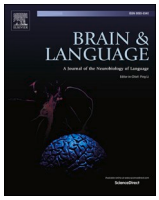
Brain & Language 238 (2023) 105243

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l



Accounting for word production, comprehension, and repetition in semantic dementia, Alzheimer's dementia, and mild cognitive impairment

Ardi Roelofs

Radboud University, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognition, Thomas van Aquinostraat 4, 6525 GD Nijmegen, The Netherlands

ARTICLE INFO

Keywords:
Comprehension
Memory
Modeling
Naming
Repetition

ABSTRACT

It has been known since Pick (1892, 1904) that word retrieval is commonly impaired in left temporal lobe degeneration. Individuals with semantic dementia (SD), Alzheimer's dementia (AD), and mild cognitive impairment (MCI) present with word retrieval difficulty, while comprehension is less affected and repetition is preserved. Whereas computational models have elucidated performance in poststroke and progressive aphasias, including SD, simulations are lacking for AD and MCI. Here, the WEAVER++/ARC model, which has provided neurocognitive computational accounts of poststroke and progressive aphasias, is extended to AD and MCI. Assuming a loss of activation capacity in semantic memory in SD, AD, and MCI, the simulations showed that severity variation accounts for 99% of the variance in naming, comprehension, and repetition at the group level and 95% at the individual patient level ($N = 49$). Other plausible assumptions do less well. This supports a unified account of performance in SD, AD, and MCI.

1. Introduction

Difficulty in word retrieval is a ubiquitous symptom of progressive neurodegenerative disease affecting the temporal lobes, which has been known since the seminal reports of Pick (1892/1977, 1904/1997). Retrieval problems are observed in semantic dementia (SD), Alzheimer's dementia (AD), and mild cognitive impairment (MCI). Picture naming, assessing word retrieval, is among the tasks showing the greatest deficits in SD (e.g., Rogers et al., 2006), AD (e.g., Laws et al., 2007), and MCI (e.g., Joubert et al., 2021; Taler et al., 2020). Naming is also a good predictor of progression from MCI to AD (Belleville et al., 2017). Individuals are more impaired on picture naming in SD than in AD, and more in AD than in MCI, and all three groups show worse performance than healthy age-matched controls (e.g., Janssen et al., 2022; Rogers et al., 2006). Semantic memory is disrupted in all three patient groups, whereas syntax and phonology tend to be preserved (e.g., Hodges, 2006; Landin-Romero et al., 2016). Therefore, a plausible hypothesis is that differences in performance between groups reflect, to an important extent, different severities of semantic disruption (but see Gallant et al., 2019; Isella et al., 2020, 2022). To examine this hypothesis, computer simulations with the WEAVER++/ARC model (Roelofs, 2014, 2022) were run. Simulation is an important tool in testing whether theoretical assumptions can account for the data, with the requirement to precisely define the nature of representation and processing, and here, also the nature of the deficit.

Computational models (e.g., Dell et al., 1997, 2013; Ueno et al., 2011; Walker & Hickok, 2016) have elucidated single-word impairments in poststroke aphasia and in one variant of primary progressive aphasia (PPA), namely SD, or have been designed to account for the deterioration of semantic memory in SD (Rogers et al., 2004). The WEAVER++/ARC model has been applied to both poststroke aphasia (Roelofs, 2014, 2021) and the three variants of PPA (Janssen et al., 2020; Roelofs, 2022). The model integrates behavioral psycholinguistic, functional neuroimaging, tractographic, and aphasiological evidence (WEAVER++/ARC is an acronym standing for Word Encoding by Activation and VERification / Arcuate Repetition and Conversation). Following a proposal by Pick (1892/1977, 1908) and modern insights (e.g., Mandelli et al., 2016; Seeley et al., 2009; Zhou et al., 2012), the impairments in PPA are assumed to arise from a progressive loss of activation capacity in portions of the language network with neurocognitive epicenters that are specific to each PPA variant. In SD, the neurocognitive epicenter involves the anterior temporal lobes (ATL) underpinning semantic memory, particularly in the left hemisphere, which is also atrophied in AD and MCI.

In this article, I first present background information on SD, AD, and MCI, including insights from work published in German between 1892 and 1926. Although basic facts about clinical presentation and relation to circumscribed brain atrophy were documented, they laid dormant for

E-mail address: Ardi.Roelofs@donders.ru.nl.

https://doi.org/10.1016/j.bandl.2023.105243
Received 7 June 2022; Received in revised form 27 January 2023; Accepted 20 February 2023
Available online 2 March 2023

half a century, only to be discovered again in the 1970s (Warrington, 1975) and later (e.g., Hodges et al., 1992; Ohm et al., 2022; Snowden et al., 1989; Zhou et al., 2012). Also, Pick (1908) proposed an account, further supported by modern research. Next, the WEAVER++/ARC model and its extension along Pickian lines to AD and MCI is outlined. In subsequent sections, I use the model to fit the naming, comprehension, and repetition performance of patients with SD, AD, and MCI (i.e., data from Janssen et al., 2022), at both group and individual patient levels. To enable evaluation of the model fits, information about the flexibility of the model, the variability of the data, and the specificity of the predictions is provided, argued to be necessary by Roberts and Pashler (2000). The specificity was examined in simulations with an additional locus of damage for AD (Isella et al., 2020, 2022) and a mixture of loci for MCI (Gallant et al., 2019).

1.1. Historical background and characteristics of SD, AD, and MCI

Pick (1892/1977, 1904/1997) described four cases with pronounced atrophy of the left temporal lobe (his life and work in Prague are portrayed by Kertesz & Kalvach, 1996). One of the patients, Anna Jirinec, 75 years old, presented with fluent speech but severe word retrieval difficulty, impaired word comprehension, and impaired conceptual knowledge, alongside possibly spared episodic memory (Spatt, 2003). Fischer (1910) reported that her brain at autopsy did not contain plaques and tangles (Alzheimer, 1907; Fischer, 1907), thus she was not afflicted with the disease named after Alzheimer. Side and coronal views of Jirinec's brain (Fig. 1) suggest that especially the anterior part of the left temporal lobe was atrophied, while Wernicke's area and the hippocampi were less affected. These observations on the distribution of atrophy have been made in other cases by Altman (1923) and Onari and Spatz (1926) using macroscopic and histopathological methods, and also in a modern case series study (Mesulam et al., 2015) and in a quantitative meta-analysis of gray matter volume reductions in SD (Yang et al., 2012). The clinical and neuroanatomical features of Jirinec resemble those of SD (e.g., Gorno-Tempini et al., 2011; Patterson et al., 2007).

Following up on Pick (1904) and Fischer (1910), Alzheimer (1911/1991) reported on a histopathological study of two cases of circumscribed temporal cortex atrophy. The pathology of one patient, 65-year-old Therese Mühlich, was described in detail, and Stertz (1926) provided a clinical description. Using silver staining, Alzheimer did not observe the plaques and tangles that he had found a few years earlier in the brain of his patient Auguste Deter (Alzheimer, 1907). Instead, he observed round inclusions in the cytoplasm of nerve cells and ballooned cells, which later came to be called Pick bodies and cells, suggesting a different disease. Later research has shown that these inclusions and cells are present in only a minority of cases and that instead TDP-43 positive pathology is characteristic of SD (e.g., Landin-Romero et al., 2016). Onari and Spatz (1926) examined the brain of Anna Bradt, 65 years old, clinically described by Stertz as presenting with fluent speech but severe word retrieval difficulty, impaired word comprehension, and impaired conceptual knowledge, alongside seemingly spared episodic memory. Degeneration was most prominent in the left temporal pole and fusiform gyrus, while sparing Wernicke's area and the hippocampi, and with atrophy predominantly present in the upper layers of the cortex. Modern large-scale research in patients ($N = 97$) has shown that the upper-layer predominance is the signature laminar distribution of TDP-43 pathology (Ohm et al., 2022).

To explain why the atrophy is circumscribed and leads to focal symptoms, Pick (1908) proposed an account in terms of a functional network, assuming that "occasionally such a systematically similar group of neurons, i.e., a system in the older sense, succumbs to atrophy earlier than the others, and as a result the function of this system fails in a completely isolated manner"[1] (p. 24). A drawing of Ramón y Cajal was used to illustrate how laminar-specific atrophy may occur. Although Alzheimer (1911/1991) rejected Pick's functional account, Gans (1923) and Onari and Spatz (1926) provided evidence that the locus and anatomical spread of the disease have a functional basis. Modern research has shown that neurodegenerative diseases, including those underlying SD and AD, target specific functional networks, starting in regions with heavy network traffic and propagating along strong functional and anatomical connections (e.g., Mandelli et al., 2016; Seeley et al., 2009; Zhou et al., 2012).

An important functional distinction regarding memory is between the learning/consolidation and storage of knowledge (see Eichenbaum, 2012, for a review). Whereas lateral temporal cortex stores factual knowledge about concepts and words as well as knowledge about personal events, together called declarative memory, the hippocampus plays a critical role in learning new declarative knowledge and consolidating it in neocortical areas. While patients with SD present with a loss of conceptual knowledge (Warrington, 1975) and atrophy of the ATLs (Hodges et al., 1992; Snowden et al., 1989; see Lombardi et al., 2021, for a large-scale study), patients with AD have a prominent disturbance of the consolidation ability and atrophy of medial temporal regions, predominantly in the hippocampus. The disease in AD begins in the entorhinal cortex, then spreads into the hippocampal formation, and later progresses to the lateral temporal cortex and other brain regions (e.g., Braak & Braak, 1991; Josephs et al., 2020). The spread to temporal cortex includes the temporal pole (e.g., Arnold et al., 1994), disrupting the stored conceptual knowledge. People with MCI present with subjective memory complaint or objective memory impairment, normal general cognitive functioning, and intact activities of daily living (e.g., Petersen et al., 1994; Petersen & Morris, 2005). In a large-scale study examining the patterns of longitudinal cortical atrophy in MCI ($N = 295$) relative to healthy controls ($N = 134$), Edmonds et al. (2020) observed that initially atrophy may be most pronounced in medial temporal regions, in both bilateral medial and lateral temporal regions, or in these regions along with atrophy in frontal, parietal, and occipital cortex, but that over time, lateral temporal cortex is affected in all cases. For a review of a century of research into AD, and MCI as a precursor of AD, I refer to Hodges (2006).

