

Research Article

The Diagnostic Value of Language Screening in Primary Progressive Aphasia: Validation and Application of the Sydney Language Battery

Nikki Janssen,^{a,b} Ardi Roelofs,^a Esther van den Berg,^c Willem S. Eikelboom,^c Meike A. Holleman,^d Dymphie M. J. M. in de Braek,^e Olivier Piguet,^{f,g} Vitória Piai,^{a,b} and Roy P. C. Kessels^{a,b,h}

Purpose: The three variants of primary progressive aphasia (PPA) differ in clinical presentation, underlying brain pathology, and clinical course, which stresses the need for early differentiation. However, brief cognitive tests that validly distinguish between all PPA variants are lacking. The Sydney Language Battery (SYDBAT) is a promising screening instrument that can be used as a first step in a comprehensive neuropsychological assessment to distinguish PPA subtypes, but evidence on its validity and reliability is to date limited. In the current study, the validation and diagnostic value of the SYDBAT are described for discriminating PPA subtypes as well as distinguishing PPA from mild cognitive impairment (MCI) or Alzheimer's dementia (AD).

Method: Forty-five patients with PPA (13 with semantic PPA, 20 with logopenic PPA, and 12 with nonfluent/agrammatic PPA), 25 MCI patients, 13 AD patients, and 50 cognitively unimpaired controls were included in this study. Both patients and controls completed the SYDBAT-NL (Dutch version). Performance on and predictive ability of the four subtests

(i.e., Naming, Word Comprehension, Repetition, and Semantic Association) were assessed. In addition, construct validity and internal consistency were examined.

Results: Different SYDBAT performance patterns were found across PPA and non-PPA patient groups. While a discriminant function analysis based on SYDBAT subtest scores could predict PPA subtype with 78% accuracy, it was more difficult to disentangle PPA from non-PPA patients based on SYDBAT scores alone. For assisting in clinical interpretation, simple rules were set up and translated into a diagnostic decision tree for subtyping PPA, which was capable of diagnosing a large proportion of the cases. Satisfying validity and reliability measures were found.

Conclusions: The SYDBAT is an easy-to-use and promising screen for assessing single-word language processes, which may contribute to the differential diagnostic process of PPA and the assessment of language impairment in MCI and AD. It can be easily implemented for initial screening of patients in a memory clinic.

^aDonders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands

^bDepartment of Medical Psychology, Radboud University Medical Center, Nijmegen, the Netherlands

^cDepartment of Neurology, Erasmus MC University Medical Center, Rotterdam, the Netherlands

^dDepartment of Medical Psychology, Jeroen Bosch Hospital, Den Bosch, the Netherlands

^eDepartment of Medical Psychology, Maastricht University Medical Centre, the Netherlands

^fSchool of Psychology, The University of Sydney, New South Wales, Australia

^gBrain and Mind Centre, The University of Sydney, New South Wales, Australia

^hVincent van Gogh Institute for Psychiatry, Venray, the Netherlands

Correspondence to Nikki Janssen: nikki.janssen10@gmail.com

Editor-in-Chief: Stephen M. Camarata

Editor: Julius Fridriksson

Received January 14, 2021

Revision received May 20, 2021

Accepted July 20, 2021

https://doi.org/10.1044/2021_JSLHR-21-00024

Primary progressive aphasia (PPA) is a syndrome characterized by progressive decline in language abilities while other cognitive functions remain relatively spared, at least in the initial phases of the disease (Matías-Guiu & García-Ramos, 2013; Mesulam, 1982). PPA is a rare disorder with a typical age of onset between the ages of 50 and 70 years (Kertesz et al., 2005). Its early onset and progressive evolution makes it devastating for

Disclosure: The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.

patients, their occupation, and their family (Tippett et al., 2015). Three PPA subtypes are generally distinguished based on their clinical manifestation and underlying neuropathology (Gorno-Tempini et al., 2011): nonfluent/agrammatic (nfv-PPA), semantic (sv-PPA), and logopenic (lv-PPA) variant. The core symptoms of nfv-PPA include agrammatism in language production and/or effortful speech with inconsistent speech sound and distortions. Patients with nfv-PPA usually have intact single-word comprehension and object knowledge. In contrast, sv-PPA is characterized by deficits in confrontation naming and single-word comprehension, accompanied by surface dyslexia or dysgraphia. Repetition and speech production are relatively spared in these patients. Finally, lv-PPA includes impairments in single-word retrieval in spontaneous speech and naming as well as impaired repetition of phrases and sentences. Phonological errors in spontaneous speech and naming are also observed. Patients with lv-PPA usually do not show deficits in single-word comprehension and object knowledge, and frank agrammatism is absent (Gorno-Tempini et al., 2011).

Given their different underlying pathologies and clinical course, early differentiation of the PPA subtypes may have important consequences for treatment and interventions (Newhart et al., 2009; Tippett et al., 2015). Nfv-PPA and sv-PPA are most frequently associated with frontotemporal lobar degeneration pathology (Josephs et al., 2006), whereas lv-PPA tends to be primarily related to Alzheimer's disease pathology (Mesulam et al., 2003). In addition, PPA subtypes differ in their clinical course, with lv-PPA patients generally showing a more rapid cognitive and functional decline compared to the other two subtypes, which stresses the importance of early differentiation (Matias-Guiu et al., 2015).

A multidisciplinary investigation including a comprehensive cognitive assessment is essential for an accurate diagnosis of PPA (Butts et al., 2015; Marshall et al., 2018). In the current memory clinical setting, however, language assessment is often neglected or very limited due to a lack of short and easy-to-administer language tests that assess more than just naming. For example, the Boston Naming Test (Kaplan et al., 1983) is often used—in many cases even in its abbreviated 15-item Consortium To Establish a Registry for Alzheimer's Disease version (Mack et al., 1992). This test, however, only examines the presence of anomia (word-retrieval deficits) and is insufficient to differentiate between all PPA subtypes since this difficulty is common to all three variants of PPA (Gorno-Tempini et al., 2011). If language is assessed more broadly, most commonly tests are used that have been primarily developed for the assessment of aphasia syndromes after stroke and examine the different language domains separately. As a result, different language measures are often combined to distinguish across PPA subtypes. Combining different tests, however, may lead to invalid interpretation due to differences in normative groups and task properties (e.g., word frequency, number of syllables) across tests (Savage et al., 2013). Moreover, the often-used standard aphasia batteries developed for use with stroke-induced aphasia may lack sensitivity to detect the subtle deficits that are present in early stages of PPA

and may result in inappropriate nomenclature (such as Broca's or Wernicke's) based on the stroke-induced aphasia classification (Henry & Grasso, 2018). Clark et al. (2020) examined the utility of a standard stroke aphasia battery, the Western Aphasia Battery–Revised (WAB-R), for classifying variants of PPA, and they concluded that “WAB-R classification did not distinguish among PPA classification determined by consensus” (p. 498).

To overcome these problems, Savage et al. (2013) developed a new language screen, the Sydney Language Battery (SYDBAT). This instrument measures the integrity of four main aspects of language proficiency: naming, single-word comprehension, semantic association, and repetition abilities, using stimuli based on the same set of target words across the four subtests. This approach eliminates the confounds of linguistic differences and varying norms that trouble the current clinical practice when combining multiple measures from different tests. They created a simple diagnostic algorithm based on the performance on the Naming and Semantic Association subtests, in which sv-PPA is diagnosed if naming and semantic association scores are 2 *SDs* or more below the control mean, lv-PPA is diagnosed if naming is below 22 but semantic association is within 2 *SDs* of the control mean, and nfv-PPA is diagnosed if naming and semantic association are both within 2 *SDs* of the control mean. While the SYDBAT has proven successful in correctly classifying 80% of the PPA patients based on this diagnostic algorithm (Savage et al., 2013), evidence on its validity and reliability is to date limited to the Australian English version.

Furthermore, little attention has been paid to the value of language screens like the SYDBAT in the diagnostic process of non-PPA patient groups, such as mild cognitive impairment (MCI) and Alzheimer's dementia (AD). While deficits in language processing have been found to manifest early in the disease course of AD patients (Taler & Phillips, 2008), the clinical assessment of language in AD and MCI is often limited compared to other cognitive domains such as memory (Emery, 2000). Moreover, comparative studies investigating the diagnostic value of language measures in PPA versus MCI and AD are scarce and inconclusive. Some studies have suggested language test performance in PPA to be equivalent to AD (Cosentino et al., 2006) or lower than AD (Perry & Hodges, 2000); however, most of these studies have only looked at sv-PPA.

