Aging-related effects on the controlled retrieval of semantic information

Ettore Ambrosini^{1,2,3}, Francesca Peressotti⁴, Marisa Gennari⁴, Silvia Benavides-Varela^{1,4}, Maria Montefinese⁵

¹ Department of Neuroscience, University of Padova

² Department of General Psychology, University of Padova

³ Padova Neuroscience Center, University of Padova

⁴ Department of Developmental and Social Psychology, University of Padova

⁵ IRCCS San Camillo Hospital, Venice, Italy

Author note

Data, code, and materials can be found in our project on the Open Science Framework (OSF): <https://osf.io/qgryv>. This research was funded by the Italian Ministry of Health under grant number GR-2019-12371166 to MM. Correspondence concerning this article should be addressed to Maria Montefinese, IRCCS San Camillo Hospital, Via Alberoni, 70 30126 Venice (Italy); email address: maria.montefinese@gmail.com

Abstract

The efficient use of knowledge requires semantic control processes to retrieve context-relevant information. So far, it is well established that semantic knowledge, as measured with vocabulary tests, does not decline in aging. Yet, it is still unclear if controlled retrieval -the context-driven retrieval of very specific aspects of semantic knowledge- declines in aging, following the same fate of other forms of cognitive control. Here, we tackled this issue by comparing the performance of younger and older native Italian speakers during a semantic feature verification task. To manipulate the control demands, we parametrically varied the semantic salience of the target feature for the cue concept. As compared to their young counterparts, older adults showed a worse performance (in terms of reaction times) as the salience of the target feature for the concept decreased. This result suggests that older people have relatively greater difficulties in regulating the activation within semantic representation when conditions pose high demands of controlled retrieval of semantic information.

Keywords: semantic representation, semantic control, controlled retrieval, aging, word

Introduction

Semantic control is critical for a productive and appropriate use of knowledge. One of the first things that comes to mind when seeing a PENCIL is that it can be used to write; this is a dominant property of the concept PENCIL that is thought to be retrieved automatically from our semantic representation. However, the fact that a PENCIL is used to write might be of little utility when we forget our hair tie and want to hold our hair back. In this scenario, retrieving the information that a PENCIL can be used to put our hair up in a bun (relevant for the context, but generally irrelevant for PENCIL's semantic representation) could solve our problem. However, the retrieval of this latter property of PENCIL might require greater control demands as it is a non-dominant property of the concept and, thus, it is only weakly activated by the concept itself (Corbett et al., 2009; Jefferies & Lambon Ralph, 2006; Montefinese, Hallam, Thompson, et al., 2020; Noonan et al., 2010; Whitney et al., 2009).

According to Lambon Ralph and colleagues (2017), semantic cognition relies on two distinct but interacting systems: semantic representation and semantic control. Semantic representation encodes the knowledge about concepts while semantic control regulates the activation within semantic representation, in line with the current goals. Semantic representation and control can be impaired separately, yielding dissociations between semantic dementia (characterised by degradation of the store of conceptual knowledge following anterior temporal lobe atrophy) and semantic aphasia (SA), which results in deficits of semantic control in the context of an intact semantic store (Corbett et al., 2009; Jefferies & Lambon Ralph, 2006). There are two types of semantic control processes (Badre et al., 2005; Hoffman, 2019; Whitney et al., 2012): (a) controlled retrieval (i.e., the identification and retrieval of weak, but task-relevant, relations among different aspects of our semantic representation) and (b) selection (i.e., dealing with competition between representations). The controlled retrieval is thought to be supported by anterior/ventral parts of the inferior prefrontal cortex within the default mode network (Badre et al., 2005; Davey et al., 2016; Wagner et al., 2001). In contrast, the selection of goal-relevant semantic knowledge over irrelevant competitors (Badre et al., 2005; Whitney et al., 2012) is thought to be supported by posterior/dorsal parts of inferior prefrontal cortex and the adjoining inferior frontal sulcus and precentral sulcus, which overlap with the multiple-demand system (Davey et al., 2015, 2016). In this investigation

we will focus on the study of the controlled retrieval of semantic information. Generally speaking, in controlled retrieval tasks, participants have to retrieve a target that is weakly related to a given cue. The extent to which a cue is effective in eliciting automatically the retrieval of task-relevant information determines the demands of controlled retrieval. For strong relations within semantic representation, task demands require minimal controlled retrieval, whereas additional control resources are required to recover the weak relations. In other words, the retrieval of semantic information may emerge through a combination of both bottom-up (automatic) and top-down (controlled) mechanisms, operating complementarily depending not only on the efficacy of the available bottom-up cues, but also on the context and task demands. In this scenario, controlled retrieval is not an all-or-none process (Badre et al., 2005). In other words, there would not be a specific point where the controlled processes step in to supersede automatic processes to retrieve semantic information.

Previous research outside of the semantic domain showed that other forms of cognitive control tend to decline in aging, such as the inhibition of irrelevant information (e.g., Darowski et al., 2008; Hasher et al., 1991, 1997; Kramer et al., 1994), the coordination of multiple simultaneous cognitive operations (e.g., McDowd & Craik, 1988; Naveh-Benjamin et al., 2005; Somberg & Salthouse, 1982), the manipulation of information within working memory (e.g., Craik, 1977; Reuter-Lorenz & Sylvester, 2005), and the control of episodic memory (e.g., Old & Naveh-Benjamin, 2008; Spencer & Raz, 1995; Tun et al., 1998).

The literature on the semantic control in aging is scanty and controversial. Indeed, a small number of behavioral studies (Alves et al., 2021; Hoffman, 2018, 2019) investigated the aging effect on the semantic control processes. These studies agreed in reporting that while older participants performed more poorly than younger participants when selecting among competing semantic representations, suggesting that semantic selection declines with aging (Hoffman, 2018, 2019) controlled retrieval processes were preserved (Alves et al., 2021; Hoffman, 2018, 2019). However, neuroimaging evidence documented a reduction in recruitment of brain areas involved in the controlled retrieval in older participants, consistent with an age-related impairment in the ability to engage controlled retrieval and monitoring processes as a function of demands on controlled retrieval processes (Giovanello & Schacter, 2012; McDonough et al., 2013; Trelle et al.,

2019). This contrasting evidence between behavioral and neuroimaging studies makes thus necessary to verify whether the controlled retrieval actually changes with age.

In this study, we tackled this issue by investigating the effect of age on the controlled retrieval process in healthy individuals. To this aim, we used a speeded feature verification task, in which participants had to decide whether a semantic feature was true or false for a given concept. This task has been used for a long time to test the distributed feature-based models of semantic representation. These models posit that concepts are described as distributed patterns of activity across sets of semantic features (i.e., the concept ELEPHANT is characterized by various features such as “has a trunk”, “is big” and so forth) which contribute, with different weights, to the concept meaning (Montefinese et al., 2014; Montefinese, Hallam, Thompson, et al., 2020). In line with the feature-based models, we manipulated controlled retrieval demands by linearly varying the semantic significance (henceforth, salience) of the target feature to the cue concept, an emergent property of semantic representation (Montefinese et al., 2013, 2014).