1.2. Single-word task performance in SD, AD, and MCI

Janssen et al. (2022) assessed single-word production, comprehension, and repetition in PPA, including SD, as well as in AD and MCI using a Dutch version of the Sydney Language Battery (SYDBAT) originally developed for English (Savage et al., 2013). Patients with PPA, AD, MCI, and education- and age-matched healthy controls were compared on picture naming, auditory word comprehension, and word repetition tasks. The patients did not have visual issues. The target stimuli were 30 imageable nouns of three or more syllables (e.g., *elephant*), which were used in each of the tasks. In the naming task, participants saw objects as colored photographs and spoke the name of each object (e.g., an elephant, say "elephant"). In the word comprehension task, the examiner spoke words (e.g., "elephant") and the participants selected for each word the matching picture from a display of seven photographs. In the repetition task, the examiner spoke words (e.g., "elephant") and the participants repeated each word (i.e., say "elephant"). For each task, the percentage of correct responses was recorded for each participant.

[1] "dass gelegentlich eine solche systematisch gleichgeartete Neurongruppe, also ein System im älteren Sinne, früher als die übrigen der Atrophie verfällt, und dadurch ganz isoliert die Funktion dieses Systems ausfällt" (Pick, 1908, p. 24).

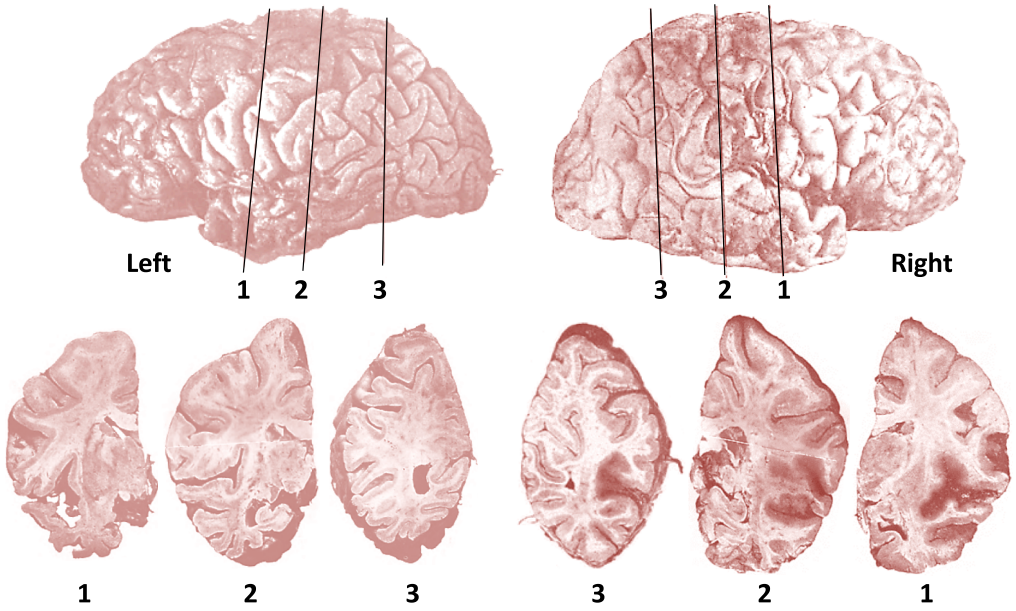


Fig. 1. Side view (top) and coronal views (bottom) of the left and right hemispheres of the brain of patient Anna Jirinec reported by Pick (1904/1997). Adapted from his Tables VII-IV (not reproduced with the English translation in 1997 or scanned as part of the digitalized journal archive).

Although only 30 items were used per task, Janssen et al. showed that the SYDBAT has good construct validity and reliability.

In line with prior studies, Janssen et al. (2022) observed that in SD ($N = 13$), AD ($N = 13$), and MCI ($N = 23$)[2], naming was more impaired than comprehension, while repetition was preserved. Naming and comprehension were more severely disrupted in SD than in AD (29% vs. 62% correct for naming, and 78% vs. 85% correct for comprehension) and more disordered in AD than in MCI (62% vs. 78% correct for naming, and 85% vs. 93% correct for comprehension), while performance for repetition was comparable between groups (i.e., 96%, 97%, and 99% correct). Using the SYDBAT, Leyton et al. (2014) compared naming, comprehension, and repetition performance between SD ($N = 13$) and AD ($N = 23$) in English. As Janssen et al., they observed that naming and comprehension were more disrupted in SD than in AD (18% vs. 67% correct for naming, and 52% vs. 83% correct for comprehension), while performance for repetition was comparably preserved (i.e., 95% vs. 97% correct). These results agree with the idea that semantic memory is more severely disordered in SD than in AD, and least in MCI (see also Rogers et al., 2006). This was confirmed by an assessment of conceptual knowledge by Janssen et al. using a fourth task of the SYDBAT that involves picture-picture matching. Participants saw a target picture and had to select a closely related picture from a set of four options. As with word comprehension, performance was worse in SD than AD (66% vs. 74% correct), and worse in AD than in MCI (74% vs. 86% correct). The similar patterns of performance for word-to-picture and picture-to-picture matching indicate that the loss of conceptual knowledge is the same across input modalities (i.e., word and picture), involving supramodal conceptual representations (cf. Patterson et al., 2007).

Although it is evident from the average test scores that overall performance across tasks is better in MCI than in AD, and worst in SD, it is not evident that the *patterns* of performance across tasks may result from different average severities. Performances do not simply proportionally differ between groups. For example, the ratio of naming and comprehension scores is 0.47 in SD but 0.92 in AD. Thus, relative to comprehension, naming has disproportionally deteriorated in SD compared to AD. Assuming a linear relationship between naming and comprehension explains only 44% of the variance in SD, 41% in AD, and 62% in MCI. Whereas naming is worse than comprehension in SD, AD, and MCI, Ueno et al. (2011) reported that naming and comprehension decline to the same extent in simulations with their Lichtheim 2 model. Simulations may clarify whether the WEAVER++/ARC model exhibits the disproportionate deterioration effects. Moreover, although variation in semantic memory damage shared between SD, AD, and MCI would be the simplest account, other investigators have argued for an additional locus of damage in AD (Isella et al., 2020, 2022) or a mixture of loci in MCI (Gallant et al., 2019). Simulations may clarify how well these other assumptions do in the model.

1.3. The WEAVER++/ARC model

Pick (1913) outlined a stage theory of language production, which foreshadowed the modern theories of Garrett and Levelt (see Levelt, 2013, for a historical account). The theory assumes that thoughts are first made ready for verbal expression ("ausdrucksfähig", p. 229) by mapping them onto propositional structures (in modern terms, a preverbal message making explicit the lexical concepts), which activate syntactic frames, followed by the insertion of words. Word production involves word selection ("Wortwahl", p. 245) followed by word formulation ("Wortformulierung"), which seems similar to lexical selection and word-form encoding in modern theories. Pick assumed that in spoken word production, concepts are not directly mapped onto motor programs (as Wernicke, 1874, assumed), but that more intermediate stages are involved. He argued that the nature of these stages should be illuminated by research into the normal process (p. 259). The multistage WEAVER++/ARC model was originally built on evidence about the normal process (e.g., Levelt et al., 1999; Roelofs, 1992, 2018), but has been extended to account for impairments during the past decade.

Semantic memory is part of the brain system for declarative knowledge, which is distinct from a brain system for procedural knowledge (e.g., Eichenbaum, 2012). The WEAVER++/ARC model assumes that an associative network realizes declarative knowledge about concepts and words, thought to be represented in temporal and inferior frontal regions, and that condition-action rules realize procedural

[2] Janssen et al. (2022) tested 25 participants with MCI, but two were excluded here and from the simulations because they had missing scores for repetition or for comprehension and repetition.

knowledge, thought to be represented in frontal regions, basal ganglia, and thalamus (Roelofs, 2014, 2021, 2022; for a review, see Roelofs & Ferreira, 2019). The structure of the model is illustrated in Fig. 2.

In naming, comprehension, and repetition, the associative network is accessed by spreading activation while condition-action rules select activated nodes that satisfy the task demands specified in working memory (i.e., name a picture, comprehend a word, repeat a word). The condition-action rules also exert top-down control in conceptually driven word production by selectively enhancing the activation of target lexical concept nodes in the network in order to achieve quick and accurate retrieval and encoding operations. Lexical concepts are assumed to be part of a hub of supramodal conceptual representations, thought to be represented in the ATL bilaterally, which integrate modality-specific features that are represented in widespread brain areas for perception and action (e.g., Lambon Ralph, 2014; Patterson et al., 2007; Pick, 1931). The lexical concepts are linked to lemmas (in the middle section of left middle temporal gyrus, MTG) that specify the syntactic properties of words (such as that the word *cat* is a noun, N), thought to be represented in left posterior superior temporal gyrus (STG) and MTG. Lemmas are linked to lexical output forms or morphemes (e.g., singular <cat>) in left posterior STG and MTG (Wernicke's area), and the lexical output forms are linked to output phonemes (e.g., /k/, /æ/, and /t/) in left posterior inferior frontal gyrus (IFG; i.e., Broca's area), which are linked to syllable motor programs (e.g., [kæt]) in ventral precentral gyrus. Input phonemes (e.g., /k/, /æ/, and /t/) and lexical input forms (e.g., <cat>) underpinning word comprehension and repetition are thought to be represented in middle to posterior STG and superior temporal sulcus (STS) bilaterally (see Kemmerer, 2022, for a review).