This study had several aims. First, this study sought to determine the validity and reliability of the Dutch version of the SYDBAT (henceforth referred to as the SYDBAT-NL) and derive a predictive diagnostic algorithm using the distinct task profiles for each PPA subtype with the goal of confirming the findings of Savage et al. (2013) in an independent sample. Based on previous research (Savage et al., 2013), each PPA subtype was expected to demonstrate a distinct profile across SYDBAT subtests, with most prominent impairments in naming as well as in verbal and visual comprehension (i.e., Word Comprehension and Semantic Association) in sv-PPA because these tests are semantically based. Conversely, nfv-PPA patients were expected to perform well on the visual and verbal comprehension subtests, but poorer on speech

production-related subtests (i.e., Naming and Repetition). Lastly, lv-PPA patients were expected to perform at an intermediate level on all subtests, yet with most evident deficits in naming. Next, this study aimed to examine the SYDBAT in patients diagnosed with MCI or with AD, thus extending previous research by including a non-PPA patient population in which language impairment also is present. This will enable a direct comparison of PPA with MCI and AD. We expected that the SYDBAT would demonstrate a broad spectrum of language deficits in MCI and AD, mainly on semantically based subtests (i.e., Naming, Word Comprehension, and Semantic Association). Since we expect language deficits to be present in both MCI and AD, the assessment of language deficits may not be sufficient to differentially diagnose MCI and AD from PPA. The construct validity was examined through comparison between performance on the SYDBAT-NL subtests and other commonly used language tests. Here, we expected good convergent validity when comparing SYDBAT subtest scores with other established language measures (Savage et al., 2013), while for divergent validity, we expected that the Naming, Word Comprehension, and Semantic Association tests would correlate due to a shared semantic component. Reliability was measured as internal consistency via Cronbach's alpha. By establishing the SYDBAT as a promising tool that can be used in memory clinics, early detection and differentiation of language impairment in neurodegenerative disease becomes possible, which has great prognostic value and possible implications for treatment (Leyton et al., 2013; Tippett et al., 2015).

Method

Participants

Forty-five patients diagnosed with PPA (13 sv-PPA, 20 lv-PPA, and 12 nvf-PPA), 35 non-PPA patients diagnosed with either AD ($n = 13$) or MCI ($n = 25$), and 50 cognitively unimpaired controls were included in this study. Demographic characteristics of all participant groups

are described in Table 1. The patients visited the memory clinic of one of the participating medical centers in the Netherlands (Radboud University Medical Center in Nijmegen, Erasmus MC University Medical Center in Rotterdam, Jeroen Bosch Hospital in 's-Hertogenbosch, and Maastricht University Medical Center in Maastricht) between September 2016 and January 2020 with cognitive complaints (self-reported or reported by significant others). For the PPA group, the SYDBAT-NL was added to the standard neuropsychological assessment when a patient presented with language problems and suspicion of neurodegenerative disease. Subsequently, the patients were included in this study if they fulfilled the clinical diagnostic criteria of PPA (Gorno-Tempini et al., 2011). Furthermore, the SYDBAT-NL was added to the neuropsychological assessment of a randomly selected group of patients visiting the memory clinic for the non-PPA patient groups. The cognitively unimpaired controls were recruited through (oral) advertisement.

Clinical diagnoses were established based on an extensive multidisciplinary assessment including neuropsychological assessment, neurological testing, and neuroimaging. PPA patients were between 52 and 82 years of age, and specific variants were diagnosed according to the guidelines by Gorno-Tempini et al. (2011). Non-PPA patients were between 52 and 80 years of age and were diagnosed according to the Dutch guidelines for MCI and AD (Nederlandse Vereniging voor Klinische Geriatrie, 2014). Specifically, all MCI patients fulfilled the MCI-core clinical criteria as defined by Albert et al. (2011) and were categorized as either amnesic MCI ($n = 15$) or nonamnesic MCI ($n = 10$; Petersen, 2004). Cognitively unimpaired controls were between 48 and 85 years of age and had neither cognitive impairments (Montreal Cognitive Assessment [MoCA] score ≥ 26) nor self-reported cognitive complaints. All participants were native speakers of Dutch. For all participants, exclusion criteria were a current psychiatric disorder or substance abuse, or a current or past severe neurological disorder. All patients were assessed by use of a neuropsychological assessment, including the SYDBAT-NL. The

Table 1. Clinical characteristics of study participants.

Variable	PPA				Non-PPA			Controls
	All	sv-PPA	lv-PPA	nfv-PPA	All	MCI	AD	All
n (%)	45 (33.8)	13 (9.8)	20 (15.0)	12 (9.0)	38 (28.6)	25 (18.8)	13 (9.8)	50 (37.6)
Age (SD)	67.2 (7.0)	65.2 (7.0)	67.5 (6.9)	68.8 (7.4)	69.2 (6.8)	70.0 (6.0)	67.7 (8.4)	66.0 (8.2)
Sex distribution (males/females)	26/19	6/7	15/5	5/7	20/18	15/10	5/8	24/26
Education level (1–7) ^a (SD)	5.1 (1.1)	5.7 (0.9)	4.8 (1.0)	5.1 (1.2)	5.1 (1.1)	5.1 (1.2)	4.9 (0.9)	5.5 (1.2)
Disease duration in years (SD)	3.2 (1.8)	3.3 (1.3)	3.2 (2.1)	3.2 (1.9)	3.6 (3.2)	3.3 (3.7)	4.0 (2.0)	N/A
MoCA score (SD) ^{***}	20.1 (3.8)	21.4 (2.6)	18.5 (4.5)	21.0 (3.6)	19.6 (6.9)	23.7 (4.5)	13.50 (5.0)	27.4 (1.3)

Note. Data are reported as mean (SD) or number (%). sv-PPA = semantic variant PPA; lv-PPA = logopenic variant PPA; nvf-PPA = nonfluent variant PPA; MCI = mild cognitive impairment; AD = Alzheimer's dementia; n = number of participants; MoCA = Montreal Cognitive Assessment; N/A = not applicable.

^a1 = less than 6 years elementary school; 2 = 6 years elementary school; 3 = more than 6 years elementary school; 4 = vocational training; 5 = community college; 6 = advanced vocational training; 7 = Bachelor of Science or higher.

Statistical significance (PPA vs. non-PPA vs. controls): ^{***} $p < .001$.

controls only completed the MoCA and the SYDBAT-NL. All data were collected in compliance with the Declaration of Helsinki.

SYDBAT-NL

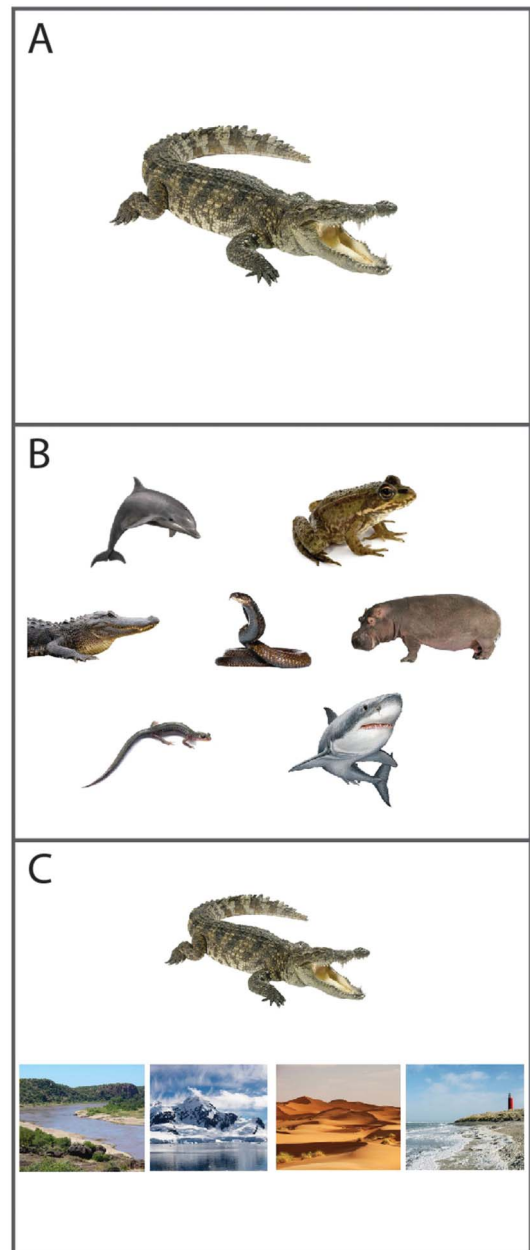
The specifics of the SYDBAT have been described in detail previously (Savage et al., 2013). In short, the Sydney Language Battery consists of four subtests (Naming, Repetition, Word Comprehension, and Semantic Association), each including the same 30 nouns of three or more syllables (e.g., elephant, thermometer). Items include living and non-living things and are graded in difficulty based on word frequency. Each subtest has a total score of 30, and the administration of a subtest has to be ended after six incorrect responses. The total administration takes approximately 15–30 min.

The translation and adaptation of the SYDBAT into Dutch was described previously in a pilot study (Eikelboom et al., 2017). In the SYDBAT-NL, characteristics of the original test were kept as similar as possible. Items with fewer than three syllables after translation into Dutch were replaced by new items that matched the category of the removed items as much as possible. For example, the item hippopotamus of the original test comprises only two syllables in Dutch (“nijlpaard”) and was therefore replaced by the three-syllable item crocodile (“krokodil”) in the SYDBAT-NL. In addition, some response options on the Semantic Association subtest were modified when the semantic relation of existing SYDBAT items was not apparent for the Dutch language or culture. Subsequently, items were ordered by decreasing word frequency and graded into three blocks of difficulty based on the SUBTLEX-NL frequency database (Keuleers et al., 2010). A pilot version of the SYDBAT-NL was tested on 10 individuals before the final version was established. An example of the Naming, Word Comprehension, and Semantic Association subtests is shown in Figure 1.

