

Supplemental Text A. Description of data preparation and recoding of leniently scored responses for patients with apraxia of speech.

The individual response data were downloaded from MAPPD.org as comma-separated text (see Supplemental Text B) and imported into MS Excel for organization and cleaning. The raw data included 276 cases, of which 20 were identified as healthy control subjects and excluded. There were five patients with missing data due to administrator error (four cases with one missing response and one case with four missing responses) and these were excluded as well. In order to verify the accuracy of the data coding for the remaining 251 cases, we compared the final prepared data file containing the dichotomized item responses for each patient to a dataset that had been independently downloaded and prepared by the third author previously in November 2011. This prior data set included 241 of the 251 patients in the present data set. In comparing them, we found 16 disagreements involving eight patients. In 10 cases, the disagreements were due to errors in data manipulation or re-coding made in preparing the 2011 data set; in the six remaining cases, the raw data appear to have been changed in the MAPPD. In all 16 cases, the data codes used in the 2012 data set were consistent with values downloaded from MAPPD.org in August 2014.

Another issue that we chose to address in cleaning the data was the lenient coding of correct responses for patients with apraxia of speech (AOS). The instructions for the Philadelphia Naming Test (PNT) scoring include the option to allow responses by patients with AOS containing a single phoneme addition, deletion, or substitution to be scored as correct. We adopted a procedure to remove the lenient coding for two reasons. First, the lenient scoring rule for the PNT evolved over the period during which the studies contributing to the MAPPD were conducted (Mirman, personal communication, 2012), with earlier applications being even more lenient and permitting responses containing multiple erred phonemes to be considered correct. Second, in our view, the lenient scoring rule, even when applied with a consistent accuracy threshold, does not appropriately address the issues raised by AOS. AOS is a phonetic-motoric disorder associated with phoneme distortions (which by themselves do not impact PNT scoring) and phoneme substitutions (which are often distorted and which do impact PNT scoring) (McNeil, Robin, & Schmidt, 2009). The critical issue is that when a sound substitution occurs, it is typically not possible to definitively determine whether it is due to AOS or to phonemic paraphasia. Given the propensity for apraxic sound substitutions and phonemic paraphasias to occur in the same individuals, we believe that adopting a lenient scoring criterion for individuals with AOS complicates rather than clarifies the analysis. It is our position that application of a consistent scoring rule with targeted comparisons of responses by aphasic speakers with and without AOS (which are beyond the scope of the present study) will ultimately be more informative.

We implemented the following steps to recode the leniently scored responses: First, we identified 1,081 responses from 116 patients that were coded as correct and also given a phonetic transcription. According to the MAPPD website, provision of a phonetic transcription for a response coded as correct indicates either application of the lenient scoring rule for individuals with AOS or a correct response containing normal dialectal variation. From this set, we further identified 752 responses from 36 individuals identified as having received the lenient apraxia coding in the MAPPD. We also

identified 14 responses from one other individual who was not identified in the MAPPD as having received lenient scoring but for whom the phonetic transcriptions clearly indicated that lenient scoring had been applied (e.g., / fædəlz/ for "saddle," / kɔrp/ for "harp," and / mɔlkeno/ for "volcano" all coded as correct). Of note, there were 10 additional individuals who were coded as having received lenient apraxia scoring but for whom there were no correct responses with phonetic transcriptions, which we took as an indication that the person had AOS but there were no responses scored as correct due to the lenient scoring rule.

In our next step, the second author re-scored the 766 identified responses based on the phonetic transcriptions, according to the standard scoring rules rather than the lenient scoring rule. Specifically, responses that differed from the target by a single phoneme were counted as incorrect, except in cases where the discrepancy could plausibly have been due to dialectal variation. The first author independently re-coded a randomly selected subset of 150 of these responses, and interrater agreement was acceptable, as indicated by a Cohen's kappa value of .73. Landis & Koch (1977) proposed the following scale for interpreting kappa values: $\leq 0 = \textit{poor}$; $.01-.20 = \textit{slight}$; $.21-.40 = \textit{fair}$; $.41-.60 = \textit{moderate}$; $.61-.80 = \textit{substantial}$; $\geq .81 = \textit{almost perfect}$.

References

- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- McNeil, M. R., Robin, D. A., & Schmidt, R. A. (2009). Apraxia of speech: Definition and differential diagnosis. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (2nd ed., pp. 249–268). New York, NY: Thieme.
- Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S., & Brecher, A. (1996). The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology*, 24, 121–133.