Published in final edited form as:

Brain Lang. 2017 May ; 168: 84–94. doi:10.1016/j.bandl.2017.01.012.

Temporal acoustic measures distinguish primary progressive apraxia of speech from primary progressive aphasia

Joseph R. Duffy¹, Holly Hanley⁴, Rene Utianski¹, Heather Clark¹, Edythe Strand¹, Keith A. Josephs², and Jennifer L. Whitwell³

¹Department of Neurology (Speech Pathology), Mayo Clinic, Rochester, Minnesota, USA

²Department of Neurology (Behavioral Neurology and Movement Disorders), Mayo Clinic, Rochester, Minnesota, USA

³Department of Radiology (Neuroradiology), Mayo Clinic, Rochester, Minnesota, USA

⁴Department of Communication Sciences and Disorders, University of North Carolina, Greensboro, North Carolina, USA

Abstract

The purpose of this study was to determine if acoustic measures of duration and syllable rate during word and sentence repetition, and a measure of within-word lexical stress, distinguish speakers with primary progressive apraxia of speech (PPAOS) from nonapraxic speakers with the agrammatic or logopenic variants of primary progressive aphasia (PPA), and control speakers. Results revealed that the PPAOS group had longer durations and reduced rate of syllable production for most words and sentences, and the measure of lexical stress. Sensitivity and specificity indices for the PPAOS versus the other groups were highest for longer multisyllabic words and sentences. For the PPAOS group, correlations between acoustic measures and perceptual ratings of AOS were moderately high to high. Several temporal measures used in this study may aid differential diagnosis and help quantify features of PPAOS that are distinct from those associated with PPA in which AOS is not present.

Keywords

primary progressive apraxia of speech; primary progressive aphasia; acoustic analysis

Corresponding author (present/permanent address): Joseph R. Duffy, Ph.D., Division of Speech Pathology, Department of Neurology, Mayo Clinic, Rochester, MN 55905, USA, jduffy@mayo.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest

None of the authors has any conflicts of interest to declare.

1. Introduction

Apraxia of speech (AOS) is a neurological speech disorder that has been of considerable theoretical and clinical interest for many years. Studies of AOS have contributed to the understanding of speech programming and the neural network that undergirds that process. Refinements in defining its clinical features have helped sharpen the diagnostic boundaries among AOS, aphasia, and dysarthria. Although AOS is usually stroke-induced, it can be the primary sign of neurodegenerative disease, in which case it has been termed primary progressive AOS (PPAOS, Duffy, 2006). Neuroimaging correlates of PPAOS include the left or left-greater-than-right superior premotor cortex and supplementary motor areas (Josephs et al., 2012; Whitwell et al., 2013). Pathologically, it most often is associated with tau biochemistry (4R tau) and progressive supranuclear palsy or corticobasal degeneration (Josephs et al., 2006). PPAOS was the primary focus of this study.

Because there is considerable variability in terminology and criteria for identifying different subtypes of primary progressive aphasia (PPA), particularly PPA subtypes that may be associated with AOS, it is important to provide at the outset the terminology we will adopt in this paper. The designation PPAOS will be used for speakers who have AOS, but no evidence of aphasia. Although the terms agrammatic PPA (agPPA) and nonfluent PPA (nfPPA) are often used together or synonymously in many studies, the term agPPA in this paper will refer to speakers who met criteria for agPPA but had no evidence of AOS. We will use the designation nfPPA when referring to the literature on the nonfluent/agrammatic variant of PPA in which AOS and agrammatism are considered core features (Gorno-Tempini et al., 2011) and one or both problems can be present. Although AOS occurs very frequently in nfPPA (Duffy, Strand, & Josephs, 2014; Ogar et al., 2007), in this study we did not examine individuals who had both AOS and agrammatism. It is generally accepted that AOS does not occur in the semantic variant of PPA and that it is uncommon in the logopenic variant of PPA (lvPPA), although AOS errors can be difficult to distinguish from the phonological errors associated with lvPPA (Croot, Ballard, Leyton, & Hodges, 2012).

Diagnosis of nfPPA may be less reliable than for other PPA variants, possibly because of reduced reliability of AOS diagnosis (Harris et al., 2013). Some suggest that AOS be abandoned as a classification feature in PPA because its diagnosis relies on expert clinical judgment (Sajjadi et al., 2012a). These concerns can be raised about many PPA features (e.g., judgments of “fluency”) but they do point to a need for greater diagnostic reliability.

Acoustic measures have potential as markers to improve diagnostic reliability and quantify features of AOS, as demonstrated in studies of stroke-induced AOS in which acoustic measures have quantified slow rate, syllable segmentation, equalized stress, and voicing abnormalities (e.g., Ballard et al., 2016; Haley et al., 2012; Kent & McNeil, 1987; Rogers & Storkel, 1999; Vergis et al., 2014; Ziegler & von Cramon, 1986). Acoustic measures have also been sensitive to temporal abnormalities in speakers with nfPPA, and to differences between groups with nfPPA versus lvPPA in which the nfPPA group included at least some speakers with AOS (Ash et al., 2009; Ballard et al., 2014; Code, Ball, Tree, & Dawe, 2013; Knibb, Woolams, Hodges, & Patterson, 2009). Wilson et al. (2009) concluded that reduced

maximum speech rate in their nfPPA group likely reflected effects of AOS, but such conclusions are qualified by the possible confounding influence of co-occurring aphasia.

Acoustic correlates of PPAOS have received little attention but they recently have documented progressive slowing of speech rate in a few cases (Duffy et al., 2015, $n = 2$; Laganaro, Croisier, Bagou, & Assal, 2012, $n = 1$). More convincing evidence would come from comparing speakers with PPAOS to neurologically normal speakers and to speakers with PPA without AOS. Inclusion of groups with agPPA without AOS, and lvPPA without AOS, would capture PPA variants that most often co-occur with AOS (nfPPA), or whose phonological errors can be difficult to distinguish from AOS sound-level errors (lvPPA and, probably to a lesser degree, nfPPA). The degree to which acoustic measures distinguish PPAOS from those comparison groups would help gauge their value as diagnostic markers of PPAOS.

This study examined temporal acoustic features that are relevant to rate and syllabic stress abnormalities commonly associated with PPAOS in order to determine their sensitivity to abnormality and their specificity relative to normal speech and variants of PPA with which PPAOS may be confused. Assuming that some acoustic measures are sensitive to perceived rate and prosodic abnormalities associated with speech programming difficulties, and not to language impairment, we hypothesized that they would not differ among normal speakers and speakers with agPPA and lvPPA without AOS, and that those groups would have shorter utterance durations, more rapid syllable production rates, and greater differences between an unstressed and stressed vowel within a multisyllabic word than PPAOS speakers. A primary goal was to identify an easily measured acoustic marker(s) that can confirm the diagnosis and quantification of PPAOS and support its distinction from PPA without AOS.

2. Methods

2.1 Participants

The participants with PPAOS or PPA were among a cohort of 169 people who participated in IRB-approved studies of PPA and PPAOS. The criteria and methods for diagnosing PPAOS and PPA and its variants have been described elsewhere (Botha *et al.*, 2015; Josephs *et al.*, 2012, 2013).