Semantic salience (called significance in Montefinese et al., 2013, 2014) is composed of two components: a local one, called semantic accessibility (i.e., a measure that represents the priority with which a feature emerges in the conceptual representation for a given concept) and a global one, called distinctiveness (i.e., a measure that indicates how important a feature is for identifying and distinguishing a certain concept among similar ones, thus taking into account the whole set of concepts in the norms; for the calculation of this measure, see the Method section).

A feature that captures the core meaning of a concept will be easily retrieved and will have both high accessibility and high distinctiveness, resulting in a high semantic salience. For example, the feature "has a trunk" (high salience) is more easily retrieved in response to the concept ELEPHANT than the feature "is big" (low salience). Thus, in terms of control demands, as semantic salience decreases, control demands on the retrieval of semantic information increase. Previous studies on healthy young individuals (Montefinese, Hallam, Stampacchia, et al., 2020) and SA patients (Montefinese, Hallam, Thompson, et al., 2020) observed that both these groups of participants had greater difficulty in the selection of low salient features, suggesting that with respect to highly salient features, selecting low salient features involves to a larger extent semantic control processes. However, previous literature on the effects of aging on controlled retrieval

operationalized the semantic dimensions as categorical variables (Hoffman, 2018; 2019), thus oversimplifying the complexity and variability of these variables and preventing the examination of a fine-grained link between age and controlled retrieval demands. In particular, Hoffman (2018) adopted the word association (a measure of the relation between words that captures different types of information, not only strictly semantic but also lexical) in a semantic decision task (Hoffman, 2018), where participants were presented with a probe word and were asked to select its associate (which was operationalized as having either high or low association to the target word) from either two or four alternatives that were presented below the probe word. The year after, the same author used synonym words (with no quantitative measurement of the association between probe and target) in a synonym judgement task (Hoffman, 2019) where participants were presented with a probe word and asked to select which of three alternatives presented below the probe word was the most closely related in meaning to the probe.

Conversely, Alves and colleagues manipulated the semantic control demands as the concept typicality for a semantic category (e.g., the concept whale for the category mammals) on a 7-point scale in a semantic categorization task (Alves et al., 2021), where participants were asked to decide whether a target concept presented at the bottom of the screen belonged to the category written above. This study thus employed a continuous measure of the strength of the semantic relationship between a category and a concept to assess in a fine-grained way the age-related differences in controlled retrieval abilities, finding preserved controlled retrieval abilities in older adults. However, it is still unknown whether aging affects the controlled retrieval of semantic features.

Here we investigated how aging affects controlled retrieval of semantic features in a fine-grained way by parametrically manipulating controlled retrieval demands by linearly varying the semantic salience. We assume that the retrieval of features with a lower semantic salience would require a greater controlled retrieval, with a consequent decline in performance compared with those with a higher semantic salience. Consequently, if controlled retrieval suffers the same fate as other forms of cognitive control, with an age-related decline, we predict that performance of older adults should be more strongly affected by the salience value of the target feature in comparison to younger adults.

Method

Transparency and openness

We report how we determined our sample size, and describe all data exclusions, manipulations, and all measures in the study (described in detail below). All deidentified data, analysis code, and research materials are made available (see author note for the links to the study materials). Data were analyzed using R, version 4.0.3 (R Development Core Team, 2012). The study design and the analyses were not pre-registered.

Participants

Sixty younger (32 females, 28 males; mean age = 26.13 years, $SD = 3.88$ years) and 60 older (33 females, 27 males; mean age = 66.75 years, $SD = 6.68$ years) native Italian speakers participated in this experiment. The participants were recruited via e-mails and social networks or identified via researchers' personal contacts. Note that since the stimuli and the procedure of this study are based on the method of our previous study (Montefinese et al., 2014), 63 additional young participants from that study (37 females, 26 males; mean age = 27.57 years, $SD = 8.12$ years; see Method section of Experiment 1 in Montefinese et al., 2014, for details on the sample) were added to the current sample of young participants in the final model (see Statistical Analysis section). We did so to increase the statistical power and precision of the experimental effects estimates after having verified that young adults' performance in this study and the one in the previous study were comparable (see Statistical Analysis and Results sections). The behavioural data of one participant from Montefinese et al.'s (2014) study went missing due to technical errors.

We used the method introduced by Westfall and colleagues (2014) to perform a sensitivity power analysis for a linear mixed-effects model, assuming conservatively participant, stimulus, participant-by-stimulus, and residual variance partitioning coefficients of .05, .05, .05, and .85, respectively. This analysis revealed that our final sample size (183 participants and 60 experimental stimuli/participant) was large enough to detect a small effect size ($d = .10$) with a power $> .99^1$.

¹ Note also that an a-priori power analysis performed with G*Power (Faul et al., 2007) indicated that a total sample size of 112 participants is required to detect the within-between interaction of interest in a Group (older vs. younger) by Salience (low vs. high) ANOVA with a small effect size ($d = .1$) and a power of .80, assuming a correlation between repeated measures of $r = .93$, as estimated from our previous study (Montefinese et al., 2014).

All participants reported to have a normal or corrected-to-normal visual acuity. They all reported to be in good health and with no history of neurological or psychiatric illness. The Mini-Mental State Examination (MMSE) (Folstein et al., 1975) of the older participants was higher than the cut-off score 24. Experience and competence in computer use were not formally assessed, but all participants completed the computerized task without major issues. The total sample of younger participants was similar to the sample of older participants in terms of proportions of males and females ($\chi^2_{(1)} < 0.020$, $p = 0.888$), but different in terms of education (Welch's t test: $t_{(88.72)} = 8.209$, $p < .001$, Cohen's $d = 1.36$), with more years of education for younger participants ($M = 14.94$, $SD = 3.34$) as compared to older ones ($M = 9.33$, $SD = 4.70$)

All participants were naïve about the purpose of the study and gave online informed consent prior to their inclusion in the study, which was approved by the Ethical Committee for the Psychological Research of the University of Padova (Protocol number: 3998) in compliance with the ethical standards of the Declaration of Helsinki.

Materials

The materials were the same as Montefinese et al.'s (2014) study. In brief, the stimulus set was selected from Montefinese, Ambrosini, Fairfield and Mammarella's (2013) database including numerous lexico-semantic measures collected from Italian native speakers using exemplar and feature production tasks. To create the stimulus set, 12 exemplars were included from each of the following 10 categories: animals, body parts, clothes, furnishings/fittings, furniture, housing buildings, kitchenware, plants, stationary and vehicles. Therefore, the stimuli consisted of words denoting 120 concrete concepts. Each concept was then paired with two short featural descriptions taken from the semantic norms so to maximize the difference in their semantic salience (i.e., one feature had higher salience compared with the other: $M = 23.24$, $SD = 6.37$ for the high salience features vs. $M = 12.21$, $SD = 4.57$ for the low salience features across all concepts) while covering the entire distribution of it. Therefore, there were 240 true concept–feature pairs, which were divided in four lists of 60 true concept–feature pairs each. The 60 concepts of these concept–feature pairs included six concepts from each of the 10 semantic categories. Each concept occurred only once in each list.

