Review article

Episodic and working memory function in Primary Progressive Aphasia: A meta-analysis

Willem S. Eikelboom\(^a,1\), Nikki Janssen\(^{b,c,1}\), Lize C. Jiskoot\(^a,d\), Esther van den Berg\(^a\), Ardi Roelofs\(^b\), Roy P.C. Kessels\(^{b,c,e}\)

\(^a\) Department of Neurology, Erasmus Medical Center, Rotterdam, The Netherlands
\(^b\) Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands
\(^c\) Department of Medical Psychology, Radboud University Medical Center, Nijmegen, The Netherlands
\(^d\) Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands
\(^e\) Vincent van Gogh Institute for Psychiatry, Venray, The Netherlands

**ARTICLE INFO**

**Keywords:**
Primary progressive aphasia
Progressive nonfluent aphasia
Logopenic aphasia
Semantic dementia
Frontotemporal dementia
Memory
Systematic review
Meta-analysis

**ABSTRACT**

**Objective:** The distinction between Primary Progressive Aphasia (PPA) variants remains challenging for clinicians, especially for the non-fluent (nfv-PPA) and the logopenic variants (lv-PPA). Previous research suggests that memory tests might aid this differentiation. This meta-analysis compares memory function among PPA variants.

**Method:** Effects sizes were extracted from 41 studies (N = 849). Random-effects models were used to compare performance on episodic and working memory tests among PPA patients and healthy controls, and between the PPA variants.

**Results:** Memory deficits were frequently observed in PPA compared to controls, with large effect sizes for lv-PPA (Hedges' \(g = -2.04 \, [-2.58 \text{ to } -1.49]\)), nfv-PPA (Hedges' \(g = -1.26 \, [-1.60 \text{ to } -0.92]\), \(p < .001\)), and the semantic variant (sv-PPA; Hedges' \(g = -1.23 \, [-1.50 \text{ to } -0.97]\)). Sv-PPA showed primarily verbal memory deficits, whereas lv-PPA showed worse performance than nfv-PPA on both verbal and non-verbal memory tests.

**Conclusions:** Memory deficits were more pronounced in lv-PPA compared to nfv-PPA. This suggests that memory tests may be helpful to distinguish between these PPA variants.

---

1. Introduction

Primary Progressive Aphasia (PPA) is a rare neurodegenerative disorder characterized by a progressive decline in language functions (Matías-Guiu and García-Ramos, 2013; Mesulam, 1982; Mesulam and Weintraub, 1992). The most recent diagnostic guidelines distinguish three PPA variants based on differences in linguistic deficits and underlying neuropathology (Gorno-Tempini et al., 2011). The semantic variant (sv-PPA) involves semantic deficits and impairments in confrontation naming and word comprehension. The logopenic variant (lv-PPA) includes difficulties with word retrieval and naming in spontaneous speech, as well as impaired repetition of sentences and phrases. The non-fluent/agrammatic-variant (nfv-PPA) consists of agrammatism in language production and effortful, slowed speech together with apraxia of speech (Gorno-Tempini et al., 2011).

Despite these criteria, the distinction between the different PPA subtypes remains complex and challenging for clinicians. This holds for lv-PPA and nfv-PPA in particular, because both subtypes overlap with respect to several linguistic deficits (Croot et al., 2012). This highlights the need to establish other clinical markers that can reliably distinguish between subtypes. Recent studies have suggested that deficits in cognitive domains other than language may be promising in this respect (Kielb et al., 2016; Ramanan et al., 2016).

The cognitive domain of memory could possibly function as such a behavioural marker to facilitate the distinction between PPA variants (e.g., Piguet et al., 2015; Ramanan et al., 2016). That is, both subjective memory complaints by patients and caregivers (Magnin et al., 2013; Weintraub et al., 2013), and objective memory impairments have been described in the literature, even in PPA patients in the early phase of the disorder (e.g., Flanagan et al., 2014; Gorno-Tempini et al., 2004). Previous research showed that the prevalence and extent of memory deficits differs across PPA variants, with evidence that both episodic...
memory and working memory deficits are prevalent in patients with lv-PPA (e.g., Butts et al., 2015; Flanagan et al., 2014; Foxe et al., 2013).

The differences in memory profile across PPA subtypes can be explained by the distinctive underlying neuropathology among these subtypes. Sv-PPA and nfv-PPA have both been related to frontotemporal lobar degeneration (FTLD) spectrum (respectively FTLD TDP-43 and FTLD-tau pathology; Grossman, 2012; Hodges & Patterson, 2007), whereas the majority of patients with lv-PPA show pathology that has been related to Alzheimer’s disease (AD; Gorno-Tempini et al., 2004; Mesulam et al., 2003).

One complicating factor in the assessment of memory is that many neuropsychological tests make use of verbal instructions, verbal stimuli and require a verbal response. As a consequence, aphasia severity negatively affects neuropsychological performance in PPA in any cognitive domain (Machulda et al., 2013). Thus, the question arises whether the subjective and objective memory difficulties observed in PPA patients can be attributed to language impairments or can be considered as an independent deficit.

Previous studies on PPA have included small numbers of patients given the low prevalence of this syndrome. In order to gain more insight into the extent of memory function in PPA patients, a quantitative meta-analytic approach is preferred (Grossman, 2010). To date, memory performance and its manifestations in different PPA subtypes have not been systematically reviewed, despite many individual studies in this area. The aim of this study is to systematically review the existing studies covering memory functioning in PPA patients, and apply meta-analytic techniques to establish and compare the nature, extent and prevalence of memory impairments among PPA variants. For each PPA variant we aimed to (i) directly compare episodic memory and working memory function, and (ii) compare the performance on both verbal and nonverbal memory tests to examine whether memory dysfunction exceeds verbal memory.