2.1.1 Primary progressive apraxia of speech (PPAOS)—Within the 169 participant cohort, 65 individuals had AOS based on the presence of speech features currently accepted as either distinctive of AOS (e.g., distorted sound substitutions or additions; increased sound distortions with increased length, complexity, or rate), or commonly associated with AOS but possibly overlapping with aphasia (e.g., audible or visible articulatory groping) or dysarthria (e.g., slow rate; sound distortions) (Ballard *et al.*, 2015; McNeil, Robin, & Schmidt, 2009; Strand, Duffy, Clark, & Josephs, 2014). After excluding those who unequivocally also had aphasia, 31 received a diagnosis of PPAOS. Among them, nine had unequivocal dysarthria and thus were excluded because of the possible influence of dysarthria on the acoustic measures. One additional individual was excluded because of a prominent non-English accent that might have biased acoustic measures. The remaining 21 individuals were the primary focus of this study (two of them were described in detail by

Duffy *et al.*, 2015). All of them had speech features distinctive of AOS. Many of them also had speech features that can overlap with aphasia or dysarthria but, as stated, none had unequivocal evidence of aphasia or dysarthria by any other measure. That is, their performance on several measures of language ability/aphasia was normal (see Results for details) and their motor speech was not characterized by any unequivocal speech features associated with dysarthria that do not overlap with AOS (e.g., strained voice quality, reduced loudness, hypernasality, rapid rate). We believe this is the largest cohort of individuals with “pure” AOS, regardless of etiology, for whom speech-language and acoustic data have been reported in a single study.

2.1.2 Agrammatic variant of primary progressive aphasia (agPPA)—Within the 169 participant cohort were 38 individuals who were classified as agPPA at initial assessment and one individual who initially was unclassifiable but was classified as agPPA at second assessment. The agPPA diagnosis was consistent with consensus criteria for nfPPA/agPPA diagnosis (Gorno-Tempini et al., 2011); that is, there was evidence of agrammatism and/or AOS, plus at least two of three additional features (impaired complex sentence comprehension, spared single word comprehension, spared object knowledge). Among those 39 individuals, only six had no AOS or dysarthria, which qualified them for this study. As documented in the Results, their overall aphasia severity was mild but ranged from near normal to moderate. In addition to evidence of agrammatism in spoken or written language, they had mild to marked impairment in sentence comprehension and normal to moderate impairment on confrontation naming. Their inclusion allowed us to examine if agPPA without AOS is associated with acoustic temporal abnormalities, and if any such abnormalities are distinguishable from those that might be associated with PPAOS. We hypothesized that the acoustic measures would be sensitive to PPAOS but not to deficits associated with agPPA in the absence of AOS.

2.1.3 Logopenic variant of primary progressive aphasia (lvPPA)—Within the 169 participant cohort were 65 individuals who were classified as lvPPA. The lvPPA diagnosis was consistent with consensus criteria; that is, there was evidence of impaired word retrieval in spontaneous speech and naming, and at least three of four additional features (phonologic speech errors, spared single word comprehension and object knowledge, absence of frank agrammatism, spared motor speech). Twenty of them were selected for study because none had dysarthria or features of AOS that do not overlap with features of phonologic paraphasias (e.g., syllable segmentation) and because a previous study (Petroi, Duffy, Strand, & Josephs, 2014) established that they made few-to-many phonological errors. As documented in the Results, their overall aphasia severity was mild-moderate but ranged from near normal to severe. Their sentence comprehension and naming ability ranged from normal to severely impaired. Their inclusion allowed us to examine if lvPPA without AOS or dysarthria, but with varying degrees of phonologic errors, is associated with acoustic temporal abnormalities, and if any such abnormalities are distinguishable from those that might be associated with PPAOS. We hypothesized that the acoustic measures would be sensitive to the PPAOS but not to deficits associated with lvPPA in the absence of AOS.

For all PPAOS, agPPA and lvPPA cases, diagnoses were based on independent agreement or consensus agreement between two speech-language pathologists with considerable experience with PPAOS and PPA, based on clinical test results and review of video recordings of the formal testing.

2.1.4 Control speakers—Eleven individuals served as normal speech controls for acoustic measure comparisons. Four of them had been entered into the study that included the PPAOS and PPA participants but detailed speech-language and neurological assessment and neuroimaging failed to reveal evidence of aphasia, motor speech disorder, or imaging abnormality. The remaining seven individuals were seen as part of the routine speech-language outpatient practice and had no evidence of speech or language difficulty.

2.2 Speech, language, and other measures

Language measures reported here included the Western Aphasia Battery (WAB, Kertesz, 2007), Part V of the Token Test (DeRenzi & Vignolo, 1962), and a 15-item version of the Boston Naming Test (Lansing, Ivnik, Cullum, & Randolph, 1999). Classification of PPA participants was based largely on the results of these tests.

Perceptual judgments of speech included: (1) the Apraxia of Speech Rating Scale (ASRS, Strand *et al.*, 2014), an index of the presence, prominence or severity of 16 speech features associated with AOS (higher scores reflect greater prominence or severity), (2) a 0-4 rating of AOS severity (1 = mild; 4 = severe), as an index of AOS severity regardless of its specific features, (3) a 1-10 rating (10 = normal) of motor speech severity (MSD severity) (adapted from Yorkston, Strand, Miller, & Hillel, 1993), which indexed intelligibility and functional speech impairment, and (4) an Articulation Error Score (AES) which indexed the prevalence of sound level errors that can be associated with AOS or aphasic phonological errors. The AES is the percentage of words on which distorted or undistorted sound substitutions, additions or repetitions, sound omissions, sound prolongations, false starts, and successful or unsuccessful attempts to correct sound errors occurred during repetition of words or sentences, totaling 56 words (Duffy *et al.*, 2015). All speech ratings were based on responses to conversational questions and the WAB picture description task, and responses to stimuli used for the AES.

Finally, to document the presence and severity of nonverbal oral apraxia (NVOA), an eight-item measure of NVOA was administered (Botha *et al.*, 2014).

Demographic features for all participants, and PPAOS and agPPA and lvPPA participants' performance on speech-language measures are summarized in Table 1.

2.3 Temporal acoustic measures

Imitative speech tasks were used to minimize the influence or frequency of pause, hesitation, and restart occurrences that can be associated with aphasia or AOS during non-imitative utterances; because those temporal events can reflect influences of aphasia, they can confound measures intended to reflect motoric aspects of speech. In addition, imitation of words of increasing length and multisyllabic words are known to be sensitive to AOS and its distinction from aphasia (e.g., Ballard *et al.*, 2016). The choice of simple temporal acoustic

measures was supported by the limited data for a few PPAOS speakers (Duffy *et al.*, 2015; Laganaro *et al.*, 2012) and a desire to identify measures that can be easily used in clinical settings.

Temporal measures were completed for three consecutive repetitions each for the words, *cat*, *catnip*, *catapult*, *catastrophe*, and *stethoscope*, and a single repetition of the sentences, “*The municipal judge sentenced the criminal*” and “*My physician wrote out a prescription.*” These stimuli were selected because words of increasing length and sentences containing multisyllabic words have proven challenging to individuals with AOS in many perceptual and acoustic studies by other investigators (e.g., Haley & Overton, 2001; Odell *et al.*, 1991; Strand & McNeil, 1996). The specific words and sentences chosen for this study represent more than 35 percent of the imitative speech stimuli used in our assessment of motor speech within our overall research protocol. The stimuli were presented face-to-face using natural loudness, rate and prosody. They were repeated on request or if they were not retained. Repetition was seldom needed for words but sometimes was necessary for sentences, especially for lvPPA participants who often failed to retain all words in the sentence stimuli.