In order to avoid cueing responses to the participants, each list also included 60 filler concept–feature pairs, in which the feature did not reasonably describe the

corresponding concept. To create these trials, the features from List 1 were combined with the concepts from List 2 (for which the features were false), and vice versa. The same method was used for Lists 3 and 4. The experimental concept–feature pairs were characterized by a number of concept- and feature-level variables taken from the Montefinese et al.’s norms (2013), which were all shown in the literature (e.g., Cree et al., 2006; McRae et al., 1999; Yap et al., 2011) to influence (or could logically influence) participants’ performance in semantic tasks. These variables are listed in Table 1, organized in the categories emerged from the hierarchical clustering analysis performed in our previous study based on their intercorrelations (Montefinese et al., 2014; For a detailed description and calculation of the measures, see study by Montefinese et al. (2013)). Specifically, there were six categories of variables: 1) a concept-level category consisting of the two lexical variables for the concept words; 2) a concept-level category consisting of the two semantic variables reflecting the richness of the semantic representation of a concept (i.e., the amount of semantic information that concept evokes in the participants’ mind); 3) a concept-level category consisting of the five semantic variables reflecting what we called the “importance” of a concept for the semantic representation of the corresponding category, that is, its familiarity and typicality, but also the readiness and likelihood with which a concept is produced in response to its semantic category; 4) a feature-level category consisting of the only lexical variable for the featural description²; 5) a feature-level category consisting of the four semantic variables reflecting what we called the “informativeness” of a feature, that is, the degree to which it facilitates the identification of a given concept; 6) a feature-level category consisting of the four semantic variables reflecting what we called the “importance” of a feature for the semantic representation of the corresponding concept, that is the readiness and likelihood with which a feature is produced in response to a concept, or the strength of the semantic relationship between a feature and concept.

These variables were all tested as potential predictors of participants’ performance in the model building procedure (see Statistical Analysis section). Since each list comprised all the 120 concepts included in Montefinese et al.’s norms (2013), all the concept-level variables were automatically matched across the four lists (see Table 1). In

² Note that we did not include the frequency of use for the featural descriptions because in most cases they consisted of multiple words

addition, we balanced the four lists for all the feature-level variables shown in Table 1 except the semantic Salience (all $t_s(118) \leq 1.09$; all $p_s \geq 0.27$, two-tailed).

Table 1. Concept- and feature-level variables characterizing the experimental stimuli.

Measure	Description
Concept level	
<i>Lexical</i>	
Length	Number of letters composing the concept word
Frequency	Natural logarithm of the frequency taken from the COLFIS (Bertinetto et al., 2005)
<i>Richness</i>	
Number of features	Number of features listed for a concept (Pexman et al., 2008)
Summed production frequency	Summed frequency of production of all the non-idiosyncratic features listed for a concept (Montefinese et al., 2013)
<i>Concept Importance</i>	
Familiarity	Ratings on a 10-point Likert scale (1 = not familiar at all; Montefinese et al., 2013)
Typicality	Ratings on a 10-point Likert scale (1 = extremely atypical; Montefinese et al., 2013)
Dominance	Proportion of participants who produced a concept in response to a category (McRae et al., 2005)
First occurrence	Proportion of participants who produced a concept as their first response to a category (Montefinese et al., 2013);
Availability	Readiness with which a concept is produced in response to a category (Izura et al., 2005)
Feature level	
<i>Lexical</i>	
Length	Number of letters composing the featural description
<i>Informativeness</i>	
Correlation percentage	Percentage of features occurring in the same concept with which a given feature is correlated (McRae et al., 1997)
Intercorrelational strength	Mean proportion of shared variance between a feature and each other correlated features (McRae et al., 1997)
Cue validity	The conditional probability of a concept, given a feature (Bourne & Restle, 1959; McRae et al., 1997)
Distinctiveness	The inverse of the number of concepts in which the feature appears in the entire norms (McRae et al., 2005)
<i>Feature Importance</i>	
Dominance	Proportion of participants who produced a feature in response to a concept (Montefinese et al., 2013)
Accessibility	Readiness with which a feature is produced in response to a concept (Montefinese et al., 2013)
Relevance	Composite measure reflecting the importance of a feature both for a given concept (Dominance) and for all the concepts (Distinctiveness; Sartori & Lombardi, 2004; Montefinese et al., 2013)
Salience	Composite measure reflecting the importance of a feature both for a given concept (Accessibility) and for all the concepts (Distinctiveness; Montefinese et al., 2013)

To manipulate the control demands, we parametrically varied the semantic salience (i.e., the strength of the semantic relationship) between the cue concept and the target feature (see Figure 1 for an example): as the value of salience decreases, the control demands increase. The semantic salience is a composite measure derived from semantic relevance (Sartori & Lombardi, 2004) and reflects the degree to which a given feature facilitates both the identification of a concept and its discrimination from other, similar concepts. The semantic salience is indeed computed for each true concept-feature pair as the product of two components: a local component reflecting the importance of that feature for that concept (for the concept identification), consisting in its accessibility value multiplied by 10, and a global component reflecting the importance of that feature for all the concepts in the norms (for the concept discrimination), consisting in the natural logarithm of its distinctiveness (see Montefinese et al., 2013, for more details).

A further set of 12 concept–feature pairs was selected for the training session, equally divided in true and false items. Concept and feature names were only presented once across the practice, filler, and experimental trials. The selection of the total stimulus set aimed to cover, as much as possible, the range of values of the above-mentioned featural variables while avoiding the choice of similar features both within and across concepts.

Procedure

The procedure used in the current study builds on the method described in a previous study (Montefinese et al., 2014). Differently from the previous study, participants were tested online, since the lockdown due to the spread of COVID-19 did not allow us to test participants in the laboratory. However, it is important to note that before performing the final analyses to assess the age-related differences of interest, we verified that the performance of the young groups of the present and previous studies was not statistically different (see Statistical Analysis and Results sections). Data was collected using the web-based version of PsyToolkit (Stoet, 2017).

All participants performed a speeded feature verification task. In this task the name of a concept (e.g., CAR) was presented, followed by a target-featural description (e.g., “has four wheels”). Participants were asked to indicate as quickly and as accurately as possible whether the feature description (the target) was reasonably true or not for the concept, by pressing the “A” or “L” key on the keyboard (counterbalanced across

participants) using their index fingers. For example, the correct response would be “yes” to the concept-feature pair CAR – “has four wheels” and “no” to the concept-feature pair FIR – “has legs”. A “reasonably true” rather than an “always true” criterion was necessary because in an “always true” decision task, for example, SKIRT – “is short” and KNIFE – “is sharp” would require a “no” response (McRae et al., 1997). The instructions about the study and the key response mapping were presented on the computer screen before and after the training phase.

Each trial began with a prime concept presented for 500 ms in the center of the screen, followed by a blank screen with a fixation cross (the interstimulus interval) for 500 ms. After this delay, the target feature appeared at the center of the screen and remained there until the participants responded or, if they failed to respond, for 4,000 ms. In total, the trial lasted 5,000 ms. The intertrial interval was 1,500 ms (see Figure 1). Concepts were presented in blue and features in black, both written in 24-point Verdana, in lowercase fonts and on a white background.

The lists were randomly assigned to participants. The order of presentation of the concept–feature pairs within a particular list was randomized as well. Participants completed 12 practice trials and 120 test trials (i.e., a stimuli list, as described above). Experiment lasted approximately 10 min. Response times (RTs) were recorded as the time interval between the onset of the feature description and the participant’s response.