In picture naming, activation traverses from lexical concepts via lemmas, lexical output forms, and output phonemes to motor programs. In word comprehension, activation spreads from input phoneme nodes via lexical input forms and lemmas to lexical concepts. In repetition, activation spreads from input phonemes to output phonemes, and from input phonemes via lexical form and lemma nodes to output phonemes, and from output phonemes to motor programs. Hanley et al. (2004) and Nozari and Dell (2013) also assumed lexical and sublexical routes in simulations of spoken word repetition in poststroke aphasia. In naming and repetition in WEAVER++/ARC, condition-action rules for phonological encoding are engaged to syllabify the output phonemes and assign a stress pattern across syllables. The resulting phonological word representation is then used to select matching motor programs for the phonological syllables.

1.4. Assumptions about SD, AD, and MCI

The model assumes that a progressive degeneration of the conceptual network is common to SD, AD, and MCI. Dissolution of the network determines the pattern of aphasic symptoms and the spatial distribution of atrophy, as proposed by Pick (1908). Atrophy may reduce the capacity of the conceptual network to transmit activation or diminish its capacity to maintain activation over time, which may be functionally implemented as a reduction of connection weights and an increased decay rate, respectively. The weight decrease concerns all connections to, within, and from the conceptual network, and the decay increase concerns all lexical concept nodes. In accounting for the effect of normal aging and stroke on language performance, researchers have assumed a loss in activation transmission (Burke et al., 1991; Dell et al., 2013). Poststroke patterns of naming errors have been argued to be better explained by a transmission deficit in some patients and by a maintenance deficit in others (Martin & Dell, 2019).

2. Methods

The simulation protocol, including the network structure and parameter values, was the same as in earlier studies (e.g., Roelofs, 2014, 2022). The target was *cat* and the other words were *dog* and *fish* (both semantically related), *fog* (phonologically related to a semantic alternative, namely *dog*), and *mat* (phonologically related to *cat*). To examine the effect of varying the size of the lexicon, the simulations were also run with a larger network that additionally contained all the animal names of the SYDBAT (i.e., *butterfly*, *elephant*, *caterpillar*, *dinosaur*, *rhinoceros*, *hippopotamus*, and *orangutan*). The simulations with the larger network yielded outcomes similar to those with the small network. The Pearson correlation between the simulated patterns of naming, comprehension, and repetition of the individual patients (i.e., 147 data points) of Janssen et al. (2022) for the small and larger networks was $r = 0.99$ for a weight lesion and $r = 0.98$ for a decay lesion. Thus, varying the size of the lexicon does not change the simulation outcomes.

Information is retrieved from the network by spreading activation according to the following activation function:

$$a(m, t + \Delta t) = a(m, t)(1 - d) + \Sigma_n\, r\, a(n, t).$$

In this equation, $a(m, t)$ denotes the activation level of node m at point in time t, d represents the decay rate, and Δt indicates the duration of a time step in ms. The sum denotes the amount of activation that m receives between t and $t + \Delta t$, where $a(n, t)$ is the output of neighbor n and r is a weight indicating the strength of the connection between nodes m and n. Atrophy severity was simulated by manipulating connection weights (r) or the decay rate (d) of the conceptual network.

A simulation began by providing external activation to lexical concepts for naming and to input phonemes for repetition and comprehension. Activation then spread in $\Delta t = 25$ ms steps for 2 sec, and the mean activation of nodes was computed. Condition-action rules were

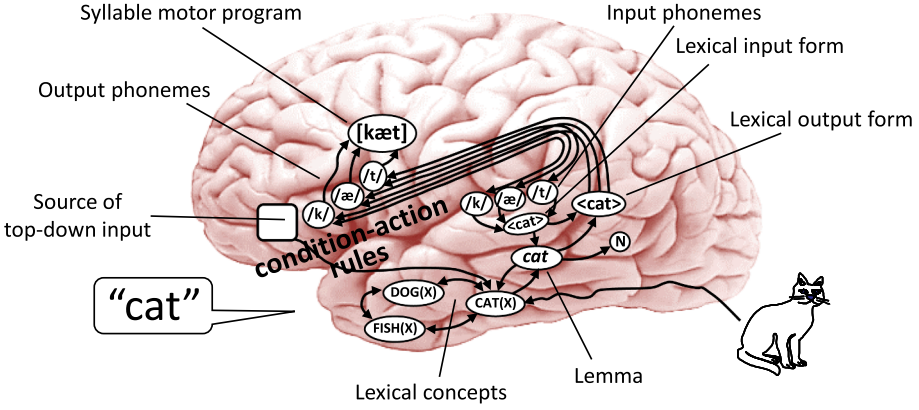


Fig. 2. Illustration of the functional architecture of the WEAVER++/ARC model and its mapping onto brain regions. N = noun. Adapted from Roelofs (2022). Copyright 2022 by the author.

assumed to select nodes depending on the task. For each of several severities of atrophy, the difference in mean activation between target and closest alternative was computed and expressed as a percentage of the normal activation difference. With smaller activation differences, selection of nodes takes longer and errors are more likely to occur. Thus, lower percentages will lead to poorer performance. The activation difference concerned syllable program nodes in naming and repetition, and lexical concept nodes in comprehension.

The values of the weight decrease and decay increase parameters that provided the best fit between model and data were obtained by an exhaustive search through the parameter space, varying between minimal and maximal damage. Weight decrease was varied between 1.0 and 0.0 ($\times r$) and decay increase between 1.0 and 1.66 ($\times d$), both in steps of 0.01. The search aimed to obtain the parameter value for each individual patient that minimizes the mean absolute difference (i.e., mean absolute error, MAE) between simulated and empirical performance for naming, comprehension, and repetition. To indicate goodness of fit between model and data, MAEs and Pearson correlation coefficients are reported. Correlations were statistically compared using the *cocor* package in R (Diedenhofen & Musch, 2015).

The C programming language and the programming environment of Microsoft Visual C++ 2022 were used to computationally implement the simulations. The simulation source code may be obtained from the Open Science Framework at https://osf.io/ue4bn/ or from the author.

3. Results and discussion

3.1. Performance accuracy as a function of atrophy severity

Fig. 3 displays how performance accuracy varies as a function of weight decrease (left panel) and decay increase (right panel) in simulations of naming, comprehension, and repetition, constraining what the model can and cannot fit (Roberts & Pashler, 2000). The predictions were derived deterministically rather than by repeated random sampling, so the error bars are zero and not displayed. The figure shows that weight decrease and decay increase tend to have similar effects, although there are also differences. For example, word production and comprehension may fully deteriorate with a weight lesion but not with a decay lesion.

Some aspects of the patterns of performance that are observed empirically seem more similar to the decreased performances that result from weight decreases (left panel of Fig. 3) than from decay increases (right panel). In discussing SD, Landin-Romero et al. (2016) stated: "Over time, many patients become essentially mute with only a limited repertoire of stereotypic phrases and a complete loss of word comprehension" (p. 2). This is consistent with an ultimate accuracy of 0% on naming and comprehension tasks, which may occur under a weight lesion but not a decay lesion in the model. In simulations of PPA reported in Roelofs (2022), performance was also generally better captured by a weight decrease than a decay increase. In what follows, the results of the weight lesion simulations are depicted, whereas those of a decay lesion are only verbally described. Details of the findings for both type of lesion are reported in the Supplementary material. It was not an aim to arbitrate between weight decrease and decay increase, which both lead to a loss of activation capacity. Still, their goodness of fit was statistically compared.

A close fit between model and data only provides support for the theoretical assumptions when the model provides substantial constraints, as just discussed (Fig. 3), and the data are constraining (Roberts & Pashler, 2000). The variability of the data is discussed in the next two sections, together with the model fits at group and individual patient levels.

3.2. Behavioral profiles at the group level

Fig. 4 shows the WEAVER++/ARC simulation averages for naming, comprehension, and repetition in SD, AD, and MCI, along with the empirical group averages observed by Janssen et al. (2022). The variability of the data is denoted by 95% confidence intervals. Naming and comprehension are more severely disrupted in SD than in AD, and more in AD than in MCI, while repetition is largely preserved in all groups. The best fit between model and data is obtained when the weight decrease is larger for SD than AD (i.e., 0.74*r* vs. 0.90*r*), and larger for AD than MCI (i.e., 0.90*r* vs. 0.98*r*). This corresponds to the assumption that semantic memory is more disrupted in SD than in AD, and more in AD than in MCI.

The MAE measures of fit between simulated and empirical group data are rather small, averaging 2.7% (see Supplementary Table 1 for details). The correlation between model and data is $r = 0.99$, $p < .001$. Similar outcomes were obtained when assuming a decay lesion (Supplementary Table 1). The best fit between model and data is obtained when the decay increase is larger for SD than AD (i.e., 1.31*d* vs. 1.14*d*), and larger for AD than MCI (i.e., 1.14*d* vs. 1.03*d*), with an average MAE of 3.3% and correlation of $r = 0.99$, $p < .001$. The correlations did not differ between weight and decay lesions ($p > .99$).