Based on previous studies, we hypothesize that episodic memory dysfunction is most pronounced in lv-PPA considering its characterization by a disruption of the temporoparietal circuitry (Gorno-Tempini et al., 2004). To a lesser extent, patients with sv-PPA can be expected to show a worse episodic memory performance on verbal tests only given the (left) anterior temporal lobe atrophy often associated with this subtype (Rohrer et al., 2010). Episodic memory function is expected to be mostly intact in patients with nfv-PPA consistent with the relatively spared temporal lobe (Hornberger and Piguet, 2012). Working memory deficits in turn may be most frequently impaired in lv-PPA due to the loss of storage and rehearsal processes of the phonological system caused by left temporoparietal atrophy (Gorno-Tempini et al., 2008). Verbal working memory is hypothesized to be more impaired compared to the non-verbal working memory given the spared right parietal and frontal regions in the early phases of lv-PPA (Gorno-Tempini et al., 2004). Nfv-PPA patients may show some working memory problems, whereas these deficits will be rare in sv-PPA patients (Carthey-Goullart et al., 2012).

2. Methods

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were followed to perform and report this meta-analysis (Moher et al., 2009).

2.1. Search strategy

With the help of a university librarian, appropriate MeSH terms and entry terms were identified. Consequently, a literature search in PubMed was conducted with the following search terms: “primary progressive aphasia”, “pick’s disease of the brain”, “frontotemporal dementia”, “memory”, “cognition”, “cognitive dysfunction”, “neuropsychological tests”. In addition, reference lists of identified papers where manually checked for potential articles. The guidelines by Gorno-Tempini et al. (2011) were used to define PPA and its variants. For articles published before 2011, characterisations of the semantic variant by Hodges et al. (1992), the non-fluent variant by Grossman et al. (1996), and the logopenic variant by Gorno-Tempini et al. (2004) were used. The last search was carried out in October 2017 and updated in May 2018.

2.2. Study selection

For this review, articles were selected only when the following criteria were met: a) the performance on a memory test had to be one of the outcome measures or had to be reported in the characteristics; b) a healthy control (HC) group was included; c) PPA patients had to be classified as having one of the three PPA subtypes; d) studies had to report sufficient information (e.g., means, standard deviations, exact p-values, or standardized effect sizes) in order to be able to perform a meta-analysis. Only a few research groups are extensively investigating PPA patients by the use of large cohort groups. Therefore, in cases in which there was a probability that the same patient sample was used by different studies, the study’s principal investigator (PI) was contacted by email with the question to comment on possible overlap. Based on the PI’s response we only included the studies that had either no or minimal (< 10%) overlap in patient sample as compared to other included studies from the concerned research group (Bown and Sutton, 2010). Study selection was done for each PPA variant, memory domain and verbal or non-verbal tests separately. Case studies (N ≤ 5) and animal studies were excluded. These criteria were examined by careful screening of the titles and abstracts of English-language research articles. Subsequently, the full-text papers were screened for eligibility by two independent raters (WE and NJ). Disagreements were resolved through discussion until consensus was reached. Only studies on which both authors agreed were included in the final systematic review and meta-analysis.

2.3. Data synthesis

First, an overall effect size (ES) was calculated including all PPA subtypes and memory domains. Next, categorical analyses were run, in order to answer all research questions. For each PPA variant 1) performance on memory tests was compared to the performance of HC; 2) performance on episodic memory tests and working memory tests was compared to that of HC; 3) the performance on non-verbal and on verbal memory tests was compared. For the classification of memory tests, Lezak et al.’s (2012) handbook for neuropsychological assessment was used. Widely reported examples of working memory tasks included span tasks such as the Digit Span, Corsi Blocks, and Spatial Span. Widely reported episodic memory tasks included word-list learning tasks such as the California Verbal Learning Test, Philadelphia Verbal Learning Test, Rey Auditory Verbal Learning Tests, and tests such as the Rey’s Complex Figure Test and subtests from the Wechsler Memory Scales (Lezak et al., 2012). Tests were considered verbal in nature if the material presented (either visually or auditory) were digits, words, sentences or stories. Tests were nonverbal in nature if stimuli were pictures of objects, scenes, line drawings or abstract figures.

2.4. Statistical analysis

In order to conduct analyses, means, standard deviations, and samples sizes for the PPA subtypes and memory tests were extracted from the studies or, if necessary, acquired through personal correspondence. All means (M) and standard deviations (SD) were converted into a summary statistics (Hedges’ g) based on the following formula: $g = \frac{M_1 - M_2}{SD_{pooled}}$, where $SD_{pooled}$ was calculated using the following formula:

$$SD_{pooled} = \sqrt{\frac{SD_1^2 + SD_2^2}{2}}.$$  

A negative effect size (ES) indicates that the performance of PPA patients is lower compared to HC. If more than one
measure was reported, an average ES was calculated. The computed ESs were interpreted according to Cohen’s (1992) convention of small (0.10), medium (0.30), and large (0.50) effects. Sample sizes were incorporated to correct for the biased ES in studies with small sample sizes (Hedges and Olkin, 1985).

Random-effects models were used since a substantial heterogeneity was expected between studies, with regard to study design and patient samples, which these models are able to account for (DerSimonian & Kacker, 2007). In addition, random-effects models are preferred when the aim is to generalize the results beyond the observed studies (Clark-Carter, 2010).

