SPEECH METRICS THAT DIFFERENTIATE VARIANTS OF PRIMARY PROGRESSIVE APHASIA DEPEND ON LANGUAGE TYPOLOGY

Wendy Elvira-García[1], Victoria Marrero-Aguiar[2], Joan Borràs-Comes[1], Miguel Ángel Santos-Santos[3, 4]

[1]Universitat de Barcelona, [2]Universidad Nacional de Educación a Distancia (UNED), [3]Hospital de la Santa Creu i Sant Pau, [4]Biomedical Research Institute Sant Pau
wendyelvira@ub.edu, vmarrero@flog.uned.es, joanborras@ub.edu, MSantosS@santpau.cat

ABSTRACT

Primary Progressive Aphasia (PPA) variants can be distinguished using speech measures. In particular, a new one, syllabic-PVI, which computes syllable duration differences between unstressed and stressed adjacent syllables, can differentiate between the logopenic (lvPPA) and non-fluent (nfvPPA) variants.

The main goal of this study is to test if syllabic-PVI differentiates PPA variants in Spanish, a syllable-timed language like it does in English.

We analysed 5 repetitions of 5 multisyllabic words in a cohort of 37 patients (svPPA, lvPPA, nfvPPA) and 15 healthy controls. Data were annotated in Praat (manually revised) and metrics were computed in R by means of Generalized Linear Mixed Models.

Our results show that while syllable duration is a useful feature to distinguish lvPPA and nfvPPA in Spanish, syllabic PVI does not show significant differences between groups, which highlights the need of taking into account linguistic typology in linguistic batteries and clinical tests.

Keywords: clinical phonetics, Primary Progressive Aphasia speech, frontotemporal dementia, rhythm

1. INTRODUCTION

The Primary Progressive Aphasias (PPA) are a group of focal neurodegenerative syndromes primarily affecting language during the initial stages of the disease [1]. Recent studies have established the association between particular patterns of neuroanatomical damage and clinical presentation leading to the publication of the 2011 International consensus criteria for diagnosis of the three most common variants: non-fluent/agrammatic (nfvPPA), semantic (svPPA) and logopenic variants (lvPPA). Most cases of PPA at autopsy display frontotemporal lobar degeneration (FTLD)-type or Alzheimer's disease (AD) pathology. The classification of PPA cases into these clinical–anatomical phenotypes is important because these syndromes have predictive value for underlying molecular pathology (FTLD-tau in nfvPPA, FTLD-TDP in svPPA, and AD pathology in lvPPA), however these associations are not absolute, posing a significant impediment for the development of disease-modifying therapies based only on clinical presentation.

The diagnosis and differentiation of PPA is a challenge both for neurologists and speech pathologists. However, depending on the clinical syndrome, patients will tend to have semantic (svPPA), phonologic (lvPPA), or motor speech (apraxia of speech (AOS) +/- dysarthria) and grammatical (nfvPPA) problems. This makes AOS a crucial hint for differential diagnosis [2]–[4].

Research has shown that rhythmic and durational speech features are interesting to further distinguish subtypes of PPA [5] even automatically [2], [6], [7]. Specifically, previous research in English states that speech rate [6], [8]–[11], vowel duration [12] and Pairwise Variability Index (PVI) [13], [14] reflect the severity of AOS and therefore are useful speech metrics for characterizing PPA. For example, relative vowel duration in a polysyllable word repetition task can distinguish between lvPPA and nfvPPA [12].

When analyzing apraxia on isolated words, [15] proved that mean syllable duration is the most useful parameter if the stress position is not controlled, and vowel duration if the stress position is controlled. In a larger study, [5] tested average syllable duration and showed that it differs between lvPPA and nfvPPA.