Fig. 4 reveals that the average comprehension score is somewhat overpredicted for MCI. However, numerically the deviation is rather small (i.e., 5.1%). Also, the overlapping error bars in Fig. 4 for average AD comprehension hide the fact that the overprediction is consistent at the individual patient level. Importantly, however, the deviations do not concern the pattern of performance across tasks and patient groups that is predicted by the model. For both tasks, scores are higher in MCI than in AD, and lowest in SD, both empirically and in the model. And for all three patient groups, scores are lower for naming than for comprehension, both empirically and in the model. Deviations are further discussion at the level of individual patients in the next section.

According to the model, the presence of a conceptual deficit leads to impaired naming and comprehension, while repetition remains spared. This is because both naming and comprehension require processing of conceptual representations, whereas repetition can be done without conceptual involvement. As a consequence, severity of damage affects naming and comprehension, but not repetition.

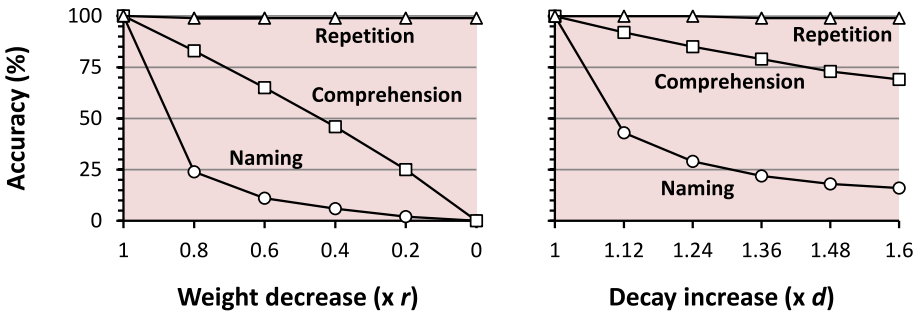


Fig. 3. Performance accuracy as a function of weight decrease (left panel) and decay increase (right panel) in the conceptual network in WEAVER++/ARC simulations of single word naming, comprehension, and repetition. Adapted from Roelofs (2022). Copyright 2022 by the author.

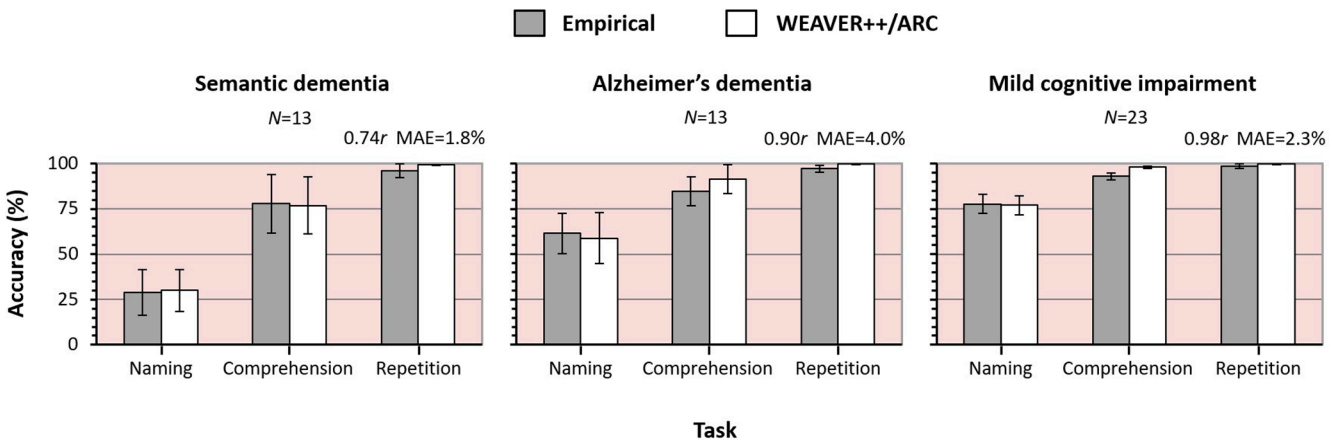


Fig. 4. Performance accuracy in semantic dementia, Alzheimer's dementia, and mild cognitive impairment for naming, comprehension, and repetition: Empirical group averages from Janssen et al. (2022) and WEAVER++/ARC simulation averages. The error bars represent 95% confidence intervals. For each panel, the estimated weight decrease (e.g., 0.74*r*) and mean absolute error (MAE) are displayed. *N* = number of patients. Partly adapted (for semantic dementia) from Roelofs (2022). Copyright 2022 by the author.

3.3. Behavioral profiles at the individual patient level

The top panel of Fig. 5 shows the variability of the data at the individual patient level. The figure shows the patterns of performance on the naming, comprehension, and repetition tasks of the 49 individual patients in the study of Janssen et al. (2022) together with the group averages for SD, AD, and MCI. Using box plots to determine outlying observations in the patient scores for each task within a group, only 6 out of the 147 data points were deemed to be outliers, denoted by numbers (e.g., #10). The figure reveals that for SD, the pattern for the group corresponds to the individual patterns, except for one patient (case 10) with extremely disrupted comprehension and another patient (case 2) with disrupted repetition. For AD, the pattern for the group also corresponds to the individual patterns, except for two patient (cases 11 and 13) presenting with exceptionally disrupted naming, and one patient (case 6) with disrupted repetition. Finally, for MCI, the pattern for the group also corresponds to the individual patterns, except for one patient (case 22) with disrupted repetition.

The bottom panel of Fig. 5 shows the WEAVER++/ARC simulation results for all 49 individual patients in the study of Janssen et al. (2022), assuming a weight lesion. For each patient, denoted by dot and number with the tasks color coded, the predicted performance scores with the lowest MAE are plotted against the observed scores (for details, see Supplementary Table 2). Overall, the model succeeds reasonably well at simulating the performance patterns of the individual cases. The average MAEs across patients for SD, AD, and MCI are 3.7%, 4.8%, and 2.9%, respectively. The overall correlation between model and individual patient data is $r = 0.97$, $p < .001$. Similar outcomes were obtained when assuming a decay lesion (Supplementary Table 3), with average MAEs across patients for SD, AD, and MCI of, respectively, 6.3%, 4.9%, and 2.7%, and an overall correlation of $r = 0.95$, $p < .001$. The correlation was stronger for a weight lesion than for a decay lesion ($p < .001$).

The model yields the outlying pattern of case 10 of SD (i.e., very poor naming and comprehension with preserved repetition) when the weight lesion is severe (MAE = 3.4%). Fig. 2 shows that this pattern cannot occur with a decay lesion (MAE = 29.3%). The model underestimates the severely disrupted naming of cases 11 and 13 of AD. In the model, naming is worse than is to be expected on the basis of the comprehension and repetition performance of these patients (their MAE is 6.8% and 5.6%, respectively).

Fig. 5 indicates that the fits are best for naming. However, as the Supplementary Table 2 shows, numerically the deviations for comprehension and repetition tend to be small. The predictions for comprehension differ from the empirical individual patient scores on average by 4.3% (range 0.0%–14.2%), 8.3% (range 0.1%–27.5%), and 5.1% (range 0.0%–10.8%), for SD, AD, and MCI, respectively. Still, the comprehension scores of patients #3 and #6 with AD are captured poorly by the model (with overestimations of 27.5% and 20.0%, respectively). The predictions for repetition differ from the empirical individual patient scores on average by 3.9% (range 0.2%–19.6%), 2.8% (range 0.1%–9.8%), and 1.5% (range 0.0%–13.1%), for SD, AD, and MCI, respectively. Here, the repetition scores of patient #2 with SD and patient #22 with MCI are captured poorly by the model (with overestimations of 19.6% and 13.1%, respectively). Thus, the model cannot fit well the performance of a number of patients. Nevertheless, although deviations are clearly present, most model predictions are rather close to the empirical observations.

Importantly, as Supplementary Table 2 shows, the deviations for the individual patients are numerical differences and do not concern the pattern of performance across tasks in any of the patients simulated by the model. In all patients, scores are lower for naming than for comprehension (47 patients) or they are the same (2 patients), both empirically and in the model. Thus, none of the performance scores violates the predicted task ordering, shown in Fig. 3.

3.4. Specificity of the atrophy effects

Good fits support the assumptions of a model only if other plausible assumptions would not fit the data equally well or better (Roberts & Pashler, 2000). The specificity of the predictions from common damage to semantic memory was examined by running simulations with a different or additional lexical locus of damage for AD (Isella et al., 2020, 2022) and a mixture of lexical and semantic loci for MCI (Gallant et al., 2019).