Neuropsychological Assessment

The MoCA was used as a short screening instrument to establish global cognitive functioning (Nasreddine et al., 2008). The Trail Making Test (TMT)–Part A was used to measure processing speed (Bowie & Harvey, 2006). Executive functioning was assessed using the TMT (dividing the TMT-Part B score by the score on Part A; Oosterman et al., 2010). Verbal fluency was examined with the Semantic Fluency Test (1-min Animal & Profession naming; Van der Elst et al., 2006) and the Letter Fluency Test (Letters tested: “D-A-T,” which is the Dutch-language equivalent of the Controlled Oral Word Association Test [FAS variant]; Lezak et al., 2012; Schmand et al., 2008). Working memory was assessed with the Digit Span (total score for Forward, Backward, and Sorting) subtest of the Dutch version of the Wechsler Adult Intelligence Scale–Fourth Edition (Wechsler, 2008). Visuospatial episodic memory was examined with the Location Learning Test–Revised (Bucks et al., 2011). The

Figure 1. An example item of the Naming (A), Word Comprehension (B), and Semantic Association (C) subtests, in which the patient has to either name or point to the correct answer. For the Repetition subtest, no visual item is presented as both the instruction and response for this subtest are given verbally.



copy trial of the Rey–Osterrieth Complex Figure Test was used to evaluate visuoconstruction (Osterrieth, 1944).

Several established language tests were used to examine the convergent validity of the SYDBAT-NL. The short form of the Boston Naming Test (BNT) with 29 line drawings, validated for the Dutch language (Van Loon-Vervoorn & Van der Velden, 2006), was used to assess naming abilities. The verbal subtest of the Semantic Association Test

(Visch-Brink et al., 1996) was used to examine deficits in semantic associations. The Word Comprehension and Word Repetition subtests from the Comprehensive Aphasia Test (CAT-NL; Visch-Brink et al., 2014) were administered to measure word comprehension and repetition abilities, respectively.

Descriptive Measurements

The level of formal education was measured in accordance with the Dutch educational system, which uses levels rather than years of education (Duits & Kessels, 2012). This ordinal scale includes the following seven categories: (a) fewer than 6 years primary education; (b) completed primary education; (c) more than 6 years of primary education, without a secondary school diploma; (d) lower vocational training; (e) advanced vocational training or lower professional education; (f) advanced professional training or upper secondary school; and (g) academic degree. The duration of cognitive decline was estimated based on the reported years since symptom onset either via self-report or reported by significant others.

Statistical Analyses

As the PPA and non-PPA groups were relatively small and scores were not normally distributed, between-group comparisons were performed using nonparametric tests (Kruskal & Wallis, 1952). Therefore, to assess the between-group differences in performance between the PPA, non-PPA, and control groups on the SYDBAT-NL and other neuropsychological tests, Kruskal–Wallis and follow-up Mann–Whitney *U* tests were used with Benjamini and Hochberg correction for multiple testing (Benjamini & Hochberg, 1995) with false discovery rate *Q* set at 0.05 (McDonald, 2009). Effect sizes were calculated using Rosenthal's (1994) formula. Also, raw scores of the neuropsychological assessment were converted into *z* scores based on sex, age, and education-adjusted norms for descriptive purposes using 1.5 *SD* below average as a cutoff for classifying a patient as impaired on that test (Petersen et al., 1999).

The specific SYDBAT impairment profile was analyzed by comparisons between each patient (sub)group on the SYDBAT subtests using Kruskal–Wallis and follow-up Mann–Whitney *U* tests. Next, discriminant function analyses were conducted to explore whether patients could be classified into diagnostic groups using either a combination of SYDBAT scores of the Naming and Semantic Association subtests (following the diagnostic decision tree by Savage et al., 2013) or a combination of all SYDBAT subtests. Subsequently, cutoff scores were determined by analyzing the profile of each PPA variant in a qualitative way and selecting cutoff scores with the highest predictive ability (i.e., the highest number of patients who were correctly classified). Finally, based on these cutoff values, sensitivity and specificity were calculated per subtype, and simple diagnostic rules for clinical practice were created.

Construct validity was examined by exploring convergent and divergent validity (Campbell & Fiske, 1959). Because of the expected ceiling performance of the language unimpaired controls, the convergent validity of the SYDBAT-NL subtests was examined only in patients by calculating Spearman correlation coefficients (r_s) between each subtest and a similar established test measuring the same construct. To illustrate, correlation coefficients were calculated between the Naming subtests and the BNT, the Word Comprehension subtest and the Word Comprehension task of the CAT-NL, the Repetition subtest and the Word Repetition subtests of the CAT-NL, and the Semantic Association subtest and the verbal task of the Semantic Association Test (SAT). Divergent validity of the SYDBAT was examined only in patients by calculating Spearman correlation coefficients (r_s) between each SYDBAT-NL subtest and between the SYDBAT subtests and the established language tests measuring a different construct. These were thus expected to be correlated to a lesser extent. Since the SYDBAT-NL scores were not normally distributed, Spearman's rank-order correlation coefficients (r_s) were used (Spearman, 1904). The correlations were interpreted according to Cohen's (1988) convention of small (0.10), moderate (0.30), and large (0.50). Discriminant function analyses based on the four (sub)tests used for the study of convergent validity were conducted to assess whether the SYDBAT outperformed these standard measures in terms of patient classification. Reliability of the SYDBAT-NL was tested in the 50 cognitively unimpaired controls by assessing the internal consistency via Cronbach's α coefficients per subtest. All statistical analyses were performed with IBM SPSS Statistics 25.0.

Results

General Cognitive Profile

There were no significant group differences for age, sex distribution, or level of education between the PPA, non-PPA, and cognitively unimpaired control groups (all *H* values < 5.3, all *p* values > .05; Table 1). MoCA scores showed significant differences between the three groups, $H(2) = 49.62$, $p < .001$, with the controls scoring significantly higher compared with the PPA ($U = 8.5$, $z = 6.91$, $p < .001$, $r = .80$) and non-PPA patient groups ($U = 32.5$, $z = 2.78$, $p = .004$, $r = .38$). The PPA and non-PPA patient groups did not differ with regard to disease duration ($U = 352$, $z = -0.30$, $p = .76$, $r = -.04$) or severity of cognitive deficits ($U = 56.5$, $z = -0.20$, $p = .84$, $r = -.04$) as measured by the MoCA (see Table 1).

Table 2 shows the results of the complete neuropsychological assessment for the PPA and non-PPA patient groups, including scores on clinical impairment based on *z* scores of 1.5 *SD* below the control mean. A Kruskal–Wallis test that compared all neuropsychological tests as reported in Table 2 between PPA-subtypes showed no significant differences after correction for multiple comparisons (all *H* values < 7.3, all *p* values > .025). A Kruskal–Wallis

Table 2. Performance on neuropsychological tests for patients with primary progressive aphasia and non-PPA patients per subtype.

Domain	PPA						Non-PPA				
	sv-PPA		lv-PPA		nfv-PPA		MCI		AD		
	Raw	CI	Raw	CI	Raw	CI	Raw	CI	Raw	CI	
Processing speed											
TMT A	47.8 (26.4)	3/10 (30%)	102.8 (93.0)	9/18 (50%)	88.6 (68.1)	4/10 (40%)	55.0 (18.2)	5/19 (26%)	71.2 (27.5)	3/8 (38%)	
Executive function											
TMT B ^a	116.9 (51.2)	4/11 (36%)	192.4 (118.5)	12/17 (71%)	200.1 (154.3)	3/9 (33%)	171.6 (140.8)	4/17 (24%)	268.6 (156.1)	4/7 (57%)	
Letter Fluency	22.2 (16.5)	4/6 (67%)	17.2 (8.6)	10/17 (59%)	20.9 (8.5)	5/8 (63%)	26.5 (11.3)	8/17 (47%)	18.5 (8.3)	4/6 (67%)	
Semantic Fluency ^b	7.8 (5.1)	6/7 (86%)	9.1 (3.6)	16/18 (89%)	13.8 (7.0)	4/10 (40%)	14.5 (5.2)	8/20 (40%)	9.8 (4.3)	3/5 (60%)	
Working memory											
Digit Span total	21.1 (5.4)	1/7 (14%)	14.1 (3.7)	10/13 (77%)	16.4 (3.8)	3/6 (50%)	22.3 (4.3)	2/19 (11%)	17.3 (4.5)	0/7 (0%)	
Visuospatial memory											
LLT-R total score	30.4 (22.9)	2/10 (20%)	42.0 (22.1)	4/17 (24%)	33.1 (18.1)	0/7 (0%)	38.4 (22.0)	1/16 (6%)	63.7 (28.6)	2/5 (40%)	
LLT-R delayed score ^c	0 (2.8)	1/11 (9%)	-1.8 (1.8)	1/18 (6%)	0.2 (3.0)	0/8 (0%)	-0.56 (3.3)	1/16 (6%)	-0.83 (6.2)	2/5 (40%)	
Visuoconstruction											
RCFT copy	29.8 (7.1)	1/5 (20%)	29.1 (4.6)	4/11 (36%)	30.5 (4.6)	1/3 (33%)	28.4 (4.4)	7/16 (86%)	27.8 (4.9)	2/4 (50%)	

Note. Data are reported as raw test scores (Raw) (SD) and number of clinically impaired (CI; $z < 1.5$ SD) for sv-PPA = semantic variant PPA; lv-PPA = logopenic variant PPA; nfv-PPA = nonfluent variant PPA; MCI = mild cognitive impairment; AD = Alzheimer's disease. TMT = Trail Making Test; LLT-R = Location Learning Test-Revised; RCFT = Rey Complex Figure Test.

