Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Comparing two picture naming tasks in primary progressive aphasia: Insights from behavioural and neural results

Cristina Polito [a,1], Francesca Conca [b,1], Gaia C. Santi [c], Valentina Esposito [b], Silvia P. Caminiti [d,e,f], Cecilia Boccalini [d,e], Valentina Berti [g], Carmen Morinelli [h], Salvatore Mazzeo [a,i], Alessandra Marcone [j], Sandro Iannaccone [j], Valentina Bessi [a,h,i], Sandro Sorbi [a,h,i], Daniela Perani [d,e,f], Stefano F. Cappa [b,c,*] and Eleonora Catricalà [c]

[a] IRCCS Fondazione Don Carlo Gnocchi, Florence, Italy
[b] IRCCS Mondino Foundation, Pavia, Italy
[c] ICoN Cognitive Neuroscience Center, Institute for Advanced Studies, IUSS, Pavia, Italy
[d] Vita-Salute San Raffaele University, 20132, Milan, Italy
[e] In Vivo Human Molecular and Structural Neuroimaging Unit, Division of Neuroscience, San Raffaele Scientific Institute, 20132, Milan, Italy
[f] Nuclear Medicine Unit, San Raffaele Hospital, 20132, Milan, Italy
[g] Nuclear Medicine, Department of Biomedical Experimental and Clinical Sciences "Mario Serio", University of Florence, Italy
[h] SOD Neurologia 1, Dipartimento Neuromuscolo-Scheletrico e Degli Organi di Senso, Azienda Ospedaliera-Universitaria Careggi, Florence, Italy
[i] NEUROFARBA, Department of Neuroscience, Psychology, Drug Research and Child Health, University of Florence, Florence, Italy
[j] Department of Rehabilitation and Functional Recovery, San Raffaele Hospital, Milan, Italy

ARTICLE INFO

Article history:
Received 29 October 2022
Reviewed 16 January
Revised 24 February 2023
Accepted 26 April 2023
Action editor Brad Dickerson
Published online 21 May 2023

ABSTRACT

Picture naming tests are widely used to evaluate language impairments in neurodegenerative diseases, especially in Primary Progressive Aphasia (PPA). The available tests differ for many factors affecting the performance, e.g. format of stimuli and their psycholinguistic properties. We aim to identify the most appropriate naming test to be used on PPA according to the clinical and research demands. We investigated the behavioural characteristics, i.e. proportion of correct responses and error type, and their neural correlates in two Italian naming tests, CaGi naming (CaGi) and naming subtest of the Screening for Aphasia in NeuroDegeneration battery (SAND), administered to 52 PPA patients who underwent an

* *Corresponding author*. IcoN Cognitive Neuroscience center, Institute for Advanced Studies, IUSS Pavia, Palazzo del Broletto, Piazza Vittoria 15, 27100, Pavia, Italy.
E-mail address: stefano.cappa@iusspavia.it (S.F. Cappa).
[1] equal contribution.
https://doi.org/10.1016/j.cortex.2023.04.014
0010-9452/© 2023 Elsevier Ltd. All rights reserved.

Keywords:
Primary progressive aphasia
Picture naming
Semantic errors

FDG-PET scan. We analysed the effectiveness of the tests in distinguishing PPA versus controls and among PPA variants, considering the psycholinguistic variables affecting performance. We explored the brain metabolic correlates of behavioural performance in the tests. SAND, differently from CaGi, has time limits for the response and its items are less frequent and acquired later. SAND and CaGi differed in terms of number of correct responses and error profile, suggesting a higher difficulty to name SAND items compared to CaGi. Semantic errors predominated in CaGi, while anomic and semantic errors were equally frequent in SAND. Both tests distinguished PPA from controls, but SAND outperformed CaGi in discriminating among PPA variants. FDG-PET imaging revealed a shared metabolic involvement of temporal areas associated with lexico-semantic processing, encompassing anterior fusiform, temporal pole, and extending to posterior fusiform in sv-PPA. Concluding, a picture naming test with response time limit and items which are less frequent and acquired later in life, as SAND, may be effective at highlighting subtle distinctions between PPA variants, improving the diagnosis. Conversely, a naming test without time limit for the response, as CaGi, may be useful for a better characterization of the nature of the naming impairment at the behavioural level, eliciting more naming errors than anomia, possibly helping in the development of rehabilitation protocols.

© 2023 Elsevier Ltd. All rights reserved.

1. Introduction

Picture naming is a multi-componential process encompassing dissociable yet interacting stages, from mapping the visual input to the semantic representation, to linking it to the word form for spoken output (Gleichgerrcht et al., 2015). It entails a widespread brain network, including the occipital regions, running along the fusiform and temporal gyri and the temporal pole, and further extending to the inferior frontal cortices (Catricalà et al., 2020; Migliaccio et al., 2016). Picture naming tests are widely used to assess language impairments in neurological patients, both after cerebrovascular accidents (McKinnon et al., 2018) and in different neurodegenerative diseases (Catricalà et al., 2020), including Primary Progressive Aphasia (PPA, Gorno-Tempini et al., 2011).

Different kinds of errors, e.g. visual, semantic, phonological, and anomia, may arise during the naming process. Although there is no one-to-one correspondence between the error type and the impaired processing stage, qualitative scoring provides relevant information about the neurocognitive components that are likely to be damaged (Balthazar et al., 2010; Schwartz et al., 2009). In our previous study on naming across neurodegenerative patients, we found: a) visual errors to be correlated with hypometabolism in the occipital lobes, involved in visual processing; b) semantic errors to be correlated with brain metabolism in the left fusiform and inferior and middle temporal gyri, extending to the temporal pole, regions which are known to be associated with conceptual knowledge; c) phonological errors to be correlated with hypometabolism in the left superior and middle temporal gyri, brain regions involved in the coding for the phonological forms (Catricalà et al., 2020).

Impaired naming is among the core diagnostic features of both the semantic variant of PPA (sv-PPA) and the logopenic variant PPA (lv-PPA) (Gorno-Tempini et al., 2011). Impaired naming qualitatively results in lexical-semantic errors and anomia. These kinds of errors are present since the early stages of the disease both in sv-PPA and in lv-PPA, but they may be attributed to different causes. In sv-PPA, naming errors have been ascribed to the progressive loss of semantic knowledge, emerging from the progressive involvement of the anterior temporal lobe (Catricalà et al., 2021; Jefferies & Lambon Ralph, 2006; Mion et al., 2010). Coordinate semantic errors reflect the initial blurring between the boundaries of concepts belonging to the same semantic category, resulting from the degradation of relevant distinctive features (Catricalà et al., 2015). As the disease progresses, additional features are lost and superordinate errors and anomia tend to predominate (Hodges & Patterson, 1995). Impairments at the lexical level have been also reported in this syndrome (Mesulam et al., 2013). In lv-PPA, lexico-semantic errors have been mainly attributed to a dysfunction in the retrieval of phonological word form, associated with cortical atrophy in the left middle and superior temporal gyri (Leyton et al., 2012, 2017, 2014; Win et al., 2017; Migliaccio et al., 2016) and resulting in anomia and phonological errors. Accordingly, lv-PPA patients may show preserved knowledge of the item, for example by producing accurate circumlocutions, and their performance is improved by cueing (Leyton & Hodges, 2013). A semantic origin of naming errors is however possible also in these patients (Rohrer et al., 2010; Teichmann et al., 2013). Phonological errors are associated with cortical thinning in the left posterior superior temporal gyrus (Leyton et al., 2014).

Naming is also affected in patients with the non-fluent variant of PPA (nfv-PPA), even though the impairment is generally less severe as compared to the other variants (Henry et al., 2016). In nfv-PPA, in addition to articulatory distortions, errors are mainly phonological, including omission, substitution, deletion, or insertion of phonemes (Migliaccio et al., 2016).

Picture naming tests are the most frequently used for the assessment of word retrieval abilities. The currently available picture naming tests differ in many factors potentially affecting the performance, such as the format of the stimuli

(e.g., coloured pictures as in Savage et al., 2013 *vs* black and white line drawings as in Kaplan et al., 1983, and Mckenna & Warrington, 1983), the number of stimuli (e.g., from 6 as in Lindeboom et al., 2002 to 80 as in Deloche & Hannequin, 1997), the presentation time (e.g., ranging from no limits, as in Mckenna & Warrington, 1983, and Savage et al., 2013, to 20 s as in Kaplan et al., 1983), as well as the psycholinguistic properties, i.e. differences in frequency, familiarity, and typicality. Previous studies in aphasia reported, although not univocally, several insights on how different variables may influence not only performance accuracy but also differences in error types (Cuetos et al., 2002; Kremin & Arabia, 2003; Kittredge et al., 2008; Nickels & Howard, 1995). Studies on neurodegenerative diseases have largely focused on sv-PPA (Barbarotto et al., 1995; Lambon Ralph et al., 1998; Woollams et al., 2008), and suggested that naming accuracy is predicted by typicality, frequency, familiarity, and age of acquisition. Anomic errors are elicited by items with low typicality, characterized by unusual features failing to activate any name, whereas semantic errors are more frequent in the case of high typicality items, whose features overlap among semantically related exemplars (Woollams et al., 2008).

Of note, different naming tests are used interchangeably, sometimes according to the usual practice of the clinical and research centres, without any clear agreement on which test is preferable in evaluating a specific performance in specific types of patients.

A recent systematic review reported in fact how not all naming tests are equally effective in discriminating, for example, lv-PPA from other PPA variants (Conca et al., 2022). Indeed, in the majority of cases, the Graded Naming Test (Mckenna & Warrington, 1983; Warrington, 1997) revealed no difference in naming accuracy between sv-PPA and lv-PPA, while other instruments, such as the DO80 naming test (Deloche & Hannequin, 1997), the Boston Naming Tests (Kaplan et al., 1983; Mack et al., 1992, pp. P154–P158), and the Northwestern Naming test (Thompson & Weintraub, 2014), showed a worse accuracy of sv-PPA patients. All these tests included black and white line drawings. While the Graded Naming Test and the Northwestern Naming test have no time limits for the response, the other two tests have time limits of 6 and 20 s, respectively. Different results have been also reported in comparing lv-PPA and nfv-PPA, with no difference in the Boston Naming Test (Kaplan et al., 1983), and a worse performance of lv-PPA than nfv-PPA in SYDBAT (Savage et al., 2013) and DO80 naming test (Deloche & Hannequin, 1997). SYDBAT, differently from the Boston Naming and DO80 tests, includes coloured photographs and has no time limit for the response. Crucially, the lack of information about the characteristics of the stimuli and the diagnostic metrics prevents a straightforward interpretation of the results. In fact, the characteristics of the naming tests may result in different weighting on components of the neural networks linked to specific aspects of language, such as visual perception, lexico-semantics, and phonology, leading to specific performance profiles in PPA variants.

The present study aims to investigate the behavioural characteristics, both in terms of proportion of correct responses and error type, of two Italian naming tests, i.e. CaGi naming subtest (CaGi, Catricalà et al., 2013) and Screening for Aphasia in NeuroDegeneration naming subtest (SAND, Catricalà et al., 2017), administered to a sample of PPA patients. We investigated the capacity of the two tests in discriminating PPA versus controls and among PPA variants, taking also into account the psycholinguistic variables affecting performances. We also explored the correlation between naming performance in both tests and regional brain metabolism as assessed by FDG-PET in PPA cases. Besides the research purpose, identifying the psycholinguistic characteristics of different naming tests, which best discriminate across PPA variants and best unveil the nature of the naming impairment, may improve the management of PPA patients in both diagnostic and rehabilitation settings.

2. Materials and methods

No part of the study procedures and analyses was preregistered prior to the research being conducted. No analysis code was used. Here we report all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. No sample size calculation was performed. The conditions of our ethics approval do not permit public archiving of raw brain imaging or behavioural data. Readers seeking access to these data can submit their queries to the corresponding author. Access may be granted in compliance with the study's procedures regulating the use of sensitive data. Any eventual reuse should undergo the completion of a formal agreement on data sharing. The materials and stimuli of CaGi and SAND naming tests are available in the Open Science Framework repository (https://osf.io/ufvge/).