Durations were measured from the initial stop release or the initial onset of noise energy or voicing for each response, to cessation or marked reduction of acoustic energy at the end of the response. Duration and syllable rates for the single words reflect the three repetition average, unless there were missing data. Sentence duration and syllable rate reflect the average duration and rates across the two sentences, unless there were missing data. Syllables per second rates were derived by dividing the number of syllables in the target response by its duration.

To quantify the abnormal syllabic stress that can characterize AOS (Odell, McNeil, Rosenbek, & Hunter, 1991; Strand & McNeil, 1996), a pairwise variability index (PVI) was computed using the average duration of the vowels in the first two syllables across three repetitions of the word *catastrophe*. The formula used was: $PVI = 100 \times (d1 - d2) / [(d1 + d2) / 2]$, where d1 and d2 are the durations of the first and second vowel, respectively). PVI values closer to zero are consistent with relatively equalized stress between compared syllables/vowels. For *catastrophe*, normal responses should have negative PVI values because the first syllable is unstressed (shorter) and the second syllable is stressed (longer). Based on PVI findings for speakers with stroke-related AOS and nfPPA with frequent co-occurring AOS (Ballard *et al.*, 2014; Vergis *et al.*, 2014), and two PPAOS speakers (Duffy *et al.*, 2015), we hypothesized that PPAOS speakers would maintain the stress distinction between the two syllables but that the distinction would be less than that achieved by Control, agPPA and lvPPA participants who presumably have no problems with lexical stress.

The speech analysis software, Praat (version 5.3.63; Boersma & Weenink, 2014), was used for the acoustic measures.

2.4 Acoustic measure reliability

Before measurements were made for all speakers, two of the authors (JRD and HH) independently measured all responses for eight of the 58 participants (5 PPAOS, 3 Controls).

Pearson correlations between the two judges across all of those measures exceeded 0.99, and all but one between-judge t-test comparison across the measures for each speaker was not significant ($p > 0.05$); the one significant difference ($p = 0.047$) was associated with only a 2% average difference between the judges. Differences in raw data measurements between the judges averaged less than 2% and ranged from 1% to 3%. Reliability was thus considered high. Consensus was sought on a few occasions when substantial measurement uncertainty arose during measures for the full data set.

2.5 Missing data

For Control speakers, all duration measures (23 measures per subject: 3 repetitions of 5 words; 2 sentences; 2 vowels with three repetitions for the PVI) could be measured validly. For the agPPA group, six responses (4% of group responses) from among three of the six speakers could not be measured validly because of false starts or filler syllables on words within sentence responses. For the PPAOS group, 14 responses (5% of group responses) from among eight of the 21 speakers could not be measured validly, primarily because of false starts or syllable additions or omissions during multisyllabic words or sentences. Adequate responses following a false start were measured from the onset of the adequate response for all participants. In spite of these missing data, analyzable responses were produced by all speakers in the agPPA and PPAOS groups for at least one of the three repetitions of the five multisyllabic words; at least one of the two sentences were analyzable for all agPPA speakers and all but three of the PPAOS speakers. Responses containing sound level errors that contained the correct number of syllables, and did not contain any of the exclusionary features noted above, were measured. Across five members of the PPAOS group there were 13 such responses; such responses were not evident in the agPPA group.

For the lvPPA group, 97 responses among 17 of the 20 participants (27% of group responses) could not be measured validly, primarily because of false starts in multisyllabic words or between words in sentences, silent pauses (> 0.5 secs), semantic errors, aborted responses before response completion, or inadequate retention of sentence stimuli, even following a second stimulus presentation. This led to an absence of any analyzable data for two to four participants across the four multisyllabic words, and the absence of data for both sentences for 13 participants. For three participants, sentence syllable rates were computed on the basis of a validly repeated string of six or seven syllables within a sentence (e.g., "...wrote out a prescription"). Responses containing sound level errors that contained the correct number of syllables, and did not contain any of the exclusionary features, were measured; across eight members of the lvPPA group there were 17 such responses.

Computed averages for the repetitions of each word and the average duration of the two sentences were adjusted appropriately when data were missing (i.e., the sum of the responses obtained were divided by the number of responses obtained).

2.6 Statistical analysis

Statistical analyses were conducted using JMP (version 8.0.0; SAS Institute Inc., Cary, NC). Continuous data were compared across groups using one-way ANOVAs; post-hoc testing used the Tukey-Kramer HSD with alpha set at $p = 0.05$. Tests for homogeneity of variance

revealed departure from homogeneity for the stimulus words *catapult* and *catastrophe*, leading to the alternative use of Welch's ANOVA which does not require assumptions about homogeneity of variance. Effect sizes were calculated using Cohen's D. Sensitivity and specificity analyses were completed using a receiver operating characteristic (ROC) analysis. The ROC curve plots sensitivity against specificity to determine the cut-off point at which maximum sensitivity and specificity are achieved. ROC analysis also provides a measure of the area under the curve (AUC), which reflects the robustness of discriminability over chance; an AUC of 0.5 corresponds to chance discriminability, whereas an AUC of 1.0 is consistent with perfect accuracy. The Wilcoxon rank sum/Mann-Whitney U test of group-wise differences was used to assess the ROC data.

3. Results

3.1 Clinical measures (language, speech, oral praxis)

Table 1 summarizes language, speech, and oral praxis performance and related statistical tests for group differences. There were no group differences for age or education. The agPPA group's mean illness duration of 1.8 years was shorter than for the lvPPA (3.4 years) and PPAOS groups (3.6 years), but the difference was not statistically significant.

The PPAOS group performed normally on language measures. On the WAB AQ, the lvPPA group was more impaired than the PPAOS and agPPA groups but only the difference between the lvPPA and PPAOS groups was statistically significant. Two marginally abnormal AQ scores (92.6 and 93.3) in the PPAOS group were attributed to reduced animal fluency scores which can occur in individuals with PPAOS and no evidence of aphasia (Josephs *et al.*, 2012). For the Token Test, the PPAOS group was normal and superior to both aphasic groups; the agPPA group was superior to the lvPPA group. For the Boston Naming Test, the PPAOS group was superior to the lvPPA group but not the agPPA group (note, however, that no PPAOS participant performed abnormally, whereas three agPPA participants did perform abnormally); the agPPA group was superior to the lvPPA group.

On the ASRS, the PPAOS group performed more poorly than both PPA groups, as expected, with no difference between the agPPA and lvPPA groups, and with all agPPA and lvPPA speakers scoring in the non-AOS range (consistent with Strand *et al.*, 2014). By definition, the agPPA and lvPPA groups were rated as normal (0) on the AOS severity ratings and normal or minimally impaired on the MSD severity ratings. Average AOS severity in the PPAOS group was mild-moderate; 13 participants were rated as mild, two as mild-moderate, four as moderate, and two as marked.

On the AES, the agPPA group made the smallest percentage of errors (7%) and the lvPPA group made the most (29%), with the PPAOS group falling intermediate between them (17%) but those differences were not statistically significant.