^aFor calculation of % clinically impaired, B/A-values were used (Oosterman et al., 2010). Patients in whom the test was discontinued were counted as clinically impaired ($n = 11$).

^bMean raw for animal and profession naming. ^cNumber of items forgotten during the delay period.

test that compared all neuropsychological tests between specific PPA subtypes, MCI and AD showed a significant difference only on the Digit Span task after correction for multiple comparisons, $H(4) = 23.11$, $p < .001$. Follow-up Mann–Whitney U tests showed that the lv-PPA patients performed significantly worse on Digit Span compared to MCI ($U = 20.0$, $z = -3.99$, $p < .001$, $r = -.71$) and AD ($U = 14.0$, $z = -2.51$, $p < .05$, $r = -.56$) patients, whereas nfv-PPA patients performed significantly worse on Digit Span compared to MCI patients ($U = 15.0$, $z = -2.70$, $p < .01$, $r = -.54$).

SYDBAT Between-Group Comparisons

SYDBAT subtest scores of all groups are shown in Figure 2, and results of the differences in performance between all PPA and non-PPA patient subtypes are shown in Table 3, including scores on clinical impairment based on z scores of 1.5 SD below average. As expected, significant group differences were found between the PPA patients, non-PPA patients, and healthy controls on Naming, $H(2) =$

69.0, $p < .001$; Repetition, $H(2) = 14.10$, $p < .001$; Word Comprehension, $H(2) = 14.8$, $p < .001$; and Semantic Association, $H(2) = 37.61$, $p < .001$. This was followed-up with Mann–Whitney U tests that showed both the PPA and non-PPA groups scored significantly lower on all subtests compared with controls (U values between 123.0 and 852.0, all p values $< .05$). Moreover, the PPA patients scored significantly lower on the Naming ($U = 548.0$, $z = -2.81$, $p < .01$, $r = -.31$) and Repetition ($U = 459.5$, $z = -3.52$, $p < .001$, $r = -.39$) subtests, but not the Word Comprehension ($U = 691.0$, $z = -1.34$, $p = .18$) and Semantic Association ($U = 796.5$, $z = -0.54$, $p = .59$) subtests, compared to non-PPA patients.

PPA Subtypes

A Kruskal–Wallis test showed significant differences in SYDBAT performance between the PPA subtypes on the Naming, Word Comprehension, and Semantic Association subtests. Sv-PPA patients performed worse on both the Naming ($U = 13.0$, $z = -4.32$, $p < .001$, $r = -.75$) and

Figure 2. Violin plots of the performance on the Sydney Language Battery per subtest for cognitively unimpaired controls (Control) and patients with Alzheimer’s dementia (AD), mild cognitive impairment (MCI), logopenic variant PPA (lv-PPA), nonfluent variant PPA (nfv-PPA), and semantic variant PPA (sv-PPA). The outer shapes represent the distribution of individual data (indicated by dots), the thick horizontal line inside the box indicates the median, and the bottom and top of the box indicate the first and third quartiles.

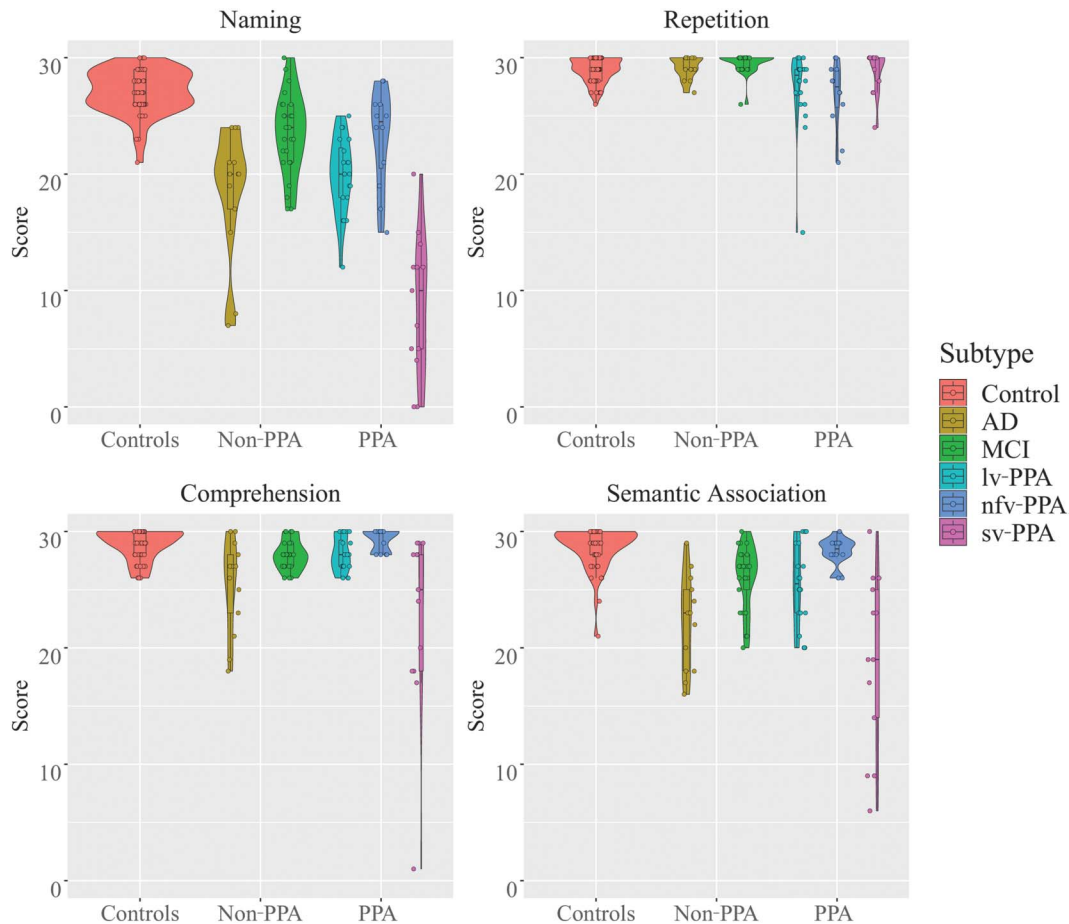


Table 3. Performance on the Sydney Language Battery (SYDBAT) per subtest by patient group.

SYDBAT-NL Subtest	PPA						Non-PPA				Controls
	sv-PPA		lv-PPA		nfv-PPA		MCI		AD		All
	Raw	CI	Raw	CI	Raw	CI	Raw	CI	Raw	CI	Raw
Naming (SD)	8.7 (6.2)	13/13 (100%)	19.9 ^{***†} (3.3)	20/20 (100%)	23.2 ^{***†} (4.2)	8/12 (67%)	23.5 ^{***††} (3.5)	18/25 (72%)	18.5 ^{**†} (5.5)	13/13 (100%)	27.1 (1.9)
Word Comprehension (SD)	23.4 (8.0)	6/13 (46%)	28.1 [‡] (1.4)	3/20 (15%)	29.3 ^{**†} (1.0)	0/12 (0%)	28.0 ^{**} (1.3)	3/24 (13%)	25.4 ^{**} (4.0)	6/13 (46%)	28.9 (1.3)
Repetition (SD)	28.9 (1.9)	3/13 (23%)	27.4 (3.4)	7/20 (35%)	26.8 [*] (2.9)	6/12 (50%)	29.6 ^{†††‡‡‡} (0.9)	1/23 (4%)	29.2 [‡] (1.0)	1/13 (8%)	29.0 (1.1)
Semantic Association (SD)	19.7 (6.9)	10/13 (77%)	25.5 ^{**} (3.6)	10/20 (50%)	28.3 ^{***} (1.2)	0/12 (0%)	26.1 ^{**†} (2.8)	7/25 (28%)	22.2 ^{‡‡‡} (4.1)	10/13 (77%)	28.6 (1.8)

Note. Data are reported as raw mean test scores (Raw) (SD) and number of clinically impaired (CI > 1.5 SD below control mean). PPA = primary progressive aphasia; sv-PPA = semantic variant PPA; lv-PPA = logopenic variant PPA; nfv-PPA = nonfluent variant PPA; MCI = mild cognitive impairment; AD = Alzheimer's dementia. SYDBAT-NL = Dutch version of the SYDBAT.

