At the present time there is no measure of phonology in aphasia. In our lab, we have constructed a standardized assessment of phonologic processes in aphasia that will be a sensitive and specific measure of phonologic function in adults with acquired aphasia. This measure will be useful to determine an appropriate course of treatment, to differentiate between types of phonologic dysfunction and to be used as a valid outcome measure. In order to begin the development of this comprehensive test, a model of phonology in aphasia was selected (Nadeau, 2000), leading to the identification of four domains of phonology: concept representation, reading, perception, and repetition. Stimuli were created for each domain and subsequently tested in individuals with aphasia. Results of the full test battery have been submitted as a platform titled, "The development of a standardized assessment of phonology in aphasia." The purpose of this paper is to discuss the results only of the repetition domain. The perception domain is being submitted as a companion poster entitled, "The development of a standardized assessment of phonology in aphasia: Creating items to test perception."

## Methods

Item Bank Development

Item Response Theory (IRT) was used as the basis for item development. IRT is a statistical approach that measures responses at an item level based on the claim that the probability of a person's response to an item is the combined function of that person's ability and the difficulty level of the item (Bond & Fox, 2001). IRT methods calibrate item difficulty and subject ability on a linear scale.

The first step in employing this approach to our repetition test was to create the items that would later be tested with individuals with aphasia and normal controls. Six tasks were identified based on the literature and currently published tests of phonology such as the Comprehensive Test of Phonologic Processes (CTOPP; Wagner, Torgesen, Rashotte, 1999), real and nonword repetition, real and nonword blending, real and nonword parsing. Items were created for each task within a hierarchy of easiest to hardest based on psycholinguistic variables. The manipulated variables were number of syllables, clusters, and phonemes. Phonotactic probability (Vitevitch & Luce, 2004) and frequency (Kucera & Francis, 1982) were controlled within and across all categories. All real word items were nouns.

*Real Word and Nonword Repetition*. For this task, items were divided into groups of 1-3 syllables. In the 1-syllable (2-4 phonemes) and 2-syllable (5-7 phonemes) groups, half the items were without clusters and half with clusters. In the three-syllable words (8-10 phonemes), half had 1-2 clusters and half had 3 clusters.

*Real Words and Nonwords Parsing and Blending*. Items were divided into 6 groups based on the division of the word for parsing or blending: compound words, 2-syllable non-compound words, onset-rime, body-coda, and individual phonemes. Only 2 syllable words were allowed to have clusters. Syllable structure was controlled within each category. Phonotactic probability and frequency (Kucera & Francis, 1982) were controlled within and across all categories.

*General Procedures of Item Bank Development.* The MRC Psycholinguistic Database (Coltheart, 1981) was used to select real words and values for the psycholinguistic variables. The Probability Calculator was used to calculate all biphone probabilities (Vitevitch & Luce, 2004). Experts in the fields of speech-pathology and neuropsychology reviewed the domains and categories within each. Their suggestions were used to adjust items to better fit the proposed hierarchy and model of phonology. Items were recorded by a male speaker in an audiologic sound booth using a Marantz Digital Audio Recorder. The final items were then tested with individuals with aphasia.

## Data Collection

*Participants.* Thirty-seven individuals with aphasia were tested. Inclusion criteria were a single left hemisphere stroke at least 6 months prior to enrollment resulting in aphasia as determined by standardized testing. In addition, all participants were required to be premorbidly right handed, monolingual English speaking adults. Exclusion criteria included other chronic or neurological illnesses as determined by a neurologist or severe impairment in hearing. Information was gathered from each participant using several standardized measures as well as a hearing screening and informal interview. Table 1 displays average demographic data and standardized test scores.

*Testing Procedures.* Items were presented to participants over two external speakers. Presentation of the items was conducted on a Dell Lattitude X1 laptop using E-prime software with ISI of 8.0 seconds. The order of the four tasks was randomized as well as the order of stimuli within tasks. Participants wore a headset microphone to record responses. Examiners used a button box to indicate correct and incorrect participant responses. Button box data was recorded in E-prime.

*Scoring.* Responses were correct if they matched the pre-determined pronunciation of the item.

*Data Analysis.* Participant responses were analyzed using WINSTEPS Rasch analysis computer software (Bond & Fox, 2001; Linacre, 1994). Results of Rasch analysis were used to determine the item-level psychometric characteristics such as unidimensionality of construct, item difficulty, floor and ceiling effects, and internal consistency of responses. Criterion for infit mean square was ≤1.4 and z-score was ≤2.0.