Figure 1. Experimental timeline, including both a “true” and a “false” trial. Participants had to decide if the target feature (in black), which may take different degrees of significance for the probe concept (in blue), was true for the concept. In this first example, when the concept ORECCHIO (EAR) was followed by the feature *sente* (*hears*, with high significance, in this case), participants had to respond “true”. In the second example, the concept ABETE (FIR) was followed by the feature (*has legs*), participants had to respond “false”.

Statistical analysis

All test trials related to incorrect responses (about 9.27% of the trials), were discarded from the analyses on RTs. In order to obtain a detailed picture of the relative influence of a number of lexical and semantic predictor variables on our semantic verification task, the effect of the conceptual and featural indexes described above was assessed by conducting linear mixed effects modeling using the open software R (version 4.0.3, R Development Core Team, 2012). Continuous predictors were centred and scaled to facilitate model convergence.

In order to identify the best model explaining participants' RTs while mitigating the high collinearity between the possible predictors (see Section 1 at <https://osf.io/j3hds> for the intercorrelations between our predictors), we performed a hierarchical model-building procedure using a log-likelihood ratio test to compare progressively more complex models with simpler models. This model-building procedure was performed on the data from the sample collected for the present study. We first built the random part of the model and then iteratively tested each of the variables composing each of the six categories of intercorrelated predictors revealed by a hierarchical cluster analysis (see Table 1 and Materials section; for a detailed description of this approach, see Montefinese et al., 2014) to select the one to be added to the model that better explained the participants' RTs. Specifically, we sequentially tested for the potential confounding effect of 1) the only feature-level lexical variable; 2) the two concept-level lexical variables; 3) the two concept-level semantic variables reflecting the richness of the concept representation; 4) the five concept-level semantic variables reflecting the importance of the concept for the representation of the corresponding semantic category; 5) the four feature-level semantic variables reflecting the informativeness of the feature; and 6) the four feature-level semantic variables reflecting the importance of the feature for the representation of the corresponding concept, which were those of primary theoretical interest and included the semantic salience. We followed the same analytical approach as in our previous study to assess the replicability of our results and their generalizability to older participants. Finally, in order to verify our hypothesis, we tested for the inclusion of the main effect of the group factor and its interaction with the salience measure. We also included a term for the effect of Education level, which was not balanced in younger

and older samples, to verify whether our results could have been biased by the observed difference in years of education between Older and Younger adults.

We assessed statistical significance of the predictors by means of t tests with Satterthwaite's approximation to degrees of freedom. The alpha level was set at .05. We reported the estimated coefficient (b), standard error (SE), t and p values for each parameter included in the final model³.

At this point, we took additional steps to verify the robustness of our results and improve the accuracy and precision of the estimate for the hypothesised age-dependent difference in the salience effect. First, we added the data from the young sample collected in our previous study (Montefinese et al., 2014) after having performed control analyses to test for possible differences in their semantic performance. To this aim, we directly compared their performance and salience effects and also ran a linear mixed model including a 3-level group factor (Older vs Present Younger vs Previous Younger).

Next, the quantile-quantile plot for the residuals was inspected for evidence of stress in the model fit. We then refitted the final model after removing data points with absolute standardized residuals exceeding 3. Given that there were severe departure from normality of the residuals and violations of the homoscedasticity assumption (see Results), we performed a distributional analysis to identify the best distribution to appropriately model RTs and we used it in a generalized linear mixed model analyses to avoid the need for RTs transformation (Lo & Andrews, 2015).

Finally, we ran a control linear mixed model analysis in which we also included a parameter for the correlation between the by-subject random slope for the Salience effect (i.e., the differential Salience effect) and the by-subject random intercept (i.e., the mean RTs). We thus inspected the variance/covariance component parameter of these competing linear mixed-effects models and assessed the significance of the correlation between the participants' intercepts and slopes by performing a likelihood ratio test to contrast the linear mixed-effects model with one that assumed this correlation to be zero (Kliegl et al., 2010). This allowed us to verify whether our results of main interest could

³ Whenever possible, the corresponding effect sizes were estimated as standardized differences based on the formula proposed by Westfall and colleagues (2014, dT) and also based on the computed t value and degrees of freedom ($d = t/\sqrt{\text{dof}}$). Note that these procedures can not be easily applied to generalized linear mixed-effects models like the one we used and, to the best of our knowledge, accurately computing standardized effect size for these models is far from trivial.

be simply due to the observed aging-related difference in general performance, as predicted by the general cognitive slowing accounts (e.g., Salthouse, 1985). Indeed, these theories assume that the experimental effects expressed in terms of RTs are proportional to the participants' general performance or, in other words, that there is an across-subjects correlation between the differential effects for RTs and the mean RTs.

We also performed a generalized linear mixed effects analysis on participants' accuracy with the same predictors included in the final analysis of RTs, but now selecting the binomial distribution and the logistic link function. Note that in this case we provided b , SE , z and p values for each parameter.

Results

The mean (SD) RTs and accuracy for the present sample of younger participants were, respectively, 1047.81 ms (311.83 ms) and 93.03% (4.13%). The performance of the current sample of younger participants was thus similar to that observed in our previous work (Montefinese et al., 2014) (respectively, 992.90 ms (209.26 ms) and 90.29% (5.17%)), suggesting that the online data collection did not compromise the quality of the data obtained with our experimental paradigm. The descriptive statistics for the older participants suggest the expected general worsening in performance, especially for the RTs ($M = 1805.28$ ms, $SD = 566.72$ ms) as compared to the accuracy ($M = 88.89\%$, $SD = 7.62\%$). We thus checked for the presence of a speed-accuracy trade off by correlating participants' mean RTs and accuracy. This analysis revealed that participants with faster RTs also tended to have better accuracy ($r = -.545$, $t(118) = -7.06$, $p < .0001$), indicating the absence of a speed-accuracy trade off in our sample. Interestingly, the same pattern was observed when performing the same analysis on both the samples of older ($r = -.568$, $t(58) = -5.26$, $p < .0001$) and younger participants separately, even if in this case the negative correlation was lower and non-significant ($r = -.123$, $t(58) = -0.94$, $p = .3495$).

For the sake of clarity, here we only report the results of the generalized linear mixed model analyses for the participants' RTs. The detailed results of the intermediate steps of our analytical approach and the related control analyses (see Statistical analysis section) are reported in the online supplementary materials at <https://osf.io/j3hds>.