2.1. *Participants*

Fifty-two PPA patients were retrospectively selected from the databases of the Neurology Unit of Careggi University Hospital of Florence ($N = 46$) and the Clinical Neuroscience Department, San Raffaele Hospital of Milan ($N = 6$). We included 15 sv-PPA, 20 lv-PPA, 7 nfv-PPA (Gorno-Tempini et al., 2011) and 10 mixed variants (specifically 6 sv/lv-PPA, 3 nfv/lv-PPA, 1 nfv/sv-PPA). The clinical diagnoses were made by expert neurologists (AM, SI, VB, SS, SFC) following the current guidelines (Gorno-Tempini et al., 2011). Amyloid biomarker data, i.e. CSF and Amyloid PET, were available only for a subset of patients, and not used for the diagnosis. In single subjects, behavioural, neuropsychological, and imaging assessment (i.e., conventional MRI and/or FDG-PET imaging) supported the clinical classification. Imaging data were visually inspected by an expert neurologist/neuroradiologist (DP), adhering to the criteria of Gorno-Tempini et al. (2011) for an imaging-supported diagnosis, i.e., involvement of the anterior temporal lobe for sv-PPA, the left posterior fronto-insular regions for nfv-PPA and the left posterior perisylvian or parietal regions for lv-PPA. See Table 1 for demographic and clinical information.

Sex, age, education, disease duration, and MMSE score did not differ between the three PPA variants (i.e. sv-PPA, nfv-PPA and lv-PPA) (all P values $> .064$).

Table 1 – Demographic and clinical information for each patients' group.

Patients group (number of patients)	sex, M/F	age, years, mean (SD)	education, years, mean (SD)	disease duration, months, mean (SD)	MMSE, Corrected score, mean (SD)
sv-PPA (15)	6/9	67.93 (6.29)	11.27 (3.54)	20.50 (9.66)	22.51 (5.61)
lv-PPA (20)	12/8	66.75 (7.07)	11.90 (3.93)	23.82 (12.31)	22.78 (3.79)
nfv-PPA (7)	6/1	71.71 (5.88)	11.43 (6.21)	26.50 (20.45)	23.18 (2.37)
PPA (all three variants) (42)	24/18	68.00 (6.69)	11.59 (4.14)	23.00 (12.80)	22.75 (4.29)
sv/lv-PPA (6)	5/1	73.33 (8.52)	12.50 (4.28)	33.67 (21.47)	22.94 (2.63)
nfv/lv-PPA (3)	0/3	71.67 (6.51)	6.67 (1.53)	28.67 (11.85)	23.95 (1.62)
nfv/sv-PPA (1)	0/1	67.00	7.00	24.00	16.90
All PPA patients (52)	29/23	68.81 (6.95)	11.33 (4.19)	24.74 (14.08)	22.73 (4.03)

sv-PPA = semantic variant of Primary Progressive Aphasia, lv-PPA = logopenic variant of Primary Progressive Aphasia, nfv-PPA = nonfluent variant of Primary Progressive Aphasia, SD = standard deviation, MMSE = Mini-Mental State Examination.

Eighty-four healthy participants were recruited as well; 42 of them were administered the CaGi (18 females, age = 66.24 ± 10.75; education = 9.74 ± 3.83) and 42 the SAND battery (18 females, age = 66.71 ± 8.58; education = 11.19 ± 4.35). The PPA group (i.e., sv-PPA, lv-PPA, nfv-PPA) and the respective control group were matched for age, education, and sex (all *P*-values at least >.446).

The study complies with all the provisions of Helsinki Declaration and received approval by the local ethics committees. All participants gave written informed consent to participate.

2.2. *Picture naming tests*

Participants were administered the picture naming tests of the CaGi (Catricalà et al., 2013) and SAND (Catricalà et al., 2017) batteries, widely used in clinical and research settings in Italy (Catricalà et al., 2014; 2015; 2013; 2020; 2021; 2022; Picillo et al., 2019; Battista et al., 2018). In both tests, participants were asked to orally name each visually presented stimulus. The modality of administration and scoring were however different (Catricalà et al., 2013; 2017). CaGi test consists of 48 coloured photographs to be named without time limits, and no cue was provided. For the calculation of the total correct responses, 1 point was given to each correctly named picture, for a maximum score of 48. SAND test consists of 14 black and white line drawings, and the participants had to respond within 6 s. A phonological cue was given only in the case in which patients produced an anomia, in all the other cases of incorrect response no cue was given. In the calculation of the total correct responses, 1 point was given if the correct name was produced within the 6 s time limit or if the patient self-corrected, .5 points if the patient named the picture correctly after the phonological cue, and 0 points if the patient made an error. The maximum score of correct responses was 14. A slightly different procedure was adopted for the calculation of the different types of error (semantic, phonological, visual, anomia), where responses produced after the cue were not considered, and the respective errors were classified as anomia.

Values of psycholinguistic variables, namely number of letters, phonemes, and syllables, phonological neighborhood, frequency (PhonItalia database, Goslin et al., 2014), Age of Acquisition (AoA) (Dell'Acqua et al., 2000), typicality (from Catricalà et al., 2013 for CaGi and from Dell'Acqua et al., 2000 for SAND), and syllabic complexity, measured as the number

Table 2 – Properties of CaGi and SAND naming tests; in bold significant differences (all *P*-values ≤.002); SD = standard deviation; AoA = Age of Acquisition;– = not tested.

Variables	CaGi	SAND	Mann-Whitney test, U (p-value)
stimuli	Coloured photographs	Black and white line drawings	-
number of items	48	14	-
time limit	No	6 seconds	-
number of letters, mean (SD)	7.08 (1.61)	6.64 (2.40)	268.50 (.248)
number of phonemes, mean (SD)	6.92 (1.54)	6.57 (2.34)	273.00 (.280)
number of syllables, mean (SD)	2.98 (.64)	2.93 (.83)	307.50 (.585)
phonological neighborhood, mean (SD)	3.13 (2.65)	3.21 (3.07)	323.00 (.823)
frequency, mean (SD)	**3.06 (1.38)**	**1.60 (.82)**	**115.00 (< .001)**
AoA, mean (SD)	**3.10 (.95)**	**4.15 (.84)**	**76.50 (.002)**
typicality, mean (SD)	5.88 (1.14)	4.06 (1.50)	-
syllabic complexity, mean (SD)	.55 (.26)	.67 (0.26)	256.50 (.171)

of consonant-vowel (CV) syllabic structures on the total number of syllables (Borleffs et al., 2017), are reported in Table 2. The items of the SAND were less frequent and acquired later than those of the CaGi (all *P*-values ≤.002) (see Table 2). Note that typicality was not considered as it was taken from two different databases.

In both CaGi and SAND, the following errors were considered and classified according to the procedures described in Catricalà et al. (2020):

- Visual (VIS) errors: responses visually similar to the target and from a different semantic category; whole-part responses where part of the target is named; gestures suggesting a visual error, for instance miming another object visually similar to the target; circumlocutions indicating an object visually similar to the target;
- Semantic (SEM) errors: semantic-within category errors, i.e. visually non-similar responses from the same category, superordinate, subordinate category errors, and circumlocutions;
- Phonological (PHON) errors: deletions, transpositions, or insertions of phonemes of the target word, phonemic conduites d'approche reaching or not reaching the target, fragments, neologisms, and verbal pharaphasias phonologically similar to target;
- Anomia (AN) errors: no responses.

Articulatory errors, mixed errors, e.g. ambiguous visual/phonologic or visual/semantic errors, perseverations, non-classifiable errors, and other errors, e.g. empty circumlocutions, were not included in the analyses.

2.3. *Analyses of behavioural data*

2.3.1. *Correct responses and error type*

Separately for CaGi and SAND, the performance of each patient was classified as impaired or spared according to the respective normative data, and the percentage of impaired cases was calculated. In the whole PPA group (i.e. the three PPA variants and mixed PPA, N = 52), in the three PPA variants considered together (N = 42), and separately for each variant (namely sv-PPA, lv-PPA, and nfv-PPA, but not mixed PPA), we compared the percentage of impaired cases between CaGi and SAND by means of Chi–Square tests. Mixed PPAs were only considered for the whole group analyses as this group was highly heterogeneous.

The proportions of correct responses (the number of correctly named stimuli), visual, semantic, phonological errors and anomia were obtained by dividing them by the number of trials, namely 48 in CaGi and 14 in SAND. In the whole PPA group, in the three PPA variants considered together, and separately for each variant, a Friedman test was performed to explore differences in the proportions of visual, semantic, phonological errors and anomia, in CaGi and SAND respectively. The Wilcoxon test using Bonferroni correction (.05/6 = .008) was used for post-hoc analyses. Differences between CaGi and SAND in terms of proportion of correct responses, visual, semantic, phonological errors and anomia were computed using Wilcoxon tests in the whole PPA group, in the three PPA variants considered together, and separately for each of the three variants.

Separately for CaGi and SAND, differences in the three PPA variants in terms of proportion of correct responses, visual, semantic, phonological errors and anomia were computed by means of Kruskall Wallis tests. Mann–Whitney test with Bonferroni correction (.05/3 = .017) was used for post-hoc analyses.

The distribution of the difference in the proportion of semantic and anomic errors between CaGi and SAND (semantic errors in CaGi-semantic errors in SAND, and anomia in CaGi-anomia in SAND, respectively) were also plotted at the single subject level.

2.3.2. *Psycholinguistic variables and test length*

Separately for CaGi and SAND, we accounted for possible influences of psycholinguistic variables on naming performance, in terms of correct responses and error type, with an item-level analysis. Considering the correct responses, we calculated the proportion of patients who named each item correctly out of the total number of patients (i.e. three PPA variants and mixed PPA, N = 52). We then used a multiple linear regression model with the proportion of correct responses as dependent variable and the psycholinguistic variables as independent ones. Considering error type, we calculated the proportion of patients who produced, respectively, a visual, semantic, phonological, and anomic error, out of the total number of patients (i.e. three PPA variants and mixed PPA, N = 52) for each item. In a general linear model, the proportions of each type of error were entered as dependent variables and the psycholinguistic variables as independent ones. Specifically, separately for CaGi and SAND, the following variables (converted into z-scores) were selected on the basis of their correlation with the proportion of correct responses and type of errors (i.e., at least with one error type), respectively, and then entered as independent variables: number of syllables, phonological neighborhood, frequency, AoA were entered as predictors of proportion of correct responses and error type for CaGi; number of syllables and typicality were entered as predictors of error type for SAND. Note that, to avoid multi-collinearity, when both the number of syllables and the number of letters were significantly correlated with proportion of correct responses and/or error type, we included only the number of syllables in the model as it displayed a higher correlation strength.

We further explored the effects of the test length on the performance, in terms of number of items, i.e. 14 items in SAND *vs* 48 items in CaGi. Specifically, we considered the proportion of correct responses for the first 14 items of CaGi and for the SAND items, to investigate possible differences. We thus used an ANCOVA model in order to account for the effect of test length on the performance, controlling for the two variables in which the tests differed, namely frequency and AoA. The proportion of the correct responses was entered as dependent variable, the test (i.e. CaGi *vs* SAND) as independent variable and the frequency and AoA as covariates.

2.3.3. *ROC analysis*

Receiver Operating Characteristic (ROC) analysis was used to assess the ability of CaGi and SAND naming tests, in terms of

proportion of correct responses (CR), and proportions of semantic, phonological, visual errors, and anomia, to discriminate between:

a) PPA (all three variants) and HC;
b) sv-PPA and nfv-PPA;
c) sv-PP and lv-PPA;
d) lv-PPA and nfv-PPA.

To directly compare the proportion of correct responses and error types, CR was converted into (1-CR) score. For each ROC analysis, we evaluated the accuracy using the area under the curve (AUC). When both CaGi and SAND AUCs were significant, or when different variables were significant within the same test, a non-parametric approach based on U-statistics was used to compare ROC curves (De Long et al., 1988). The Youden *J* Index (J), i.e. sensitivity + specificity −1, was computed in order to identify the optimal cut-off as a trade-off between sensitivity and specificity. We considered as acceptable the values above .5 (Fluss et al., 2005).