3.4.1. Shared locus versus a different or additional locus for AD

Simulations were run to test whether a semantic locus of atrophy for AD, shared with SD and MCI, does better than a different or additional locus for AD. Isella et al. (2020, 2022) argued that left inferior parietal cortex is an additional locus. Indeed, angular gyrus may be atrophied in AD (e.g., Zhou et al., 2012) but is not affected in SD (Yang et al., 2012). However, neurodegenerative diseases target specific functional networks (e.g., Seeley et al., 2009). The atrophy epicenter in the angular gyrus (predominantly right) in AD concerns the default mode network, also involved in episodic memory retrieval (e.g., Eichenbaum, 2012), rather than the language network. Isella et al. observed that hypometabolism in left supramarginal gyrus contributes to naming errors in AD *spectrum* (i.e., 46 patients with classic amnestic AD, 16 with a

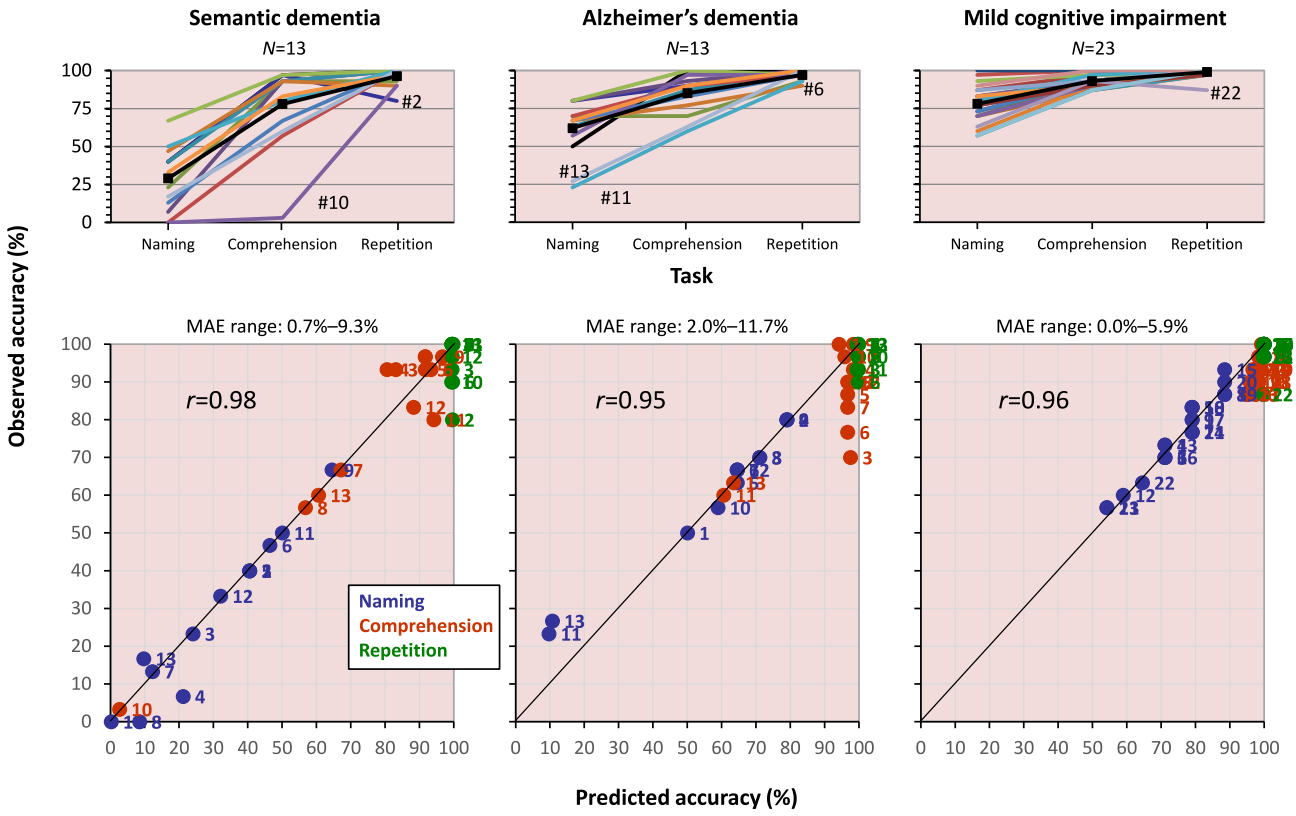


Fig. 5. Performance accuracy in semantic dementia, Alzheimer's dementia, and mild cognitive impairment for naming, comprehension, and repetition. The top panel shows the performance of individual patients from Janssen et al. (2022) denoted by different colored lines. The group averages are represented by black lines and squares. Numbers (e.g., #10) indicate boxplot-determined outliers in the patient scores. The bottom panel shows for each patient the predicted performance accuracy with the lowest MAE plotted against the observed accuracy. Individual patients are denoted by dot and number, and the tasks are color coded. *N* = number of patients. Partly adapted (for semantic dementia) from Roelofs (2022). Copyright 2022 by the author.

visuospatial deficit, and 8 with PPA, including 5 logopenic ones). However, the clinical heterogeneity precludes claims about classic AD specifically. Moreover, the contribution was small, affecting less than one percent of all responses. Furthermore, the effect concerned nonword phonemic errors, which are atypical for classic AD (e.g., Hodges, 2006). Thus, a possible role of left inferior parietal cortex in AD is small and atypical at most, if it exists at all. Still, atrophy in left temporoparietal cortex and corresponding white matter, including the arcuate fasciculus running underneath inferior parietal cortex, contributes to naming in the logopenic variant of PPA, associated with Alzheimer's pathology. In the model, the atrophy is taken to affect the network centered on lexical output forms and the connections between input and output phonemes (Roelofs, 2022). Simulations were run for AD assuming atrophy only in temporoparietal cortex (i.e., a different locus, as in logopenic PPA) or both in the ATL and temporoparietal cortex (i.e., an additional locus).

Assuming a different temporoparietal locus of damage for AD, weight decrease accounted for 40% of the variance in the individual AD patient data rather than the 91% assuming a semantic memory impairment shared with SD and MCI. Similarly, decay increase accounted for 64% of the variance rather than 91%. These differences were statistically significant (both *ps* < 0.001). Thus, the assumption of a different locus of damage for AD (i.e., the locus of logopenic PPA) receives less support from the data than a shared semantic locus. This agrees with the fact that the patients of Janssen et al. (2022) discussed here were diagnosed with AD rather than logopenic PPA.

Assuming ATL and temporoparietal loci of damage for AD (i.e., damage to the temporoparietal region in addition to the ATL damage shared with SD and MCI), weight decrease accounted for 80% of the variance rather than the 91% assuming a semantic memory impairment only ($p < .001$). Decay increase accounted for 90% of the variance, similar to the 91% assuming a semantic locus only ($p > .22$). Considering both lesion types, the assumption of an additional locus of damage for AD receives less support from the data than a shared semantic locus. It is important to note that these simulations concerned an addition locus of atrophy but not an additional free severity parameter for this additional locus. Adding another free parameter to the model for the additional location of damage could never make the behavioral fit worse, as the new free parameter could be set to the original value to obtain the previous result. However, if the new parameter takes the original value for all patients, then there is no additional locus of damage. Isella et al. (2022) observed that hypometabolism in left inferior parietal cortex affected less than one percent of all naming responses in AD spectrum (with the remaining uncertainty to what extent this small effect could be attributed to classic AD, which is the issue at stake). Introducing an extra free parameter only to capture this possible effect in AD, which is minor at most, would be overfitting the data.

To conclude, the simulations show that, compared to a semantic memory locus of damage shared with SD and MCI, a different lexical locus in AD (i.e., the locus of logopenic PPA) reduces the fit, whereas an additional lexical locus (with the same severity of damage) reduces or does not improve the fit. This demonstrates that the good fit (explaining 91% of the variance) depends on the assumption of semantic memory impairment in AD.

3.4.2. *Shared locus versus a mixture of loci for MCI*

Gallant et al. (2019) maintained that the functional origin of problems in word retrieval differs between groups, being mixed lexical and semantic in MCI and mostly semantic in AD. This claim was based on the type of naming errors and the efficacy of phonological cues. A lexical impairment was taken to be associated with coordinate semantic errors

and precise circumlocutions as well as good efficacy of phonological cues, whereas a semantic impairment was associated with a predominance of non-responses, coordinate and superordinate semantic errors, vague circumlocutions, and a poor efficacy of phonological cues. However, these performance profiles may also reflect severity of semantic memory impairment, which is larger for AD than MCI (Janssen et al., 2022). When damage of the semantic system is moderate, circumlocutions may be precise and lexical activation may still be enough for the word to become available with the help of phonological cues. When damage of the semantic system is more severe, circumlocutions will be vague due to a lack of semantic information and lexical activation may be too low for phonological cues to be effective.

Simulations were run to evaluate the mixed loci hypothesis of Gallant et al. (2019) for MCI. They stated: "In mild cognitive impairment, the origin of anomia was lexical for 60% and semantic for 40% of participants" (p. 95). The lexical origin was taken to concern "the phonological output lexicon and its access" (p. 104). This would correspond to the level of lexical output forms in WEAVER++/ARC. The simulations therefore assumed damage of lexical output forms and corresponding connections. Each patient with MCI was fitted twice, once with a semantic locus and once with a lexical locus, and the best fitting locus was selected. Assuming a weight decrease, the locus was lexical for 11 patients, semantic for 11 patients, and either lexical or semantic for one patient. The mixture assumption accounted for 92% of the variance in the individual patient data, which does not constitute an improvement compared to assuming a semantic locus only (91%). The same results were obtained with a decay increase. For both lesion types, there was no statistical difference between a common locus and mixed loci (both $ps > 0.28$).

The simulations show that allowing for a mixture of loci of damage in MCI does not improve the fit compared to what a semantic locus only already explains. The more constraint (i.e., the narrower the prediction), the stronger the support provided by a close fit (Roberts & Pashler, 2000). Given that the assumption of a mixture of loci provides less constraint than assuming a single locus, the data more strongly support a shared semantic memory locus than a mixture of lexical and semantic loci.

4. General discussion

It has been known since Pick (1892/1977, 1904/1997) that word retrieval is often impaired in left temporal lobe degeneration. Individuals with SD, AD, and MCI present, to different degrees, with problems in word production and comprehension, while repetition is largely preserved. To examine whether differential loss of activation capacity in semantic memory accounts for the difference in performance between patient groups, WEAVER++/ARC simulations were run. The simulations showed that weight decrease accounts for 99% of the variance in naming, comprehension, and repetition at the group level and 95% at the individual patient level. For decay increase, the percentages are 99% and 90%, respectively, which is lower at the individual patient level. Other plausible assumptions did less well. These results lend computational support to a unified account of word production, comprehension, and repetition in SD, AD, and MCI, along Pickian (1908) lines.