Heterogeneity was checked for each analysis by the use of the chi-square homogeneity test (Q) and the inconsistency statistic (I²). To check the possibility of a publication bias (the degree of unpublished null-findings), the fail-safe N was calculated and a funnel plot was made (Rosenthal, 1991). To rule out a possible publication bias, the fail-safe N must be larger than \((5 \times k) + 10\), where \(k\) is the number of studies included in the meta-analysis (Clark-Carter, 2010). The funnel plot should reveal the studies included as distributed around the mean ES in a funnel shape. Studies that fall outside the funnel shape have a high risk of bias (Borenstein et al., 2009). All analyses were performed using Comprehensive Meta-Analysis version 2.0 (Engelwood, NJ, USA, 2005).

3. Results

3.1. Study characteristics

The literature search resulted in a total of 1546 articles published between 1979 and 2018. Of these, 1062 were excluded after reviewing the titles and abstracts for eligibility. Full versions were retrieved for 484 articles, of which 40 articles were eligible for inclusion. Based on the responses of PIs, we excluded three studies because of possible overlap and included four studies with no or minimal overlap, resulting in a total of 41 studies. Fig. 1 shows the flowchart of this search and Table 1 lists the characteristics of the included studies. Studies labeled with an * in the references were included in the meta-analysis.

3.2. Overall effect

Twenty-nine studies included a total of 450 sv-PPA patients, and the analysis showed a large ES of \(-1.23\) \((-1.50\) to \(-0.97\), \(p < .001\)). The analysis of the twelve studies including 212 nf-PPA patients resulted in a large ES of \(-1.26\) \((-1.60\) to \(-0.92\), \(p < .001\)). Eleven studies included a total of 187 lv-PPA patients, with the analysis showing a large overall ES of \(-2.04\) \((-2.58\) to \(-1.49\), \(p < .001\)), which was significantly lower compared to the other PPA variants \((p < .05)\). For these analyses, the heterogeneity indices (Q) were significant \((p < .05)\), indicating heterogeneity in study outcomes. Table 2 shows the results of the meta-analyses.

3.3. Episodic memory

Sv-PPA patients \((k = 19, g = -1.79, [-2.15 to -1.44], p < .001)\), lv-PPA patients \((k = 8, g = -1.52, [-1.88 to -1.15], p < .001)\) and nf-PPA patients \((k = 8, g = -0.87, [-1.18 to -0.56], p < .001)\) performed significantly worse on episodic memory tests compared to HC. Categorical analysis showed a significant difference among PPA subtypes with sv-PPA = lv-PPA < nf-PPA < HC \((p < .05)\). Lv-PPA patients thus performed similar to sv-PPA patients, but significantly worse than nf-PPA patients on episodic memory tests.

3.4. Working memory

All PPA subtypes performed significantly lower than HC on working memory \((all\ p-values < .05)\). Categorical analysis showed a significant difference between PPA subtypes, with lv-PPA patients \((k = 7, g = -2.83, [-3.73 to -1.93], p < .001)\) performing worse than nf-PPA patients \((k = 9, g = -1.21, [-1.94 to -1.47], p < .001)\), and nf-PPA patients performing worse on working memory tests compared to patients with sv-PPA \((k = 15, g = 0.51, [-0.75 to 0.26], p < .001)\).

3.5. Performance on verbal vs. nonverbal tests

For sv-PPA patients, the performance on verbal episodic memory tests \((k = 11, g = -2.50, p < .01)\) was significantly lower compared to non-verbal episodic memory tests \((k = 16, g = -1.40, p < .001)\). However, the performance of sv-PPA on non-verbal episodic memory tests was still significantly lower compared to HC (Fig. 2). Both nf-PPA patients and lv-PPA patients performed worse than HC on verbal \((k = 7, g = -0.87, p < .001\) and \(k = 5, g = -1.47, p < .001\)) and non-verbal \((k = 7, g = -0.90, p < .001\) and \(k = 6, g = -1.48, p < .001\)) episodic memory tests. There was no significant difference between performance on verbal and non-verbal episodic memory tests for both nf-PPA and lv-PPA patients, however \((p > .05; Figs. 3 and 4)\).

Patients with sv-PPA performed significantly worse compared to HC on verbal working memory tests \((k = 13, g = -0.61, p < .001)\) and non-verbal working memory tests \((k = 3, g = -1.67, p < .001)\). There was no difference between verbal and non-verbal test performance in nf-PPA patients \((p > .05; Fig. 5)\).

Patients with nf-PPA had significantly worse scores compared to HC on both verbal working memory tests \((k = 8, g = -1.76, p < .001)\) and non-verbal working memory tests \((k = 3, g = -1.67, p < .001)\). There was no difference between verbal and non-verbal test performance in nf-PPA patients \((p > .05; Fig. 6)\).

Patients with lv-PPA performed significantly worse on verbal working memory tests \((k = 5, g = -2.15, p < .001)\) and on non-verbal working memory test compared to HC \((k = 3, g = -4.71, p < .001)\). The performance on non-verbal tests was similar compared to that on verbal tests \((p > .05, Fig. 7)\).

3.6. Risk of publication bias

The Fail-safe N was calculated for each analysis, in order to estimate the number of unpublished studies with effect size zero that could be added to the meta-analysis before the result lost statistical significance. As shown in Table 2, the number of studies needed ranged from 744 for nf-PPA to 2702 for sv-PPA for the overall effects. For the sub-analysis the number of studies needed ranged from 129 to 2227 for sv-PPA. The estimated fail-safe N was thereby larger than \((5 \times k) + 10\) for all studies. The funnel plots show the relation between sample size and ES (Fig. 8). Visual inspection of the funnel plots reveals an asymmetry in the distribution of the included studies in sv-PPA and lv-PPA. This asymmetry might be due to heterogeneity in outcome measures \(\text{e.g., non-verbal or verbal tests, episodic memory or working memory tests}\) and therefore show a larger or smaller ES independent of the included sample size, since differences in memory performance are due to the tests used.