2.4. *Imaging methods*

2.4.1. *FDG-PET acquisition and data analysis*

Forty out of fifty-two PPA patients underwent [18 F]FDG-PET scans. The maximum distance between neuropsychological assessment and [18 F]FDG-PET was three months [18 F]FDG-PET acquisition was performed according to the EANM guidelines (Varrone et al., 2009) and following the same procedures adopted in previous studies of our group (see Catricalà et al., 2020; 2022). General Electric's Medical Systems Discovery-STE PETscan (in Milan) or Gemini TF PET/CT (Philips Medical Systems) scan (in Florence) were used. We previously adopted a multi-center data collection approach (Catricalà et al., 2020; 2022), based on the reported evidence that the validated optimized FDG-PET-based procedure is not affected by heterogeneity in the scanners adopted for image acquisition, thus representing a reliable method for longitudinal and multi-center studies (Presotto et al., 2017). The obtained FDG-PET images were then analysed using a standardized single-subject procedure implemented in statistical parametric mapping (SPM), involving: 1) spatial normalization adopting a dementia-specific template; 2) smoothing (isotropic 3D Gaussian kernel with a FWHM of 8 mm in each direction); and 3) intensity-normalization with a whole-brain reference region (Perani et al., 2014).

2.4.2. *Regions of Interest*

Six Regions of Interest (ROIs) were selected considering existing literature on naming and PPA: left occipital gyrus (Catricalà et al., 2020), posterior superior temporal gyrus (Migliaccio et al., 2016), posterior fusiform gyrus (Bruffaerts et al., 2020), anterior fusiform gyrus (Mion et al., 2010), temporal pole (Catricalà et al., 2020), inferior frontal gyrus (IFG) pars triangularis (BA 45) (Migliaccio et al., 2016). ROIs were created from the Automated Anatomical Labelling (AAL) atlas (Tzourio-Mazoyer et al., 2002), and adopted in previous studies of our group (Catricalà et al., 2020; 2022). Anterior and posterior sections of the fusiform gyrus and the posterior section of the superior temporal gyrus were defined in line with the boundaries in the AAL atlas and according to Visser et al. (2012), namely the MNI coordinates of posterior sections extended from $y = -24$ to -60 for the fusiform gyrus, and $y = -19$ to -51 for the superior temporal area. See Table 3 and Fig. 1 for the selected ROIs.

Correlation analyses (one-tailed) were used to explore the relation between the variables resulting as significant in the ROC analyses, separately for CaGi and SAND, and the metabolic values extracted from each brain ROI. Analyses were performed in the whole patients' sample (i.e. three PPA variants and mixed PPA, N = 40), and then separately for each PPA variant (sv-PPA, lv-PPA and nfv-PPA), in the latter with the application of bootstrapping (number of samples: 1000, CI: 95%). Bonferroni correction (.05/6 = .008) was adopted.

3. Results

3.1. *Behavioural results*

3.1.1. *Proportion of correct responses and error type*

The percentage of impaired cases, and the proportions of correct responses, visual, semantic, phonological errors, and anomia in the whole PPA group, in the three PPA variants considered together, and in each variant are shown in Table 4. In the whole PPA group ($p = .002$) and in the three PPA variants considered together ($p = .011$), a higher percentage of impaired cases was found in SAND as compared to CaGi, while no difference emerged when taking into account each variant separately (all *P*-values $\geq .070$).

Proportion of correct responses. Comparing SAND and CaGi, the proportion of correct responses was lower in SAND than CaGi in the whole PPA group, in the three PPA variants considered together, and also in sv-PPA and lv-PPA (all *P*-values $\leq .001$). SAND, but not CaGi, showed a lower proportion of correct responses in sv-PPA compared to lv-PPA ($p = .009$) and nfv-PPA ($p = .014$).

Error types. Using CaGi, in the whole PPA group and in the three variants considered together, semantic errors was the

Table 3 – List of the 6 ROIs with the respective MNI coordinates taken from AAL atlas.

laterality	Label	X	Y	Z
Left	Middle Occipital Gyrus	−32.39	−80.73	16.11
Left	Superior Temporal Gyrus (posterior section)*	−53.16	−20.68	7.13
Left	Middle Temporal Pole	−36.32	14.59	−34.08
Left	Fusiform Gyrus (posterior and anterior sections)*	−31.16	−40.3	−20.23
Left	Inferior Frontal Gyrus (triangularis)	−45.58	29.91	13.99

* = ROIs divided into anterior and posterior sections, see text for details.

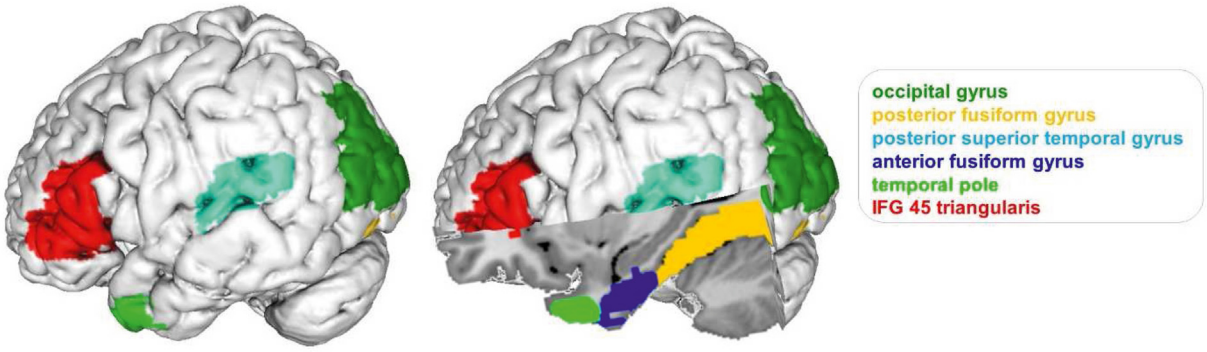

Fig. 1 – Anatomical 3D rendering of the 6 ROIs.

most frequent error type (all *P*-values ≤.001), while using SAND anomic and semantic errors were equally frequent (all *P*-values >.072), and more frequent than both visual and phonological errors (all *P*-values ≤.001); phonological errors were more frequent than visual errors (all *P*-values ≤.002). Phonological errors and anomias were more frequent in SAND than CaGi (all *P*-values ≤.041), while visual errors were more frequent in CaGi than SAND (all *P*-values ≤.014).

Considering separately the three variants using CaGi, semantic errors were the most frequent error type in sv-PPA (*P*-values ≤.001), while using SAND, anomias and semantic errors were equally frequent ($p = .379$), and more frequent than phonological errors (*P*-values ≤.004). A similar pattern was reported in lv-PPA, where semantic errors were the most frequent error type in CaGi (*P*-values <.001), whereas in SAND anomia and semantic errors were equally frequent ($p = .556$), and more frequent than visual, and in case of anomia, also than phonological errors (*P*-values ≤.001). In nfv-PPA, no differences emerged between error types. Comparing SAND and CaGi, both sv-PPA and lv-PPA showed a higher proportion of anomia in SAND than CaGi ($p \leq .001$), while a higher proportion of semantic errors in CaGi compared to SAND ($p = .042$) was found in the case of nfv-PPA.

Comparing the three groups using CaGi, sv-PPA showed a greater proportion of semantic errors than nfv-PPA ($p = .005$); using SAND, sv-PPA had a higher proportion of anomia ($p = .013$) and semantic errors ($p = .004$) than nfv-PPA, see Fig. 2.

Fig. 3 reports the difference between CaGi and SAND in the proportion of semantic and anomic errors at the single subject level in sv-PPA and lv-PPA. The same subject showed predominantly semantic errors in CaGi, while in SAND anomic errors were predominant.

3.1.2. *Psycholinguistic variables*

The proportion of correct responses of CaGi was significantly predicted by the number of syllables ($\beta = -.368$, $p = .003$) and AoA ($\beta = -.650$, $p < .001$).

Semantic errors in CaGi were predicted by number of syllables ($\beta = .463$, $p = .002$) and AoA ($\beta = .560$, $p = .001$); anomic errors were predicted by AoA ($\beta = .563$, $p = .002$) and syllabic complexity ($\beta = .287$, $p = .025$). AoA predicted visual ($\beta = .485$, $p = .016$) and phonological errors ($\beta = .581$, $p = .003$). In SAND, phonological errors were predicted by the number of syllables ($\beta = .419$, $p = .022$), and visual errors by the typicality ($\beta = -.697$, $p = .013$).

In addition, concerning the test length, the first 14 items of CaGi were more frequent and acquired earlier than the items of SAND (all *P*-values ≤.002), and they did not differ from the remaining CaGi items (all *P*-values ≥.186). A higher proportion of correct responses was reported in the first 14 items of CaGi than SAND ($p < .001$), in line with what we found considering all CaGi items (see above), and thus speaking against potential length effects. The difference was thus significant also after controlling for the effect of frequency ($p = .953$) and AoA ($p = .013$), with this last variable showing a significant effect.

3.1.3. *ROC*

All PPA (three variants) vs controls. PPA were discriminated from HC using the proportion of correct responses (CaGi, sensitivity: .83, specificity: .86; SAND, sensitivity: .71, specificity: .98), and semantic errors of both tests (CaGi, sensitivity: .76, specificity: .98; SAND, sensitivity: .74, specificity: .95) and anomia of SAND (sensitivity: .74, specificity: .88), see Table 5. In CaGi, the proportion of correct responses and semantic errors equally discriminated PPA from HC ($p = .218$) and both performed better than visual and anomic errors (all *P*-values <.001). In SAND, the proportion of correct responses, semantic errors, and anomia discriminated PPA from HC better than phonological errors (all *P* values < .001), without difference among each other (all *P*-values at least >.101), see Table 5.

Sv-PPA vs lv-PPA. Sv-PPA were discriminated from lv-PPA using the SAND proportion of correct responses (sensitivity: .67, specificity: .85), see Table 5.

Sv-PPA vs nfv-PPA. Sv-PPA were discriminated from nfv-PPA using the proportion of correct responses (CaGi, sensitivity: .73, specificity: .86; SAND, sensitivity: .67, specificity: .86) and semantic errors (CaGi, sensitivity: .87, specificity: .86; SAND, sensitivity: .67, specificity: 1) of both CaGi and SAND and with anomic errors of SAND (sensitivity: .87, specificity: .71), without differences within tests, see Table 5.

Lv-PPA vs nfv-PPA. Lv-PPA were discriminated from nfv-PPA only using SAND semantic errors (sensitivity: .70, specificity: .86), see Table 5.

See Table 5 for the results of the ROC analyses. Significant AUCs in both tests are represented in Fig. 4.

Table 4 – Percentage of impaired cases, and proportions of correct responses, visual, semantic, phonological errors and anomia in the whole PPA group, in the three PPA variants considered together, and in each variant for CaGi and SAND.