There was no statistically significant difference among the groups on the measure of nonverbal oral apraxia (NVOA). Each group contained members who performed abnormally. The percentage of participants in the agPPA, lvPPA, and PPAOS groups with NVOA was 50%, 50%, and 33%, respectively (using norms reported by Botha *et al.*, 2014).

3.2 Acoustic measures

Table 2 summarizes results for word and sentence durations, syllable rate per second for words and sentences, and the PVI measure. The box plots in Figure 1 illustrate the rate data results (duration data would look nearly identical to rate data, relative to group differences).

3.2.1 Word and sentence duration—ANOVA identified no differences among the four groups for duration of *cat* ($p = 0.33$).

ANOVAs were significant for all of the multisyllabic words and the sentences (all p values were 0.0006 or smaller). On post hoc testing, for *catnip*, there were no significant differences among the Control, agPPA, and lvPPA groups, but the PPAOS group had significantly longer durations than the Control ($p = 0.02$), agPPA ($p = 0.04$), and lvPPA ($p = 0.0008$) groups. For *catapult*, there were no significant differences among the Control, agPPA, and lvPPA groups. The PPAOS group had significantly longer durations than the Control ($p = 0.004$) and lvPPA ($p = 0.0001$) groups, but not the agPPA group ($p = 0.08$). For *catastrophe*, there were no differences among Control, agPPA, and lvPPA groups. The PPAOS group had significantly longer durations than the Control ($p = 0.001$), agPPA ($p = 0.02$), and lvPPA ($p = 0.001$) groups. For *stethoscope*, there were no differences among Control, agPPA, and lvPPA groups. The PPAOS group had longer durations than the Control and lvPPA groups ($p < 0.0001$) and the agPPA group ($p = 0.006$). For *sentences*, there were no differences among Control, agPPA, and lvPPA groups. The PPAOS group had longer durations than the Control and lvPPA groups ($p < 0.0001$) and the agPPA group ($p = 0.0004$).

Effect sizes for the significant differences in duration between the PPAOS group and each of the other groups for the multisyllabic words and the sentences were nearly always large or very large, and averaged 1.21 (Table 3). Among multisyllabic words, effect sizes increased in magnitude as length or complexity increased (*catnip* < *catapult* < *catastrophe* < *stethoscope*). For sentences, effect sizes were intermediate in magnitude relative to the multisyllabic words.

To summarize, the PPAOS group had longer durations than the Control and both of the PPA groups for all words of two or more syllables, and the sentences. For one three-syllable word (*catapult*), the difference between the PPAOS and agPPA group was not statistically significant.

3.2.2 Syllable rate measures—ANOVA showed no differences among the four groups for syllable rate for *cat* ($p = 0.29$).

ANOVAs were significant for all of the multisyllabic words and the sentences (p values were 0.0002 or smaller). On post hoc testing, for *catnip*, there were no differences among Control, agPPA, and lvPPA groups. The PPAOS group had significantly slower syllable rate than the Control and agPPA groups ($p = 0.02$) and the lvPPA group ($p = 0.0001$). For *catapult*, there were no differences among Control, agPPA, and lvPPA groups. The PPAOS group had significantly slower rate than the Control ($p = 0.002$) and lvPPA ($p = 0.0005$) groups, but not the agPPA group ($p = 0.08$). For *catastrophe*, there were no differences among the Control,

agPPA, and lvPPA groups. The PPAOS group had slower rate than the Control ($p < 0.0001$), agPPA ($p = 0.0004$), and lvPPA ($p < 0.0001$) groups. For *stethoscope*, there were no differences among Control, agPPA, and lvPPA groups. The PPAOS group had significantly slower rate than the Control and lvPPA groups ($p < 0.0001$) and the agPPA group ($p = 0.002$). For the *sentences*, the PPAOS group had a slower syllable rate than the Control ($p = 0.0001$), agPPA ($p = 0.0003$), and lvPPA ($p = 0.0009$) groups. The lvPPA group had a slower syllable rate than the Control group ($p = 0.0008$) but did not differ from the agPPA group, and the agPPA group did not differ from the Control group.

Effect sizes for the significant differences in syllable rate were nearly always large or very large, and averaged 1.38. Among multisyllabic words, effect sizes tended to increase in magnitude as length or complexity increased. For sentences, effect sizes were intermediate or of lower magnitude relative to the multisyllabic words.

To summarize, the PPAOS group had slower syllable rates than the Control and both of the PPA groups for all words of two or more syllables, and the sentences. For one three-syllable word (*catapult*), the difference between the PPAOS and agPPA group was not statistically significant. For sentences, the lvPPA group also had slower syllable rate than the Control group.

3.2.3 Pairwise variability index—All groups had negative PVI values (Figure 2), consistent with the expected shorter vowel duration in the unstressed initial syllable versus the stressed second syllable. ANOVA for group differences was significant ($p = 0.0003$). Post hoc tests revealed no differences among Control, agPPA, and lvPPA groups. The PPAOS group had smaller PVI values than the Control ($p = 0.008$), agPPA ($p = 0.001$), and lvPPA ($p = 0.02$) groups. Effect sizes for comparison of the PPAOS group to each of the other groups were large or very large, and averaged 1.44, but were of somewhat lower magnitude than the duration and syllable rate measures for *catastrophe*.

Closer examination of the PVI for PPAOS speakers was consistent with equalization of stress due to relatively increased stress on the unstressed syllable. For example, although average PPAOS vowel duration within the initial unstressed syllable and the following stressed syllable were both abnormally lengthened in comparison to the Control group, the initial unstressed syllable vowel was disproportionately lengthened (2.1 times greater than the Control group mean) compared with the stressed syllable vowel (1.4 times greater than the Control group mean).

3.2.4 Sensitivity and specificity—Given the predominant absence of significant differences among the Control, agPPA, and lvPPA groups across all acoustic measures, their data were collapsed to form a single group against which to examine sensitivity and specificity of the rate measures to PPAOS¹. Duration data will not be presented here because the results for rate and duration are identical or nearly identical. For this study, sensitivity is

¹To further assess the validity of this assumption, we conducted sensitivity and specificity analyses comparing the PPAOS group to the Control group only and another analysis comparing the PPAOS group to the combined agPPA and lvPPA groups. The differences in sensitivity, specificity, and AUC among the comparisons ranged from nil to small differences that would not require a different interpretation of any of the results in which the Control, agPPA, and lvPPA data were combined.

the percentage of patients who are correctly identified as having PPAOS (i.e. true positive). Specificity is the percentage of patients without PPAOS correctly identified as such (i.e. true negative). Sensitivity and specificity were assessed for the PVI for *catastrophe* (Figure 2) and the rate measures for each word and the sentences (Figure 3).

ROC curves, sensitivity and specificity measures, and AUC data for rate are summarized in Figure 3 which shows increased sensitivity and specificity as number of syllables increases. Thus, rate for *cat* yields limited sensitivity and specificity for identification of PPAOS patients, with an AUC at chance level. This is compatible with the comparatively minimal programming demands for single versus multisyllabic words. Sensitivity and specificity increase substantially and are significant ($p < 0.001$) for all other stimuli. The word *catastrophe* probably can be considered the “most effective” stimulus, in that it correctly identified 90% of PPAOS members as having PPAOS, and correctly identified 91% of the combined non-AOS group as not having PPAOS, with an AUC of 0.90 (considered good-excellent). The results for the *sentences* are comparably good but, as already noted, the sentences had frequent unanalyzable data, especially for the lvPPA group.