4. Discussion

In this meta-analysis, we investigated and compared the prevalence, nature and extent of episodic memory and working memory impairments in PPA and its variants. In addition, to examine whether this memory dysfunction might be only a secondary manifestation of the prominent language deficits, performance of PPA patients on both verbal and non-verbal memory tests was compared.

4.1. Differences in episodic memory

With regard to episodic memory, the test performance was found to
be compromised in all PPA variants compared to HC. However, somewhat different from what we expected, the categorical comparison of episodic memory performance between the different PPA variants showed sv-PPA patients being impaired to a similar extent as lv-PPA patients, which in turn were more impaired than nfv-PPA patients. Yet, the significant impairment in episodic memory in sv-PPA patients appears to be mainly driven by verbal memory test performance, as was expected. As already proposed by Hornberger and Piguet (2012), patients with sv-PPA may perform poorly on verbal episodic memory tests since these tests require verbal output that is hampered by the loss of semantic knowledge and anomia in these patients. The left-sided atrophy of the anterior temporal regions often observed in sv-PPA may account for this. Indeed, this is supported by Scahill et al. (2005), who showed that sv-PPA patients with predominant left-sided atrophy performed poorly on verbal memory tests, but within the normal range on nonverbal memory tasks, whereas sv-PPA patients with predominant right-sided atrophy did perform poorly on the non-verbal tests.

The episodic memory deficits in lv-PPA and nfv-PPA, on the other hand, are revealed in both verbal and non-verbal measures. Compared with nfv-PPA, lv-PPA patients have lower verbal as well as non-verbal episodic memory scores, as expected based on neuroanatomical differences involving more temporoparietal disruption in lv-PPA (Gorno-Tempini et al., 2004; Hornberger and Piguet, 2012).

A longstanding view holds that the hippocampus and surrounding medial temporal lobe structures are critical for episodic memory performance. Recent evidence, however, points to a more widely distributed neural network underlying episodic memory (Simons and Spiers, 2003). The presence of episodic memory deficits in various neurodegenerative disorders can therefore be based on different underlying neural substrates, as shown by several studies comparing patients with behavioral variant frontotemporal dementia (bv-FTD) and patients with AD (Irish et al., 2014; Papma et al., 2013; Poos et al., 2016; Tan et al., 2014; Win et al., 2017). Our results, together with the very few studies that have investigated the neural correlates of episodic memory deficits in PPA patients are also in line with this notion. In general, prior studies have revealed that the presence of episodic memory deficits in lv-PPA and sv-PPA patients might be more dependent on disrupted frontal and partial regions and to a lesser degree on hippocampal damage (Irish et al., 2016; Tan et al., 2014; Win et al., 2017).

### 4.2. Differences in working memory

In the working memory domain, performance was also impaired in all PPA variants compared to HC. However, the comparison between performances of PPA variants shows a different profile than observed for episodic memory, with working memory in lv-PPA being more affected than in nfv-PPA, while impairments were less prominent in sv-PPA. The finding of impaired verbal working memory performance in sv-PPA is somewhat surprising, since the most frequently used verbal working memory test (i.e., Digit Span) does not rely heavily on semantic representations. Moreover, in contrast to object knowledge, concepts of quantity such as numbers have been shown to be relatively preserved in sv-PPA (Rascovsky and Grossman, 2013). However, as a
consequence of advanced lexical-semantic degradation in later stages of the disease, impairments of number knowledge have been documented in sv-PPA (Jeannerod et al., 2005). Furthermore, studies investigating white matter tracts in sv-PPA have shown involvement of dorsal tracts in later stages of the disease (Schwindt et al., 2013), which are known to be involved in verbal working memory (Hickok, 2009; Saur et al., 2008). In addition to this, executive dysfunction has been documented as PPA progresses and neuropathology becomes more widespread (Harciarek and Cosentino, 2013). Updating and monitoring of working memory is a crucial component of executive function (Miyake et al., 2000). The reported verbal working memory deficits in sv-PPA could thus be the result of more widespread brain changes and consequently more extensive cognitive deficits.

As hypothesized, lv-PPA patients showed deficits on working memory tasks, not only in the verbal but also in the non-verbal domain. As expected, nfv-PPA patients also showed deficits in working memory, which held for both verbal and non-verbal tests. Overall, our results indicate that working memory performance does differ between lv-PPA and nfv-PPA patients.

Studies investigating the neural correlates of working memory have shown the importance of frontoparietal networks, a set of brain regions encompassing dorsomedial prefrontal, lateral prefrontal, and superior parietal regions of the human cortex. Furthermore, the dorsal white matter pathway connecting these regions appears to be implicated as

Table 1

Characteristics of studies included in the meta-analysis.