Patient group	CAGI						SAND					
	% of impaired cases	Correct responses, mean (SD)	SEM, mean (SD)	PHON, mean (SD)	VIS, mean (SD)	AN, mean (SD)	% of impaired cases	Correct responses, mean (SD)	SEM, mean (SD)	PHON, mean (SD)	VIS, mean (SD)	AN, mean (SD)
sv-PPA	86.67	.713 (.219)	.222 (.193)	.006 (.015)	.008 (.017)	.029 (.059)	60	.4 (.315)	.229 (.208)	.029 (.045)	.000 (.000)	.352 (.293)
lv-PPA	45	.828 (.132)	.129 (.092)	.008 (.028)	.011 (.014)	.022 (.030)	50	.695 (.250)	.114 (.140)	.032 (.043)	.004 (.016)	.161 (.184)
nfv-PPA	42.86	.833 (.238)	.068 (.064)	.039 (.094)	.009 (.011)	.033 (.069)	28.57	.827 (.314)	.010 (.027)	.051 (.135)	.000 (.000)	.082 (.162)
PPA (all 3 variants)	40.48	.788 (.190)	.152 (.143)	.012 (.043)	.010 (.015)	.026 (.048)	50	.611 (.324)	.354 (.312)	.034 (.065)	.002 (.011)	.216 (.245)
sv/lv-PPA	50	.663 (.316)	.253 (.294)	.024 (.043)	.003 (.009)	.069 (.114)	66.67	.518 (.390)	.167 (.246)	.024 (.058)	.024 (.058)	.274 (.291)
nfv/lv-PPA	66.67	.840 (.032)	.056 (.024)	.076 (.032)	.000 (.000)	.028 (.032)	100	.476 (.144)	.048 (.041)	.048 (.041)	.000 (.000)	.405 (.180)
nfv/sv-PPA	100	.250	.542	.000	.000	.021	100	.214	.714	.000	.000	.071
All PPA patients	59.62	.766 (.214)	.166 (.172)	.017 (.044)	.008 (.138)	.031 (.057)	79.31	.585 (.323)	.147 (.193)	.033 (.062)	.004 (.022)	.231 (.246)

% of Impaired cases = Percentage of impaired cases according to the published normative cut-offs (Catricalà et al., 2013; 2017), Correct responses = proportion of correct responses, SEM mean, PHON mean, VIS mean, AN mean = mean proportion of semantic, phonological, visual errors and anomia, respectively, sv-PPA = semantic variant of primary progressive aphasia, lv-PPA = logopenic variant of primary progressive aphasia, nfv-PPA = nonfluent variant of primary progressive aphasia, SD = standard deviation.

3.2. *FDG-PET imaging results*

Correlation analyses in the whole PPA sample revealed that the proportion of correct responses of both CaGi and SAND positively correlated with the brain metabolism in the anterior fusiform (CAGI: $r = .497$, $p = .001$; SAND: $r = .569$, $p < .001$) and temporal pole (CAGI: $r = .449$, $p = .002$; SAND: $r = .414$, $p = .004$). Semantic errors in both tests negatively correlated with the brain metabolism in the anterior fusiform gyrus (CaGi: $r = -.527$, $p < .001$; SAND: $r = -.416$, $p = .004$), extending to the temporal pole in the case of CaGi ($r = -.486$, $p = .001$). Anomias in SAND negatively correlated with brain metabolism in the anterior fusiform gyrus ($r = -.416$, $p = .004$). Considering only the sv-PPA, the proportion of correct responses ($r = .694$, $p = .006$) and semantic errors of CaGi ($r = -.706$, $p = .005$) correlated with the brain metabolism in posterior fusiform gyrus, respectively positively and negatively. No other effects were significant for the other variants. See Fig. 5 for the significant results.

4. Discussion

The present study aimed at identifying the psycholinguistic characteristics of different naming tests, which were most effective in discriminating among PPA variants and which could contribute to unveil the nature of the naming impairment in PPA for both research and clinical purposes. To this aim, we explored the behavioural characteristics and the neural underpinnings of the performance of PPA patients in two different tests developed by our group, the CaGi (Catricalà et al., 2013) and the SAND naming tests (Catricalà et al., 2017).

Taken together, the behavioural results indicated that SAND and CaGi differed in terms of proportion of correct responses and error profile. SAND is more difficult than CaGi, as revealed by a reduced proportion of correct responses; moreover, while in CaGi semantic errors predominated, in SAND both semantic and anomic errors were equally frequent. Both tests achieved high sensitivity and specificity in differentiating between PPA and healthy individuals. However, when considering the distinction among PPA variants, SAND outperformed CaGi: only SAND indeed allowed to differentiate between lv-PPA and sv-PPA and nfv-PPA using the proportion of correct responses and semantic errors, respectively.

There are several differences between the two tests, possibly accounting for the observed effects, which we consider in sequence: the format of the stimuli, i.e. coloured picture in CaGi *vs* black and white line drawings in SAND, the time limits for the response, i.e. 6 s in SAND and no time limits in CaGi; and the psycholinguistic properties of the pictures/words. Of note, it was not possible to disentangle the specific contribution of the effects of time limit and of colour, given the intrinsic characteristics of the two tests.

4.1. *Colour*

The stimuli are black and white line drawings in SAND and coloured photographs in CaGi. The amount of visuo-perceptual information conveyed by the stimuli, e.g. colour, texture details, has been suggested to affect naming

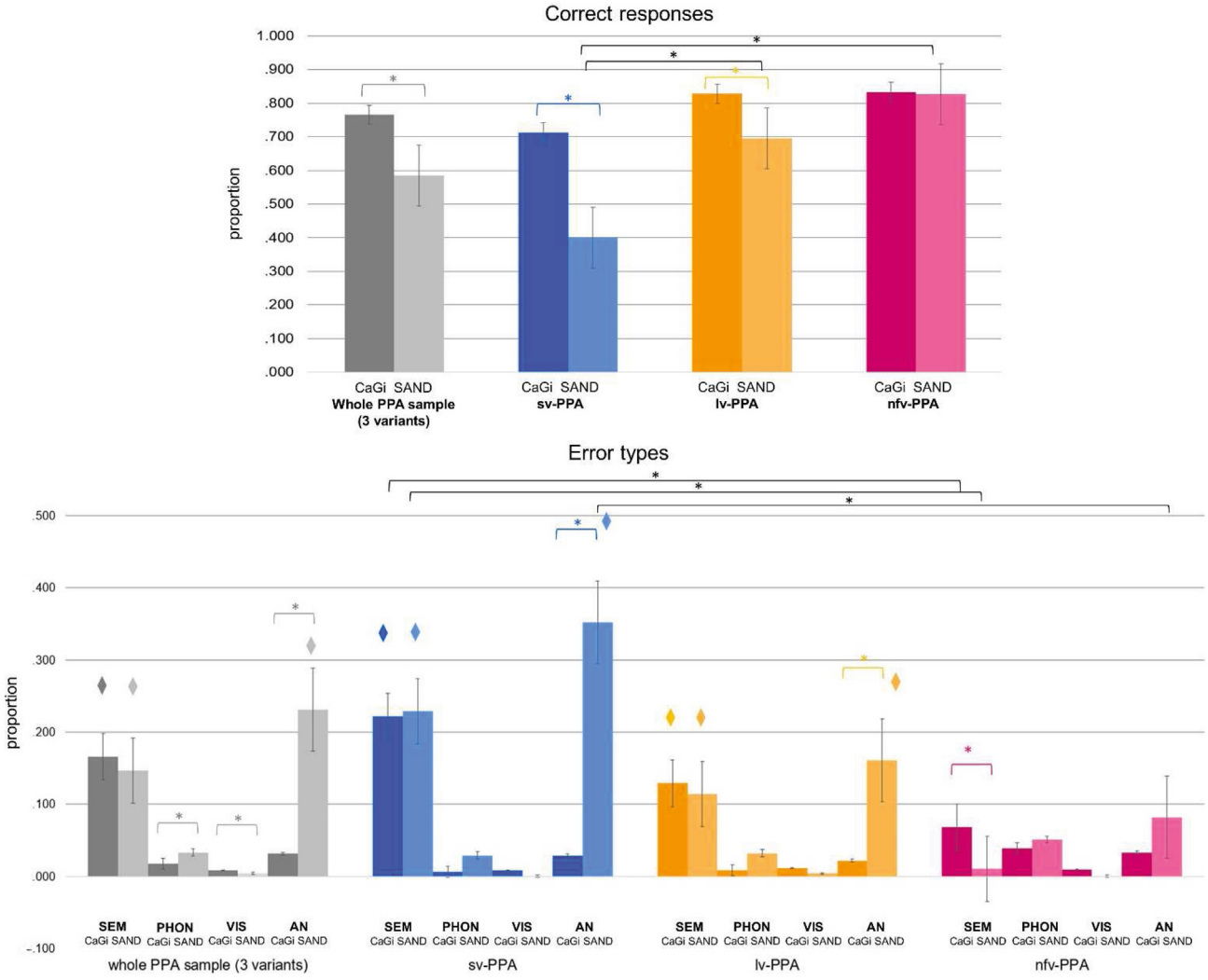

Fig. 2 – Proportions of correct responses (above) and error types (below) in CaGi and SAND in the whole PPA group (i.e., the three variants considered) and separately for each variant. Note that the results in the whole PPA group including the mixed cases ($n = 52$) are comparable to the ones found in the PPA sample including only the three variants, and consequently only the latter are reported in the figure. Diamonds indicate the most frequent error type in each PPA variant, separately for CaGi and SAND; bars indicate standard errors; * = significant, see text for details, sv-PPA = semantic variant of primary progressive aphasia, lv-PPA = logopenic variant of primary progressive aphasia, nfv-PPA = nonfluent variant of primary progressive aphasia, SEM = semantic errors, PHON = phonological errors, VIS = visual errors, AN = anomia.

performance, albeit not consistently (Zannino et al., 2010). Colour is thought to aid the process of perceptual differentiation among exemplars (Humphreys & Price, 2001), especially in cases where it is diagnostic for identifying the target among competitors endowed with a similar shape, as in the cases of fruits (e.g. plum *vs* apricot; Price & Humphreys, 1989). In healthy participants the addition of colour to Snodgrass and Vanderwart's stimuli entailed higher name agreement and reduced naming latencies, resulting in an advantage of 100 ms compared to the original black and white line drawings (Rossion & Pourtois, 2004). Mixed evidence emerged from patients' studies. In post-stroke individuals, line drawings depicting only the outline of the item elicited more anomic responses compared to coloured figures (Bisiach, 1966). No difference was reported between the latter type of stimuli and real objects by Corlew and Nation (1975), while an advantage in naming real objects compared to coloured photographs was reported by Fonseca et al. (2021). Surprisingly, a meta-analysis suggested that in Alzheimer's Disease (AD) the use of colour worsened naming performance of living items, while not affecting non-living ones (Laws et al., 2007), whereas, more recently, an advantage of colour for living items has been reported (Moreno-Martínez & Rodríguez-Rojo, 2015). As mentioned by the authors, these discrepancies may arise from methodological issues. In the studies included in Laws et al. (2007), the presence of ceiling effects in controls may have led to an overestimation of the impairments for living items in AD patients. Coloured images, being easier to name, may be especially susceptible to this effect. On the other hand, additional confounds may derive from uncontrolled differences between coloured and black-and-white stimuli in psycholinguistic variables. The higher proportion of correct responses