More recent fine-grained research has focused on the duration of pre-stressed and stressed vowels' duration in apraxia. [16] found that apraxic patients' productions showed smaller differences in vowel duration between a first unstressed syllable and a following stressed one than controls which, perceptually, translates into an unclear lexical stressed syllable. Following this research a new metric that had proved valuable for patients suffering apraxia following a stroke, the syllable pairwise variability index (syl-PVI), [17] has been used for the differentiation between lvPPA and nfvPPA [5], [18] and has become a valuable metric in speech characterization of aphasic patients.

The Syl-PVI index is a metric that computes the difference between stressed and unstressed syllable pairs (weak-strong pattern), which linguistically highlights inaccuracies to express lexical stress.

As has been noted, Syl-PVI, highlights differences in duration between stressed and unstressed syllables, those differences are noticeably bigger in stress-timed languages than in syllable-timed languages but, to date, this metric has only been tested in stress-timed speaking populations. The goal of this study is to test

the syl-PVI index with a prototypical syllable-timed lanauge (Spanish) in order to test if this metric can be used in any language or its use should be limited to stress-timed languages.

2. METHODS

2.1. Materials

The linguistic battery used for this study is based on [19] which was adapted to Spanish in [20]. The protocol includes repetition of syllables, pseudowords, words, sentences, a reading of a phonetically balanced text (the Spanish version of "The North Wind and the Sun") the description of a cookie theft sheet and a picnic scene sheet. This study is centered on the word repetition task, which is known to yield less promising results than elicitation techniques. However, it is a faster and less demanding task for patients [5] [21] and has shown good results for characterizing AOS [15]. Of those materials, this study uses the repetition task of polysyllabic words (3 and 4 syllables), which includes words with the lexical stress in different positions (*patata* 'potato' [paˈt̪at̪a], *bicicleta 'bicycle'* [biθiˈkleta], *paquistaní* 'pakistani' [pakis̪t̪aˈni], *cúpula* 'Dome' [ˈkupula], *depósito* 'diposit' [deˈpos̪it̪o]).

2.2. Participants

52 participants were evaluated for this study (27 male, 25 female, mean age= 71.04 years). All of them were patients at the Hospital Clínic Sant Pau (Barcelona, Spain). All participants were tested with neurological and neuropsychological tests, and image diagnostics (computed tomography or MRI) and tested for biomarkers of neurodegeneration using the cerebrospinal fluid using the usual method at the Memory Unit of Sant Pau Hospital. The results of these tests were used to diagnose the patients by a neurologist of the Unit. Of the 52 participants, 14 were diagnosed with lvPPA, 16 with nfvPPA and 7 svPPA, 15 were controls. Since Barcelona is a social bilingual community the participants also completed a detailed bilingualism questionnaire.

2.3. Procedure

The speech sample was collected in one session (their first visit to the neurologist) using a Redenlab application and a microphone connected to an iPad. The sessions were carried out between 2019 and September 2022. Productions were semiautomatically annotated in Praat to segment by sound and by syllable. The by-sound annotation was done automatically and then manually corrected by 2 expert phoneticians. Once the by-sound annotation was done, a script was used to create the syllable boundaries [22].

Computation of duration and different metrics was done using R using the TextGrid as input. In order to compute the metrics only correct productions of the words were analysed. For example, if out of the 5 repetitions of *paquistaní*, a patient had produced [patitaˈni] in one instance, that instance has not been included in the database. This means that, while in general the analyses are based on 5 repetitions of the word, some patients may have less. In addition, as usual in clinical studies, some participants were unable to produce some words, the data of which PPA variants failed to repeat which words has been collected but its analysis is out of the scope of this paper.

The temporal parameters analysed were:

- Speech rate: computed as the number of sounds per second.
- Average syllable duration: mean duration of every syllable in each word.
- Syl-PVI: Syllabic pairwise-variability-index computes the normalized difference in duration between the stressed syllable of a word and the preceding one. Therefore, the word *cúpula* was excluded from the analysis. The final result is the mean of the 5 repetitions (when present) of each word [12].