<table>
<thead>
<tr>
<th>Study</th>
<th>Patients</th>
<th>Years of symptoms</th>
<th>Healthy controls</th>
<th>Memory domain tested</th>
<th>Memory tests used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adlam et al. (2010)</td>
<td>15 sv-PPA</td>
<td>N/A</td>
<td>20</td>
<td>Episodic &amp; WM</td>
<td>WMS-R LM, DS backward</td>
</tr>
<tr>
<td>Ash et al. (2016)</td>
<td>19 sv-PPA</td>
<td>4.3 ± 2.2</td>
<td>16</td>
<td>WM</td>
<td>DS backward</td>
</tr>
<tr>
<td>Auclair-Ouellet et al. (2016)</td>
<td>10 sv-PPA</td>
<td>N/A</td>
<td>20</td>
<td>Episodic</td>
<td>DS backward</td>
</tr>
<tr>
<td>Binney et al. (2016)</td>
<td>33 sv-PPA</td>
<td>4.2 ± 2.9</td>
<td>14</td>
<td>Episodic &amp; WM</td>
<td>CVLT-SF, RCF, DS backward</td>
</tr>
<tr>
<td>Charles et al. (2013)</td>
<td>12 sv-PPA</td>
<td>N/A</td>
<td>12</td>
<td>Episodic &amp; WM</td>
<td>PVLT, DS backward</td>
</tr>
<tr>
<td>Downey et al. (2015)</td>
<td>15 sv-PPA</td>
<td>6.2 ± 1.9</td>
<td>37</td>
<td>Episodic</td>
<td>RMT faces &amp; words</td>
</tr>
<tr>
<td>Duval et al. (2012)</td>
<td>6–8 sv-PPA</td>
<td>3.3 ± 1.9</td>
<td>36</td>
<td>Episodic</td>
<td>RCF, TdR, WMS-III LM</td>
</tr>
<tr>
<td>Fozz et al. (2016)</td>
<td>15 lv-PPA</td>
<td>4.3 ± 2.9</td>
<td>15</td>
<td>Episodic &amp; WM</td>
<td>RCF, DMS faces &amp; words, WMS-R LM</td>
</tr>
<tr>
<td>Galton et al. (2001)</td>
<td>18 sv-PPA</td>
<td>4.0 ± 2.4</td>
<td>21</td>
<td>Episodic</td>
<td>RCF, DMS faces &amp; words, WMS-R LM</td>
</tr>
<tr>
<td>Gold et al. (2005)</td>
<td>6 sv-PPA</td>
<td>N/A</td>
<td>14</td>
<td>WM</td>
<td>DS backward</td>
</tr>
<tr>
<td>Goll et al. (2011)</td>
<td>7 lv-PPA</td>
<td>4.1 ± 0.9</td>
<td>20</td>
<td>WM</td>
<td>DS backward, WMS-III SS</td>
</tr>
<tr>
<td>Gorno-Tempini et al. (2004)</td>
<td>10 sv-PPA</td>
<td>4.0 ± 1.2</td>
<td>10</td>
<td>Episodic &amp; WM</td>
<td>CVLT-M, RCF, WMS-III faces, DS backward</td>
</tr>
<tr>
<td>Graham et al. (2004)</td>
<td>14 sv-PPA</td>
<td>3.5 ± 1.6</td>
<td>11</td>
<td>WM</td>
<td>DS total</td>
</tr>
<tr>
<td>Halilstone et al. (2012)</td>
<td>6 sv-PPA</td>
<td>3.5 ± 1.3</td>
<td>15</td>
<td>WM</td>
<td>DS backward, WMS-III SS backward</td>
</tr>
<tr>
<td>Hardy et al. (2016)</td>
<td>14 sv-PPA</td>
<td>6.7 ± 4.1</td>
<td>24</td>
<td>Episodic &amp; WM</td>
<td>RMT words, DS backward, SS backward</td>
</tr>
<tr>
<td>Hazleton et al. (2017)</td>
<td>21 sv-PPA</td>
<td>4.3 ± 2.8</td>
<td>24</td>
<td>WM</td>
<td>DS backward</td>
</tr>
<tr>
<td>Hodges et al. (1999)</td>
<td>8 sv-PPA</td>
<td>2.0 – 5.0</td>
<td>8</td>
<td>WM</td>
<td>DS backward</td>
</tr>
<tr>
<td>Hoffman et al. (2009)</td>
<td>6 sv-PPA</td>
<td>3.8 ± 1.2</td>
<td>11</td>
<td>WM</td>
<td>DS backward</td>
</tr>
<tr>
<td>Irish et al. (2016)</td>
<td>20 sv-PPA</td>
<td>4.7 ± 1.7</td>
<td>35</td>
<td>WM</td>
<td>DS total</td>
</tr>
<tr>
<td>Johnson et al. (2011)</td>
<td>20 sv-PPA</td>
<td>N/A</td>
<td>17</td>
<td>WM</td>
<td>SS backward</td>
</tr>
<tr>
<td>Juli et al. (2010)</td>
<td>14 sv-PPA</td>
<td>5.3 ± 1.9</td>
<td>10</td>
<td>Episodic &amp; WM</td>
<td>VOM, DS backward</td>
</tr>
<tr>
<td>Kamminga et al. (2015)</td>
<td>12 sv-PPA</td>
<td>3.4 ± 2.1</td>
<td>20</td>
<td>Episodic</td>
<td>Doors A</td>
</tr>
<tr>
<td>Laisney et al. (2009)</td>
<td>18 sv-PPA</td>
<td>3.4 ± 1.8</td>
<td>18</td>
<td>WM</td>
<td>DS backward, SS backward</td>
</tr>
<tr>
<td>Leyton et al. (2017)</td>
<td>22 sv-PPA</td>
<td>4.0 ± 2.8</td>
<td>29</td>
<td>WM</td>
<td>DS backward</td>
</tr>
<tr>
<td>Mack et al. (2013)</td>
<td>6 lv-PPA</td>
<td>2.8 ± 1.1</td>
<td>17</td>
<td>WM</td>
<td>DS backward</td>
</tr>
<tr>
<td>Magerova et al. (2014)</td>
<td>6 sv-PPA</td>
<td>N/A</td>
<td>15</td>
<td>Episodic &amp; WM</td>
<td>AVLT, RCSR, DS backward</td>
</tr>
<tr>
<td>Magnin et al. (2015)</td>
<td>7 nfv-PPA</td>
<td>N/A</td>
<td></td>
<td>Episodic</td>
<td>DMS-4B, RCF</td>
</tr>
<tr>
<td>Mandelli et al. (2016)</td>
<td>25 nfv-PPA</td>
<td>0 – 0.5</td>
<td>34</td>
<td>Episodic &amp; WM</td>
<td>CVLT-SF, RCF, DS backward</td>
</tr>
<tr>
<td>Matuszewski et al. (2009)</td>
<td>14 sv-PPA</td>
<td>3.57</td>
<td>21</td>
<td>Episodic</td>
<td>AMIPPB</td>
</tr>
<tr>
<td>Mckay et al. (2007)</td>
<td>7 sv-PPA</td>
<td>N/A</td>
<td>19</td>
<td>Episodic &amp; WM</td>
<td>RCF, DS total</td>
</tr>
<tr>
<td>Montembeault et al. (2017)</td>
<td>9 sv-PPA</td>
<td>N/A</td>
<td>12</td>
<td>Episodic</td>
<td>RALVT, RCF</td>
</tr>
<tr>
<td>Nestor et al. (2003)</td>
<td>7 nfv-PPA</td>
<td>3.4 ± 1.4</td>
<td>10-31</td>
<td>Episodic</td>
<td>RCF, DMS faces &amp; words</td>
</tr>
<tr>
<td>Perges et al. (2010)</td>
<td>15 sv-PPA</td>
<td>4.8 ± 2.4</td>
<td>35</td>
<td>Episodic</td>
<td>RAVLT, RCF</td>
</tr>
<tr>
<td>Piolino et al. (2003)</td>
<td>10 sv-PPA</td>
<td>0 – 2.0</td>
<td>18</td>
<td>Episodic</td>
<td>AVLT</td>
</tr>
<tr>
<td>Ramanan et al. (2016)</td>
<td>25 lv-PPA</td>
<td>4.0 ± 2.7</td>
<td>90</td>
<td>Episodic</td>
<td>Doors A, RAVLT, RCF</td>
</tr>
<tr>
<td>Rohrer et al. (2010)</td>
<td>9 lv-PPA</td>
<td>5.3 ± 2.1</td>
<td>18</td>
<td>Episodic</td>
<td>CPRMT</td>
</tr>
<tr>
<td>Rosen et al. (2002)</td>
<td>14 nfv-PPA</td>
<td>4.2 ± 0.9</td>
<td>18</td>
<td>Episodic</td>
<td>WMS vis, Benson figure, SS</td>
</tr>
<tr>
<td>Savage et al. (2013)</td>
<td>20 sv-PPA</td>
<td>4.2</td>
<td>54</td>
<td>Episodic</td>
<td>RCF</td>
</tr>
<tr>
<td>Scallil et al. (2005)</td>
<td>16-18 sv-PPA</td>
<td>0 – 1.0</td>
<td>9</td>
<td>Episodic</td>
<td>RCF, DMS faces &amp; words, WMS-III LM</td>
</tr>
<tr>
<td>Watson et al. (2018)</td>
<td>74 sv-PPA</td>
<td>N/A</td>
<td>79</td>
<td>Episodic &amp; WM</td>
<td>WMS vis, Benson figure, SS</td>
</tr>
<tr>
<td>Wiman et al. (2015)</td>
<td>34 lv-PPA</td>
<td>N/A</td>
<td>48</td>
<td>Episodic</td>
<td>WMS vis, Benson figure, SS</td>
</tr>
<tr>
<td>Whitwell et al. (2015)</td>
<td>24 lv-PPA</td>
<td>3.5 ± 1.4</td>
<td>24</td>
<td>Episodic</td>
<td>AVLT</td>
</tr>
</tbody>
</table>