In tandem with the increase in sensitivity and specificity with increasing number of syllables, there is also an increase in the AUC, which is a composite reflection of the measures. The *sentences* yielded the largest AUC but again there were missing sentence data for many lvPPA and some PPAOS participants.

3.2.5 Additional observations—The magnitude of differences in duration between the PPAOS group and the _{no-AOS} groups seems amplified by increasing word length and/or complexity. This is best illustrated by comparing the mean durations between the PPAOS and Control speakers. For *catnip*, PPAOS duration was 35% longer than the Control group mean. For *catapult*, PPAOS duration was 45% longer, for *stethoscope* 78% longer, for *catastrophe* 82% longer, and for *sentences* 94% longer. This disproportionate increase in duration cannot be explained by a simple increase in duration alone because syllable rate for the PPAOS group actually increased, relative to the single syllable word *cat*, by 6%, 34%, and 10% for *catnip*, *catapult*, and *catastrophe*, respectively. Although these rate increases are far less than those for the Control group (increases of 50%, 99%, and 93%, respectively), the PPAOS group did follow the “rule” for increased syllable rate across one through three/four syllable words, similar to the reduction in duration of stressed vowels that occurs in normal speakers and speakers with stroke-induced AOS as the number of syllables in words increases (e.g., Collins, Rosenbek, & Wertz, 1983; Haley & Overton, 2001).

3.3 Relationships between Acoustic and Perceptual Measures

Table 4 shows correlations between the acoustic rate measures and the ASRS, AOS severity rating, MSD severity rating, AES, and the NVOA measures for the PPAOS group. Because the rate and duration measures were highly correlated with each other, correlations between duration and the clinical measures are not reported in Table 1 because they are very similar to those for rate and are interpreted identically. Correlations between the NVOA and syllable duration and rate measures were low and suggest an absence of any notable relationship between nonverbal oral apraxia and the acoustic measures. In contrast, correlations between

the ASRS and AOS severity ratings and syllable duration and rate measures were mostly moderately high to high, and suggest that the acoustic measures, especially for multisyllabic words, share common variance with clinical perceptual measures of AOS severity and/or prominence of features associated with AOS. Correlations between the MSD severity rating and the acoustic measures were significant only for *catnip* (and *catastrophe* for duration, 0.38) and never higher than moderate, suggesting that the duration and rate measures, and PVI, are not strongly predictive of intelligibility.

4. Discussion

The primary finding of this study is that temporal acoustic measures of imitated word and sentence duration, syllable rate, and lexical stress are sensitive to the slow speech rate and abnormal lexical stress that are perceived in speakers with PPAOS. These acoustic measures distinguish speakers with PPAOS not only from neurologically normal speakers but also from those with PPA who do not have AOS. It is noteworthy that the temporal measures were sensitive to relatively mild AOS in that the majority of PPAOS speakers (15/21) were rated as mildly or mild-moderately impaired. The validity of these temporal measures as markers of PPAOS is also supported by their fairly high correlations with a scale measuring the pervasiveness/severity of features associated with AOS (i.e., ASRS) and a simple clinical rating of AOS severity. The temporal measure differences, beyond the single syllable word level, between the PPAOS group and the Control, agPPA and lvPPA groups are substantial enough to support their usefulness in differential diagnosis and quantification of some of the features of AOS. The use of such acoustic measures may reduce concerns about the reliability of perceptual judgments alone for PPAOS diagnosis, and AOS diagnosis in general (e.g., Haley *et al.*, 2012; Sajjadi *et al.*, 2012a). For example, at least when dysarthria is not suspected, using the cut-off values provided in Figures 2 and 3 could provide confirmatory evidence for the presence versus absence of AOS derived from perceptual judgments made during clinical evaluation. Cross-validation of these cut-off values must occur before the data can be applied confidently in this way, however.

The results of this study concur with those of Laganaro *et al.* (2012) who found that slow speech rate was a primary feature in one case with PPAOS, and they are consistent with findings of longer speech durations in groups with nfPPA in which at least some members had AOS (Ballard *et al.*, 2014; Knibb *et al.*, 2009; Wilson *et al.*, 2009). The predominant absence of temporal abnormalities in our agPPA speakers without AOS suggests that increased utterance duration or reduced syllable rates found in previous studies of nfPPA may indeed reflect effects of AOS and not agrammatism or other aphasic features (cf. Wilson *et al.*, 2009). We acknowledge that we measured imitated words and sentences, and thus do not claim that durational abnormalities might not exist in people with nfPPA without AOS on non-imitative speech tasks, as may have been the case in some studies (Sajjadi, Patterson, Arnold, & Nestor, 2012b). However, such tasks are more likely influenced by linguistic (e.g., pauses for word retrieval) or related factors (e.g., working memory) that can confound measures intended to capture difficulties with the motor aspects of speech. Because multisyllabic words fared about as well or better than sentence imitation in distinguishing PPAOS from normal, agrammatic, and logopenic speech, we suggest using multisyllabic words instead of sentences for acoustic temporal analysis when the purpose is

diagnosis and quantification of PPAOS. Word rather than sentence imitation likely is less susceptible to lexical, phonological, and sentence-level language influences. Our data provide support for this, in that sentence imitation yielded a much higher rate of responses that could not be validly measured, at least in part because of such non-motor influences. This was particularly evident for the lvPPA group, in which valid measurement of sentences was possible for only 7/20 lvPPA group members.

The duration, rate and PVI measures were similarly effective in distinguishing the PPAOS group from the lvPPA group that had some members who made a substantial number of phonological errors, as reflected in the AES. This suggests that the temporal measures, at least of the word stimuli we used, are not sensitive to phonological errors, as long as they are considered measurable by the criteria used in this study. This is important for measures intended to mark and quantify features of AOS that are separable from errors made by speakers without AOS who make phonological errors.

The duration and rate data revealed no differences among any of the groups for the word *cat*, which suggests that duration and rate of simple single syllable words is not sensitive to PPAOS, at least within the severity range of our PPAOS group. Beyond the single syllable word level, however, most members of the PPAOS group were abnormal, and the magnitude of abnormality increased with increasing word length and/or complexity. This was the case in spite of the fact that syllable rate increased as word length increased from one to three syllables for the PPAOS group, as it did for the other groups. Although the degree of rate increase with increasing length/complexity was far less in the PPAOS group than for the other groups, it does indicate that mild-to-moderately impaired speakers with PPAOS have some capacity, albeit limited, to increase speed as number of syllables increases. This suggests that their programming mechanism recognizes length as a relevant programming parameter (i.e., it attempts to program more than one syllable at a time) but that its capacity to meet that programming demand is reduced as length and/or complexity increases. We acknowledge that our stimuli were not controlled for factors related to utterance motor programming load or error frequency in AOS (e.g., number and complexity of phonemes, syllables, and consonant clusters; syllable frequency; metrical factors) (cf, Ziegler, 2009). Nor did we do a finer grained temporal acoustic analysis to identify the segment locus within word or sentence responses of reduced rate (e.g., the degree to which reduced rate might reflect prolonged syllables, increased intersyllable intervals, or both). Thus, our data cannot address the likely specific source(s) of duration and rate differences among and within our stimuli. These variables clearly deserve attention in future studies.