The resulting metrics were normalized by word. The normalisation was performed using the orderNorm function from the bestNormalize package [23], which enforces a normal distribution. Then, the transformation was approximated using a cubic and logarithmic function. Finally, the obtained equation was applied to the original values. Then, data were analysed using Generalised Mixed Models [24] using the diagnosis group as a fixed factor, language and word as additional factors and the word by subject as a random slope.

3. RESULTS

The following section reports the results for syllable duration, speech rate and Syl-PVI.

Syllable duration behaves as expected, the control group had the smallest average syllable duration for the five polysyllabic words, followed by the lvPPA, nfvPPA and svPPA groups (Fig. 1). The durations are longer in nfvPPA compared to lvPPA (Cohen's d = 2.249, p=0.02). And the biggest effect is found for the word *depósito* (Cohen's d = 3.365, p<0.05). The rest of comparisons are not significant.

The results differ among words, svPPA shows the highest syllabic duration only for the word *depósito* 'diposit' a noun that forms a minimal pair with *deposito* "[I] deposit' (verb). *Paquistaní*, an oxytone

four-syllable word with a coda in the 3rd syllable shows longest syllables, especially in nfvPPA patients.

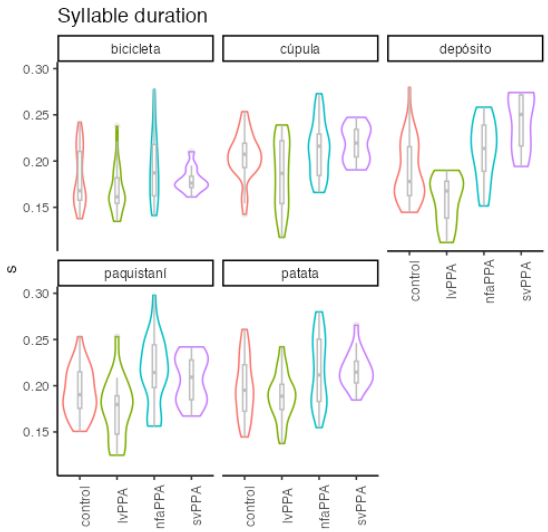

Figure 1: Average duration (s) box and violin plots for syllable duration for each diagnostic group by word.

Speech rate results (Fig. 2) follow the same line, which is expected given that all participants had repeated the same words (i.e. have produced the same number of sounds), the highest effect is found between lvPPA and nfvPPA but it does not reach significance (Cohen's d = 1.434, p=0.07). As noted by a reviewer, speech rate and average syllable duration should (and do) exhibit a strong inverse correlation. However, in the analysis conducted, we observed that only the average syllable duration showed statistical significance. This result can be attributed to the method used to calculate speech rate, the number of sounds produced per second. The duration of individual sounds can vary widely, as exemplified in the contrast between the length of a final vowel and a spirant approximant. In contrast, the duration of syllables exhibits relatively lower variation, contributing to the significant results observed for average syllable duration.

Fig. 3 illustrates the syl-PVI by word. According to the formula, lower (negative) values imply more syllable reduction at the pre-stressed position, whereas a value of 0 indicates the equal duration of both syllables. *Syllable reduction at the pre-stressed position is a phenomenon of "normal" speech in English which is absent or reduced in patients with AOS.* Therefore the lowest values should be expected for the control group, whereas apraxic patients should show values closer to 0. Our results show that nfvPPA tends to have the value closest to 0, but the effect size does not reach significance for any word.