The model building procedure (see code at <https://osf.io/9s4kv>) resulted in a model that included, in the random part, the parameters for both the by-features and by-

subjects random intercepts, as well as the uncorrelated by-subjects random slopes for both the Trial and Salience effects, which varied across participants as expected. In the fixed part, the model included the parameters for the effects of Intercept, Trial order, Salience, Feature Length, Concept Frequency of use, Correlation Percentage, Education (as continuous predictors), Group (Younger vs. Older participants), and the Group by Salience interaction (see Section 3 at <https://osf.io/j3hds>). After having verified that there were no significant differences in the semantic performance and Salience effects between the groups of younger adults in our present and previous study (see Section 4 at <https://osf.io/j3hds>), we combined them for subsequent analyses to improve the accuracy and precision of the estimate for the age-dependent difference in the salience effect (see Section 5 at <https://osf.io/j3hds>). Based on the results of a distributional analysis (see Section 6 at <https://osf.io/j3hds>), we then performed a generalized mixed-effects model using the inverse Gaussian distribution family. The statistics for the trimmed version of this model are shown in Table 2. The main effect of Trial was significant, revealing that participants' RTs decreased during the experiment. Moreover, RTs decreased for target features that were more intercorrelated with the other features of that concept and for concept with higher word frequency, and increased for longer target features. As expected, the main effect of Group was highly significant, showing the strong unspecific aging-related effect on participants' RTs. The analysis also revealed that participants' RTs linearly decreased as the semantic salience of the target feature increased⁴. Finally, the Group by Salience interaction further qualified this result, showing that the effect of Salience was stronger in older participants (see Figure 2). Indeed, there was a significant effect of Salience for Younger participants ($b = -72.29$, $SE = 4.33$, $t = -16.70$, $p < .001$), with lower RTs for trials with higher Salience, and an even larger effect of Salience for Older adults ($b = -36.82$, $SE = 14.32$, $t = -2.57$, $p = .010$)⁵, corresponding to approximately a 50% increase in the raw effect of Salience on RTs for Older as compared to Younger participants. Note also that the inclusion of the Group by Salience interaction led to an increase of the R^2 value of .108 (with a total R^2 of .389), which corresponds to

⁴ We examined the scatterplots for the reported effects to exclude the presence of non-linear effects (see Section 2 at <https://osf.io/j3hds>).

⁵ Note that this result was confirmed by the same analysis conducted on the data from the present sample only (see Section 8 at <https://osf.io/j3hds>)

a relative increase of about 47% with respect to the R^2 of .232 for the model without this interaction (see <https://osf.io/9s4kv>).

Table 2. Fixed effects parameter estimates and statistics of the generalized linear mixed-effects analysis of response times using the inverse Gaussian distribution.

Effects	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Intercept	1046.71	5.388	194.273	< .001
Trial	-24.93	4.103	-6.075	< .001
FeatLen	45.25	4.218	10.728	< .001
ConceptFr	-9.93	4.415	-2.248	.025
CorrPerc	-13.90	4.245	-3.275	.001
Education	-84.74	5.798	-14.616	< .001
Salience	-72.30	4.329	-16.701	< .001
GROUP _{Older}	641.03	16.692	38.403	< .001
Salience*GROUP _{Older}	-36.82	14.321	-2.571	.010

Notes: FeatLen, feature length; ConceptFr, concept frequency of use; CorrPerc, correlation percentage. The reported statistics are for the trimmed model (see Results for details).

Figure 2. Response time results, group by significance interaction. The plot shows the participants' response times (y axis) predicted by the generalized linear mixed-effect model as a function of the (scaled) target feature significance (x axis) and the group (Younger vs. Older participants, in pink and blue respectively). Shaded regions indicate the 95% confidence interval.

The quantile-quantile plot for the residuals and the residuals vs. predicted plot indicated no severe violations of the normality and homoscedasticity assumptions after having removed outliers with extreme standardized residuals (0.55% of the data, see Section 6 at <https://osf.io/j3hds>).

Moreover, the inspection of the variance/covariance component parameters revealed that the estimated correlation between participants' mean RTs and Saliency effect was very close to zero ($r < .01$), as also indicated by the likelihood ratio test contrasting this model with a model with correlated by-subjects random intercepts and slopes ($\chi^2(1) < 1, p > .99$), thus showing that our interaction of interest could not be simply explained by Older participants' general slowing.

Finally, the analysis on participants' accuracy confirmed the statistical significance of the Group by Saliency interaction (see Section 9 at <https://osf.io/j3hds>), further indicating that our results were not due to a speed-accuracy trade off. However, it

should be noted that this model did not converge properly and severely violated the statistical assumptions, so this result should be taken with extreme caution.

Discussion

In this study, we investigated the effect of age on the controlled retrieval of semantic information by comparing the performance of older and young healthy individuals in a feature verification task in which control demands were manipulated by linearly varying the salience of the feature-concept semantic relation (i.e., a measure of the strength of the relationship between a feature and the corresponding concept), differently from previous studies on the topic (Alves et al., 2021; Hoffman, 2018, 2019). This experimental manipulation, together with the use of a linear mixed-effects model analysis, allowed us to assess in a fine-grained way the demands of controlled retrieval: as semantic salience decreased (i.e., the importance of a feature for a concept representation decreased), control demands on the retrieval of semantic information increased. In other words, trials with lower feature-concept salience were those in which additional control resources would be necessary in order to retrieve semantic information, while trials with higher feature-concept salience were those which required minimal controlled retrieval. This is the first time in which semantic salience has been used to manipulate controlled retrieval demands in older individuals. Importantly, our analyses suggested that these results were not simply due to the age-related general cognitive slowing (e.g., Salthouse, 1985) (see Statistical Analysis and Results section).

Firstly, consistent with a previous study (Montefinese et al., 2014) we found that feature length was a significant predictor of participants' response latencies, because the speeded feature verification task we used requires, first of all, that the featural description is read and understood before verifying it as a true characteristic of the corresponding concept. These basic cognitive processes, such as reading, are related to the number of letters that compose the target description: it takes more time to read and process a long word or sentence. Therefore, the influence of feature length on verification performance would not concern the verification processes per se, but rather the simple, and necessary, reading processes.

Moreover, in line with previous findings (McRae et al., 1997, 1999; Montefinese et al., 2014; Montefinese, Hallam, Thompson, et al., 2020), the correlation percentage of target features also significantly predicted response times, with faster reaction times as

the correlation percentage increased. In line with the feature-based distributed view of semantic representation (McRae et al., 1997, 1999; Montefinese et al., 2014, 2015), correlated features are activated more quickly and automatically in speeded online tasks with healthy participants. This finding confirms the important role played by correlated features in the organization of semantic representation. Given these results, correlation percentage is expected to generally benefit performance (e.g., by supporting the automatic spreading activation between different semantic features within semantic representation).

Importantly, and as hypothesised, we found an age-related effect on the participants' controlled semantic retrieval. In particular, the results revealed that participants' RTs linearly increased as the semantic salience of the feature-concept relation decreased, but this effect was stronger for the group of older adults compared to the younger one. Thus, older participants had a greater difficulty to efficiently retrieve features of concepts that were less salient to the conceptual identity, conditions which require high levels of controlled processing as proposed by the controlled semantic cognition framework (Lambon Ralph et al., 2017). As far as we know, only just a few studies have investigated the age-related effect on semantic control and found mixed evidence: while it seems that the selection process declines (Hoffman, 2018, 2019), the very few studies on the controlled retrieval in aging agree for its sparing in older adults (Alves et al., 2021; Hoffman, 2018). In particular, these latter studies observed that older participants had a similar performance to their young counterparts, suggesting a preserved controlled retrieval of weaker information from the semantic store (Alves et al., 2021; Hoffman, 2018, 2019). Our results speak against previous evidence on controlled retrieval, suggesting that the controlled retrieval follows the same fate of the semantic selection process in aging. This inconsistency may be due to the methodological differences between our and previous studies. First, as we noted in the Introduction, previous studies investigated the age-related differences in the controlled retrieval of concept-concept associations (Hoffman, 2018), synonyms (Hoffman, 2019), or concept-category relationship (Alves et al., 2021), while we investigated the age-related differences in the controlled retrieval of concept features; this type of semantic information could thus be more sensitive to manipulations of control demands and age-related changes in control abilities. Moreover, related to this point, unlike Hoffman and