Statistical significance (vs. sv-PPA): * = $p < .05$. ** = $p < .01$. *** = $p < .001$ (corrected for multiple comparisons).

Statistical significance (vs. lv-PPA): † = $p < .05$. †† = $p < .01$. ††† = $p < .001$ (corrected for multiple comparisons).

Statistical significance (vs. nfv-PPA): ‡ = $p < .05$. ‡‡ = $p < .01$. ‡‡‡ = $p < .001$ (corrected for multiple comparisons).

Semantic Association ($U = 59.5, z = -2.61, p = .009, r = -.45$) subtests compared to lv-PPA patients. Compared to nfv-PPA, sv-PPA patients scored significantly lower on Naming ($U = 3.5, z = -4.06, p < .001, r = -.81$), Semantic Association ($U = 13.5, z = -3.54, p < .001, r = -.71$), and Word Comprehension ($U = 21.5, z = -3.17, p = .002, r = -.63$). Lv-PPA patients performed worse on both the Naming ($U = 60.5, z = -2.33, p = .02, r = -.41$) and Word Comprehension ($U = 61.5, z = -2.37, p = .018, r = -.42$) subtests compared to nfv-PPA patients.

PPA Subtypes Versus MCI and AD

A Kruskal–Wallis test that compared the SYDBAT subtests between PPA and non-PPA patients showed significant differences on Naming, $H(4) = 39.45, p < .001$; Repetition, $H(4) = 20.48, p < .001$; Word Comprehension, $H(4) = 16.95, p < .01$; and Semantic Association, $H(4) = 25.57, p < .001$. On follow-up Mann–Whitney U tests that compared the specific PPA and non-PPA patient subtypes, significant profile differences were observed.

When comparing lv-PPA with MCI, scores on the Word Comprehension and Semantic Association subtests were similar for both groups, yet lv-PPA patients scored significantly lower on the Naming ($U = 111.5, z = -3.18, p = .001, r = -.47$) and Repetition ($U = 90.0, z = -3.64, p < .001, r = -.55$) subtests. No significant differences in SYDBAT performance were observed between patients with lv-PPA and AD after correction for multiple comparisons.

Sv-PPA patients had significantly lower scores than MCI patients on the Naming ($U = 4.0, z = -4.89, p < .001, r = -.79$) and Semantic Association ($U = 58.0, z = -3.24, p = .001, r = -.52$) subtests. Compared to AD patients, sv-PPA patients performed worse only on the Naming subtest ($U = 19.5, z = -3.35, p = .001, r = -.66$).

In contrast, nfv-PPA patients had a lower performance only on the Repetition subtest compared to MCI patients ($U = 42.5, z = -3.60, p < .001, r = -.61$), whereas MCI patients had a lower score on Word Comprehension ($U = 63.5, z = -2.80, p = .005, r = -.47$) and Semantic Association ($U = 73.5, z = -2.52, p = .012, r = -.41$). This same pattern of differences was present when comparing nfv-PPA to AD, where AD patients scored lower on Word Comprehension ($U = 23.5, z = -3.05, p = .002, r = -.61$), Semantic Association ($U = 11.5, z = -3.65, p < .001, r = -.73$), but also Naming ($U = 35.5, z = -2.33, p = .020, r = -.47$). Nfv-PPA patients, however, had significantly lower scores on the Repetition subtest ($U = 35.0, z = -2.40, p = .016, r = -.48$).

Discriminant Function Analysis

PPA Subtypes

Next, a discriminant function analysis based upon the Naming and Semantic association subtests was performed, which significantly differentiated the PPA variants, $\Lambda = 0.33, \chi^2(4) = 46.4, p < .001$, with the cross-validated classification resulting in 77.8% of cases (35/45) correctly classified (canonical $R^2 = .67$). The majority of classification errors involved the distinction between lv-PPA and nfv-PPA, where

50% of the nfv-PPA cases were incorrectly classified as lv-PPA based on this analysis. Repeating the discriminant function analysis including all SYDBAT subtests yielded a lower cross-validated classification result (68.9%), showing the two extra subtests (Repetition and Word Comprehension) to be redundant for this analysis.

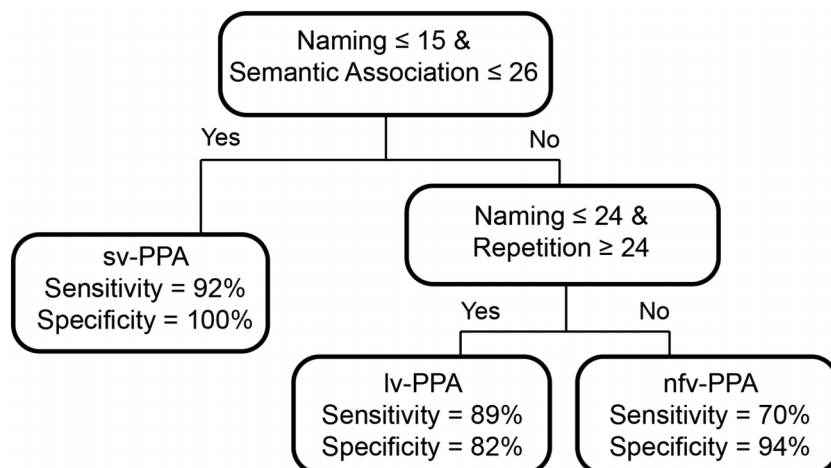
Despite its ability to establish classification accuracy in a statistically sound way, the clinical interpretability of a discriminant function analysis may not be very straightforward. For this purpose, diagnostic decision rules were created based on cutoff scores with the highest predictive ability (i.e., 39/45 patients who were correctly classified), as these outperformed rules based on clinical impairment scores (z scores of 1.5 SDs below average). This resulted in the following diagnostic decision rules that were used to set up a diagnostic decision tree (see Figure 3).

1. If the Semantic Association score is 26 or lower and the Naming score is 15 or lower, the most likely diagnosis is sv-PPA. This simple rule correctly identified 12 out of 13 (92%) sv-PPA cases, whereas no patients diagnosed with the other variants were misclassified as having sv-PPA (i.e., 0% false positives). A diagnosis of sv-PPA is further supported by preserved repetition and a relatively normal performance on neuropsychological tests such as the TMT and nonverbal memory measures like the Location Learning Test–Revised.
2. If the Naming score is 24 or lower and Repetition score is 24 or higher, the diagnosis is most likely lv-PPA. With this rule, 18 out of 20 (90%) lv-PPA patients were correctly classified. However, it also resulted in the highest false-positive rate (one sv-PPA and four nfv-PPA cases), leading to a specificity of 80%. Impaired scores on other neuropsychological tests such as digit span and letter fluency further support the diagnosis of lv-PPA.
3. If either the Naming score is higher than 24 or Repetition is severely impaired (i.e., score < 24), the diagnosis is most likely nfv-PPA. Application of this rule correctly classified eight out of 12 nfv-PPA patients and resulted in only two false-positive cases (6%). This diagnosis is supported by relatively high scores on the Word Comprehension and Semantic Association subtests.

PPA Versus MCI and AD

Again, a discriminant function analysis based upon the Naming and Semantic Association subtests was conducted, only this time for all PPA and non-PPA patient subtypes. While the predictive capacity of this function was significant, the cross-validated classification resulted in only 50.6% of cases (42/83) correctly classified, $\Lambda = 0.37, \chi^2(8) = 77.2, p < .001$, canonical $R^2 = .59$. The majority of classification errors were made in the nfv-PPA group, where cases were incorrectly classified as having either MCI (66.7%) or lv-PPA (33.3%) based on this discriminant function analysis. Also, with this algorithm, 77% of the AD patients were

Figure 3. Decision tree for diagnosing primary progressive aphasia (PPA) subtypes for when PPA is an established diagnosis. sv-PPA = semantic variant PPA; lv-PPA = logopenic variant PPA; nfv-PPA = nonfluent variant PPA.



misclassified as having either MCI (46%), sv-PPA (15.4%), or lv-PPA (15.4%).

Construct Validity and Reliability

The assessment of construct validity over the whole patient sample showed a good convergent validity for Naming (SYDBAT-Naming vs. BNT: $r_s = .84, p < .001$), Repetition (SYDBAT-Repetition vs. CAT-NL-Repetition: $r_s = .56, p < .001$), and Semantic Association (SYDBAT-Semantic Association vs. SAT: $r_s = .59, p < .001$). A moderate-to-good convergent validity was found for Word Comprehension (SYDBAT-Word Comprehension vs. CAT-NL-Word Comprehension: $r_s = .38, p < .001$).