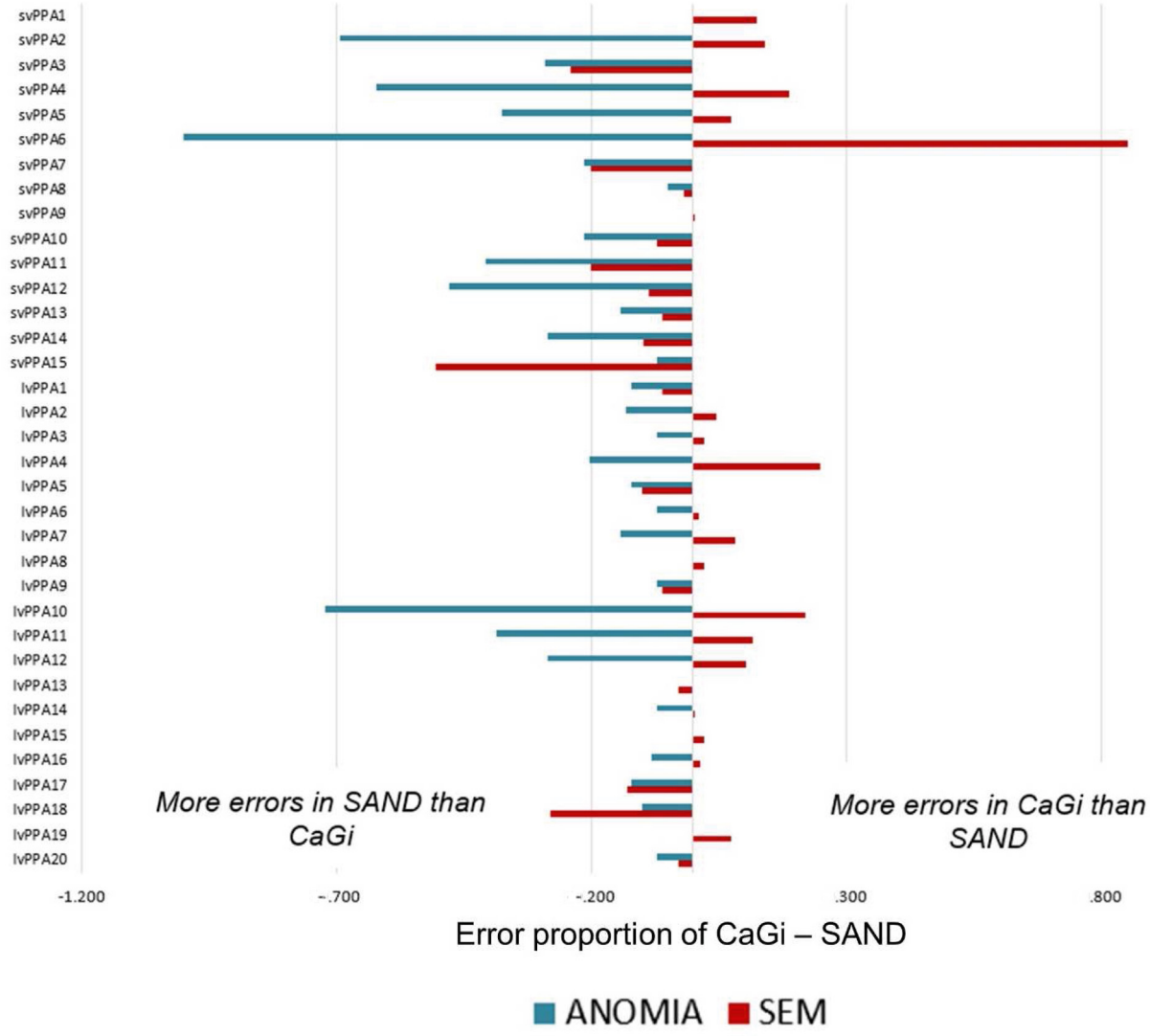

Fig. 3 – Distribution of the difference in error proportion between CaGi and SAND, considering semantic and anomic errors at the single subject level. Specifically, the bars in the right side of the graph indicated the presence of more errors in CaGi than SAND, while the ones in the left side indicated the presence of more errors in SAND than CaGi. SEM = semantic errors.

in naming CaGi coloured stimuli as compared to SAND black and white stimuli suggests that the presence of color stimuli in a naming test may represent a factor improving the performance.

4.1.1. *Time limit*

Studies in healthy participants suggested a decreased naming accuracy when time pressure or constraints for response production were introduced, either in the form of deadline time limits, e.g. maximum of 600 ms (Lloyd-Jones & Nettlemill, 2007; Vitkovitch et al., 1993; Vitkovitch & Humphreys, 1991), or when target name had to be produced according to a fast compared to slow tempo (i.e. tempo picture naming task, Hodgson and Lambon Ralph, 2008). According to some proposals, time limits are thought to constrain lexical selection (Vitkovitch & Humphreys, 1991), while others suggested an additional impact on phonological encoding, by reducing the processing phases devoted to checking and monitoring spoken output (Cai et al., 2020). Interestingly, in post-stroke patients a heterogeneous latency of error detection in picture naming has been found, lasting up to 4.5 s, in turn inducing either a successful or unsuccessful correction (Schuchard et al., 2017). Time limit in a naming test, as in the case of SAND, may prevent the production of the response (thus inducing anomia), and/or any possible attempts of correction which, on the contrary, may be possible when there are no time limits, as in the case of CaGi.

4.1.2. *Psycholinguistic variables*

The items of SAND are more frequent and acquired earlier in life than the ones of CaGi.

In CaGi, patients produced a higher number of correct responses for stimuli acquired early and with a short length. No variable influenced the proportion of correct responses in SAND. A greater accuracy and/or faster RTs were found for those items which were frequent, acquired early and short in

Table 5 – Results of the ROC analyses comparing all PPA with controls.

Comparisons	Error	CaGi			SAND			CaGi *vs* SAND
		AUC (*P* value)	Cut off	J (sens-spec)	AUC (*P* value)	Cut off	J (sens-spec)	Delta AUC (*P* value)
All PPA *vs* HC	CR	.903 (<.001)	44.5	.690 (.833–.857)	.852 (<.001)	11.5	.690 (.714–.976)	NT
	SEM	.880 (<.001)	3.4	.738 (.762–.976)	.861 (<.001)	1.5	.690 (.738–.952)	NT
	AN	.662 (.001)	.5*	.309 (.452–.857)	.944 (<.001)	.5	.619 (.738–.881)	NT
	PHON	.NS			.667 (<.001)	.5*	.333 (.333–1.000)	NT
	VIS	.642 (.001)	.5*	.286 (.381–.905)	NS			NT
sv-PPA *vs* lv-PPA	CR	NS			.760 (.003)	6.5	.517 (.667–.850)	NT
	AN	NS			.702 (.037)	2.5*	.417 (.667–.750)	NT
sv-PPA *vs* nfv-PPA	CR	.781 (.034)	39.5	.590 (.733–.857)	.829 (.002)	7.5	.524 (.667–.857)	−.048 (.448)
	SEM	.867 (<.001)	5.0	.724 (.867–.857)	.876 (<.001)	1.5	.667 (.667–1.000)	−.010 (.931)
	AN	NS			.829 (<.001)	.5	.581 (.867–.714)	NT
lv-PPA *vs* nfv-PPA	SEM	NS			.807 (<.001)	.5	.557 (.700–.857)	NT

sv-PPA with nfv-PPA, sv-PPA with lv-PPA, and between lv-PPA and nfv-PPA; AUC = area under the curve, NS = not significant, NT = not tested, *J* = Youden *J* index, sens = sensitivity, spec = specificity, * = explorative cut-off, i.e. when *J* < .5 (see Methods section for details); CR = correct responses, SEM = semantic errors, AN = anomias, sv-PPA = semantic variant of primary progressive aphasia, lv-PPA = logopenic variant of primary progressive aphasia, nfv-PPA = nonfluent variant of primary progressive aphasia.

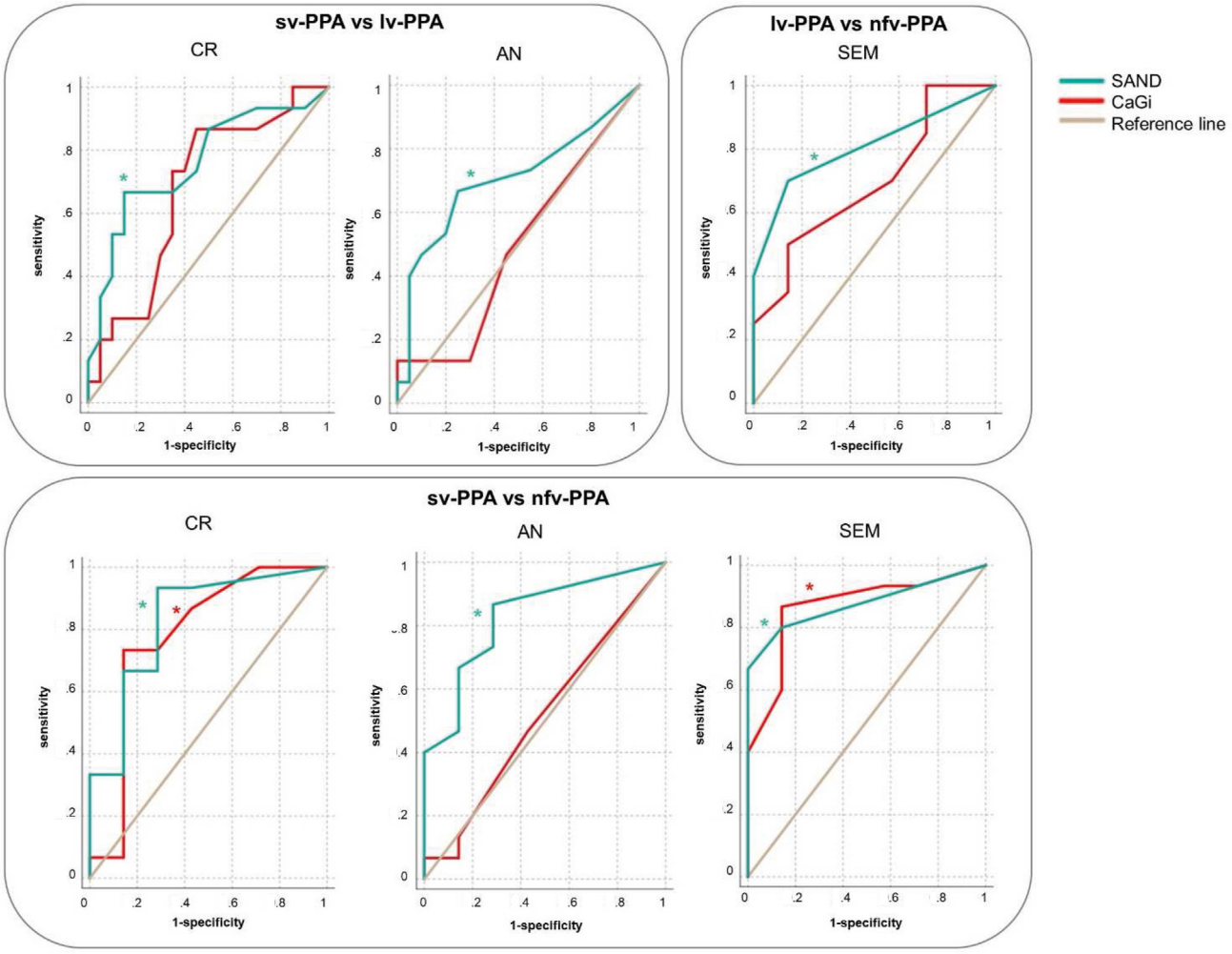

Fig. 4 – Graphics of the ROC comparing: sv-PPA with lv-PPA, sv-PPA with nfv-PPA, and lv-PPA with nfv-PPA, using correct responses, semantic errors, and anomia. * = significant AUC; CR = correct responses, SEM = semantic errors, AN = anomia, sv-PPA = semantic variant of primary progressive aphasia, lv-PPA = logopenic variant of primary progressive aphasia, nfv-PPA = nonfluent variant of primary progressive aphasia.