colleagues' studies, in which the semantic control demands were operationalized using a categorical manipulation (2018, 2019), here we used a continuous measure of control demands as the strength of the relationship between a feature and the corresponding concept, which has been proven to be able to modulate semantic performance in a fine-grained way (Montefinese et al., 2014; see also CIT and CIT). Finally, previous studies presented the probe stimulus together with the target stimulus (Alves et al., 2021) and also the distractor stimuli (Hoffman, 2018, 2019) in the same screen; this could have biased the measure of participants' performance (e.g., because participants had to scan the screen to read the different stimuli), making it harder to detect the experimental effect of interest. Related to this point, it is important to note that the age-related difference in controlled retrieval that we found in the present study, albeit robust, had a small estimated effect size ($dT = -0.194$).

The controlled retrieval has been closely associated with the anterior ventral prefrontal cortex, which is thought to provide top-down constraint over the information activation within the semantic store in the temporal cortices (Badre et al., 2005). Neuroimaging studies observed a less involvement of brain areas implicated in the controlled retrieval (e.g., the ventral prefrontal cortex) in older participants, suggesting a difficulty in the ability to engage controlled retrieval in the older participants (Giovanello & Schacter, 2012; McDonough et al., 2013; Trelle et al., 2019). Our result is also consistent with further neural changes in later life. For example, the inferior prefrontal cortex showed large declines in volume (Fjell et al., 2009; Raz et al., 1997, 2005), and changes in functional activation in later life (Cabeza, 2002; Grady, 2012). A recent neuroimaging meta-analysis indicated that older adults displayed a reduced activation in core areas of the semantic control network, including the left inferior prefrontal cortex (Hoffman & Morcom, 2018). Controlled retrieval deficits were also found in young participants when a virtual lesion was applied to their inferior frontal lobe by an inhibitory transcranial magnetic stimulation (TMS) to simulate the semantic control deficit of SA patients (Krieger-Redwood & Jefferies, 2014; Whitney et al., 2011, 2012).

Generally speaking, the pattern of results in this study is similar to that observed in SA patients in controlled retrieval tasks, who present deficits in retrieving distant relationships between concepts and less relevant meaning dimensions even in the context of an intact semantic store (Jefferies et al., 2008, 2009; Montefinese, Hallam,

Stampacchia, et al., 2020). A recent study on semantic control in aging showed that older adults exhibited a worse performance in semantic control compared to that of young adults. Interestingly, while younger adults' performance was less accurate on the trials which patients with semantic dementia (i.e., with a degradation of the semantic representation) found more difficult, older adults' performance was worse on the trials in which SA patients were least accurate. The results suggested that, similarly to SA patients, older individuals have difficulties to exercise control over their unimpaired semantic representation (Hoffman, 2019). The participants' performance in our study was also mediated by the education level. In the literature, the education level is often used as a proxy for cognitive reserve (the knowledge and experience an individual acquires throughout the lifespan) (Stern, 2009), which may be a protective factor to age-related or pathological cerebral deterioration and is strongly correlated to an individual's semantic knowledge (Darby et al., 2017; Montemurro et al., 2020). For this reason, it is not surprising that this factor affected the semantic processing performance of the participants of the present study (with a better general performance for participants with a higher education level).

Still, it may be that our results are exacerbated by a difference in the structure of semantic representation between younger and older participants. As an example, we do not know if older and younger adults displayed differences in the automatic spreading of activation of information within semantic representation and contradictory results have been also observed in the literature (see Laver, 2000, 2009; Laver & Burke, 1993; Spaniol et al., 2006). Graph-based approaches showed that although network size appears to grow continuously across the life span (in terms of vocabulary size), older participants' semantic networks are less connected, organized, and efficient (Dubossarsky et al., 2017; Wulff et al., 2018, 2019; Zortea et al., 2014). However, this age-dependent weakening of semantic connections could explain our results only if it is assumed to affect weaker semantic connections more than stronger ones, an aspect that is currently unclear. Otherwise, we should not have found a significant interaction with the salience. Importantly, so far, there seems to be no clear consensus concerning the extent to which representation and semantic control processes are entangled in this type of task (Wulff et al., 2022). It is likely that both bottom-up, automatic spreading activation and top-down, controlled retrieval processes can be both engaged depending not only on the efficacy of

the available bottom-up cues, but also on the context and task demands. Nevertheless, a single-paradigm approach cannot provide a conclusive solution to this conundrum. One potential way to distinguish the role of structure versus control processes could be to use different assessment methods of the retrieval of semantic information. Further studies are clearly needed to clarify this issue.

Conclusion

The major contribution of this study is to indicate that the controlled retrieval of information can vary across age groups. In particular, our results show that older adults have worse performance in retrieving meaningful relations between words when conditions require high demands of semantic control, suggesting that conditions requiring more controlled retrieval are an area of particular weakness in later life. These results are in line with previous neuroimaging literature on the topic showing changes in the structure and function of the brain areas implicated in the controlled retrieval (such as, for example, the ventral prefrontal cortex) in older participants, determining a difficulty in the ability to engage controlled retrieval in the older participants. Of course, future research would be needed to link more directly behavioural results and neural observations to delineate better how age-related cerebral modifications impact the retrieval process of information. Understanding more about this aspect is essential, given that this can have significant implications in real life situations. For instance, being less efficient in the retrieval of weakly-activated (but context relevant) information may impact the efficiency and the confidence of elderly in daily life communications and social interactions. This may influence their decision making in real life, when they will utilize their knowledge to guide their day to day behaviour. The results of this study might have implications for designing a training program that promotes an efficient control of semantic knowledge in older adults. Indeed, as we mentioned before, although semantic representation remains stable or even improves in aging, this ability alone is not sufficient to allow for an efficient retrieval and an appropriate use of the acquired knowledge in older adults. We strongly believe the results of this study can move us closer to understanding how the controlled retrieval of information works in later life and its implications for cognitive aging.