Notes: Years of symptoms presented as mean (± SD) or as range. sv-PPA = semantic variant; lv-PPA = logopenic variant; nfv-PPA = non-fluent variant; WM = working memory; WMS-R LM = Wechsler Memory Scale-Revised logical memory subtest; DS = Digit Span; CVLT-SF = California Verbal Learning Test – Short Form; RCF = Rey Complex Figure; PVLT = Philadelphia Verbal Learning Test; RMT = Recognition Memory Test; TdR = Test de la Ruche; Doors A = Doors Test A from the Doors and People memory battery; WMS vis = Wechsler Memory Scale – Visual Reproductions; WMS-III SS = Wechsler Memory Scale-III Spatial Span; CVLT-MS = California Verbal Learning Test – Mental Status; VOM = Visual Object Memory; AVLT = Auditory Verbal Learning Test; RCSR = Free and Cued Selective Reminding Test; DMS-48 = Delayed Matching to Sample – 48 items; AMIPPB = Adult Memory and Information Processing Battery; RAVLT = Rey Auditory Verbal Learning Test; CPRMT = Camden Pictorial Recognition Memory Test.
Interestingly, in both nfv-PPA and lv-PPA these frontoparietal networks are affected, namely posterior fronto-insular regions in nfv-PPA and posterior perisylvian or parietal regions in lv-PPA (Gorno-Tempini et al., 2011). Moreover, damage of the dorsal language pathway has been found in both nfv-PPA and lv-PPA, but not in sv-PPA (Galantucci et al., 2011). This overlap of implicated brain regions may explain why working memory deficits were mainly found in nfv-PPA and lv-PPA, and not in sv-PPA.