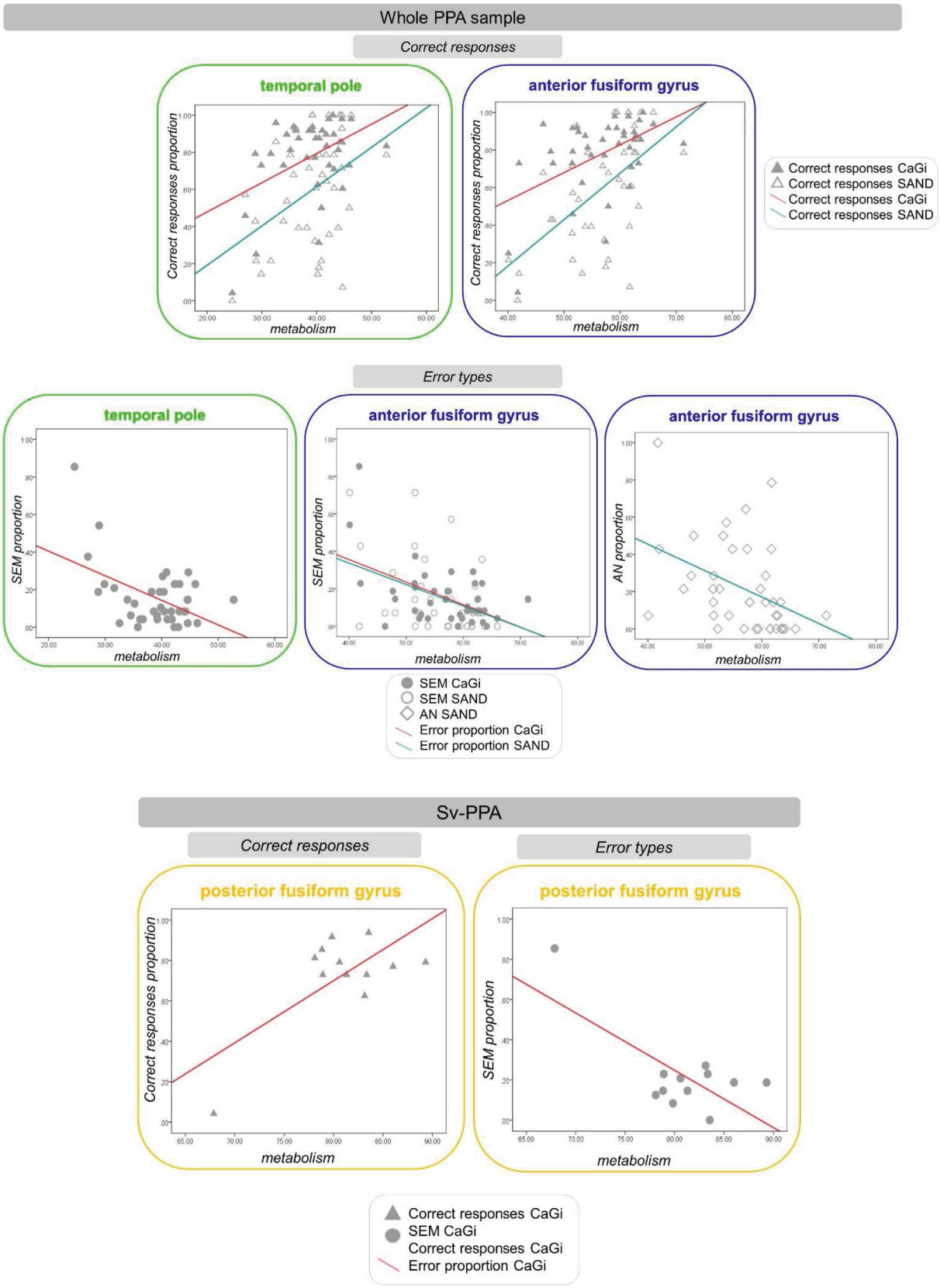

Fig. 5 – Results of the correlation analysis between behavioural data, i.e. proportions of correct responses, semantic and anomic errors in CaGi and SAND naming tests, and FDG-PET metabolism in the six Regions of Interest (ROIs); the colour of the border of boxes corresponds to the respective ROI; sv-PPA = semantic variant of Primary Progressive Aphasia; SEM = semantic errors; AN = anomias.

several studies in both healthy subjects and patients (Alario et al., 2004; Krautz & Keuleers, 2021; Barbarotto et al., 1995; Lambon Ralph et al., 1998; Kremin et al., 2001; Alyahaya et al., 2020).

There is additional evidence that the psycholinguistic properties of words/pictures may affect the qualitative aspects of naming performance, eliciting one type of error over the other(s). Our results suggested that in CaGi AoA predicted all the types of errors, namely visual, phonological, semantic and anomic errors, with the latter two also predicted by, respectively, word length and syllabic complexity. In the case of SAND, phonological errors were predicted by word length, visual errors by typicality. Notably, we did not find an effect of frequency in predicting naming performance.

The main discussion focuses on which variable most influences the naming performance and at which level in the naming process its effect occurs. Doubts have been cast regarding the presence of a genuine frequency effect on naming when the effect of AoA was controlled (Kremin et al., 2001; Kremin and Arabia 2003; Carrol and ,White 1973; Nickels & Howard, 1995), whereas other evidence suggested an effect of both variables (Alario et al., 2004; Cuetos et al., 2002). These heterogeneous results may arise from a combination of factors. For instance, frequency effects tend to be more commonly reported for more updated frequency norms and larger numbers of items (Alario et al., 2004; Snodgrass & Yuditsky, 1996). This issue appears however not to be relevant in our case, because frequency values for CaGi and SAND items were both taken from the recent paper by Goslin et al. (2014).

Moreover, it is debated at which level in the naming process the effects of frequency and AoA operate. Previous evidence in literature is inconsistent. An effect of AoA has indeed been reported on omissions but not on semantic, phonological and visual errors (Kremin & Arabia, 2003), or on omissions and phonological errors but not on semantic errors (Kittredge et al., 2008), or on semantic and phonological errors but not on omissions (Cuetos et al., 2002), or on semantic but not on phonological errors (Nickels & Howard, 1995). Moreover, an effect of frequency has been found on omissions but not on visual, semantic and phonological errors (Kremin & Arabia, 2003), or on omissions, semantic and phonological errors (Kittredge et al., 2008). Some proposals suggested a role of frequency, but not of AoA, on all the stages of naming process (Kittredge et al., 2008), while others limited the contribution of frequency to phonological retrieval (Jescheniak & Levelt, 1994). In our study, an effect of AoA (independent of the contribution of frequency) on the proportion of correct responses and error types, i.e., visual, semantic, anomic, and phonological errors was reported. In agreement with our results, AoA has been suggested to operate on different naming stages, i.e. object recognition (Cuetos et al., 2002), semantic processing (Nickels & Howard, 1995) and lexical-phonological retrieval (Lambon Ralph et al., 1998).

Of note, the role of AoA and frequency is not easy to be disambiguated as the two variables are correlated, i.e. words acquired earlier are usually also more frequent (Montefinese et al., 2019), and the two factors tend to be grouped together in dimensionality reduction studies of psycholinguistic properties (Alyahaya et al., 2020). Exploring the correlation between AoA and frequency in our two naming tests, a significant negative correlation was present in CaGi ($r = -.574$, $p < .001$) but not in SAND, suggesting that the selection of different sub-samples of items may lead to different relationships, to be explored and not just assumed for each new sample.

The effects of word length on semantic errors may be interpreted in light of the qualitative observation of Cuetos et al. (2002) in three aphasic patients who made a high proportion of semantic errors when presented with items with a high number of syllables. The authors suggested that when faced with long and difficult names, patients did not attempt to say the correct name, but indicated their recognition using a semantic circumlocution, possibly producing an easier related word resulting in a semantic error.

The effect of syllabic complexity is controversial, with some studies reporting a greater number of phonological errors for complex words in naming tasks (Schwartz et al., 2015) and others claiming that syllabic complexity effects are spurious, deriving from the lack of control over other psycholinguistic variables (Nickels & Howard, 2004). Specifically, while an initial study in healthy participants suggested faster naming latencies for words with a low onset complexity (Santiago et al., 2000), the result has been subsequently criticised and attributed to the confounding effect of word length (Roelofs, 2002).

Finally, the effect of word length on phonological errors is in line with previous data in stroke patients (Kittredge et al., 2008) and it is compatible with the view that longer words are more susceptible to errors in the correct organization of phonemes for spoken output. To the best of our knowledge, no studies reported the effect of typicality on visual errors, although converging evidence suggested that typical compared to less typical items are more accurately named, thanks to the availability of a higher number of overlapping and shared features (Rossiter & Best, 2013).

Summarizing, our results suggest that the psycholinguistic properties of the two naming tests affected patients' performance, both quantitatively and qualitatively. The effects were limited to the number of syllables and the typicality in the case of SAND. The low number of items ($n = 14$) may have reduced the possibility of finding additional significant results. Notably, we ruled out any potential effect of test length. Accordingly, in post-stroke aphasics, adding more than 20 items in a picture naming test did not increase the accuracy in predicting aphasia severity (Walker et al., 2022).

FDG-PET imaging findings revealed a partially shared metabolic involvement of temporal areas between the two naming tests, i.e. the anterior fusiform gyrus and the temporal pole, extending to the posterior fusiform gyrus in the case of sv-PPA.

The fusiform gyrus and temporal pole are typically affected by the pathological conditions responsible for PPA, particularly in sv-PPA (Bisenius et al., 2107; Brambati et al., 2009) and to a lesser extent also in lv-PPA, nfv-PPA (Lombardi et al., 2021; Matias-Guiu et al., 2015; Reyes et al., 2018; Tetzloff et al., 2018) and in mixed cases (Mazzeo et al., 2020). Both regions were associated with naming accuracy in PPA (Breining et al., 2022; Iaccarino et al., 2015; Migliaccio et al., 2016; Mion et al., 2010), with fusiform gyrus implied in semantic errors (Bruffaerts

et al., 2020) and omissions (Snowden et al., 2019), and temporal pole in semantic errors (Snowden et al., 2019).

These brain regions are crucial components of the network supporting conceptual knowledge (Catricalà et al., 2020), with a progressive increase in computational complexity and specificity proceeding anteriorly through the ventral pathway along the temporal lobe (Tyler et al., 2004). Posterior fusiform has been considered part of the network responsible for visual processing and colour perception (Howard et al., 1998; Zhang et al., 2016), and is associated with the features sharedness of concepts (Tyler et al., 2013). Its involvement in CaGi, but not in SAND, may be ascribed to the visual characteristics of the stimuli, which, in the case of CaGi, were endowed with rich visuo-perceptual details. Anterior fusiform has been involved in verbal semantics (Mion et al., 2010) as the representational hub integrating different kinds of information (Ralph, Jefferies, Patterson, & Rogers, 2017). In turn, the temporal pole has been implied in the representation of concepts at a high level of specificity (Bright et al., 2004) and in differentiating among similar exemplars (Rogers et al., 2006), thus accounting for its specific involvement in semantic errors in our results.

5. Conclusions

The assessment of language abilities still represents a challenge in clinical settings. Multiple factors, including the limited time resources, make the choice of a dedicated test of pivotal importance. Naming tests have been identified as the most adequate instrument to differentiate among PPA variants in some studies, while other failed to support these findings (Stockbridge et al., 2021). Our results suggest that these discrepancies may arise from the overlooked properties of the available naming tests.

The results of the present study, although deriving from two country-specific (Italian) naming tests, may be generalized to other tests sharing similar characteristics. Naming tests characterized by the presence of time limit (as in SAND) seem to be more prone to induce anomia, by reducing the possibility to produce a response, either correct or incorrect. Although anomic errors may originate from impairments in multiple stages of the naming process, in this study the brain metabolic correlates in the whole PPA sample involved the left anterior fusiform gyrus, suggesting a prevalent lexico-semantic origin.

Conversely, naming tests characterized by the absence of time limit for the response (as in CaGi) allowed participants to name the picture, either correctly or incorrectly. Brain metabolic correlates, selectively involving the posterior regions of the temporal lobe, suggest the recruitment of shared color and conceptual features processing (Tyler et al., 2013), possibly responsible for the production of a semantic error. Accordingly, behavioural data at the single patient level reveal that the same individual shows predominantly semantic errors in the test based on colour pictures and without time constraints, while anomic errors are more easily induced by black and white stimuli presented with time constraints to respond.

In conclusion, when we take into account the imaging data, the hypometabolic pattern is comparable, although not completely overlapping, between the two tests. Considering instead the behavioural data, in line with previous evidence (Staffaroni et al., 2021), naming tests which are more demanding given their intrinsic characteristics, i.e. presence of response time limits and psycholinguistic properties, may better discriminate among PPA variants, even if the number of items is limited. As a direct implication in clinical practice, this type of tests may be particularly suited for initial screening and diagnosis. On the other hand, naming tests with the availability of more time for the response offer the possibility to detect more representative error types (Catricalà et al., 2020), thus leading to a better characterization of the nature of the naming impairment, in contrast to the anomia produced in SAND. Although anomia is the most frequent error in general, its nature is also the least interpretable, as it could be the consequence of a disruption of several stages of the naming processes, i.e. phonological, lexical, semantic, visual. In our case, the uninterpretable anomia in SAND may become, for example, a clear semantic error in CaGi, shedding light into the nature of the impairment subtending single patients' performance. The aforementioned characteristics make tests similar to CaGi the most adequate for an in-depth characterization of naming deficits, as needed for example in the case of planning a rehabilitation protocol. Specifically, rehabilitation treatments aimed at ameliorating naming abilities in PPA usually target a specific naming sub-process according to the individual patient difficulties (Pagnoni et al., 2021).