References

- Alves, M., Figueiredo, P., Roberto, M. S., & Raposo, A. (2021). Using concept typicality to explore semantic representation and control in healthy ageing. *Cognitive Processing*, 22(3), 539–552. <https://doi.org/10.1007/s10339-021-01024-7>
- Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable Controlled Retrieval and Generalized Selection Mechanisms in Ventrolateral Prefrontal Cortex. *Neuron*, 47(6), 907–918. <https://doi.org/10.1016/j.neuron.2005.07.023>
- Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., & Thornton, A. M. (2005). *Corpus e Lessico di Frequenza Dell'italiano Scritto (CoLFIS)* [Corpus and frequency lexicon for written Italian].
- Bourne Jr., L. E., & Restle, F. (1959). Mathematical theory of concept identification. *Psychological Review*, 66(5), 278–296. <https://doi.org/10.1037/h0041365>
- Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: The HAROLD model. *Psychology and aging*, 17, 85–100. <https://doi.org/10.1037/0882-7974.17.1.85>
- Corbett, F., Jefferies, E., & Lambon Ralph, M. A. (2009). Exploring multimodal semantic control impairments in semantic aphasia: Evidence from naturalistic object use. *Neuropsychologia*, 47(13), 2721–2731. <https://doi.org/10.1016/j.neuropsychologia.2009.05.020>
- Craik, F. I. (1977). Age differences in human memory. *Handbook of the psychology of aging*.
- Cree, G. S., McNorgan, C., & McRae, K. (2006). Distinctive features hold a privileged status in the computation of word meaning: Implications for theories of semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 643–658. <https://doi.org/10.1037/0278-7393.32.4.643>
- Darby, R. R., Brickhouse, M., Wolk, D. A., & Dickerson, B. C. (2017). Effects of cognitive reserve depend on executive and semantic demands of the task. *Journal of Neurology, Neurosurgery & Psychiatry*, 88(9), 794–802. <https://doi.org/10.1136/jnnp-2017-315719>

- Darowski, E. S., Helder, E., Zacks, R. T., Hasher, L., & Hambrick, D. Z. (2008). Age-related differences in cognition: The role of distraction control. *Neuropsychology*, 22(5), 638–644. <https://doi.org/10.1037/0894-4105.22.5.638>
- Davey, J., Cornelissen, P. L., Thompson, H. E., Sonkusare, S., Hallam, G., Smallwood, J., & Jefferies, E. (2015). Automatic and Controlled Semantic Retrieval: TMS Reveals Distinct Contributions of Posterior Middle Temporal Gyrus and Angular Gyrus. *Journal of Neuroscience*, 35(46), 15230–15239. <https://doi.org/10.1523/JNEUROSCI.4705-14.2015>
- Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., Krieger-Redwood, K., Bernhardt, B. C., Smallwood, J., & Jefferies, E. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *NeuroImage*, 137, 165–177. <https://doi.org/10.1016/j.neuroimage.2016.05.051>
- Dubossarsky, H., De Deyne, S., & Hills, T. T. (2017). Quantifying the structure of free association networks across the life span. *Developmental Psychology*, 53(8), 1560–1570. <https://doi.org/10.1037/dev0000347>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
- Fjell, A. M., Westlye, L. T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., Agartz, I., Salat, D. H., Greve, D. N., & Fischl, B. (2009). High consistency of regional cortical thinning in aging across multiple samples. *Cerebral cortex*, 19(9), 2001–2012.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). «Mini-mental state». A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)
- Giovanello, K. S., & Schacter, D. L. (2012). Reduced Specificity of Hippocampal and Posterior Ventrolateral Prefrontal Activity during Relational Retrieval in Normal Aging. *Journal of Cognitive Neuroscience*, 24(1), 159–170. https://doi.org/10.1162/jocn_a_00113

- Grady, C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience*, 13(7), 491–505.
- Hasher, L., Quig, M. B., & May, C. P. (1997). Inhibitory control over no-longer-relevant information: Adult age differences. *Memory & Cognition*, 25(3), 286–295. <https://doi.org/10.3758/bf03211284>
- Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(1), 163–169. <https://doi.org/10.1037/0278-7393.17.1.163>
- Hoffman, P. (2018). An individual differences approach to semantic cognition: Divergent effects of age on representation, retrieval and selection. *Scientific Reports*, 8(1), 8145. <https://doi.org/10.1038/s41598-018-26569-0>
- Hoffman, P. (2019). Divergent effects of healthy ageing on semantic knowledge and control: Evidence from novel comparisons with semantically impaired patients. *Journal of Neuropsychology*, 13(3), 462–484. <https://doi.org/10.1111/jnp.12159>
- Hoffman, P., & Morcom, A. M. (2018). Age-related changes in the neural networks supporting semantic cognition: A meta-analysis of 47 functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, 84, 134–150. <https://doi.org/10.1016/j.neubiorev.2017.11.010>
- Izura, C., Hernández-Muñoz, N., & Ellis, A. W. (2005). Category norms for 500 Spanish words in five semantic categories. *Behavior Research Methods*, 37(3), 385–397. <https://doi.org/10.3758/BF03192708>
- Jefferies, E., Hoffman, P., Jones, R., & Lambon Ralph, M. A. (2008). The impact of semantic impairment on verbal short-term memory in stroke aphasia and semantic dementia: A comparative study. *Journal of Memory and Language*, 58(1), 66–87. <https://doi.org/10.1016/j.jml.2007.06.004>
- Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: A case-series comparison. *Brain*, 129(8), 2132–2147. <https://doi.org/10.1093/brain/awl153>
- Jefferies, E., Patterson, K., Jones, R. W., & Lambon Ralph, M. A. (2009). Comprehension of concrete and abstract words in semantic dementia. *Neuropsychology*, 23(4), 492–499. <https://doi.org/10.1037/a0015452>

- Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition*, 18(5), 655–681. <https://doi.org/10.1080/13506280902986058>
- Kramer, A. F., Humphrey, D. G., Larish, J. F., & Logan, G. D. (1994). Aging and inhibition: Beyond a unitary view of inhibitory processing in attention. *Psychology and Aging*, 9(4), 491–512. <https://doi.org/10.1037/0882-7974.9.4.491>
- Krieger-Redwood, K., & Jefferies, E. (2014). TMS interferes with lexical-semantic retrieval in left inferior frontal gyrus and posterior middle temporal gyrus: Evidence from cyclical picture naming. *Neuropsychologia*, 64, 24–32. <https://doi.org/10.1016/j.neuropsychologia.2014.09.014>
- Laver, G. D. (2000). A speed-accuracy analysis of word recognition in young and older adults. *Psychology and Aging*, 15(4), 705–709. <https://doi.org/10.1037/0882-7974.15.4.705>
- Laver, G. D. (2009). Adult aging effects on semantic and episodic priming in word recognition. *Psychology and Aging*, 24(1), 28–39. <https://doi.org/10.1037/a0014642>
- Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychology and Aging*, 8(1), 34–43. <https://doi.org/10.1037/0882-7974.8.1.34>
- Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, 6, 1171. <https://doi.org/10.3389/fpsyg.2015.01171>
- McDonough, I. M., Wong, J. T., & Gallo, D. A. (2013). Age-Related Differences in Prefrontal Cortex Activity during Retrieval Monitoring: Testing the Compensation and Dysfunction Accounts. *Cerebral Cortex*, 23(5), 1049–1060. <https://doi.org/10.1093/cercor/bhs064>
- McDowd, J. M., & Craik, F. I. (1988). Effects of aging and task difficulty on divided attention performance. *Journal of experimental psychology: human perception and performance*, 14(2), 267.
- McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, 37(4), 547–559. <https://doi.org/10.3758/BF03192726>