Although the temporal measures correlated strongly with the ASRS and AOS severity ratings, they were not strongly correlated with the MSD severity rating or the AES. This suggests that the temporal measures, by themselves, do not predict speech intelligibility or frequency of apraxic articulatory errors. Our temporal measures obviously do not capture all abnormalities within the full cluster of features that can characterize PPAOS, at least some of which are probably more strongly related to intelligibility and sound level errors. This is compatible with our clinical experience that some people with PPAOS who have noticeably slow speech rate and segmentation of syllables are highly intelligible.

The good diagnostic discriminative value of the PVI measure in this study is consistent with findings for stroke-induced aphasia with versus without AOS (Ballard *et al.*, 2016; Vergis *et al.*, 2014), and for speakers with lvPPA and nfPPA in which PVI was a strong predictor of AOS (Ballard *et al.*, 2014). Our finding that abnormal PVI values were primarily due to disproportionate lengthening of vowels in the word-initial weak syllable as opposed to lengthening of the vowel in the subsequent stressed syllable also concurs with those studies' findings. Our results thus support the usefulness of the PVI for both stroke-induced AOS with co-occurring aphasia, and neurodegenerative AOS when it occurs in isolation or with PPA. Our results did not establish that the PVI is diagnostically superior to more straightforward measures of word duration and syllable rate, however. In fact, PVI effect sizes for differences between PPAOS and the non-PPAOS groups, while high, were roughly comparable to those for the duration and syllable rate for the word *catastrophe*, the stimulus from which the PVI was derived, and sensitivity and specificity measures were lower in comparison to those for duration and rate. Thus, for differential diagnosis, the simpler duration and rate measures may be sufficient or perhaps superior to the PVI. Nonetheless, word and sentence duration and rate measures do not directly quantify abnormal equalization of stress that is captured by the PVI. The need to use both kinds of measures for diagnostic, descriptive, or quantification purposes must be determined by further study.

We did not examine individuals with dysarthria in this study. Comparing PPAOS to dysarthria is important to further determine the diagnostic value of these acoustic measures because dysarthria is sometimes difficult to distinguish from AOS perceptually, and because dysarthria not infrequently co-occurs with neurodegenerative and stroke-related AOS. We suspect the acoustic measures used here will not, by themselves, discriminate AOS from dysarthria as well as they discriminate PPAOS from PPA (cf., Ballard *et al.*, 2016); that is, these acoustic measures may simply mark the presence of some type of motor speech disorder, but are not specific to any particular motor speech disorder (i.e., AOS versus a type(s) of dysarthria). It may be that a combination of perceptual measures and these or additional acoustic measures will be necessary to maximize sensitivity and specificity for differentiating PPAOS (or AOS regardless of etiology) from dysarthria. At this point, it would be premature to conclude that these temporal measures, by themselves, can be diagnostic of PPAOS when dysarthria is present. The measures do, however, provide good support for a PPAOS diagnosis when there is no clinical evidence of dysarthria and the distinctions to be made are between PPAOS and normal, agrammatic, or logopenic speech without AOS.

In summary, our results indicate that temporal measures of multisyllabic word and sentence duration, syllable rate, and lexical stress can distinguish PPAOS from normal speech and speech of individuals with agrammatic and logopenic PPA, but without AOS, with good sensitivity and specificity. The validity of the acoustic temporal measures is supported further by their fairly high correlations with perceptual measures of PPAOS. These simple acoustic measures may aid clinical differential diagnosis between AOS and aphasia without AOS, and help quantify the perceptual features of increased utterance duration, slow rate of syllable production, and equalized lexical stress that are among core features of PPAOS.

Acknowledgments

The authors would like to thank Stephen D. Weigand (statistician) for assistance with data analysis.

Funding

This work was supported by National Institute on Deafness and Communication Disorders Grants R01-DC010367 (Keith A. Josephs, P.I.) and R01-DC012519 (Jennifer L. Whitwell, P.I.). These funding sources played no role in the design, data collection, analysis or interpretation of this study.

References

- Ash S, Moore P, Vesely L, Gunawardena D, McMillan C, Anderson C, Grossman M. Non-fluent speech in frontotemporal lobar degeneration. *Journal of Neurolinguistics*. 2009; 22:370–383. [PubMed: 22180700]
- Ballard KJ, Azizi L, Duffy JR, McNeil MR, Halaki M, O'Dwyer N, Robin DR. A predictive model for diagnosing stroke-related apraxia of speech. *Neuropsychologia*. 2016; 81:129–139. [PubMed: 26707715]
- Ballard KJ, Savage S, Leyton CE, Vogel AP, Hornberger M, Hodges JR. Logopenic and nonfluent variants of primary progressive aphasia are differentiated by acoustic measures of speech production. *PLOS ONE*. 2014; 9:1–14.
- Ballard KJ, Wambaugh JL, Duffy JR, Layfield C, Maas E, McNeil MR. Treatment for acquired apraxia of speech: a systematic review of intervention research between 2004 and 2012. *American Journal of Speech-Language Pathology*. 2015; 24:316–337. [PubMed: 25815778]
- Boersma, P., Weenink, D. Praat: Doing phonetics by computer (Version 5.3.63) [Computer software]. Institute of Phonetic Sciences; Amsterdam, the Netherlands: 2014. <http://www.praat.org/>
- Botha H, Duffy JR, Strand EA, Machulda MM, Whitwell JL, Josephs KA. Nonverbal oral apraxia in primary progressive aphasia and apraxia of speech. *Neurology*. 2014; 82:1729–1735. [PubMed: 24727315]
- Botha H, Duffy JR, Whitwell JL, Strand EA, Machulda MM, Schwarz CG, Josephs KA. Classification and clinicoradiologic features of primary progressive aphasia (PPA) and apraxia of speech. *Cortex*. 2015; 69:220–236. [PubMed: 26103600]
- Code C, Ball M, Tree J, Dawe K. The effects of initiation, termination and inhibition impairments on speech rate in a case of progressive nonfluent aphasia with progressive apraxia of speech with frontotemporal degeneration. *Journal of Neurolinguistics*. 2013; 26:602–618.
- Collins MJ, Rosenbek JC, Wertz RT. Spectrographic analysis of vowel and word duration in apraxia of speech. *Journal of Speech and Hearing Research*. 1983; 26:224–230. [PubMed: 6887809]
- Croot K, Ballard KJ, Leyton CE, Hodges JR. Apraxia of speech and phonological errors in the diagnosis of nonfluent/agrammatic and logopenic variants of primary progressive aphasia. *Journal of Speech, Language and Hearing Research*. 2012; 55:S1562–S1572.
- De Renzi E, Vignolo LA. The Token Test: A sensitive test to detect receptive disturbances in aphasics. *Brain*. 1962; 85:665–678. [PubMed: 14026018]
- Duffy JR. Apraxia of speech in degenerative neurologic disease. *Aphasiology*. 2006; 20:511–527.
- Duffy JR, Strand EA, Clark H, Machulda M, Whitwell JL, Josephs KA. Primary progressive apraxia of speech: Clinical features and acoustic and neurologic correlates. *American Journal of Speech-Language Pathology*. 2015; 24:88–100. [PubMed: 25654422]
- Duffy JR, Strand EA, Josephs KA. Motor speech disorders associated with primary progressive aphasia. *Aphasiology*. 2014; 28:1004–1017. [PubMed: 25309017]
- Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, Grossman M. Classification of primary progressive aphasia and its variants. *Neurology*. 2011; 76:1006–1014. [PubMed: 21325651]
- Haley KL, Jacks A, de Riesthal M, Abou-Khalil R, Roth HL. Toward a quantitative basis for assessment and diagnosis of apraxia of speech. *Journal of Speech, Language and Hearing Research*. 2012; 55:S1502–S1517.