Direct evidence on the effect of time limits in the characterization of the type of error may be obtained from future research introducing the same time limit in both coloured and black and white naming tests. Further studies are also needed to evaluate the different capacities of picture naming tests to longitudinally monitor language functions, as well as to extend the present findings to other neurodegenerative diseases. In addition, although, as previously mentioned, our results may largely apply outside one language-specific context, cross-linguistic differences need to be taken into consideration. For example, some factors, such as frequency, have been suggested to influence naming performance in several languages, while others may be associated with language-specific effects, e.g. word length effect is found Spanish, Italian, Hungarian, but not in English, German, Bulgarian (Bates et al., 2003).

Open practices section

The study in this article earned Open Material badge for transparent practices. The data and materials for this study are available at https://osf.io/ufvge/.

Funding

Italian Ministry of Health grant RCR-2021-23671214.

CRediT roles

Polito Cristina: Conceptualization, Writing–Original Draft. **Conca Francesca**: Formal analysis, Writing–Original Draft.

Santi Gaia Chiara: Formal analysis. **Esposito Valentina**: Data Curation.

Caminiti Silvia Paola: Formal analysis. **Boccalini Cecilia**: Formal analysis. **Berti Valentina**: Data Curation. **Morinelli Carmen**: Data Curation. **Mazzeo Salvatore**: Resources. **Marcone Alessandra**: Resources. **Iannaccone Sandro**: Resources. **Bessi Valentina**: Resources. **Sorbi Sandro**: Resources. **Perani Daniela**: Writing–Review & Editing. **Cappa Stefano Francesco**: Supervision, Writing–Review & Editing. **Catricalà Eleonora**: Conceptualization, Supervision, Writing–Review & Editing.

Declaration of competing interest

None.

REFERENCES

Alario, F., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers, 36*(1), 140–155.

Alyahya, R. S., Halai, A. D., Conroy, P., & Ralph, M. A. L. (2020). Mapping psycholinguistic features to the neuropsychological and lesion profiles in aphasia. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 124*, 260–273.

Balthazar, M. L. F., Yasuda, C. L., Pereira, F. R. S., Bergo, F. P. G., Cendes, F., & Damasceno, B. P. (2010). Coordinated and circumlocutory semantic naming errors are related to anterolateral temporal lobes in mild AD, amnestic mild cognitive impairment, and normal aging. *Journal of the International Neuropsychological Society: JINS, 16*(6), 1099–1107.

Barbarotto, R., Capitani, E., Spinnler, H., & Trivelli, C. (1995). Slowly progressive semantic impairment with category specificity. *Neurocase, 1*(2), 107–119.

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., … Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic bulletin & review, 10*(2), 344–380.

Battista, P., Catricalà, E., Piccininni, M., Copetti, M., Esposito, V., Polito, C., … Cappa, S. F. (2018). Screening for aphasia in neurodegeneration for the diagnosis of patients with primary progressive aphasia: Clinical validity and psychometric properties. *Dementia and Geriatric Cognitive Disorders, 46*(3–4), 243–252.

Bisenius, S., Mueller, K., Diehl-Schmid, J., Fassbender, K., Grimmer, T., Jessen, F., & FTLDc Study Group. (2017). Predicting primary progressive aphasias with support vector machine approaches in structural MRI data. *NeuroImage: Clinical, 14*, 334–343.

Bisiach, E. (1966). Perceptual factors in the pathogenesis of anomia. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior*, 90–95.

Borleffs, E., Maassen, B. A., Lyytinen, H., & Zwarts, F. (2017). Measuring orthographic transparency and morphological-syllabic complexity in alphabetic orthographies: A narrative review. *Reading and writing, 30*(8), 1617–1638.

Brambati, S. M., Rankin, K. P., Narvid, J., Seeley, W. W., Dean, D., Rosen, H. J., … Gorno-Tempini, M. L. (2009). Atrophy progression in semantic dementia with asymmetric temporal involvement: A tensor-based morphometry study. *Neurobiology of Aging, 30*(1), 103–111.

Breining, B. L., Faria, A. V., Caffo, B., Meier, E. L., Sheppard, S. M., Sebastian, R., … Hillis, A. E. (2022). Neural regions underlying object and action naming: Complementary evidence from acute stroke and primary progressive aphasia. *Aphasiology, 36*(6), 732–760.

Bright, P., Moss, H., & Tyler, L. K. (2004). Unitary vs multiple semantics: PET studies of word and picture processing. *Brain and language, 89*(3), 417–432.

Bruffaerts, R., Schaeverbeke, J., De Weer, A. S., Nelissen, N., Dries, E., Van Bouwel, K., … Vandenberghe, R. (2020). Multivariate analysis reveals anatomical correlates of naming errors in primary progressive aphasia. *Neurobiology of Aging, 88*, 71–82.

Cai, X., Ouyang, M., Yin, Y., & Zhang, Q. (2020). The effect of time pressure and semantic relatedness in spoken word production: A topographic ERP study. *Behavioural Brain Research, 387*, Article 112587.

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *The Quarterly Journal of Experimental Psychology: QJEP, 25*(1), 85–95.

Catricalà, E., Conca, F., Borsa, V. M., Cotelli, M., Manenti, R., Gobbi, E., … Cappa, S. F. (2021). Different types of abstract concepts: Evidence from two neurodegenerative patients. *Neurocase, 27*(3), 270–280.

Catricala, E., Della Rosa, P. A., Ginex, V., Mussetti, Z., Plebani, V., & Cappa, S. F. (2013). An Italian battery for the assessment of semantic memory disorders. *Neurological Sciences, 34*(6), 985–993.

Catricalà, E., Della Rosa, P. A., Plebani, V., Perani, D., Garrard, P., & Cappa, S. F. (2015). Semantic feature degradation and naming performance. Evidence from neurodegenerative disorders. *Brain and language, 147*, 58–65.

Catricalà, E., Della Rosa, P. A., Plebani, V., Vigliocco, G., & Cappa, S. F. (2014). Abstract and concrete categories? Evidences from neurodegenerative diseases. *Neuropsychologia, 64*, 271–281.

Catricalà, E., Gobbi, E., Battista, P., Miozzo, A., Polito, C., Boschi, V., … Garrard, P. (2017). SAND: A screening for aphasia in NeuroDegeneration. Development and normative data. *Neurological Sciences, 38*(8), 1469–1483.

Catricalà, E., Polito, C., Presotto, L., Esposito, V., Sala, A., Conca, F., … Perani, D. (2020). Neural correlates of naming errors across different neurodegenerative diseases: An FDG-PET study. *Neurology, 95*(20), e2816–e2830.

Catricalà, E., Santi, G. C., Polito, C., Conca, F., Esposito, V., Caminiti, S. P., … Cappa, S. F. (2022). Comprehensive qualitative characterization of linguistic performance profiles in primary progressive aphasia: A multivariate study with FDG-PET. *Neurobiology of Aging*, 137–148.

Conca, F., Esposito, V., Giusto, G., Cappa, S. F., & Catricalà, E. (2022). Characterization of the logopenic variant of primary progressive aphasia: A systematic review and meta-analysis. *Ageing Research Reviews*. , Article 101760. https://doi.org/10.1016/j.arr.2022.101760. Oct. 13.

Corlew, M. M., & Nation, J. E. (1975). Characteristics of visual stimuli and naming performance in aphasic adults. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 11*(2), 186–191.

Cuetos, F., Aguado, G., Izura, C., & Ellis, A. W. (2002). Aphasic naming in Spanish: Predictors and errors. *Brain and language, 82*(3), 344–365.

Dell'acqua, R., Lotto, L., & Job, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers, 32*(4), 588–615.

Deloche, G., & Hannequin, D. (1997). *Test de dénomination orale d'image: DO 80*. ECPA, Les éditions du Centre de psychologie appliquée.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver

operating characteristic curves: A nonparametric approach. *Biometrics*, 837–845.
Fluss, R., Faraggi, D., & Reiser, B. (2005). Estimation of the Youden Index and its associated cutoff point. *Biometrical Journal: Journal of Mathematical Methods in Biosciences, 47*(4), 458–472.
Fonseca, J., Miranda, F. D., Leal, G., Melo, T. P., & Martins, I. P. (2021). Aphasia assessment: Impact of material on naming performance. *Arquivos de Neuro-Psiquiatria, 79*, 774–780.
Gleichgerrcht, E., Fridriksson, J., & Bonilha, L. (2015). Neuroanatomical foundations of naming impairments across different neurologic conditions. *Neurology, 85*(3), 284–292.
Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., … Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology, 76*(11), 1006–1014.
Goslin, J., Galluzzi, C., & Romani, C. (2014). PhonItalia: A phonological lexicon for Italian. *Behavior Research Methods, 46*(3), 872–886.
Henry, M. L., Wilson, S. M., Babiak, M. C., Mandelli, M. L., Beeson, P. M., Miller, Z. A., & Gorno-Tempini, M. L. (2016). Phonological processing in primary progressive aphasia. *Journal of cognitive neuroscience, 28*(2), 210–222.
Hodges, J. R., & Patterson, K. (1995). Is semantic memory consistently impaired early in the course of alzheimer's disease? Neuroanatomical and diagnostic implications. *Neuropsychologia, 33*(4), 441–459.
Hodgson, C., & Ralph, M. A. L. (2008). Mimicking aphasic semantic errors in normal speech production: Evidence from a novel experimental paradigm. *Brain and Language, 104*(1), 89–101.
Howard, R. J., Barnes, J., McKeefry, D., Ha, Y., Woodruff, P. W., Bullmore, E. T., … Brammer, M. (1998). The functional anatomy of imagining and perceiving colour. *Neuroreport, 9*(6), 1019–1023.
Humphreys, G. W., & Price, C. J. (2001). Cognitive neuropsychology and functional brain imaging: Implications for functional and anatomical models of cognition. *Acta Psychologica, 107*(1–3), 119–153.
Iaccarino, L., Crespi, C., Della Rosa, P. A., Catricalà, E., Guidi, L., Marcone, A., … Perani, D. (2015). The semantic variant of primary progressive aphasia: Clinical and neuroimaging evidence in single subjects. *Plos One, 10*(3), Article e0120197.
Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: A case-series comparison. *Brain: a Journal of Neurology, 129*(8), 2132–2147.
Jescheniak, J. D., & Levelt, W. J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 824.
Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston naming test*. Philadelphia: Lea & Febiger.
Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive neuropsychology, 25*(4), 463–492.
Krautz, A. E., & Keuleers, E. (2021). LinguaPix database: A megastudy of picture-naming norms. *Behavior Research Methods*, 1–14.
Kremin, H., & Arabia, C. (2003). The relative effects of imageability and age-of-acquisition on aphasic misnaming. *Brain and langage*, (87), 33–34.
Kremin, H., Perrier, D., De Wilde, M., Dordain, M., Le Bayon, A., Gatignol, P., … Arabia, C. (2001). Factors predicting success in picture naming in Alzheimer's disease and primary progressive aphasia. *Brain and cognition, 46*(1–2), 180–183.
Lambon Ralph, M. A., Howard, D., Nightingale, G., & Ellis, A. W. (1998). Are living and non-living category-specific deficits causally linked to impaired perceptual or associative knowledge? Evidence from a category-specific double dissociation. *Neurocase, 4*(4–5), 311–338.
Laws, K. R., Adlington, R. L., Gale, T. M., Moreno-Martínez, F. J., & Sartori, G. (2007). A meta-analytic review of category naming in Alzheimer's disease. *Neuropsychologia, 45*(12), 2674–2682.
Leyton, C. E., Ballard, K. J., Piguet, O., & Hodges, J. R. (2014). Phonologic errors as a clinical marker of the logopenic variant of PPA. *Neurology, 82*(18), 1620–1627.
Leyton, C. E., & Hodges, J. R. (2013). Towards a clearer definition of logopenic progressive aphasia. *Current Neurology and Neuroscience Reports, 13*(11), 1–7.
Leyton, C. E., Hodges, J. R., Piguet, O., & Ballard, K. J. (2017). Common and divergent neural correlates of anomia in amnestic and logopenic presentations of Alzheimer's disease. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 86*, 45–54.
Leyton, C. E., Piguet, O., Savage, S., Burrell, J., & Hodges, J. R. (2012). The neural basis of logopenic progressive aphasia. *Journal of Alzheimer's Disease, 32*(4), 1051–1059.
Lindeboom, J., Schmand, B., Tulner, L., Walstra, G., & Jonker, C. (2002). Visual association test to detect early dementia of the Alzheimer type. *Journal of Neurology, Neurosurgery, and Psychiatry, 73*(2), 126–133.
Lloyd-Jones, T. J., & Nettlemill, M. (2007). Sources of error in picture naming under time pressure. *Memory & cognition, 35*(4), 816–836.
Lombardi, J., Mayer, B., Semler, E., Anderl-Straub, S., Uttner, I., Kassubek, J., … FTLD consortium.. (2021). Quantifying progression in primary progressive aphasia with structural neuroimaging. *Alzheimer's & Dementia, 17*(10), 1595–1609.
Mack, W. J., Freed, D. M., Williams, B. W., & Henderson, V. W. (1992). Boston naming test: Shortened versions for use in alzheimer's disease. *Journal of gerontology, 47*(3), P154–P158.
Matias-Guiu, J. A., Cabrera-Martín, M. N., Moreno-Ramos, T., García-Ramos, R., Porta-Etessam, J., Carreras, J. L., & Matías-Guiu, J. (2015). Clinical course of primary progressive aphasia: Clinical and FDG-PET patterns. *Journal of Neurology, 262*(3), 570–577.
Mazzeo, S., Polito, C., Padiglioni, S., Berti, V., Bagnoli, S., Lombardi, G., … Bessi, V. (2020). Linguistic profiles, brain metabolic patterns and rates of amyloid-β biomarker positivity in patients with mixed primary progressive aphasia. *Neurobiology of Aging, 96*, 155–164.
Mckenna, P., & Warrington, E. K. (1983). *Graded naming test: Manual*. NFER-Nelson.
McKinnon, E. T., Fridriksson, J., Basilakos, A., Hickok, G., Hillis, A. E., Spampinato, M. V., … Bonilha, L. (2018). Types of naming errors in chronic post-stroke aphasia are dissociated by dual stream axonal loss. *Scientific reports, 8*(1), 1–12.
Mesulam, M. M., Wieneke, C., Hurley, R., Rademaker, A., Thompson, C. K., Weintraub, S., & Rogalski, E. J. (2013). Words and objects at the tip of the left temporal lobe in primary progressive aphasia. *Brain: a Journal of Neurology, 136*(2), 601–618.
Migliaccio, R., Boutet, C., Valabregue, R., Ferrieux, S., Nogues, M., Lehéricy, S., … Teichmann, M. (2016). The brain network of naming: A lesson from primary progressive aphasia. *Plos One, 11*(2), Article e0148707.
Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y. T., … Nestor, P. J. (2010). What the left and right anterior fusiform gyri tell us about semantic memory. *Brain: a Journal of Neurology, 133*(11), 3256–3268.
Montefinese, M., Vinson, D., Vigliocco, G., & Ambrosini, E. (2019). Italian age of acquisition norms for a large set of words (ItAoA). *Frontiers in psychology, 10*, 278.
Nickels, L., & Howard, D. (1995). Aphasic naming: What matters? *Neuropsychologia, 33*(10), 1281–1303.

