Evaluating the use of *An Object and Action Naming Battery* for Spanish/English bilingual adults: An item and participant analysis

BACKGROUND

The U.S. has 28 million Spanish-speaking citizens, the majority of which are bilingual (U.S. Census, 2000). As a result, the incidence of bilingual aphasia is projected to grow to 45,000 new cases per year by 2050 (Paradis, 2001). Impairment in lexical retrieval is a hallmark of aphasia, and picture naming is widely used to examine lexical retrieval in normal and clinical populations. Naming of nouns and verbs should be evaluated, since noun and verb production can be differentially impaired (Druks, 2002).

There are currently no psychometrically validated naming tests for Spanish/English bilingual adults. A number of studies (Gollan et al., 2007; Kohnert et al., 1998; Roberts et al., 2002) have evaluated the appropriateness of the *The Boston Naming Test* (Kaplan, Goodglass, & Weintraub, 1983) on Spanish/English bilinguals. However, the authors concluded that the BNT is not appropriate for this population, largely due to its psychometric organization (items decrease in English frequency). To our knowledge, there are no normative data for action naming in Spanish/English bilingual adults.

Without normative data on appropriate materials, the ability to evaluate cross-linguistic differences in naming performance and the effect of aphasia on naming is limited in bilinguals with aphasia. Thus, the current study aims to collect normative data on, and evaluate the appropriateness of, *An Object and Action Naming Battery* (*An O&A Naming Battery:* Druks & Masterson, 2002) with 90 Spanish/English normal young bilingual adults. *An O&A Naming Battery* was chosen because it has a large number (N=262) and wide range of object and action stimuli. The specific aims are to determine:

1) which items have acceptable item difficulty and discrimination to distinguish naming proficiency levels among participants in Spanish and English.

2) which demographic, language use, and language self-rating data from participants best predict lexical retrieval abilities in English and Spanish.

METHOD

Participants. Seventy-five neurologically normal Spanish/English bilingual adults (*M*=22.29 (4.68) years) from Florida participated. (Data from 90 participants have been collected and would be included for the conference). (Young participants were chosen to evaluate the test's appropriateness). Future studies will include older adults. Participants were more proficient in English, used more English, and were more educated in English, although there was variability due to our attempt to recruit a range of participants. (See Table 1). Participants were born in 13 different countries and reported their Spanish being influenced by 18 different countries (some participants reported two countries), reflecting a highly diverse group.

Design. Participants were tested over two sessions, 5-7 days apart. During the first session, participants completed a language use questionnaire (adapted from Muñoz, Marquardt, & Copeland, 1999). Order of language testing (English or Spanish) was counterbalanced across participants.

Procedure. The pictures from *An O&A Battery* were shown on a 17" computer monitor with DirectRT software. Participants wore headphones with a microphone to record responses. Half of the participants named verbs first and half named nouns first. Participants named verbs in the present progressive form (e.g., eating/*comiendo*) (practice items were provided)**.** Each picture was presented until the participant named it and pushed the space bar to self-advance the picture.

Scoring and reliability. Scoring was conducted by bilingual research assistants. All appropriate responses were accepted as correct, including dialectical and lexical variations. (e.g., *hongo, champiñon, seta* are acceptable for "mushroom"). A number of sources were consulted, including 1) bilingual research assistants, 2) various dictionaries, and 3) faculty in Spanish Linguistics. Scoring reliability for 100% of responses was 98% in English and 95% in Spanish.

RESULTS

To be conservative in our item analysis, both classical item analysis and item response theory were conducted. First a classical item analysis was conducted. Table 2 shows that the noun and verb stimuli for English and Spanish have high test reliability (estimates exceeding 0.9). According to mean item difficulties for all items (Table 2), English items were easier to name than Spanish items ($S = 516$; $p = 0.0056$).

Item discriminations show the extent to which an item differentiates high and low scoring examinees by measuring the correlation between the item and the total score after removing the item. If an item has an acceptable discrimination value (>0.3), the item is considered "good" and can be used even if its difficulty has an extreme value (<10% or >90% correct responses) (Crocker & Algina, 1986). In this sample, Spanish items discriminated better than English items, and the mean discrimination of Spanish noun items was the greatest among the four tests. Among the good items, there were a greater number of Spanish items than English items with acceptable item statistics.