- Haley KL, Overton HB. Word length and vowel duration in apraxia of speech: the use of relative measures. *Brain and Language*. 2001; 79:397–406. [PubMed: 11781050]
- Harris JM, Gall C, Thompson JC, Richardson AMT, Neary D, du Plessis D, Jones M. Classification and pathology of primary progressive aphasia. *Neurology*. 2013; 81:1–8.
- Josephs KA, Duffy JR, Strand EA, Machulda MM, Senjem ML, Lowe VJ, Whitwell JL. Syndromes dominated by apraxia of speech show distinct characteristics from agrammatic PPA. *Neurology*. 2013; 81:1–9.
- Josephs KA, Duffy JR, Strand EA, Machulda MM, Senjem ML, Master AV, Whitwell JL. Characterizing a neurodegeneration syndrome: Primary progressive apraxia of speech. *Brain*. 2012; 135:1522–1536. [PubMed: 22382356]
- Josephs KA, Duffy JR, Strand EA, Whitwell JL, Layton KF, Parisi JE, Petersen RC. Clinicopathological and imaging correlates of progressive aphasia and apraxia of speech. *Brain*. 2006; 129:1385–1398. [PubMed: 16613895]
- Kent, RD., McNeil, MR. Relative timing of sentence repetition in apraxia of speech and conduction aphasia.. In: Ryalls, JH., editor. *Phonetic approaches to speech production in aphasia and related disorders*. College-Hill Press; Boston, MA: 1987. p. 181-220.
- Kertesz, A. *Western Aphasia Battery (Revised)*. The Psychological Corporation; San Antonio, TX: 2007.
- Knibb JA, Woollams AM, Hodges JR, Patterson K. Making sense of progressive non-fluent aphasia: An analysis of conversational speech. *Brain*. 2009; 132:2734–2746. [PubMed: 19696033]
- Laganaro M, Croisier M, Bagou O, Assal F. Progressive apraxia of speech as a window into the study of speech planning processes. *Cortex*. 2012; 48:963–971. [PubMed: 21513930]
- Lansing AE, Ivnik RJ, Cullum CM, Randolph C. An empirically derived short form of the Boston Naming Test. *Archives of Clinical Neuropsychology*. 1999; 14:481–487. [PubMed: 14590575]
- McNeil, MR., Robin, DA., Schmidt, RA. Apraxia of speech.. In: McNeil, MR., editor. *Clinical management of sensorimotor speech disorders*. Thieme; New York, NY: 2009. p. 249-268.
- Odell K, McNeil MR, Rosenbek JC, Hunter L. Perceptual characteristics of vowel and prosody production in apraxic, aphasic, and dysarthric speakers. *Journal of Speech and Hearing Research*. 1991; 34:67–80. [PubMed: 2008083]
- Ogar JM, Dronkers NF, Brambati SM, Miller BL, Gorno-Tempini ML. Progressive nonfluent aphasia and its characteristic motor speech deficits. *Alzheimer's disease and Associated Disorders*. 2007; 21:S23–S30.
- Petrot D, Duffy JR, Strand EA, Josephs KA. Phonologic errors in the logopenic variant of primary progressive aphasia. *Aphasiology*. 2014; 28:1223–1243.
- Rogers MA, Storkel HL. Planning speech one syllable at a time: the reduced buffer capacity hypothesis in apraxia of speech. *Aphasiology*. 1999; 13:793–805.
- Sajjadi SA, Patterson KP, Arnold RJ, Watson PC, Nestor PJ. Primary progressive aphasia: a tale of two syndromes and the rest. *Neurology*. 2012a; 78:1670–1677. [PubMed: 22573633]
- Sajjadi SA, Patterson KP, Tomek M, Nestor PJ. Abnormalities of connected speech in the non-semantic variants of primary progressive aphasia. *Aphasiology*. 2012b; 26:1219–1237. 2012b.
- Strand EA, Duffy JR, Clark H, Josephs KA. The Apraxia of Speech Rating Scale: A new tool for diagnosis and description of AOS. *Journal of Communication Disorders*. 2014; 51:43–50. [PubMed: 25092638]
- Strand EA, McNeil MR. Effects of length and linguistic complexity on temporal acoustic measures in apraxia of speech. *Journal of Speech and Hearing Research*. 1996; 39:1018–1033. [PubMed: 8898255]
- Vergis MK, Ballard KJ, Duffy JR, McNeil MR, Scholl D, Layfield C. An acoustic measure of lexical stress differentiates aphasia and aphasia plus apraxia of speech after stroke. *Aphasiology*. 2014; 28:554–575.
- Whitwell JL, Duffy JR, Strand EA, Xia R, Mandrekar J, Machulda MM, Josephs KA. Distinct regional anatomic and functional correlates of neurodegenerative apraxia of speech and aphasia: An MRI and FDG-PET study. *Brain and Language*. 2013; 125:245–252. [PubMed: 23542727]

- Wilson SM, Ogar JM, Laluz V, Growdon M, Jang J, Glenn S, Gorno-Tempini ML. Automated MRI-based classification of primary progressive aphasia variants. *Neuroimage*. 2009; 47:1558–1567. [PubMed: 19501654]
- Yorkston K, Strand EA, Miller R, Hillel A. Speech deterioration in amyotrophic lateral sclerosis: Implications for timing of intervention. *Journal of Medical Speech-Language Pathology*. 1993; 1:35–46.
- Ziegler W. Modeling the architecture of phonetic plans: Evidence from apraxia of speech. *Language and Cognitive Processes*. 2009; 24:631–661.
- Ziegler W, von Cramon D. Disturbed coarticulation in apraxia of speech: acoustic evidence. *Brain and Language*. 1986; 29:34–47. [PubMed: 3756460]

HIGHLIGHTS

- Diagnosis of primary progressive apraxia of speech (PPAOS) can be challenging
- Acoustic measures' ability to distinguish PPAOS from PPA was assessed
- Measures of duration, rate and lexical stress distinguished PPAOS from PPA
- Acoustic measures can serve as a confirmatory diagnostic marker for PPAOS

Figure 1.

Rate measure data for each group and the combined Control, agPPA, and lvPPA groups (Combined). The Box plots show the median, quartiles, and “whiskers out to the furthest data point within 1.5 interquartile ranges. Individual data points are noted by points along the boxplot.

Figure 2.

Box plot and ROC curve for PVI for *catastrophe*. Numeric data for cut-off value, sensitivity, specificity, and AUC are shown within the ROC curve.

Figure 3.

ROC curves for rate measures for each word and the PVI comparing the PPAOS with the remaining Combined groups. Numeric data for cut-off values, sensitivity, specificity, and AUC are shown within each ROC curve.

