of that dialect.

CONCLUSION

The inclusion of a NWRT in the assessment process has many advantages over relying exclusively on standardized tests. Specifically, this task has been regarded as less-culturally biased and more sensitive to detecting word learning differences resulting from difficulties holding phonological information in memory. Therefore, NWRT has been heralded as an indicator of the process, rather than the product, of word learning (Campbell et al., 1997). Given its relationship with vocabulary ability and word learning (Bowey, 2001; Gathercole, &

Baddeley, 1990b; Gathercole et al., 1992; Gathercole et al., 1997), nonword repetition appears to be a robust supplement to standardized vocabulary tests; however, it should not be regarded as a primary means of evaluating word learning differences in children with language impairment. While some children with language impairment have poor performance on NWRT, others do not, thus nonword repetition fails to identify all children with language impairment (Bishop et al., 2006). Rather, it may be more sensitive in detecting differences in children who have difficulties with the phonological aspects of word learning. In addition to performance on the NWRT, clinicians should continue to consider performance on standardized vocabulary and language tests, parent-teacher reports, and observation of the child in his or her natural environment in making a final clinical diagnosis and identifying intervention goals.

References

Baddeley, A.D. (1986). *Working Memory.* Oxford: Oxford University Press.

Biemiller, A. & Slonin, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence from a common sequence of vocabulary acquisition. *Journal of Educational Psychology, 93,* 498-520.

Bishop, D.V.M., Adams, C.V., & Norbury, C.F. (2006). Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes, Brain, and Behavior, 5,* 158-169.

Bowey, J.A. (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Pscholinguistics, 22,* 441-469.

Campbell, T., Dollaghan, C., Needleman, H., & Janosky, J. (1997). Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech, Language, and Hearing Research, 40,* 519-525.

Dollaghan, C., Biber, M.E., & Campbell, T.F. (1995). Lexical influences on nonword repetition. *Applied Psycholinguistics, 16,* 211-222.

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1136-1146.

Edwards, J. & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics, 19,* 279-309.

Ellis Weismer, S., Tomblin, J.B., Zhang, X., Buckwalter, P., Gaura Chynoweth, J., Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 43,* 865-878.

Gathercole, S.E. & Adams, A.M. (1994). Children's phonological working memory: Contributions of long-term knowledge and rehearsal. *Journal of Memory and Language, 33,* 672-688.

Gathercole, S.E. & Baddeley, A.D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28,* 200-213.

Gathercole, S.E. & Baddeley, A.D. (1990a). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29,* 336-360.

Gathercole, S.E. & Baddeley, A.D. (1990b). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology, 81,* 439-454.

Gathercole, S.E. & Baddeley, A.D. (1996). *The Children's Test of Nonword Repetition.* London, U.K.: The Psychological Corporation.

Gathercole, S.E., Hitch, G.J., Service, E., & Martin, A.J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33,* 966-979.

Gathercole, S.E., Tiffany, C., Briscoe, J., Thorn, A., & The ALSPAC team. (2005). Developmental consequences of poor phonological short-term memory function in childhood: A longitudinal study. *Journal of Child Psychology and Psychiatry, 46,* 598-611.

Gathercole, S.E. Willis, C.S., Emslie, H. & Baddeley, A.D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28,* 887-898.

Gerken, K. (1979). The ability of listeners to report oral responses of Black and White children. *Language, Speech, and Hearing Services in the Schools, 10,* 35-46.

Gray, S. Plante, E., Vance, R., Henrichsen, M. (1999). The diagnostic accuracy of four vocabulary tests administered to preschool-age children. *Language, Speech, and Hearing Services in Schools, 30,* 196-206.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology, 56A,* 1213-1236.

Leonard, L.B., (1998). *Children with Specific Language Impairment.* Cambridge, MA: The MIT Press.

Munson, B., Kurtz, B.A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 48,* 1033-1047.

NICHD Early Child Care Research Network (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology, 41,* 428-442.

Rice, M.L., Buhr, J.C. & Nemeth, M. (1990). Fast mapping word-learning abilities of language-impaired preschoolers. *Journal of Speech and Hearing Disorders, 55,* 33-42.

Rodekohr, R.K. & Haynes, W.O. (2001). Differentiating dialect from disorder: A comparison of two processing tasks and a standardized language test. *Journal of Communication Disorders, 34,* 255-272.

Roy, P. & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2- to 4-year olds: Evidence from typically developing children. *Journal of Speech, Language, and Hearing Research, 47,* 223-234.

Wagner, R.K., Torgesen, J.K., & Rashotte, C.A., (1999). *Comprehensive Test of Phonological Processing*. Austin, TX: Pro-Ed