The simulations with WEAVER++/ARC assumed disruption of semantic memory consisting of a hub of supramodal concepts that are linked to widely distributed modality-specific features, referred to as a hub-and-spoke view (e.g., Lambon Ralph, 2014; Patterson et al., 2007; Pick, 1931). Two spokes were implemented in the present simulations, namely the connections of visual representations to the conceptual network (i.e., picture input) and the connections to lemmas, relevant for naming and word comprehension. Although the hub-and-spoke view is supported by much evidence (e.g., Kemmerer, 2022, for a review), it also has opponents. According to an alternative view, semantic memory consists of widely distributed modality-specific features without a central hub, as originally proposed by Wernicke (1874) and revived by Geschwind (1974). For a historical account of Wernicke's distributed-only view on concepts, I refer to Gage and Hickok (2005). According to Geschwind, in naming a seen object, the visual representation in occipital cortex activates, via the angular gyrus, an auditory image for the object name in Wernicke's area, which activates by way of the arcuate fasciculus the corresponding motor image in Broca's area, followed by articulation. The ATLs play no role in this account. Damage to the visual connections would impair the naming of visually perceived objects but spare the naming of objects identified through other modalities, like touch or smell. For domain-general loss of semantic knowledge to occur, as observed in SD, AD, and MCI, several domain-specific representations or connections among them should simultaneously be interrupted. Snowden et al. (2019) argued that "there may be no domain-general hub in which concepts are *represented* and which can be disrupted by selective damage. The semantic loss may be a product of widespread loss of *connections* across the semantic network" (p. 32). They argued that atrophy of the left fusiform gyrus in SD hampers visual access to the rest of the widely distributed network of sensory and motor features making up conceptual knowledge.

However, this modality-specific distributed-only view on concepts does not readily explain the evidence that word retrieval difficulty occurs across input modalities (i.e., not only vision, but also touch and other modalities). For example, word retrieval in SD is also impaired in naming objects (e.g., a knife) from touch and without vision (Coccia et al., 2004). Moreover, the distributed-only view does not explain why word retrieval difficulty may result from circumscribed ATL atrophy rather than wide-spread atrophy in areas coding for modality-specific features or connections among them. The evidence is exemplified by Pick's (1892/1977) first reported case of circumscribed left temporal atrophy, 71-year-old August H.: "He partially recognizes objects shown to him, but often describes them incorrectly: … Given a woolen glove, he rubs the palm of his hand and says: Wool"[3] (p. 37). That is, naming failed regardless of whether the object was seen of touched.

Roberts and Pashler (2000) argued that it is problematic to only focus on a model's fit as support for its assumptions. They stated: "A good fit reveals nothing about the flexibility of the theory (how much it cannot fit), the variability of the data (how firmly the data rule out what the theory cannot fit), or the likelihood of other outcomes (perhaps the theory could have fit any plausible result), and a reader needs all 3 pieces of information to decide how much the fit should increase belief in the theory" (p. 358). The present article provides the first type of information in Fig. 3, showing the predictions of the model. The figure shows exactly how performance accuracy varies with severity of damage in the model, constraining what the model can and cannot fit. The second type of information is provided by Figs. 4 and 5, showing the variability of the data. The third type of information is provided by the specificity analyses. The best evidence for a model concerns a fit to outcomes that would be otherwise unlikely (i.e., when other models predict different outcomes). If other plausible assumptions would fit the data equally well, then a good fit provides little support for the proposed model, whereas if good fits depend specifically on the assumptions of the model, they support the model.

As concerns the flexibility issue, it is true that WEAVER++/ARC is a complex model with a severity parameter that could be adjusted so that the output of the model resembles the patient scores. However, what the adjustment of the severity parameter allowed was constrained, as Fig. 3 shows. Moreover, although the model is complex, other assumptions have been independently motivated and empirically supported by previous research, and they were all held constant in the present study. The independent support comes from a wealth of behavioral

[3] "Vorgezeigte Gegenstände erkennt er theilweise, bezeichnet sie oft falsch…. Mit einem gereichten Wollhandschuh macht er reibende Bewegungen auf der Hohlhand und sagt Wolle" (Pick, 1892, p. 165).

psycholinguistic findings as well as evidence from functional neuroimaging, tractography, and aphasiology (e.g., Levelt et al., 1999; Janssen et al., 2020; Piai et al., 2014; Roelofs, 1992, 2014, 2018, 2021, 2022; Roelofs & Ferreira, 2019).

As concerns the variability issue, Fig. 4 shows the variability of the data at the group level and the top panel of Fig. 5 shows precisely how performance varies between individual patients. In the bottom panel of Fig. 5, the horizontal and vertical axes cover the entire range of possible predicted and observed scores, and the diagonal denotes a perfect fit. The figure shows that all 147 predicted scores fall close to the diagonal, with only a few exceptions, indicating that the predictions generally agree with the observations. The shared variance is 95%.

As concerns the likelihood issue (i.e., how likely is the model, given the observed performance scores), the specificity analyses revealed that predictions derived from other plausible assumptions are less well supported by the empirical observations than the assumption of damage to semantic memory shared between SD, AD, and MCI. First, assuming a different lexical locus of damage for AD rather than a semantic locus shared with SD and MCI significantly reduced the goodness of fit. Second, overall, an additional lexical locus of damage for AD reduced the fit. Third, a mixture of semantic and lexical loci for MCI did not improve the fit compared to what a shared semantic locus only already explained. The more constraint, the stronger the support provided by a good fit. Thus, the data more strongly support the more constrained assumption of a single locus of damage in MCI that is shared with SD and AD than the more flexible assumption of a mixture of loci for MCI. Other published computational models have simulated naming and repetition performance, but make no claims about comprehension (e.g., Dell et al., 1997; Hanley et al., 2004; Nozari & Dell, 2013; Walker & Hickok, 2016), or have simulated naming and comprehension, but make no claims about repetition (e.g., Rogers et al., 2004). The Lichtheim 2 model (Ueno et al., 2011) has been tested in simulations of naming, comprehension, and repetition in SD. WEAVER++/ARC shares with Lichtheim 2 the assumption of a semantic hub that is damaged in SD, which may be extended to AD and MCI. Ueno et al. reported that naming and comprehension decline to the same extent in simulations with the Lichtheim 2 model, which disagrees with the observation that naming is worse than comprehension in SD, AD, and MCI (see Roelofs, 2022, for discussion).[4]

The present simulation study has a number of limitations. First, the support for the hypothesis that differences in performance between SD, AD, and MCI reflect to a large extent different severities of disruption to the same mechanism in the ATL comes from simulations with a particular model, namely WEAVER++/ARC. It remains possible that a new, still to be developed model that assumes multiple loci of damage will explain more data than the presented model with a single locus in the ATL. Nevertheless, the current account is important, because it provides a benchmark for future modeling. Second, although the model fits are overall good, there are a number of discrepancies between model and data at both group and individual patient levels. This means that there is room for improvement. As Dell et al. (2000) stated, "*Every* model, including the very best ones, fails to accord perfectly with the data in its domain" (p. 637). However, one should not dismiss a model when some discrepancies exist, which would be "at odds with how models are used in psychology. It is also incompatible with the complex realities of research with brain-damaged patients" (p. 644). I refer to Dell et al. for a thorough discussion of the role of computational models in neuropsychological investigations of language.

To summarize, the endorsed model is constrained in what it can fit, the empirical observations tend to rule out what the model cannot fit, and the observations provide less support for other plausible assumptions. The good fits indicate the viability of the hypothesis that differences in performance between groups and participants reflect, to an important extent, different severities of disruption of semantic memory.

To conclude, WEAVER++/ARC simulations of word production, comprehension, and repetition showed that the model succeeds rather well in capturing the patterns of performance in SD, AD, and MCI. Other plausible assumptions do less well. This supports the central hypothesis of this article that differences in performance between SD, AD, and MCI reflect, to an important extent, different severities of a common disruption of semantic memory. In future research, the model may be extended further and tested in targeted studies.

[4] An anonymous reviewer suggested that this discrepancy could simply be remedied by manipulating the number of hidden layers in Lichtheim 2. But this would then demonstrate that worse naming than comprehension confirms a prediction of WEAVER++/ARC but not of Lichtheim 2. In the latter model, naming and comprehension scores can differ or be the same.

Declaration of Competing Interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The link is in the paper

Appendix A. Supplementary data

Supplementary data to this article are available from the Open Science Framework at https://osf.io/ue4bn/ or from Brain & Language online at https://doi.org/10.1016/j.bandl.2023.105243.

References

Altman, E. (1923). Über die umschriebene Gehirnatrophie des späteren Alters [About the circumscribed cerebral atrophy of later age]. *Zeitschrift für die gesamte Neurologie und Psychiatrie, 83*, 610–643. https://doi.org/10.1007/BF02869096

Alzheimer, A. (1907). Über eine eigenartige Erkrankung der Hirnrinde [About a peculiar disease of the cerebral cortex]. *Allgemeine Zeitschrift für Psychiatrie und psychisch-gerichtliche Medizin, 64*, 146–148.