Nickels, L., & Howard, D. (2004). Dissociating effects of number of phonemes, number of syllables, and syllabic complexity on word production in aphasia: It's the number of phonemes that counts. *Cognitive Neuropsychology, 21*(1), 57–78.

Pagnoni, I., Gobbi, E., Premi, E., Borroni, B., Binetti, G., Cotelli, M., & Manenti, R. (2021). Language training for oral and written naming impairment in primary progressive aphasia: A review. *Translational Neurodegeneration, 10*(1), 1–34.

Perani, D., Della Rosa, Cerami, C., Gallivanone, F., Fallanca, F., , … Vanoli, E. G., & Eadc-pet Consortium. (2014). Validation of an optimized SPM procedure for FDG-PET in dementia diagnosis in a clinical setting. *NeuroImage: Clinical, 6*, 445–454.

Picillo, M., Cuoco, S., Carotenuto, I., Abate, F., Erro, R., Volpe, G., … Barone, P. (2019). Clinical use of SAND battery to evaluate language in patients with progressive supranuclear palsy. *Plos One, 14*(10), Article e0223621.

Presotto, L., Ballarini, T., Caminiti, S. P., Bettinardi, V., Gianolli, L., & Perani, D. (2017). Validation of 18F–FDG-PET Single-subject optimized SPM procedure with different PET scanners. *Neuroinformatics, 15*(2), 151–163.

Price, C. J., & Humphreys, G. W. (1989). The effects of surface detail on object categorization and naming. *The Quarterly Journal of Experimental Psychology: QJEP, 41*(4), 797–828.

Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience, 18*(1), 42–55.

Roelofs, A. (2002). Syllable structure effects turn out to be word length effects: Comment on Santiago et al.(2000). *Language and Cognitive Processes, 17*(1), 1–13.

Rogers, T. T., Hocking, J., Noppeney, U. T. A., Mechelli, A., Gorno-Tempini, M. L., Patterson, K., & Price, C. J. (2006). Anterior temporal cortex and semantic memory: Reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective & Behavioral Neuroscience, 6*(3), 201–213.

Rohrer, J. D., Ridgway, G. R., Crutch, S. J., Hailstone, J., Goll, J. C., Clarkson, M. J., … Warren, J. D. (2010). Progressive logopenic/phonological aphasia: Erosion of the language network. *Neuroimage, 49*(1), 984–993.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception, 33*(2), 217–236.

Rossiter, C., & Best, W. (2013). "Penguins don't fly": An investigation into the effect of typicality on picture naming in people with aphasia. *Aphasiology, 27*(7), 784–798.

Santiago, J., MacKay, D. G., Palma, A., & Rho, C. (2000). Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language and Cognitive Processes, 15*(1), 1–44.

Savage, S., Hsieh, S., Leslie, F., Foxe, D., Piguet, O., & Hodges, J. R. (2013). Distinguishing subtypes in primary progressive aphasia: Application of the sydney language battery. *Dementia and Geriatric Cognitive Disorders, 35*(3–4), 208–218.

Schuchard, J., Middleton, E. L., & Schwartz, M. F. (2017). The timing of spontaneous detection and repair of naming errors in aphasia. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 93*, 79–91.

Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Faseyitan, O., Brecher, A., Dell, G. S., & Coslett, H. B. (2009). Anterior temporal involvement in semantic word retrieval: Voxel-based lesion-symptom mapping evidence from aphasia. *Brain: a Journal of Neurology, 132*(12), 3411–3427.

Schwartz, M. F., Romani, C., Brown, D., & Brecher, A. (2015). Syllabic complexity effects in phonological speech errors: The role of articulatory-phonetic impairment. In *Academy of aphasia 53rd annual meeting, (Tucson)*.

Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments & Computers, 28*(4), 516–536.

Snowden, J. S., Harris, J. M., Saxon, J. A., Thompson, J. C., Richardson, A. M., Jones, M., & Kobylecki, C. (2019). Naming and conceptual understanding in frontotemporal dementia. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 120*, 22–35.

Staffaroni, A. M., Weintraub, S., Rascovsky, K., Rankin, K. P., Taylor, J., Fields, J. A., … Kramer, J. H. (2021). Uniform data set language measures for bvFTD and PPA diagnosis and monitoring. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 13*(1), Article e12148.

Stockbridge, M. D., Tippett, D. C., Breining, B. L., Vitti, E., & Hillis, A. E. (2021). Task performance to discriminate among variants of primary progressive aphasia. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 145*, 201–211.

Teichmann, M., Kas, A., Boutet, C., Ferrieux, S., Nogues, M., Samri, D., … Migliaccio, R. (2013). Deciphering logopenic primary progressive aphasia: A clinical, imaging and biomarker investigation. *Brain: a Journal of Neurology, 136*(11), 3474–3488.

Tetzloff, K. A., Whitwell, J. L., Utianski, R. L., Duffy, J. R., Clark, H. M., Machulda, M. M., … Josephs, K. A. (2018). Quantitative assessment of grammar in amyloid-negative logopenic aphasia. *Brain and language, 186*, 26–31.

Thompson, C. K., & Weintraub, S. (2014). *Northwestern naming battery (Unpublished experimental version*.

Tyler, L. K., Chiu, S., Zhuang, J., Randall, B., Devereux, B. J., Wright, P., … Taylor, K. I. (2013). Objects and categories: Feature statistics and object processing in the ventral stream. *Journal of cognitive neuroscience, 25*(10), 1723–1735.

Tyler, L. K., Stamatakis, E. A., Bright, P., Acres, K., Abdallah, S., Rodd, J. M., & Moss, H. E. (2004). Processing objects at different levels of specificity. *Journal of cognitive neuroscience, 16*(3), 351–362.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., & Delcroix, N. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15*(1), 273–289.

Varrone, A., Asenbaum, S., Vander Borght, T., Booij, J., Nobili, F., Någren, K., … Van Laere, K. (2009). EANM procedure guidelines for PET brain imaging using [18F] FDG, version 2. *European Journal of Nuclear Medicine and Molecular Imaging, 36*(12), 2103–2110.

Visser, M., Jefferies, E., Embleton, K. V., & Lambon Ralph, M. A. (2012). Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: Distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *Journal of Cognitive Neuroscience, 24*(8), 1766–1778.

Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: It's in the links. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 17*(4), 664.

Vitkovitch, M., Humphreys, G. W., & Lloyd-Jones, T. J. (1993). On naming a giraffe a zebra: Picture naming errors across different object categories. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 19*(2), 243.

Walker, G., Fridriksson, J., Hillis, A., den Ouden, D., Bonilha, L., & Hickok, G. (2022). *The severity-calibrated aphasia naming test*.

Warrington, E. K. (1997). The graded naming test: A restandardisation. *Neuropsychological Rehabilitation, 7*(2), 143–146.

Win, K. T., Pluta, J., Yushkevich, P., Irwin, D. J., McMillan, C. T., Rascovsky, K., … Grossman, M. (2017). Neural correlates of verbal episodic memory and lexical retrieval in logopenic variant primary progressive aphasia. *The Florida Nurse, 11*, 330.

Woollams, A. M., Cooper-Pye, E., Hodges, J. R., & Patterson, K. (2008). Anomia: A doubly typical signature of semantic dementia. *Neuropsychologia, 46*(10), 2503–2514.

Zannino, G. D., Perri, R., Salamone, G., Di Lorenzo, C., Caltagirone, C., & Carlesimo, G. A. (2010). Manipulating color and other visual information influences picture naming at different levels of processing: Evidence from Alzheimer subjects and normal controls. *Neuropsychologia, 48*(9), 2571–2578.

Zhang, W., Wang, J., Fan, L., Zhang, Y., Fox, P. T., Eickhoff, S. B., … Jiang, T. (2016). Functional organization of the fusiform gyrus revealed with connectivity profiles. *Human brain mapping, 37*(8), 3003–3016.

Reyes, P., Ortega-Merchan, M. P., Rueda, A., Uriza, F., Santamaria-García, H., Rojas-Serrano, N., … Matallana, D. (2018). Functional connectivity changes in behavioral, semantic, and nonfluent variants of frontotemporal dementia. *Behavioural neurology, 2018*, 1–10.

Moreno-Martínez, F. J., & Rodríguez-Rojo, I. C. (2015). On colour, category effects, and Alzheimer's disease: A critical review of studies and further longitudinal evidence. *Behavioural Neurology, 2015*, 1–12.