Divergent validity analyses showed significant correlations between the SYDBAT-NL Naming, Word Comprehension, and Semantic Association subtests (r_s between .44 and .74, $p < .001$). Also, significant correlations were found between these three SYDBAT-NL subtests and the SAT, CAT-NL Word Comprehension subtest, and the BNT (r_s between .4 and .66, $p < .05$). No significant correlations were found between the SYDBAT-NL Repetition subtest and the SAT, the CAT-NL Word Comprehension subtest, and the BNT after correction for multiple comparisons (all p values $> .046$). All correlations for convergent and divergent validity are shown in Table 4.

A discriminant function analysis based upon the BNT and SAT was performed, which significantly differentiated the PPA variants, $\Lambda = 0.29, \chi^2(4) = 34.1, p < .001$, with the cross-validated classification, resulting in 61.3% of cases correctly classified (canonical $R^2 = .69$). Repeating the discriminant function analysis including all construct validity (sub) tests yielded a slightly higher cross-validated classification result (68.2%), yet not higher than both the cross-validated classification results based on the SYDBAT subtests.

The Naming, Repetition, and Semantic Association subtests of the SYDBAT-NL all had high reliabilities with

Cronbach's α of .81, .84 and .78, respectively. The Word Comprehension subtest had lower reliability, Cronbach's $\alpha = .64$.

Discussion

In this study, the diagnostic value of the SYDBAT-NL was assessed in samples of patients with PPA or non-PPA cognitive decline (AD and MCI patients) and cognitively unimpaired controls. As hypothesized, different SYDBAT performance patterns were found across PPA and non-PPA patient groups. While a discriminant function analysis based on SYDBAT subtest scores could predict PPA subtype with 77.8% accuracy, it was more difficult to disentangle PPA from non-PPA patients on SYDBAT scores alone. To assist clinical interpretation, simple rules were set up and translated into a diagnostic decision tree for subtyping PPA, which was capable of diagnosing a large proportion of the cases. This diagnostic tree resulted in very high specificity for the sv-PPA (100%) and nfv-PPA (94%) subtypes, whereas the lv-PPA subtype showed a higher sensitivity (90%) compared to specificity (84%).

To examine impairment profiles per patient group, SYDBAT scores were classified as either clinically impaired or unimpaired, based on a threshold of 1.5 SDs below the cognitively unimpaired control group mean. As expected, none of the nfv-PPA patients showed impaired scores on the Semantic Association or Word Comprehension subtests (Gorno-Tempini et al., 2011). It is worth noting that a third of the nfv-PPA patients scored within the normal limits on the Naming subtest, providing a clear distinction from the two other PPA variants. Even though nfv-PPA patients typically present with effortful speech and production errors, their performance on naming tasks that require single-word production in response to a target can be relatively preserved (Graham et al., 2004).

Table 4. Construct validity of the Dutch version of the SYDBAT.

SYDBAT subtest	BNT (n = 71)	Word Comprehension CAT-NL (n = 58)	Word Repetition CAT-NL (n = 49)	SAT (n = 53)
Naming	.84**	.40**	.3	.38*
Word Comprehension	.49**	.38*	-.22	.48**
Repetition	-.21	-.36*	.56**	-.25
Semantic Association	.49**	.66**	-.33	.65**

Note. Data are reported as Spearman's rank correlations (r_s). BNT = Boston Naming Test; CAT-NL = Comprehensive Aphasia Test; SAT = Semantic Association Test.

Statistical significance: * $p < .05$. ** $p < .01$.

In contrast, sv-PPA patients showed the opposite pattern in which tasks that require semantic knowledge were affected most. All had a clinically impaired Naming score, and 10 out of 13 (77%) also had an impaired score on Semantic Association. Based on the diagnostic rules, these two subtests together correctly predicted 92% of the sv-PPA cases. Only one sv-PPA patient was misclassified as lv-PPA due to mild naming and semantic deficits. This patient may have been identified with tests involving the naming and comprehension of very low frequency nouns (e.g., proper names) not included here.

Most predictive errors were made between nfv-PPA and lv-PPA variants, as three nfv-PPA patients were misclassified as lv-PPA and two lv-PPA patients were misclassified as nfv-PPA. This is in line with previous studies reporting the distinction between these two variants to be the most challenging, as they present with overlapping characteristics and the differentiation of key linguistic features requires considerable expertise (Croot et al., 2012). At the same time, the distinction between lv-PPA and nfv-PPA is relevant to clinical practice, since the majority of lv-PPA patients have underlying Alzheimer's disease pathology, while nfv-PPA is strongly linked to frontotemporal lobar degeneration tau pathology (Gorno-Tempini et al., 2011). This stresses the importance of a more extensive neuropsychological investigation, including tasks of episodic memory, emotion processing, and executive functioning (Eikelboom et al., 2018; Piguet et al., 2015) and the development of new tests for discriminating between these variants. The current study provides preliminary recommendations regarding the use of these broad neuropsychological tests, which will require further validation in future research. Especially in the early stages of the disease, when global cognition is typically still preserved, more sensitive measures than global cognitive measures such as the MoCA will need to be developed.

We aimed to create a short and easy-to-administer language screener that can be widely used in memory clinics. As such, the SYDBAT measures language at the single-word level and arguably does not capture all aspects of language and communication. Specifically, while no differences were found between lv-PPA and nfv-PPA on word comprehension, a difference is likely to emerge at the sentence level, with difficulties arising from different aspects of the sentences—grammatical complexity in nfv-PPA and length of utterances

in lv-PPA. Utilizing this difference in underlying mechanisms to create a sentence comprehension test in which these differences result in different scores will be fruitful in distinguishing nfv-PPA from lv-PPA.

Similarly, a more extensive assessment of repetition deficits may shed light on the differences between lv-PPA and nfv-PPA. Performance on the SYDBAT Repetition subtest showed no difference between lv-PPA and nfv-PPA subtypes, with half of the nfv-PPA patients and less than half of the lv-PPA patients (7/20) showing a clinical impairment on this subtest. While both subtypes thus show repetition errors, the underlying mechanism causing these errors is likely to be different, since the impairments in word repetition in lv-PPA patients are likely to reflect a phonological rather than the articulatory impairment that typifies nfv-PPA (Leyton, Savage, et al., 2014). Moreover, the repetition impairments in lv-PPA may be more prominent in tasks that measure repetition at the sentence level, as a phonologic short-term memory deficit is thought to be a key cognitive mechanism underlying deficits in lv-PPA. For the SYDBAT Repetition subtest, both articulatory as well as phonological errors are counted as errors (see Supplemental Material S1 for the SYDBAT scoring instructions). Future studies are necessary to determine whether a qualitative analysis of the SYDBAT Repetition errors in both lv-PPA and nfv-PPA patients or the addition of a sentence repetition test might aid in the differentiation of these variants, especially because previous studies suggest that this could possibly be used as a clinical marker for underlying amyloid burden in PPA (Leyton, Ballard, et al., 2014).

Lastly, subtest items of the SYDBAT are divided in three blocks of difficulty based on lexical frequency. Previous studies (Savage et al., 2013) and the current study do not use difficulty based on lexical frequency as a predictor in their analyses due to the goal of creating a short screen that can be easily administered, scored, and interpreted in clinical practice (in which cutoff scores are often preferred). However, it would be interesting for a future study to assess whether item difficulty based on lexical frequency is reflected in patient scores and whether sensitivity and specificity would be improved by including, for example, only the most difficult items of the SYDBAT. In addition, using only the most difficult items of the SYDBAT could be beneficial for improving internal consistency of

the SYDBAT subtests, especially for the Word Comprehension subtest.

In addition to the PPA profiles, this study is one of the first to report on the SYDBAT in a non-PPA patient sample. The SYDBAT profiles of the non-PPA patient group (MCI and AD), showing Naming and Semantic Association scores to be lowest with relatively preserved Repetition, are in agreement with previous findings reporting deficits in confrontation naming as well as semantic tests like category fluency, semantic categorization, and lexical decision in this group (Taler & Phillips, 2008). Even though Naming and Semantic Association scores were significantly lower in the AD group compared to the MCI group, MCI patients performed significantly lower than cognitively unimpaired controls. This confirms increasing evidence that language deficits begin several years before the onset of AD dementia (Auriacombe et al., 2006) and that differences in language features between AD and MCI reflect predominantly quantitative instead of qualitative differences (Jokel et al., 2019). Earlier studies (Alexopoulos et al., 2006) have shown that individuals with multidomain MCI, including language deficits, are more likely to develop AD than those with an isolated memory deficit. Clarification of the nature of language impairment and development of sensitive measures for language impairment in these patient groups, therefore, constitute essential tools for the early detection of AD (Taler & Phillips, 2008; Vonk et al., 2020). The application of the SYDBAT could therefore possibly be useful in the diagnostic process of MCI and AD.