Alzheimer, A. (1911). Über eigenartige Krankheitsfälle des späteren Alters [About peculiar cases of illness of later age]. *Zeitschrift für die gesamte Neurologie und Psychiatrie, 4*, 356–385. https://doi.org/10.1007/BF02866241

Alzheimer, A. (1991). On certain peculiar diseases of old age (H. Förstl and R. Levy, Trans.) History of Psychiatry, 2(5), 71–73. (Original work published 1911) doi: 10.1177/0957154X9100200505.

Arnold, S. E., Hyman, B. T., & Van Hoesen, G. W. (1994). Neuropathologic changes of the temporal pole in Alzheimer's disease and Pick's disease. *Archives of Neurology, 51*(2), 145–150. https://doi.org/10.1001/archneur.1994.00540140051014

Belleville, S., Fouquet, C., Hudon, C., Zomahoun, H., Croteau, J., & Consortium for the Early Identification of Alzheimer's disease-Quebec (2017). Neuropsychological measures that predict progression from mild cognitive impairment to Alzheimer's type dementia in older adults: A systematic review and meta-analysis. *Neuropsychology Review*, *27*(4), 328–353. https://doi.org/10.1007/s11065-017-9361-5.

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica, 82*(4), 239–259. https://doi.org/10.1007/BF00308809

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30*(5), 542–579. https://doi.org/10.1016/0749-596X(91)90026-G

Coccia, M., Bartolini, M., Luzzi, S., Provinciali, L., & Ralph, M. A. (2004). Semantic memory is an amodal, dynamic system: Evidence from the interaction of naming and object use in semantic dementia. *Cognitive Neuropsychology, 21*(5), 513–527. https://doi.org/10.1080/02643290342000113

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (2000). The role of computational models in neuropsychological investigations of language: Reply to Ruml and Caramazza (2000). *Psychological Review, 107*(3), 635–645. https://doi.org/10.1037/0033-295x.107.3.635

Dell, G. S., Schwartz, M. F., Nozari, N., Faseyitan, O., & Coslett, H. B. (2013). Voxel-based lesion-parameter mapping: Identifying the neural correlates of a computational model of word production. *Cognition, 128*(3), 380–396. https://doi.org/10.1016/j.cognition.2013.05.007

Diedenhofen, B., & Musch, J. (2015). Cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS One, 10*(3), Article e0121945. https://doi.org/10.1371/journal.pone.0121945

Edmonds, E. C., Weigand, A. J., Hatton, S. N., Marshall, A. J., Thomas, K. R., Ayala, D. A., … ADN Initiative. (2020). Patterns of longitudinal cortical atrophy over 3 years in empirically derived MCI subtypes. e2532–e2544 *Neurology, 94*(24). https://doi.org/10.1212/WNL.0000000000009462.

Eichenbaum, H. (2012). *The cognitive neuroscience of memory: An introduction* (2nd Ed.). Oxford, UK: Oxford University Press.

Fischer, O. (1907). Miliare Nekrosen mit drusigen Wucherungen der Neurofibrillen, eine regelmässige Veränderung der Hirnrinde bei seniler Demenz [Miliary necrosis with geoidal growths of the neurofibrillary tangles, a regular change in the cerebral cortex in senile dementia]. *Monatsschrift für Psychiatrie und Neurologie, 22*, 361–372. https://doi.org/10.1159/000211873

Fischer, O. (1910). Die presbyophrene Demenz, deren anatomische Grundlage und klinische Abgrenzung [Presbyophrenic dementia, its anatomical basis and clinical differentiation]. *Zeitschrift für die gesamte Neurologie und Psychiatrie, 3*, 371–471. https://doi.org/10.1007/BF02893605

Gage, N., & Hickok, G. (2005). Multiregional cell assemblies, temporal binding and the representation of conceptual knowledge in cortex: A modern theory by a "classical" neurologist, Carl Wernicke. *Cortex, 41*(6), 823–832. https://doi.org/10.1016/s0010-9452(08)70301-0

Gallant, M., Lavoie, M., Hudon, C., & Monetta, L. (2019). Analysis of naming errors in healthy aging, mild cognitive impairment, and Alzheimer's disease. *Canadian Journal of Speech-Language Pathology and Audiology, 43*(2), 95–108. http://hdl.handle.net/20.500.11794/39569.

Gans, A. (1923). Betrachtungen über Art und Ausbreitung des krankhaften Prozesses in einem Fall von Pickscher Atrophie des Stirnhirns [Reflections on the nature and extent of the morbid process in a case of Pick's atrophy of the frontal lobe]. *Zeitschrift für die gesamte Neurologie und Psychiatrie, 80*, 10–28. https://doi.org/10.1007/BF02866679

Geschwind, N. (1974). Language and the brain. *Scientific American, 226*(4), 76–83. https://doi.org/10.1038/scientificamerican0472-76

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

Hanley, J. R., Dell, G. S., Kay, J., & Baron, R. (2004). Evidence for the involvement of a nonlexical route in the repetition of familiar words: A comparison of single and dual route models of auditory repetition. *Cognitive Neuropsychology, 21*(2), 147–158. https://doi.org/10.1080/02643290342000339

Hodges, J. R. (2006). Alzheimer's centennial legacy: Origins, landmarks and the current status of knowledge concerning cognitive aspects. *Brain, 129*(11), 2811–2822. https://doi.org/10.1093/brain/awl275

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain, 115*, 1783–1806. https://doi.org/10.1093/brain/115.6.1783

Isella, V., Rosazza, C., Gazzotti, M., Sala, J., Morzenti, S., Crivellaro, C., Appollonio, I. M., Ferrarese, C., & Luzzatti, C. (2020). A metabolic imaging study of lexical and phonological naming errors in Alzheimer Disease. *American Journal of Alzheimer's Disease and Other Dementias*, *35*, 1533317520922390. https://doi.org/10.1177/1533317520922390.

Isella, V., Rosazza, C., Ferri, F., Gazzotti, M., Impagnatiello, V., Mapelli, C., Morzenti, S., Crivellaro, C., Appollonio, I. M., & Ferrarese, C. (2022). Learning from mistakes: Cognitive and metabolic correlates of errors on picture naming in the Alzheimer's Disease Spectrum. *Journal of Alzheimer's Disease, 87*(3), 1033–1053. https://doi.org/10.3233/JAD-220053

Janssen, N., Roelofs, A., Mangnus, M., Sierpowska, J., Kessels, R., & Piai, V. (2020). How the speed of word finding depends on ventral tract integrity in primary progressive aphasia. *NeuroImage: Clinical, 28*, Article 102450. https://doi.org/10.1016/j.nicl.2020.102450

Janssen, N., Roelofs, A., Van den Berg, E., Holleman, M. A., In de Braek, D. M. J. M., Piguet, O., Piai, V., & Kessels, R. P. C. (2022). The diagnostic value of language screening in primary progressive aphasia: Validation and application of the Sydney Language Battery. *Journal of Speech, Language, and Hearing Research, 65*(1), 200–214. https://doi.org/10.1044/2021_JSLHR-21-00024

Josephs, K. A., Martin, P. R., Weigand, S. D., Tosakulwong, N., Buciuc, M., Murray, M. E., Petrucelli, L., Senjem, M. L., Spychalla, A. J., Knopman, D. S., Boeve, B. F., Petersen, R. C., Parisi, J. E., Dickson, D. W., Jack, C. R., & Whitwell, J. L. (2020). Protein contributions to brain atrophy acceleration in Alzheimer's disease and primary age-related tauopathy. *Brain, 143*(11), 3463–3476. https://doi.org/10.1093/brain/awaa299

Joubert, S., Gardy, L., Didic, M., Rouleau, I., & Barbeau, E. J. (2021). A meta-analysis of semantic memory in mild cognitive impairment. *Neuropsychology Review, 31*(2), 221–232. https://doi.org/10.1007/s11065-020-09453-5

Kemmerer, D. (2022). *Cognitive neuroscience of language* (2nd Ed.). New York, NJ: Psychology Press.

Kertesz, A., & Kalvach, P. (1996). Arnold Pick and German neuropsychiatry in Prague. *Archives of Neurology, 53*(9), 935–938. https://doi.org/10.1001/archneur.1996.00550090147021

Lambon Ralph M. A. (2014). Neurocognitive insights on conceptual knowledge and its breakdown. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 369*(1634), 20120392. https://doi.org/10.1098/rstb.2012.0392.

Landin-Romero, R., Tan, R., Hodges, J. R., & Kumfor, F. (2016). An update on semantic dementia: Genetics, imaging, and pathology. *Alzheimer's Research & Therapy, 8*(1), 52. https://doi.org/10.1186/s13195-016-0219-5

Laws, K. R., Adlington, R. L., Gale, T. M., Moreno-Martínez, F. J., & Sartori, G. (2007). A meta-analytic review of category naming in Alzheimer's disease. *Neuropsychologia, 45*(12), 2674–2682. https://doi.org/10.1016/j.neuropsychologia.2007.04.003

Levelt, W. J. M. (2013). *A history of psycholinguistics: The pre-Chomskyan era*. Oxford University Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

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

Lombardi, J., Mayer, B., Semler, E., Anderl-Straub, S., Uttner, I., Kassubek, J., Diehl-Schmid, J., Danek, A., Levin, J., Fassbender, K., Fliessbach, K., Schneider, A., Huppertz, H. J., Jahn, H., Volk, A., Kornhuber, J., Landwehrmeyer, B., Lauer, M., Prudlo, J., Wiltfang, J., … FTLD consortium (2021). Quantifying progression in primary progressive aphasia with structural neuroimaging. *Alzheimer's & Dementia*, *17*(10), 1595–1609. https://doi.org/10.1002/alz.12323.