Table 1Demographic and speech-language data.^a

	Control (11)	agPPA (6)	lvPPA (20)	PPAOS (21)
<u>Demographic</u>				
Age @ assessment (yrs.) [*]	67.5 (10.8)	67.5 (11.0)	66.0 (10.4)	68.1 (9.1)
Sex (M/F)	8/3	2/4	10/10	10/11
Education (yrs.) [*]	15.5 (2.3)	15.0 (4.7)	14.7 (2.6)	16.2 (2.6)
Illness duration (yrs.) [*]	NA	1.8 (1.2)	3.4 (1.2)	3.6 (2.2)
<u>Language-Speech-Praxis</u>				
WAB AQ (100) ^{**}	NA	87.8 (5.5)	72.3 (18.4)	97.9 (2.0)
Token Test (22) ^{***}	NA	14.5 (3.9)	9.2 (6.1)	20.6 (1.4)
Boston Naming Test (15) ^{****}	NA	11.8 (3.3)	7.4 (5.1)	14.3 (0.9)
ASRS (64) ^{b*****}	NA	1.8 (1.5)	2.0 (1.8)	16.9 (6.9)
AOS Severity (0-4)	NA	0 (0)	0 (0)	1.4 (0.7)
MSD severity (10) ^c	NA	9.7 (0.5)	10 (0)	7.4 (1.0)
AES (% errors) [*]	NA	6.8 (5.1)	29.0 (28.8)	16.5 (19.5)
NVOA (32) ^{d *}	NA	28.8 (2.5)	25.7 (7.0)	28.1(5.7)

Abbreviations: WAB AQ = Western Aphasia Battery Aphasia Quotient; BNT = 15-item Boston Naming Test; ASRS = Apraxia of Speech Rating Scale (higher scores reflect greater prominence or severity); MSD severity = Motor Speech Disorder severity (10 = normal; 1 = nonvocal); AES = Articulatory Error Score; NVOA = Nonverbal Oral Apraxia; NA = not applicable/not assessed

^{*} ANOVA not significant ($p > 0.05$)

^{**} ANOVA significant ($p < 0.0001$); post hoc PPAOS > lvPPA

^{***} ANOVA significant ($p < 0.0001$); post hoc PPAOS > lvPPA & agPPA; agPPA > lvPPA

^{****} ANOVA significant ($p < 0.0001$); post hoc PPAOS > lvPPA; agPPA > lvPPA

^{*****} ANOVA significant ($p < 0.0001$); post hoc PPAOS < lvPPA & agPPA

^a Values expressed as the mean (standard deviation)

^b Scores above 8 are associated with the presence of AOS (Strand, Duffy, Clark, and Josephs, 2014).

^c MSD severity (score descriptor examples: 1 = nonvocal; 5 = frequent repetition or a "translator" commonly needed; 7 = speech consistently impaired but easily understood; 9 = nominal speech abnormality, only patient or significant other notices that speech has changed; 10 = normal.

^d Scores below 29 are below the range of normal performance (Botha et al., 2014).

Table 2

Acoustic temporal measures; values expressed as mean (standard deviation).

	Controls (N=11)	agPPA (N=6)	lvPPA N=20)	PPAOS (N=21)
<u>Duration (sec.)</u>				
Cat	0.41 (0.07)	0.34 (0.07)	0.36 (0.08)	0.38 (0.11)
Catnip	0.54 (0.07)	0.52 (0.07)	0.51 (0.19) ^a	0.73 (0.25) [*]
Catapult	0.61 (0.06)	0.65 (0.11)	0.62 (0.26) ^a	0.88 (0.31) ^{**}
Catastrophe	0.83 (0.07)	0.87 (0.08)	0.91 (0.34) ^a	1.52 (0.74) [*]
Stethoscope	0.83 (0.10)	0.91 (0.09)	0.89 (0.30) ^a	1.47 (0.54) [*]
Sentences	2.46 (0.41)	2.97 (0.55)	2.48 (1.30) ^a	4.77 (1.15) ^{a*}
<u>Syllable per sec.</u>				
Cat	2.50 (0.38)	3.04 (0.70)	2.96 (0.75)	2.80 (0.73)
Catnip	3.74 (0.46)	3.93 (0.53)	4.11 (1.45) ^a	2.96 (0.74) [*]
Catapult	4.95 (0.47)	4.72 (0.75)	4.95 (2.12) ^a	3.75 (1.10) ^{**}
Catastrophe	4.84 (0.41)	4.62 (0.43)	4.46 (1.70) ^a	3.08 (1.09) [*]
Stethoscope	3.68 (0.47)	3.32 (0.34)	3.43 (1.16) ^a	2.28 (0.74) [*]
Sentences	4.61 (0.68)	3.97 (1.90)	3.50 (1.75) ^{a***}	2.49 (0.65) ^{a*}
<u>PVI</u>				
Ca ta strophe	-91.77 (38.12)	-104.60 (13.58)	-82.73 (38.30)	-58.00 (22.94) [*]

Abbreviation: PVI = Pairwise Variability Index

^aBecause of missing data, for the lvPPA group n = 18, 16, 17, 18, and 7 for *catnip*, *catapult*, *catastrophe*, *stethoscope*, and *sentences*, respectively; n = 17 for PVI for *catastrophe*. For the PPAOS group n = 18 for *sentences*.

^{*}Significantly different from Control, agPPA, and lvPPA groups on post hoc tests ($p < 0.05$; see text for specific value).

^{**}Significantly different from Control and lvPPA groups on post hoc tests ($p < 0.05$; see text for specific value).

^{***}Significantly slower than Control group on post hoc test ($p < 0.05$; see text for specific value)

Table 3

Effect sizes (Cohen's d)^a for duration and rate measures for PPAOS versus Control, agPPA, and lvPPA groups.

	<u>Catnip</u>	<u>Catapult</u>	<u>Catastrophe</u>	<u>Stethoscope</u>	<u>Sentences</u>
<u>Duration</u>					
PPAOS – Control	1.05	1.24	1.30	1.66	1.60
PPAOS – agPPA	1.19	1.01	1.22	1.44	1.23
PPAOS – lvPPA	0.97	0.93	1.06	1.32	0.86
<u>Syllable Rate</u>					
PPAOS – Control	1.28	1.43	2.13	2.26	2.38
PPAOS – agPPA	1.51	1.03	1.85	1.80	0.97
PPAOS – lvPPA	0.76	0.72	0.97	1.18	0.45
<u>PVI</u>					
PPAOS – Control	NA	NA	1.07	NA	NA
PPAOS – agPPA	NA	NA	2.47	NA	NA
PPAOS – lvPPA	NA	NA	0.78	NA	NA

^aEffect sizes between 0.50-0.79 are generally considered medium, and above 0.80 large.

Table 4

Correlations between acoustic rate measures and PVI, and language, motor speech and nonverbal oral apraxia ratings for the PPAOS group.

	<u>Catnip</u>	<u>Catapult</u>	<u>Catastrophe</u>	<u>Stethoscope</u>	<u>Sentences</u>	<u>PVI</u>
ASRS	-0.70 *	-0.68 *	-0.61 *	-0.83 *	-0.64 *	0.31
AOS Severity	-0.62 *	-0.52 *	-0.61 *	-0.53 *	-0.26	0.47 *
MSD Severity	-0.44 *	-0.32	-0.35	-0.25	-0.03	0.27
AES	-0.46 *	-0.40 *	-0.36	-0.33	-0.17	0.36
NVOA	-0.24	-0.07	-0.14	-0.12	-0.10	0.10

*
P < 0.05