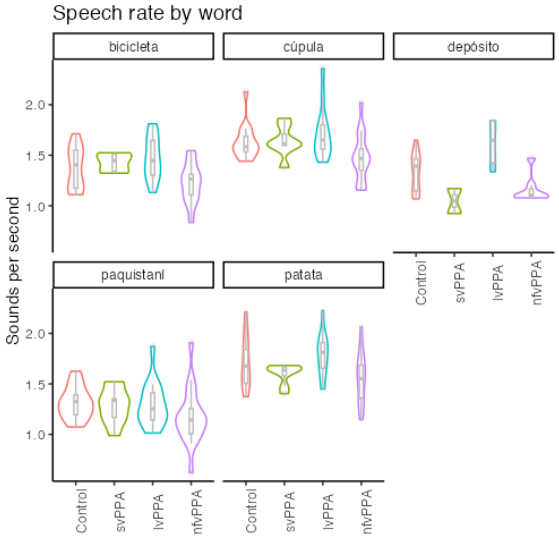

Figure 2: Speech rate box and violin plots for syllable duration for each diagnostic group by word.

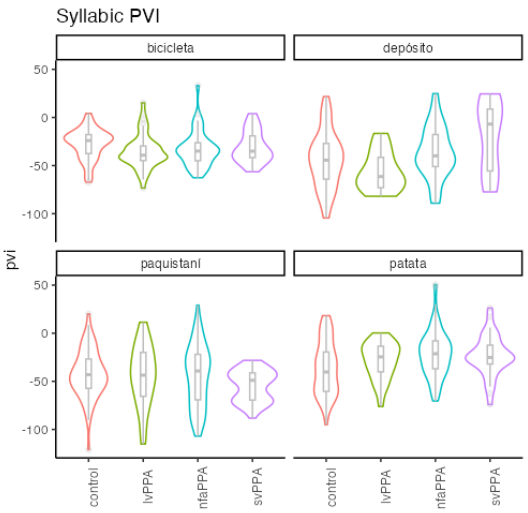

Figure 3: Average syl-PVI box and violin plots for each diagnostic group by word.

3. DISCUSSION

The results of the previous section show that in a polysyllabic repetition task, syllable duration is a relevant parameter while the Syl-PVI has no effect. The result is striking because previous studies have found that a diminished Syl-PVI is more specific to apraxia of speech whereas diminished speech rate can be found in more PPA variants. [5] found in their study that Syl-PVI was the most determinant factor for AOS even with fewer participants than in the present study: they included 10 patients with nfvPPA, 10 lvPPA and 5svPPA.

However, from a typological point of view, it is not surprising that there are no differences in syllabic-PVI in Spanish. While English is one of the most prototypical stressed-timed languages with big

durational differences between stressed and unstressed differences, Spanish is one of the most prototypical syllable-timed languages (sometimes called *pure*) in which duration and stress are always aligned [25].

Therefore, English apraxic patients who struggle with the expression of stress will have durational inaccuracies, failing to shorten unstressed syllables. Spanish syllables do not undergo this phonological process, each syllable being approximately equal. Moreover, we know that syllable-time rhythm production is easier as children produce it earlier [26] and it is less difficult to learn for foreign students than stress-timed timing (for example, we are seeing a change towards a syllable-timed rhythm in World Englishes [27]). In that sense, we could claim that apraxic patients make the "default" rhythmic pattern but that change is more noticeable in stress-timed languages.

However, if we take into account syllable duration instead of the syl-PVI we can see differences between the diagnosis groups. Those differences could be due to differences in timing or rhythm but also due to differences in speech rate given that all participants produced the same items. This is expected given that most research dealing with lvPPA and nfvPPA finds differences in speech rate notwithstanding the task (see section 1).

From a clinical point of view, the explanation why Syl-PVI is not significant in our results could lie in differences among the cohort and specifically their apraxia level. For this study, we have analysed the patient's first visit to the neurologist, which means that most patients are in an incipient stage of the disease and hence they could show mild apraxia in comparison with other studies. In order to minimize this and given that syl-PVI is primarily an indicator of the presence of apraxia, a follow-up study should be conducted correlating the level of AOS in these patients with their alteration of Syl-PVI. This way we could discard that our negative results are due to an incipient stage of the disease in our patients.