Nevertheless, it is exactly these differences in patterns of both grey and white matter damage with focal involvement of specific portions of the language areas and pathways that could explain the differences in working memory performance between lv-PPA and nfv-PPA (Galantucci et al., 2011). Moreover, damage of the dorsal language pathway has been found in both nfv-PPA and lv-PPA, but not in sv-PPA (Galantucci et al., 2011). This overlap of implicated brain regions may explain why working memory deficits were mainly found in nfv-PPA and lv-PPA, and not in sv-PPA.

4.3. Performance on verbal vs. nonverbal memory tests

The results of the current meta-analysis show PPA patients to have both verbal and non-verbal memory impairments. Only in sv-PPA, lower performance was found on verbal episodic memory tests compared to non-verbal episodic memory tests, whereas for nfv-PPA and lv-PPA such a difference was not found. Furthermore, nfv-PPA and lv-PPA patients performed worse on verbal and non-verbal working memory tests, while sv-PPA were only impaired on the verbal working memory tests. This suggests that the working memory problems in sv-PPA are dependent on the use of verbal or non-verbal tests and might therefore be secondary to their language deficits.

In current literature, a selective loss of verbal memory function has frequently been mentioned in PPA (e.g. Kielb et al., 2016; Zakzanis, 1999). This pattern of performance is consistent with the notion that memory deficits in PPA are a secondary manifestation of the aphasia. However, our results show that memory deficits are also pronounced

Table 2

<table>
<thead>
<tr>
<th></th>
<th>k</th>
<th>N</th>
<th>ES (g)</th>
<th>95% CI</th>
<th>Q</th>
<th>p (Q)</th>
<th>I²</th>
<th>Fail-safe N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>41</td>
<td>849</td>
<td>-1.23</td>
<td>-1.50 to -0.97</td>
<td>95.90</td>
<td>.01</td>
<td>70.80</td>
<td>2702</td>
</tr>
<tr>
<td>sv-PPA</td>
<td>29</td>
<td>450</td>
<td>-1.23</td>
<td>-1.50 to -0.97</td>
<td>95.90</td>
<td>.01</td>
<td>70.80</td>
<td>2702</td>
</tr>
<tr>
<td>lv-PPA</td>
<td>11</td>
<td>187</td>
<td>-2.04</td>
<td>-2.58 to -1.49</td>
<td>53.76</td>
<td>.01</td>
<td>81.40</td>
<td>899</td>
</tr>
<tr>
<td>nfv-PPA</td>
<td>12</td>
<td>212</td>
<td>-1.26</td>
<td>-1.60 to -0.92</td>
<td>30.81</td>
<td>.01</td>
<td>64.30</td>
<td>744</td>
</tr>
<tr>
<td>Episodic memory</td>
<td>sv-PPA</td>
<td>19</td>
<td>315</td>
<td>-1.79</td>
<td>-2.15 to -1.44</td>
<td>66.68</td>
<td>.01</td>
<td>73.00</td>
</tr>
<tr>
<td>lv-PPA</td>
<td>8</td>
<td>152</td>
<td>-1.52</td>
<td>-1.88 to -1.15</td>
<td>15.64</td>
<td>.03</td>
<td>55.23</td>
<td>398</td>
</tr>
<tr>
<td>nfv-PPA</td>
<td>8</td>
<td>171</td>
<td>-0.90</td>
<td>-1.26 to -0.55</td>
<td>15.27</td>
<td>.03</td>
<td>54.15</td>
<td>296</td>
</tr>
<tr>
<td>Working Memory</td>
<td>sv-PPA</td>
<td>15</td>
<td>256</td>
<td>-0.51</td>
<td>-0.75 to -0.26</td>
<td>27.39</td>
<td>.02</td>
<td>48.89</td>
</tr>
<tr>
<td>lv-PPA</td>
<td>7</td>
<td>95</td>
<td>-2.83</td>
<td>-3.73 to -1.93</td>
<td>38.09</td>
<td>.01</td>
<td>84.25</td>
<td>433</td>
</tr>
<tr>
<td>nfv-PPA</td>
<td>9</td>
<td>162</td>
<td>-1.71</td>
<td>-1.94 to -1.47</td>
<td>8.08</td>
<td>.43</td>
<td>0.98</td>
<td>493</td>
</tr>
</tbody>
</table>
| Notes: k = number of studies; N = number of patients; ES = effect size; 95% CI = 95% confidence interval; Q = heterogeneity statistic; p (Q) = p-value for heterogeneity; Fail-safe N = number of studies needed to be published to attain a non-significant effect; sv-PPA = semantic variant; lv-PPA = logopenic variant; nfv-PPA = non-fluent variant; HC = healthy controls. Statistical significance: p < .001.

Fig. 2. Performance of sv-PPA on episodic memory tests.
Note. Filled circles indicates verbal episodic memory tests and the open circles indicates nonverbal episodic memory tests.

well (Saur et al., 2008). Interestingly, in both nfv-PPA and lv-PPA these frontotoparietal networks are affected, namely posterior fronto-insular regions in nfv-PPA and posterior perisylvian or parietal regions in lv-PPA (Gorno-Tempini et al., 2011). Moreover, damage of the dorsal language pathway has been found in both nfv-PPA and lv-PPA, but not in sv-PPA (Galantucci et al., 2011). This overlap of implicated brain regions may explain why working memory deficits were mainly found in nfv-PPA and lv-PPA, and not in sv-PPA.