# Results

## Item Bank Development

The reviewers' suggested the addition of easier parsing and blending items at the individual phoneme level in order to capture speakers of lower ability. The final item bank consisted of 113 repetition items.

## Item Analysis

Table 2 displays the results of each repetition task for percent items misfit, person separation reliability (strata), Cronbach's alpha, and floor and/or ceiling effect. In general, each of the repetition tasks demonstrated positive measurement qualities except for person separation and floor/ceiling effects. Notably, the real word repetition task did not capture speakers of higher ability (23% ceiling effect) and the real and nonword parsking tasks did not capture speakers of lower ability (27% and 43%, respectively). The number of misfitting items ranged from 0-2. The items for each construct demonstrated good point measure correlation as evidenced by Cronbach's alpha (.76 to .92). The average strata, or number of ability groups stratified by the test, was 2.26 with a notable minimum of .84 for the nonword blending task. Item map data and predicted hierarchies will be displayed and discussed on the poster.

## Discussion

As a repetition domain, the 113 items demonstrated acceptable psychometric properties based on Rasch analysis. Specifically, the low number of misfitting items indicates that the items appropriately measured person ability; that is, a predicted difficult item was responded to correctly by mostly higher ability speakers and a predicted easier item was responded to correctly by most speakers. All of the tasks except nonword parsing and blending separated participants into at least 2 groups based on ability level, suggesting this group of items can be used to delineate ability level of speakers with aphasia in regards to phonology. However, all tasks had a floor effect indicating a need for easier items to capture the lower ability level of speakers with aphasia. Additionally, the real word repetition task had a ceiling effect suggesting this task did not measure the highest ability level speakers and therefore needs more difficult items. The next step in completing the repetition domain will be creating a short version. In doing so, redundant items will be deleted and new items will be created to resolve the ceiling and floor effects to ultimately capture the range of ability levels in aphasia.

## References


Bond, T.G., & Fox, C.M. 2001. Applying the Rasch model: Fundamental measurement in the human sciences. Mahwah, NJ: Lawrence Erlbaum.

Coltheart, M. 1981. The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 33A, 497-505.

Francis, W. N. & Kucera, H. Frequency Analysis of English Usage: Lexicon and Grammar. Houghton-Mifflin, Boston, 1982.

Linacre, J.M. 1994. Constructing measurement with a many-facet Rasch model. In M. Wilson (Ed.), Objective measurement: Theory into practice (Vol. 2). Norwood, NJ: Ablex

Nadeau, S. 2000. Phonology. In S.E. Nadeau, L.J. Gonzalez-Rothi & B. Crosson (Eds). Aphasia and Language. (pp.40-81). New York: Guilford Press.

Vitevitch, M.S. & Luce, P.A. 2004. A web-based interface to calculate phonotactic probability for words and nonwords in English. Behavior Research Methods, Instruments, and Computers, 36, 481-487.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. 1999. Comprehensive test of phonological processing. Austin, TX: PRO-ED.

Table 1. Participant Demographics

| | **Age** | **Months Post Onset of Stroke** | **Education** | **WAB AQ** | **BNT (spontaneous correct)** |
|---|---|---|---|---|---|
| **Average (SD)** | 65.2 (10.6) | 59.8 (41.9) | 13.7 (2.9) | 80.0 (11.4) | 34.1 (13.0) |

Table 2. Summary of Item Analysis

| Construct | Task | # items misfit (total # items) | Person separation reliability (strata) | Cronbach's alpha | Floor and/or ceiling effect | Point Measure Correlation: Percent <.30 |
|---|---|---|---|---|---|---|
| REPETITION, PARSING, BLENDING N=113 items N=37 aphasics | Real word | 1 (18) | .78 (2.81) | .92 | Yes Ceiling = 23% Floor = 7% | 0% |
| | Nonword | 2 (18) | .82 (3.2) | .92 | Yes Floor = 17% | 0% |
| | Parsing real words | 1 (18) | .76 (2.73) | .90 | Yes Floor = 27% | 6% |
| | Parsing nonwords | 1 (15) | .29 (1.17) | .78 | Yes Floor = 43% | 7% |
| | Blending real words | 2 (18) | .78 (2.81) | .85 | Yes Floor = 17% | 11% |
| | Blending nonwords | No misfit items | .13 (.84) | .76 | Yes Floor = 50% | 0% |