An item analysis was also conducted using item response theory (IRT). The primary advantage of IRT is that the item and participant parameters are sample independent, so the parameter estimates remain stable across different groups of subjects and test items (Reeve & Fayers, 2004). There are no exact guidelines about sample size with IRT, though some research

suggests a Rasch, or a one parameter logistic model, can be used with a relatively small sample sizes (50 to 100) (Linacre, 1994).

Table 3 shows the mean item difficulty parameter estimates and their standard deviations of fitted items using the Rasch model. The number of acceptable items is similar to the results of the classical item analysis. Thus, IRT and CTT give similar results.

Predictors of performance. Multiple regressions with stepwise selection were used to determine which factors best predicted proficiency in English and Spanish (on the items that had acceptable item discrimination). The explanatory variables on the English usage questionnaire (e.g., self-ratings in each language) were used to model accuracy of naming English items. The factors retained in the stepwise regression were casual English comprehension (b = .113, SE = .034) and reading in English (b = .139, SE = .056). The model R^2 was .23.

The explanatory variables on the Spanish language questionnaire were used to model accuracy of naming Spanish items. The factors retained in the stepwise regression were age (b = .009, SE = .003), self-rating of overall Spanish proficiency (b = .089 , SE = .020), ability to comprehend in formal environments (b = .039, SE = .025), percent of time using Spanish in social situations (b = .002, SE = .001) and in their current (b = .001, SE = .001) and permanent homes (b = .001, SE = .0004). The model R^2 was .67.

DISCUSSION

The preliminary results of the classical and Rasch analyses revealed a large number of noun and verb items able to distinguish between different levels of naming proficiency in English and Spanish for Spanish/English bilinguals. There were more good items in Spanish than English. The difference across languages may be due to the sample, which contained more English proficient participants.

The participant analysis, which evaluated the correlation between participant and language variables on naming accuracy, yielded a much higher R^2 for Spanish accuracy than for English accuracy and more factors figured into the model for Spanish than for English. Similar findings have been observed in previous studies (e.g., Kohnert et al., 2004; Marian, et al., 2007), and it has been suggested that there may be more predictive values for the less proficient language due to inherent variability in how that language was learned (Marian et al., 2007). We have collected data from 15 more participants who are more Spanish proficient, so those data may provide more variability on English factors.

REFERENCES

Crocker, L. M., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt Rinehart and Winston.

Druks, J. (2002). Verbs and nouns – A review of the literature. *Journal of Neurolinguistics, 15*(3-5), 289-315.

Druks, J., & Masterson, J. (2000). *An Object and Action Naming Battery.* Hove: Psychology Press.

Gollan, T.H., Fennema-Notestine, C., Montoya, R.I., & Jernigan (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society, 13,* 197-208.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test* (2nd ed.). Baltimore: Lippincott Williams & Wilkins.

Kohnert, K.J., Hernandez, A.E., & Bates, E. (1998). Bilingual performance on the Boston Naming Test: Preliminary norms in Spanish and English. *Brain and Language, 65,* 422-440.

Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions. 7, 4,* 328.

Marian, V., Blumenfeld, H.K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50,* 940-967.

Muñoz, M.L., Marquardt, T.P., & Copeland, G. (1999). A comparison of the codeswitching patterns of aphasic and neurologically normal bilingual speakers of English and Spanish. *Brain and Language, 66,* 249-274.

Paradis, M. (2001). Bilingual and polyglot aphasia. In R.S. Berndt (Ed.), *Language and aphasia* (pp. 69-91). Amsterdam: Elsevier Science.

Reeve, B. B., & Fayers, P. (2004). Applying item response theory modeling for evaluating questionnaire item and scale properties. In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 55-73). New York: Wiley.

Roberts, P.M. (2008). Issues in assessment and treatment for bilingual and culturally diverse patients. In R. Chapey (Ed.), *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders* (5th ed., pp. 245-275). Wolters Kluwer/Lippincott Williams and Wilkins: Philadelphia.

Roberts, P.M., Garcia, L.J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology, 16*(4/5/6), 635-645.