Nevertheless, it is exactly these differences in patterns of both grey and white matter damage with focal involvement of specific portions of the language areas and pathways that could explain the differences in working memory performance between lv-PPA and nfv-PPA (Galantucci et al., 2011). On the one hand, the breakdown of the storage and rehearsal processes of the phonological system in lv-PPA has been related to temporal-parietal atrophy and may explain the more pronounced working memory deficits in lv-PPA (Foxe et al., 2013). On the other hand, damage to the posterior fronto-insular areas and more frontal components of the dorsal language tracts in nfv-PPA could induce different underlying mechanisms of working memory deficits (Galantucci et al., 2011), which may explain the less pronounced working memory deficits in nfv-PPA compared to lv-PPA. The precise relationship between the different neuro-anatomical and pathological substrates seen in both PPA variants and its effects on working memory performance is, however, something that future studies should address.
when using non-verbal memory measures suggesting memory impairments that cannot be explained by language deficits alone.

However, it is important to note that the performance on non-verbal memory tests is rarely fully independent of language function. Even on memory tests that are typically considered to be non-verbal in nature, such as the Rey Complex Figure Test (RCFT), patients may use verbal strategies (e.g., to remember the locations or forms of parts of the figure) and have to understand verbal instruction in order to complete the test. Study designs for investigating memory function in PPA should therefore make sure that memory tests are used that only minimally rely on language function, for instance by using memory tests that consist of difficult-to-verbalize stimuli, such as the Continuous Visual Memory Test (Trahan and Larrabee, 1988), or by statistically adjusting for the extent of the language impairment. Until now, only very few studies on memory in PPA have controlled for language deficits in such a way. For example, Ramanan et al. (2016) showed that even after statistical adjustment for the performance on language tests, PPA patients do show significant memory deficits and that these measures are still able to discriminate between PPA variants.

Furthermore, other cognitive deficits that arise as the disease progresses may also underlie the deficits in non-verbal episodic memory. Because PPA is caused by progressive neurodegeneration, patients eventually exhibit deficits in other cognitive domains. Previous research has shown PPA-related atrophy to spread beyond the initial distinctive locations into the medial temporal lobe as well as the frontal lobe (Rogalski et al., 2011; Mesulam et al., 2014). The resulting executive dysfunction that can occur in PPA can affect both encoding and retrieval in non-verbal episodic memory. In addition, PPA patients may fail to implement sophisticated organizational strategies during learning as a result of executive impairments.
4.4. Additional factors of consideration and limitations

The present meta-analysis is the first quantitative summary of the literature on memory performance and its manifestations in different PPA subtypes. As such, it offers insight into memory dysfunction in PPA and its extension beyond the verbal memory domain. In light of the current diagnostic criteria for PPA (Gorno-Tempini et al., 2011), the outcomes of our meta-analysis offer evidence suggesting extension of these criteria might be necessary, since these include memory deficits as an exclusion criteria in the initial phase of the disorder, while we show that memory dysfunction is frequently observed in PPA patients. Unfortunately, we were not able to investigate the prevalence of memory deficits in especially the initial stage since only a part of the studies that were used reported illness duration or years from first symptom as an outcome measure. Of these, only some reported sufficient information to allow for statistical analyses, making the use of symptom duration as a confounding variable in our meta-analyses impossible. The variety in illness duration within the included studies may therefore have contributed to the size of the ESs that we found. However, the ESs we found were large and the majority of the utilized studies reported to have included patients in the beginning stages of their disease (< 5-year symptom duration; see Table 1), suggesting that this influence cannot explain all of the found effects. Future studies should, however, adequately report measures of illness duration in order to study the prevalence of memory impairments across disease stages and to provide evidence to retain memory deficits as an exclusion criterion for a PPA diagnosis.

The current meta-analysis has some more limitations and caveats that should be kept in mind when considering our findings. Although the risk of a publication bias was found to be low, the included studies showed a large heterogeneity in ESs. This might also have resulted in the asymmetrical funnel plots (Sterne et al., 2011). The heterogeneity...
Fig. 7. Performance of lv-PPA on working memory tests.

Note. Filled circles indicates verbal working memory tests and the open circles indicates nonverbal working memory tests.

possibly arises because of the substantial differences in the studies’ patient samples, such as variation in symptom duration or in diagnostic criteria. In addition, heterogeneity might be caused by differences in task demands across the memory tests that are being used. Although heterogeneity was substantially reduced when we investigated the effects for the different PPA types, the different memory systems, and for verbal and non-verbal tests separately, heterogeneity was still present. However, it should be noted that we aimed to summarize the literature...
on the underlying memory constructs in PPA rather than examine the performances on individual tests.

Furthermore, even though the ESs found in this meta-analysis are robust, it was not possible to establish how many patients performed in the clinically impaired range (i.e., < 2 SD below the clinical norm). Therefore, based on our results, it cannot be established whether the found significant ESs are also clinically relevant in all PPA patients.

4.5. Implications for the diagnosis of PPA variants

This meta-analysis was conducted in order to provide a possible clinical marker to differentiate between nfv-PPA and lv-PPA, something that remains very challenging in clinical practice (Crook et al., 2012).

Our results show that patients with lv-PPA tend to perform worse on both episodic memory as well as working memory tasks compared to nfv-PPA patients. This might not be completely explained by language deficits since this was observed in both verbal and non-verbal tests.

5. Conclusion

Taken together, this meta-analysis showed that impairments in both episodic and working memory are observed in all PPA variants. However, different patterns of memory performance were observed, with more pronounced episodic and working memory deficits in lv-PPA when compared to nfv-PPA. These findings highlight the potential benefit of using memory tests in addition to language assessment to better differentiate nfv-PPA and lv-PPA.

Acknowledgments

The authors wish to thank Laura Klabbers for assistance with the literature search. 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) supporting RPCK and AR. The authors declare that they had no conflicts of interest with respect to their authorship or the publication of this article.

References


syndrome. Lancet Neurol. 6 (11), 1004–1014.