Mandelli, M. L., Vilaplana, E., Brown, J. A., Hubbard, H. I., Binney, R. J., Attygalle, S., Santos-Santos, M. A., Miller, Z. A., Pakvasa, M., Henry, M. L., Rosen, H. J., Henry, R. G., Rabinovici, G. D., Miller, B. L., Seeley, W. W., & Gorno-Tempini, M. L. (2016). Healthy brain connectivity predicts atrophy progression in non-fluent variant of primary progressive aphasia. *Brain, 139*(10), 2778–2791. https://doi.org/10.1093/brain/aww195

Martin, N., & Dell, G. S. (2019). Maintenance versus transmission deficits: The effect of delay on naming performance in aphasia. *Frontiers in Human Neuroscience, 13, Article 406*. https://doi.org/10.3389/fnhum.2019.00406

Mesulam, M. M., Thompson, C. K., Weintraub, S., & Rogalski, E. J. (2015). The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. *Brain, 138*(8), 2423–2437. https://doi.org/10.1093/brain/awv154

Nozari, N., & Dell, G. S. (2013). How damaged brains repeat words: A computational approach. *Brain and Language, 126*(3), 327–337. https://doi.org/10.1016/j.bandl.2013.07.005

Ohm, D. T., Cousins, K., Xie, S. X., Peterson, C., McMillan, C. T., Massimo, L., Raskovsky, K., Wolk, D. A., Van Deerlin, V. M., Elman, L., Spindler, M., Deik, A., Trojanowski, J. Q., Lee, E. B., Grossman, M., & Irwin, D. J. (2022). Signature laminar distributions of pathology in frontotemporal lobar degeneration. *Acta Neuropathologica, 143*(3), 363–382. https://doi.org/10.1007/s00401-021-02402-3

Onari, K., & Spatz, H. (1926). Anatomische Beiträge zur Lehre von Pickschen umschriebenen Großhirnrinden-Atrophie ("Picksche Krankheit") [Anatomical contributions to the theory of Pick's circumscribed cerebral cortex atrophy ("Pick's disease")]. *Zeitschrift für die gesamte Neurologie und Psychiatrie, 101*, 470–511. https://doi.org/10.1007/BF02878348

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews: Neuroscience, 8*(12), 976–987. https://doi.org/10.1038/nrn2277

Petersen, R. C., & Morris, J. C. (2005). Mild cognitive impairment as a clinical entity and treatment target. *Archives of Neurology, 62*(7), 1160–1167. https://doi.org/10.1001/archneur.62.7.1160

Petersen, R. C., Smith, G. E., Ivnik, R. J., Kokmen, E., & Tangalos, E. G. (1994). Memory function in very early Alzheimer's disease. *Neurology, 44*(5), 867–872. https://doi.org/10.1212/wnl.44.5.867

Piai, V., Roelofs, A., Jensen, O., Schoffelen, J. M., & Bonnefond, M. (2014). Distinct patterns of brain activity characterise lexical activation and competition in spoken word production. *PLoS One, 9*(2). https://doi.org/10.1371/journal.pone.0088674. e88674.

Pick, A. (1892). Ueber die Beziehungen der senilen Hirnatrophie zur Aphasie [On the relationships between senile brain atrophy and aphasia]. *Prager Medicinische Wochenschrift, 17*(16), 165–167.

Pick, A. (1904). Zur Symptomatologie der linksseitigen Schläfenlappenatrophie [On the symptomatology of left-sided temporal lobe atrophy]. *Monatsschrift für Psychiatrie und Neurologie, 16*, 378–388. https://doi.org/10.1159/000219215

Pick, A. (1908). Die umschriebene senile Hirnatrophie als Gegenstand klinischer und anatomischer Forschung [Circumscribed senile cerebral atrophy as an object of clinical and anatomical research]. In A. Pick (Ed.), *Arbeiten aus der deutschen psychiatrischen Universitäts-Klinik in Prag* (pp. 20–28). Berlin: Karger.

Pick, A. (1913). *Die agrammatischen Sprachstörungen: Studien zur psychologischen Grundlegung der Aphasielehre* [The agrammatic speech disorders: Studies on the psychological basis of the theory of aphasia]. Berlin, Germany: Springer-Verlag.

Pick, A. (1931). Aphasie [aphasia]. In A. Bethe, W. Buddenbrock, E. Fischer, M. H. Fischer, K. Goldstein, E. Hansen, A. Pick, E. Scharrer, R. Sokolowsky, W. Sulze, R. Thiele, & O. Weiss (Eds.), *Arbeitsphysiologie II Orientierung: Plastizität, Stimme und Sprache* (pp. 1416–1524). Berlin: Springer. https://doi.org/10.1007/978-3-642-47376-0_11.

Pick, A. (1977). On the relation between aphasia and senile atrophy of the brain (W. C. Schoene, Trans.). In D. A. Rottenberg & F. H. Hochberg (Eds.), *Neurological classics in modern translation* (pp. 35–40). London, UK: Macmillan. (Original work published 1892).

Pick, A. (1997). On the symptomatology of left-sided temporal lobe atrophy (D. M. Girling and G. E. Berrios, Trans.). *History of Psychiatry*, *8*(29), 149–159. (Original work published 1904) https://doi.org/10.1177/0957154X9700802910.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107*(2), 358–367. https://doi.org/10.1037/0033-295x.107.2.358

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42*(1–3), 107–142. https://doi.org/10.1016/0010-0277(92)90041-F

Roelofs, A. (2014). A dorsal-pathway account of aphasic language production: The WEAVER++/ARC model. *Cortex, 59*, 33–48. https://doi.org/10.1016/j.cortex.2014.07.001

Roelofs, A. (2018). A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition, 172*, 59–72. https://doi.org/10.1016/j.cognition.2017.12.007

Roelofs, A. (2021). Phonological cueing of word finding in aphasia: Insights from simulations of immediate and treatment effects. *Aphasiology, 35*(2), 169–185. https://doi.org/10.1080/02687038.2019.1686748

Roelofs, A. (2022). A neurocognitive computational account of word production, comprehension, and repetition in primary progressive aphasia. *Brain and Language, 227*, Article 105094. https://doi.org/10.1016/j.bandl.2022.105094

Roelofs, A., & Ferreira, V. S. (2019). The architecture of speaking. In P. Hagoort (Ed.), *Human language: From genes and brains to behavior* (pp. 35–50). Cambridge, MA: MIT Press.

Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review, 111*(1), 205–235. https://doi.org/10.1037/0033-295X.111.1.205

Rogers, T. T., Ivanoiu, A., Patterson, K., & Hodges, J. R. (2006). Semantic memory in Alzheimer's disease and the frontotemporal dementias: A longitudinal study of 236 patients. *Neuropsychology, 20*(3), 319–335. https://doi.org/10.1037/0894-4105.20.3.319

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

Seeley, W. W., Crawford, R. K., Zhou, J., Miller, B. L., & Greicius, M. D. (2009). Neurodegenerative diseases target large-scale human brain networks. *Neuron, 62*(1), 42–52. https://doi.org/10.1016/j.neuron.2009.03.024

Snowden, J. S., Goulding, P. J., & Neary, D. (1989). Semantic dementia: A form of circumscribed cerebral atrophy. *Behavioural Neurology, 2*(3), 167–182. https://doi.org/10.1155/1989/124043

Snowden, J. S., Harris, J. M., Saxon, J. A., Thompson, J. C., Richardson, A. M., Jones, M., & Kobylecki, C. (2019). Naming and conceptual understanding in frontotemporal dementia. *Cortex, 120*, 22–35. https://doi.org/10.1016/j.cortex.2019.04.027

Spatt, J. (2003). Arnold Pick's concept of dementia. *Cortex, 39*(3), 525–531. https://doi.org/10.1016/s0010-9452(08)70262-4

Stertz, G. (1926). Über die Picksche Atrophie [On Pick's atrophy]. *Zeitschrift für die gesamte Neurologie und Psychiatrie, 101*, 729–747. https://doi.org/10.1007/BF02878359

Taler, V., Monetta, L., Sheppard, C., & Ohman, A. (2020). Semantic function in mild cognitive impairment. *Frontiers in Psychology, 10*, Article 3041. https://doi.org/10.3389/fpsyg.2019.03041

Ueno, T., Saito, S., Rogers, T. T., & Lambon Ralph, M. A. (2011). Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron, 72*(2), 385–396. https://doi.org/10.1016/j.neuron.2011.09.013

Walker, G. M., & Hickok, G. (2016). Bridging computational approaches to speech production: The semantic-lexical-auditory-motor model (SLAM). *Psychonomic Bulletin & Review, 23*(2), 339–352. https://doi.org/10.3758/s13423-015-0903-7

Warrington, E. K. (1975). The selective impairment of semantic memory. *The Quarterly Journal of Experimental Psychology, 27*(4), 635–657. https://doi.org/10.1080/14640747508400525

Wernicke, C. (1874). *Der aphasische Symptomencomplex: Eine psychologische Studie auf anatomischer Basis* [The aphasic symptom complex: A psychological study on anatomical basis]. Breslau: Cohn und Weigert.

Yang, J., Pan, P., Song, W., & Shang, H. F. (2012). Quantitative meta-analysis of gray matter abnormalities in semantic dementia. *Journal of Alzheimer's Disease, 31*(4), 827–833. https://doi.org/10.3233/JAD-2012-120736

Zhou, J., Gennatas, E. D., Kramer, J. H., Miller, B. L., & Seeley, W. W. (2012). Predicting regional neurodegeneration from the healthy brain functional connectome. *Neuron, 73*(6), 1216–1227. https://doi.org/10.1016/j.neuron.2012.03.004