4. CONCLUSION

Our results show that syllable duration is useful for distinguishing the logopenic and non-fluent variants of PPA in Spanish, but syllabic PVI does not show significant differences between groups.

This is probably due to a typological difference between English (a prototypical stress-timed language) and Spanish (a prototypical syllable-timed language) given that the syllabic-PVI computes differences between the duration of unstressed and stressed syllables and those differences are smaller in syllable-timed languages. Our results highlight the necessity of taking into account language particularities when using clinical tests and metrics. And, in particular, this case shows the inadequacy of using syllabic-PVI in prototypical syllable-timed languages such as Spanish.

6. REFERENCES

[1] M. L. Gorno-Tempini *et al.*, "Classification of primary progressive aphasia and its variants," *Neurology*, vol. 76, no. 11, 2011, doi: 10.1212/WNL.0b013e31821103e6.

[2] C. Themistocleous, B. Ficek, K. Webster, D. B. den Ouden, A. E. Hillis, and K. Tsapkini, "Automatic subtyping of individuals with primary progressive aphasia," *Journal of Alzheimer's Disease*, vol. 79, no. 3, 2021, doi: 10.3233/JAD-201101.

[3] M. Á. Santos Santos and Gorno Tempini Maria L, "Primary Progressive Aphasia," *Brain Mapping: An Encyclopedic Reference.* . Elsevier, pp. 653–663, 2015.

[4] K. Croot, K. Ballard, C. E. Leyton, and J. R. Hodges, "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*, vol. 55, no. 5, 2012, doi: 10.1044/1092-4388(2012/11-0323).

[5] K. L. Haley *et al.*, "Speech metrics and samples that differentiate between nonfluent/ agrammatic and logopenic variants of primary progressive aphasia," *Journal of Speech, Language, and Hearing Research*, vol. 64, no. 3, 2021, doi: 10.1044/2020_JSLHR-20-00445.

[6] K. C. Fraser *et al.*, "Automated classification of primary progressive aphasia subtypes from narrative speech transcripts," *Cortex*, vol. 55, no. 1, 2014, doi: 10.1016/j.cortex.2012.12.006.

[7] C. Themistocleous *et al.*, "Acoustic markers of PPA variants using machine learning," *Front Hum Neurosci*, vol. 12, 2018, doi: 10.3389/conf.fnhum.2018.228.00092.

[8] M. Grossman *et al.*, "Disruption of large-scale neural networks in non-fluent/agrammatic variant primary progressive aphasia associated with frontotemporal degeneration pathology," *Brain Lang*, vol. 127, no. 2, 2013, doi: 10.1016/j.bandl.2012.10.005.

[9] C. Cordella, B. C. Dickerson, M. Quimby, Y. Yunusova, and J. R. Green, "Slowed articulation rate is a sensitive diagnostic marker for identifying non-fluent primary progressive aphasia," *Aphasiology*, vol. 31, no. 2, 2017, doi: 10.1080/02687038.2016.1191054.

[10] D. Gunawardena, S. Ash, C. McMillan, B. Avants, J. Gee, and M. Grossman, "Why are patients with progressive nonfluent aphasia nonfluent?," *Neurology*, vol. 75, no. 7, 2010, doi: 10.1212/WNL.0b013e3181ed9c7d.

[11] J. D. Rohrer, M. N. Rossor, and J. D. Warren, "Syndromes of nonfluent primary progressive aphasia: A clinical and neurolinguistic analysis," *Neurology*, vol. 75, no. 7, 2010, doi: 10.1212/WNL.0b013e3181ed9c6b.

[12] K. J. Ballard, S. Savage, C. E. Leyton, A. P. Vogel, M. Hornberger, and J. R. Hodges, "Logopenic and nonfluent variants of primary progressive aphasia are differentiated by acoustic measures of speech production," *PLoS One*, vol. 9, no. 2, 2014, doi: 10.1371/journal.pone.0089864.