The comparison between the SYDBAT profiles of our PPA and non-PPA patients showed a large amount of overlap between groups. MCI patients presented with a similar profile as lv-PPA patients on the semantically based subtests (Naming, Word Comprehension, and Semantic Association). The profile of AD patients greatly overlapped with that of sv-PPA patients, with the exception of Naming, on which sv-PPA patients scored significantly lower than AD patients. As a result, it may be difficult to differentiate MCI, AD, and PPA based on cross-sectional SYDBAT scores alone, which was also confirmed by the discriminative analysis for all groups resulting in a cross-validated classification accuracy of 50.6%. A future study that includes both cognitive and language data in a discriminant analysis may be useful to determine whether this results in a higher classification accuracy. It should be noted, however, that the SYDBAT is intended for use as a first screen in patients visiting memory clinics, in which language is generally not assessed in detail. While the ease of use of the SYDBAT will facilitate its use in clinical practice, SYDBAT scores below cutoff warrant the use of additional, more in depth, approaches to language assessment and the assessment of other cognitive domains in these groups, which may prove more fruitful in the early differential diagnosis. For example, longitudinal assessment using the SYDBAT might be more informative, as language abilities have been shown to decline faster in PPA compared to AD (Blair et al., 2007). In addition, tests that measure semantic abilities at a more precise level than standardized semantic tasks, for example,

by looking at semantic interference (Vandenberghe et al., 2005), might show more fine-grained distinctions between AD, MCI, and PPA.

An examination of psychometric properties showed good reliability and validity measures for the SYDBAT, yet future studies are needed to confirm these properties in larger samples. Cronbach's alpha indicates that the SYDBAT has high levels of internal consistency, especially for the Naming, Repetition, and Semantic association subtests. Convergent construct validity was supported by moderate-to-large positive correlations between SYDBAT performance and performance on the standardized tests: the BNT, the CAT-NL, and the SAT. Divergent validity analyses showed all measures of Naming, Word Comprehension, and Semantic Association to be positively correlated, which could be explained by the fact that all these subtests contain a semantic component. As expected, however, the Repetition subtest was unrelated to any of the semantically based subtests, confirming the divergent validity to be acceptable. As hypothesized, in terms of PPA classification, the SYDBAT outperforms commonly used language tests that have been primarily developed for the assessment of aphasia syndromes after stroke. This suggests tests that are specifically designed and validated for the PPA population, like the SYDBAT, to be a useful and desirable element in the diagnostic process of PPA in a clinical setting.

The limitations of our study lie in the relatively small sample sizes and the reliance on clinical diagnosis that can limit the validity because of potential misclassification. However, it should be noted that large study samples of PPA patients are relatively rare given the low prevalence of PPA (Matías-Guiu & García-Ramos, 2013) and the sample size of the current study is similar to that used in previous studies (Savage et al., 2013). Also, the combination of complementary clinical diagnostic tools such as MRI, cerebrospinal fluid analysis, and neuropsychological assessment, as used in our patient sample, can assure a good diagnostic reliability in the clinical assessment of PPA (Grand et al., 2011). The risk of misclassification was thus minimized as much as possible in the current study. Finally, the SYDBAT and other neuropsychological tests were administered concurrently for the majority of the patients. One could argue that this bears the risk of circularity. However, the clinical diagnoses were primarily based on the elaborative language assessment using established language tests, available neuroimaging, and/or other biomarkers in a multidisciplinary way.

It should be noted that the SYDBAT is not designed to determine an individual's full linguistic profile. An extensive diagnostic assessment, therefore, is still required to measure and observe all the important characteristics of speech, including agrammatism and apraxia of speech. In addition, since the items consist of nouns only, the comprehension and repetition of single words as measured by the subtests of the SYDBAT might be intact in the early stages of the disease (Leyton et al., 2013). Therefore, the comprehension of more complex grammatical structures and the repetition of phrases and sentences should be included in a

subsequent more extensive neuropsychological or linguistic assessment.

To conclude, the SYDBAT is a promising screening instrument developed to address the paucity of simple and short tools suitable for neuropsychological assessment to help discriminate between the three variants of PPA. A diagnostic decision tree was developed to assist clinicians to determine PPA subtype in a straightforward way. The SYDBAT profiles of AD and MCI patients support previous work that suggests language impairments to be frequent and primarily semantic in nature, yet based on the SYDBAT alone it was difficult to differentiate PPA from MCI and AD. Despite the SYDBAT's limitations, combining it with the outcome of other language and cognitive tests may aid to improve the diagnostic accuracy of PPA and its subtypes in a memory clinic setting.

Acknowledgments

The authors would like to thank Ilja Klabbers-Helsper for her help in selecting and evaluating patients and William M. van der Veld for his statistical advice. This study was funded by the Gravitation Grant 024.001.006 of the Language in Interaction Consortium from the Netherlands Organization for Scientific Research (NWO) awarded to R. P. C. K., and A. R. O. P. is supported by a National Health and Medical Research Council of Australia Senior Research Fellowship (APP1103258). V. P. is supported by a grant from the Netherlands Organization for Scientific Research (NWO) under award number 451-17-003.

References

- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the national institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 270–279. <https://doi.org/10.1016/j.jalz.2011.03.008>
- Alexopoulos, P., Grimmer, T., Pernecky, R., Domes, G., & Kurz, A. (2006). Progression to dementia in clinical subtypes of mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, 22(1), 27–34. <https://doi.org/10.1159/000093101>
- Auriacombe, S., Lechevallier, N., Amieva, H., Harston, S., Raoux, N., & Dartigues, J. F. (2006). A longitudinal study of quantitative and qualitative features of category verbal fluency in incident Alzheimer's disease subjects: Results from the PAQUID study. *Dementia and Geriatric Cognitive Disorders*, 21(4), 260–266. <https://doi.org/10.1159/000091407>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Blair, M., Marczyński, C. A., Davis-Faroque, N., & Kertesz, A. (2007). A longitudinal study of language decline in Alzheimer's disease and frontotemporal dementia. *Journal of the International Neuropsychological Society*, 13(2), 237–245. <https://doi.org/10.1017/S1355617707070269>
- Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the Trail Making Test. *Nature Protocols*, 1(5), 2277–2281. <https://doi.org/10.1038/nprot.2006.390>
- Bucks, R. S., Willison, J. R., Byrne, L. M. T., & Kessels, R. P. C. (2011). *Location Learning Test—Revised Edition (English Version)*. The University of Manchester.
- Butts, A. M., Machulda, M. M., Duffy, J. R., Strand, E. A., Whitwell, J. L., & Josephs, K. A. (2015). Neuropsychological profiles differ among the three variants of primary progressive aphasia. *Journal of the International Neuropsychological Society*, 21(6), 429–435. <https://doi.org/10.1017/S1355617715000399>
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81–105. <https://doi.org/10.1037/h0046016>
- Clark, H. M., Utianski, R. L., Duffy, J. R., Strand, E. A., Botha, H., Josephs, K. A., & Whitwell, J. L. (2020). Western Aphasia Battery—Revised profiles in primary progressive aphasia and primary progressive apraxia of speech. *American Journal of Speech-Language Pathology*, 29(1S), 498–510. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0217
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.
- Cosentino, S., Chute, D., Libon, D., Moore, P., & Grossman, M. (2006). How does the brain support script comprehension? A study of executive processes and semantic knowledge in dementia. *Neuropsychology*, 20(3), 307–318. <https://doi.org/10.1037/0894-4105.20.3.307>
- Croot, K., Ballard, K., Leyton, C. E., & Hodges, J. R. (2012). 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*, 55(5), 1562–1572. [https://doi.org/10.1044/1092-4388\(2012/11-0323\)](https://doi.org/10.1044/1092-4388(2012/11-0323))
- Duits, A., & Kessels, R. (2012). Schatten van het premorbide functioneren [Estimating premorbid functioning]. In M. Hendriks, R. Kessels, M. Gorissen, & B. Schmand (Eds.), *Neuropsychologische diagnostiek* [Neuropsychological assessment] (pp. 119–129). Boom.
- Eikelboom, W. S., Janssen, N., Jiskoot, L. C., van den Berg, E., Roelofs, A., & Kessels, R. P. C. (2018). Episodic and working memory function in primary progressive aphasia: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, 92, 243–254. <https://doi.org/10.1016/j.neubiorev.2018.06.015>
- Eikelboom, W. S., Janssen, N., van den Berg, E., Roelofs, A., & Kessels, R. (2017). Differentiatie van primair-progressieve afasie varianten: De Nederlandse bewerking van de Sydney Language Battery (SYDBAT-NL) [Differentiation of primary progressive aphasia variants: The Dutch version of the Sydney Language Battery (SYDBAT-NL)]. *Tijdschrift voor Neuropsychologie*, 12, 189–202.
- Emery, V. O. B. (2000). Language impairment in dementia of the Alzheimer type: A hierarchical decline? *International Journal of Psychiatry in Medicine*, 30(2), 145–164. <https://doi.org/10.2190/X09P-N7AU-UCHA-VW08>
- Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M., & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, 76(11), 1006–1014. <https://doi.org/10.1212/WNL.0b013e31821103e6>
- Graham, N. L., Patterson, K., & Hodges, J. R. (2004). When more yields less: Speaking and writing deficits in nonfluent progressive aphasia. *Neurocase*, 10(2), 141–155. <https://doi.org/10.1080/13554790409609945>