- McRae, K., Cree, G. S., Westmacott, R., & Sa, V. R. D. (1999). Further evidence for feature correlations in semantic memory. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 53(4), 360–373. <https://doi.org/10.1037/h0087323>
- McRae, K., deSa, V., & Seidenberg, M. (1997). On the nature and scope of featural representation of word meaning. *Journal of Experimental Psychology General*, 126, 99–130. <https://doi.org/10.1037/0096-3445.126.2.99>
- Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2013). Semantic memory: A feature-based analysis and new norms for Italian. *Behavior Research Methods*, 45(2), 440–461. <https://doi.org/10.3758/s13428-012-0263-4>
- Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). Semantic significance: A new measure of feature salience. *Memory & Cognition*, 42(3), 355–369. <https://doi.org/10.3758/s13421-013-0365-y>
- Montefinese, M., Hallam, G., Stampacchia, S., Thompson, H. E., & Jefferies, E. (2020). Deficits of semantic control disproportionately affect low-relevance conceptual features: Evidence from semantic aphasia. *Aphasiology*, 0(0), 1–15. <https://doi.org/10.1080/02687038.2020.1814950>
- Montefinese, M., Hallam, G., Thompson, H. E., & Jefferies, E. (2020). The interplay between control processes and feature relevance: Evidence from dual-task methodology. *Quarterly Journal of Experimental Psychology*, 73(3), 384–395. <https://doi.org/10.1177/1747021819877163>
- Montefinese, M., Zannino, G. D., & Ambrosini, E. (2015). Semantic similarity between old and new items produces false alarms in recognition memory. *Psychological Research*, 79(5), 785–794. <https://doi.org/10.1007/s00426-014-0615-z>
- Montemurro, S., Montefinese, M., Serena, M., Pucci, V., Mondini, S., & Semenza, C. (2020). *The interactive support of Cognitive Reserve and Semantic Knowledge in Proper Name retrieval*. OSF Preprints. <https://doi.org/10.31219/osf.io/ak2mf>
- Naveh-Benjamin, M., Craik, F. I., Guez, J., & Kreuger, S. (2005). Divided attention in younger and older adults: Effects of strategy and relatedness on memory performance and secondary task costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 520.

- Noonan, K. A., Jefferies, E., Corbett, F., & Lambon Ralph, M. A. (2010). Elucidating the Nature of Deregulated Semantic Cognition in Semantic Aphasia: Evidence for the Roles of Prefrontal and Temporo-parietal Cortices. *Journal of Cognitive Neuroscience*, 22(7), 1597–1613. <https://doi.org/10.1162/jocn.2009.21289>
- Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, 23(1), 104–118. <https://doi.org/10.1037/0882-7974.23.1.104>
- Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review*, 15(1), 161–167. <https://doi.org/10.3758/PBR.15.1.161>
- R Development Core Team. (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <https://www.R-project.org/>
- Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42–55. <https://doi.org/10.1038/nrn.2016.150>
- Raz, N., Gunning, F. M., Head, D., Dupuis, J. H., McQuain, J., Briggs, S. D., Loken, W. J., Thornton, A. E., & Acker, J. D. (1997). Selective aging of the human cerebral cortex observed in vivo: Differential vulnerability of the prefrontal gray matter. *Cerebral Cortex*, 7(3), 268–282. <https://doi.org/10.1093/cercor/7.3.268>
- Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., Dahle, C., Gerstorf, D., & Acker, J. D. (2005). Regional Brain Changes in Aging Healthy Adults: General Trends, Individual Differences and Modifiers. *Cerebral Cortex*, 15(11), 1676–1689. <https://doi.org/10.1093/cercor/bhi044>
- Reuter-Lorenz, P. A., & Sylvester, C.-Y. C. (2005). The Cognitive Neuroscience of Working Memory and Aging. In *Cognitive neuroscience of aging: Linking cognitive and cerebral aging* (pp. 186–217). Oxford University Press.
- Salthouse, T. A. (1985). Speed of behavior and its implications for cognition. In J. E. Birren & K. W. Schaie, *Handbook of the psychology of aging*, 2nd ed (pp. 400–426). Van Nostrand Reinhold.

- Sartori, G., & Lombardi, L. (2004). Semantic Relevance and Semantic Disorders. *Journal of Cognitive Neuroscience*, 16(3), 439–452. <https://doi.org/10.1162/089892904322926773>
- Somberg, B. L., & Salthouse, T. A. (1982). Divided attention abilities in young and old adults. *Journal of Experimental Psychology: Human Perception and Performance*, 8(5), 651–663. <https://doi.org/10.1037/0096-1523.8.5.651>
- Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(1), 101–117. <https://doi.org/10.1037/0278-7393.32.1.101>
- Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and aging*, 10(4), 527.
- Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, 47(10), 2015–2028. <https://doi.org/10.1016/j.neuropsychologia.2009.03.004>
- Stoet, G. (2017). PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teaching of Psychology*, 44(1), 24–31. <https://doi.org/10.1177/0098628316677643>
- Trelle, A. N., Henson, R. N., & Simons, J. S. (2019). Neural evidence for age-related differences in representational quality and strategic retrieval processes. *Neurobiology of Aging*, 84, 50–60. <https://doi.org/10.1016/j.neurobiolaging.2019.07.012>
- Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and aging*, 13(2), 230.
- Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron*, 31(2), 329–338.
- Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, 143(5), 2020.
- Whitney, C., Grossman, M., & Kircher, T. T. J. (2009). The Influence of Multiple Primes on Bottom-Up and Top-Down Regulation during Meaning Retrieval: Evidence

- for 2 Distinct Neural Networks. *Cerebral Cortex*, 19(11), 2548–2560. <https://doi.org/10.1093/cercor/bhp007>
- Whitney, C., Kirk, M., O’Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2011). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cerebral cortex*, 21(5), 1066–1075.
- Whitney, C., Kirk, M., O’Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2012). Executive Semantic Processing Is Underpinned by a Large-scale Neural Network: Revealing the Contribution of Left Prefrontal, Posterior Temporal, and Parietal Cortex to Controlled Retrieval and Selection Using TMS. *Journal of Cognitive Neuroscience*, 24(1), 133–147. https://doi.org/10.1162/jocn_a_00123
- Wulff, D. U., De Deyne, S., Aeschbach, S., & Mata, R. (2022). Using Network Science to Understand the Aging Lexicon: Linking Individuals’ Experience, Semantic Networks, and Cognitive Performance. *Topics in Cognitive Science*, 14(1), 93–110. <https://doi.org/10.1111/tops.12586>
- Wulff, D. U., De Deyne, S., Jones, M. N., & Mata, R. (2019). New Perspectives on the Aging Lexicon. *Trends in Cognitive Sciences*, 23(8), 686–698. <https://doi.org/10.1016/j.tics.2019.05.003>
- Wulff, D. U., Hills, T., & Mata, R. (2018). *Structural differences in the semantic networks of younger and older adults*. PsyArXiv. <https://doi.org/10.31234/osf.io/s73dp>
- Yap, M. J., Tan, S. E., Pexman, P. M., & Hargreaves, I. S. (2011). Is more always better? Effects of semantic richness on lexical decision, speeded pronunciation, and semantic classification. *Psychonomic Bulletin & Review*, 18(4), 742–750. <https://doi.org/10.3758/s13423-011-0092-y>
- Zortea, M., Menegola, B., Villavicencio, A., & Salles, J. F. de. (2014). Graph analysis of semantic word association among children, adults, and the elderly. *Psicologia: Reflexão e Crítica*, 27, 90–99. <https://doi.org/10.1590/S0102-79722014000100011>