[13] J. R. Duffy *et al.*, "Temporal acoustic measures distinguish primary progressive apraxia of speech from primary progressive aphasia," *Brain Lang*, vol. 168, 2017, doi: 10.1016/j.bandl.2017.01.012.

[14] J. R. Duffy, E. A. Strand, H. Clark, M. Machulda, J. L. Whitwell, and K. A. Josephs, "Primary progressive apraxia of speech: Clinical features and acoustic and neurologic correlates," *Am J Speech Lang Pathol*, vol. 24, no. 2, 2015, doi: 10.1044/2015_AJSLP-14-0174.

[15] K. L. Haley and A. Jacks, "Word-level prosodic measures and the differential diagnosis of apraxia of speech," *Clin Linguist Phon*, vol. 33, no. 5, 2019, doi: 10.1080/02699206.2018.1550813.

[16] D. I. Scholl, P. J. McCabe, R. Heard, and K. J. Ballard, "Segmental and prosodic variability on repeated polysyllabic word production in acquired apraxia of speech plus aphasia," *Aphasiology*, vol. 32, no. 5, 2018, doi: 10.1080/02687038.2017.1381876.

[17] A. Basilakos *et al.*, "A multivariate analytic approach to the differential diagnosis of apraxia of speech," *Journal of Speech, Language, and Hearing Research*, vol. 60, no. 12, 2017, doi: 10.1044/2017_JSLHR-S-16-0443.

[18] R. Landin-Romero *et al.*, "Brain changes underlying progression of speech motor programming impairment," *Brain Commun*, vol. 3, no. 3, 2021, doi: 10.1093/braincomms/fcab205.

[19] A. P. Vogel *et al.*, "Motor speech signature of behavioral variant frontotemporal dementia: Refining the phenotype," *Neurology*, vol. 89, no. 8, 2017, doi: 10.1212/WNL.0000000000004248.

[20] W. Elvira-García and V. M. Aguiar, "El Ritmo y la Tasa de Habla Como Ayuda Diagnóstica en Síndromes Neurodegenerativos de Los Lóbulos Fronto-Temporales," *Loquens*, vol. 7, no. 1, 2021, doi: 10.3989/LOQUENS.2020.068.

[21] C. Potagas *et al.*, "Silent Pauses and Speech Indices as Biomarkers for Primary Progressive Aphasia," *Medicina (Lithuania)*, vol. 58, no. 10, 2022, doi: 10.3390/medicina58101352.

[22] J. M. Garrido Almiñana, "SegProso: A Praat-Based Tool for the Automatic Detection a nd Annotation of Prosodic Boundaries," in *Proceedings of TRASP (Interspeech satellite)*, B. Bigi and D. Hirst, Eds., Aix-en-Provence, 2013, pp. 74–77.

[23] R. A. Peterson and J. E. Cavanaugh, "Ordered quantile normalization: a semiparametric transformation built for the cross-validation era," *J Appl Stat*, vol. 47, no. 13–15, 2020, doi: 10.1080/02664763.2019.1630372.

[24] B. Bolker, "Getting started with the glmmTMB package," *Cran.R-Project Vignette*, no. 2009, 2019.

[25] M. Nespor, M. Shukla, and J. Mehler, " Stress-Timed vs . Syllable-Timed Languages ," in *The Blackwell Companion to Phonology*, 2011. doi: 10.1002/9781444335262.wbctp0048.

[26] M. Canault *et al.*, "Syllable duration changes during babbling: A longitudinal study of French infant productions," *J Child Lang*, vol. 47, no. 6, 2020, doi: 10.1017/S030500092000015X.

[27] T. Nishihara and J. van de Weijer, "On Syllable-Timed Rhythm and Stress-Timed Rhythm in World Englishes: Revisited," *Miyagi kyōiku daigaku kiyō [Bulletin of Miyagi University of Education]*, vol. 46. 2011.