- Grand, J. H. G., Caspar, S., & MacDonald, S. W. S. (2011). Clinical features and multidisciplinary approaches to dementia care. *Journal of Multidisciplinary Healthcare*, 2011(4), 125–147. <https://doi.org/10.2147/JMDH.S17773>
- Henry, M. L., & Grasso, S. M. (2018). Assessment of individuals with primary progressive aphasia. *Seminars in Speech and Language*, 39(3), 231–241. <https://doi.org/10.1055/s-0038-1660782>
- Jokel, R., Lima, B. S., Fernandez, A., & Murphy, K. J. (2019). Language in amnesic mild cognitive impairment and dementia of Alzheimer's type: Quantitatively or qualitatively different? *Dementia and Geriatric Cognitive Disorders Extra*, 9(1), 136–151. <https://doi.org/10.1159/000496824>
- Josephs, K. A., Duffy, J. R., Strand, E. A., Whitwell, J. L., Layton, K. F., Parsi, J. E., Hauser, M. F., Witte, R. J., Boeve, B. F., Knopman, D. S., Dickson, D. W., Jack, C. R., Jr., & Petersen, R. C. (2006). Clinicopathological and imaging correlates of progressive aphasia and apraxia of speech. *Brain*, 129(6), 1385–1398. <https://doi.org/10.1093/brain/awl078>
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test*. Lea & Febiger.
- Kertesz, A., McMonagle, P., Blair, M., Davidson, W., & Munoz, D. G. (2005). The evolution and pathology of frontotemporal dementia. *Brain*, 128(9), 1996–2005. <https://doi.org/10.1093/brain/awh598>
- Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, 42(3), 643–650. <https://doi.org/10.3758/BRM.42.3.643>
- Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583–621. <https://doi.org/10.1080/01621459.1952.10483441>
- 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. <https://doi.org/10.1212/WNL.0000000000000387>
- Leyton, C. E., Hsieh, S., Mioshi, E., & Hodges, J. R. (2013). Cognitive decline in logopenic aphasia: More than losing words. *Neurology*, 80(10), 897–903. <https://doi.org/10.1212/WNL.0b013e318285c15b>
- Leyton, C. E., Savage, S., Irish, M., Schubert, S., Piguet, O., Ballard, K. J., & Hodges, J. R. (2014). Verbal repetition in primary progressive aphasia and Alzheimer's disease. *Journal of Alzheimer's Disease*, 41(2), 575–585. <https://doi.org/10.3233/JAD-132468>
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). Oxford University Press.
- 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), 154–158. <https://doi.org/10.1093/geronj/47.3.P154>
- Marshall, C. R., Hardy, C. J. D., Volkmer, A., Russell, L. L., Bond, R. L., Fletcher, P. D., Clark, C. N., Mummery, C. J., Schott, J. M., Rossor, M. N., Fox, N. C., Crutch, S. J., Rohrer, S. D., & Warren, J. D. (2018). Primary progressive aphasia: A clinical approach. *Journal of Neurology*, 265(6), 1474–1490. <https://doi.org/10.1007/s00415-018-8762-6>
- Matías-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. <https://doi.org/10.1007/s00415-014-7608-0>
- Matías-Guiu, J. A., & García-Ramos, R. (2013). Primary progressive aphasia: from syndrome to disease. *Neurologia (English Edition)*, 28(6), 366–374. <https://doi.org/10.1016/j.nrleng.2012.04.018>
- McDonald, J. H. (2009). *Handbook of biological statistics* (Vol. 2, pp. 6–59). Sparky House Publishing. <http://www.biostathandbook.com/>
- Mesulam, M. M. (1982). Slowly progressive aphasia without generalized dementia. *Annals of Neurology*, 11(6), 592–598. <https://doi.org/10.1002/ana.410110607>
- Mesulam, M. M., Grossman, M., Hillis, A., Kertesz, A., & Weintraub, S. (2003). The core and halo of primary progressive aphasia and semantic dementia. *Annals of Neurology*, 54(S5), S11–S14. <https://doi.org/10.1002/ana.10569>
- Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2008). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Nederlandse Vereniging voor Klinische Geriatrie. (2014). *Richtlijn diagnostiek en behandeling van dementie [Guideline for diagnosis and treatment of dementia]*. https://www.richtlijnen database.nl/richtlijn/dementie/dementie_-_korte_beschrijving.html
- Newhart, M., Davis, C., Kannan, V., Heidler-Gary, J., Cloutman, L., & Hillis, A. E. (2009). Therapy for naming deficits in two variants of primary progressive aphasia. *Aphasiology*, 23(7–8), 823–834. <https://doi.org/10.1080/02687030802661762>
- Oosterman, J. M., Vogels, R. L. C., van Harten, B., Gouw, A. A., Poggesi, A., Scheltens, P., Kessels, R. P. C., & Scherder, E. J. A. (2010). Assessing mental flexibility: Neuroanatomical and neuropsychological correlates of the trail making test in elderly people. *The Clinical Neuropsychologist*, 24(2), 203–219. <https://doi.org/10.1080/13854040903482848>
- Osterrieth, P. A. (1944). Le test de copie d'une figure complexe: Contribution à l'étude de la perception et de la mémoire [The test of copying a complex figure: A contribution to the study of perception and memory]. *Archives de Psychologie*, 30, 206–356.
- Perry, R. J., & Hodges, J. R. (2000). Differentiating frontal and temporal variant frontotemporal dementia from Alzheimer's disease. *Neurology*, 54(12), 2277–2284. <https://doi.org/10.1212/WNL.54.12.2277>
- Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, 256(3), 183–194. <https://doi.org/10.1111/j.1365-2796.2004.01388.x>
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, 56(3), 303–308. <https://doi.org/10.1001/archneur.56.3.303>
- Piguet, O., Leyton, C. E., Gleeson, L. D., Hoon, C., & Hodges, J. R. (2015). Memory and emotion processing performance contributes to the diagnosis of non-semantic primary progressive aphasia syndromes. *Journal of Alzheimer's Disease*, 44(2), 541–547. <https://doi.org/10.3233/JAD-141854>
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). Russell Sage Foundation.
- 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. <https://doi.org/10.1159/000346389>
- Schmand, B., Groenink, S. C., & Van den Dungen, M. (2008). Letter fluency: Psychometric properties and Dutch normative

- data. *Tijdschrift voor Gerontologie en Geriatrie*, 39(2), 64–76. <https://doi.org/10.1007/BF03078128>
- Spearman, C.** (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 15(1), 72–101. <https://doi.org/10.2307/1412159>
- Taler, V., & Phillips, N. A.** (2008). Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *Journal of Clinical and Experimental Neuropsychology*, 30(5), 501–556. <https://doi.org/10.1080/13803390701550128>
- Tippett, D. C., Hillis, A. E., & Tsapkini, K.** (2015). Treatment of primary progressive aphasia. *Current Treatment Options in Neurology*, 17(8), Article 34 <https://doi.org/10.1007/s11940-015-0362-5>
- Vandenbergh, R. R., Vandenbulcke, M., Weintraub, S., Johnson, N., Porke, K., Thompson, C. K., & Mesulam, M. M.** (2005). Paradoxical features of word finding difficulty in primary progressive aphasia. *Annals of Neurology*, 57(2), 204–209. <https://doi.org/10.1002/ana.20362>
- Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J.** (2006). Normative data for the animal, profession and letter m naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex. *Journal of the International Neuropsychological Society*, 12(1), 80–89. <https://doi.org/10.1017/S1355617706060115>
- Van Loon-Vervoorn, W. A., & Van der Velden, H.** (2006). De boston benoemtaak: Normen voor gezonde en hersenbeschadigde personen [Boston Naming Task: Norms for healthy and brain-injured individuals]. *Tijdschrift voor neuropsychologie*, 1, 30–35.
- Visch-Brink, E. G., Denes, G., & Stronks, D.** (1996). Visual and verbal semantic processing in aphasia. *Brain and Language*, 55, 130–132.
- Visch-Brink, E. G., Van den Borre, D., de Smet, H. J., & Mariën, P.** (2014). *CAT-NL. Dutch adaptation of K. Swinburn, G. Porter, D. Howard (2004). Comprehensive Aphasia Test.* Pearson Assessment and Information.
- Vonk, J. M. J., Bouteloup, V., Mangin, J. F., Dubois, B., Blanc, F., Gabelle, A., Ceccaldi, M., Annweiler, C., Krolak-Salmon, P., Belin, C., Rivasseau-Jonveaux, T., Julian, A., Sellal, F., Magnin, E., Chupin, M., Habert, M.-O., Chêne, G., Dufouil, C., & MEMENTO cohort Study Group.** (2020). Semantic loss marks early Alzheimer's disease-related neurodegeneration in older adults without dementia. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 12(1), e12066. <https://doi.org/10.1002/dad2.12066>
- Wechsler, D.** (2008). *Wechsler Adult Intelligent Scale* (4th ed.). APA PsycTests.