Annals of the New York Academy of Sciences

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

ORIGINAL ARTICLE OPEN ACCESS

Exploring Factors Affecting Verbal Fluency in Healthy Aging

Pascale Tremblay[1,2,3] | Lydia Gagnon[1,2] | Edith Durand[4] | Joël Macoir[1,2]

[1]École des sciences de la Réadaptation, Faculté de Médecine, Université Laval, Quebec City, Canada | [2]CERVO Brain Research Center, Quebec City, Canada | [3]Center for Research on Language, Music and the Brain (CRLMB), Montreal, Canada | [4]Département D'orthophonie, Université du Québec à Trois-Rivières, Trois-Rivières, Canada

Correspondence: Pascale Tremblay (pascale.tremblay@fmed.ulaval.ca)

Funding: This research was supported by Pascale Tremblay's grants from the Fonds de la recherche du Québec–Nature et Technologies [FRQNT; # 2019-PR-254714]; the Social Sciences and Humanities Research Council (#435-2023-0578), and the Canadian Foundation for Innovation [#31408]. Pascale Tremblay holds a Canada Research Chair in the Neurobiology of Speech and Hearing (CRC #2022-00090).

Keywords: aging | cognitive reserve | disfluencies | hearing | lexical access | private speech | singing

ABSTRACT

Word-finding difficulties are a hallmark of aging. Verbal fluency tasks are broadly used to measure word-finding difficulties in adults due to their sensitivity and ease of use. However, several questions remain regarding verbal fluency in aging, particularly the strategies that may evolve over time and the role of potentially protective factors, such as cognitive engagement and social interactions, in mitigating age-related cognitive decline. In this study, we investigated verbal fluency in 144 healthy, community-dwelling adults aged 20–87 years. Participants completed both semantic and phonemic fluency tasks and several questionnaires and assessments. We analyzed accuracy and error types, as well as the occurrence of self-talk and filled hesitations and their impact on lexical access. Further, eight factors previously associated with cognitive reserve were examined: education, practice of group singing, social participation, cognitive level, self-reported health, multilingualism, positive outlook, and hearing. While filled hesitations were stable across age, self-talk increased nonlinearly with age. Singing experience, higher educational attainment, and better global cognition were associated with better fluency. However, these factors showed minimal evidence of protection against age-related decline in word retrieval.

1 | Introduction

Aging is often accompanied by changes in cognitive abilities, including language processing and executive functions [1]. These changes can have a negative impact on independent living and the ability to execute everyday activities, such as scheduling appointments, filling out forms, or taking medication [2, 3]. Understanding the nature of these changes and being able to distinguish normal from abnormal patterns is therefore a priority as the aging population continues to grow [4].

Verbal fluency is one ability that has been shown to change as a function of age [5]. Verbal fluency refers to the ability to generate words rapidly within specific constraints and is a key cognitive function providing insights into lexical access, working memory, and cognitive flexibility [6]. It is typically assessed through two types of tasks: *phonemic fluency* (e.g., generating words beginning with a specific letter, often termed letter fluency) and *semantic fluency* (e.g., naming items within a category, often called category fluency). Semantic fluency is often seen as closely resembling everyday language production activities, such as creating a grocery shopping list, as it leverages pre-existing links between related concepts in semantic memory—both between the category label and its members, and among the category members themselves—for word retrieval. Phonemic fluency requires accessing words based on a phonemic criterion, a process less frequently encountered in everyday speech production, and which does not benefit from the organization of semantic memory. As a result, participants must suppress semantically or associatively related words and rely on less familiar retrieval strategies.

Given the nature of verbal fluency tasks, age-related changes are usually seen as reflecting changes in both lexical retrieval processes and executive functions, such as attention and cognitive control [7, 8]. There is some uncertainty as to whether one type of fluency is more affected than the other by aging. While some have suggested that only semantic fluency is vulnerable to aging, for example, refs. [9, 10], reflecting slower and less accurate lexical retrieval despite well-preserved vocabulary; other have found that semantic fluency is more sensitive to age than phonemic fluency, with an impact of age on both [11–13].

While fluency tests primarily assess lexical access and executive functions, they can also reveal a broad range of language-related phenomena—such as overt self-talk (also referred to as "private" or "self-directed" speech), disfluencies, and error patterns—that offer additional perspectives on language use in aging. *Self-talk* refers to spoken language directed to oneself, often used to guide, monitor, or regulate one's behavior (e.g., "Now think of animals" or "I think I've already said that"). In contrast, *inner speech* is the silent, internalized form of self-directed language—a kind of verbal thinking that cannot be observed but only self-reported. One long-standing idea is that language, particularly in the form of inner speech, supports self-regulation and cognitive control. This concept was first developed in the foundational work of Luriia and Vygotsky [14–16]. More recently, Gade and Paelecke [17] found that certain forms of self-reported inner speech were associated with better performance in cognitive control tasks such as the Simon and arrow flanker tasks. Likewise, Nedergaard and Lupyan [18] reported that adult participants who report less inner speech performed worse when judging whether the names of two images rhymed, and they had poorer verbal working memory. Importantly, in both tasks, the group difference (low vs. high inner speech) disappeared when participants reported talking out loud to solve the problems.

Although inner speech and self-talk regularly occurs during fluency tasks, we could not identify any prior studies that have explicitly examined either phenomenon in this context. Nevertheless, recent findings suggest that audible self-talk may be functionally relevant during verbal tasks. Guo and Dobkins [19] showed that the degree to which young adults use overt self-talk—when instructed to do so—was associated with better performance on a visual-spatial working memory task. This suggests that overt self-talk can support cognitive performance. At the same time, other research suggests that older adults tend to be more verbose and produce more off-topic speech than younger adults, particularly in autobiographical narratives [20, 21] a tendency often linked to age-related declines in inhibitory control. These two strands of evidence raise the question of whether self-talk during fluency tasks serves a supportive function or instead reflects age-related processing difficulties. Whether self-talk during fluency tasks plays a beneficial, compensatory, or maladaptive role—and whether it changes with age—remains unknown.

Another nonreferential language-related behavior that occurs in fluency tasks is filled hesitation, which is a form of disfluency. An age-related increase in disfluencies has been documented in a variety of language tasks including picture descriptions [20, 22] but most often in less constrained tasks involving continuous speech such as narratives and conversations [23, 24].

While the most common measure of verbal fluency is the total number of correct responses, qualitative aspects of lexical retrieval can provide additional insight into the cognitive processes that support performance [25]. Among the most frequently observed error types are intrusions (i.e., words that do not belong to the target category) and repetitions (also referred to as perseveration errors). Some studies have reported an increase in perseveration with age, for example, refs. [26, 27], while others have found that both perseveration and intrusions are relatively rare and not reliably associated with aging [28]. Other types of errors, such as phonological errors or nonwords, are seldom reported in the literature. In sum, although verbal fluency tasks have been widely used to study aging—generally showing lower performance in older adults—little is known about the occurrence of specific behaviors such as self-talk, filled hesitations, and detailed error patterns in this population.

Another unanswered question is whether protective factors, such as cognitive engagement, may significantly mitigate age-related cognitive and linguistic declines. The cognitive reserve hypothesis proposes that environmental factors can predict—to a certain degree and in interaction with a persons' genetic makeup—responses to brain disease, with higher reserve associated with reduced cognitive decline in aging [29–31]. For instance, one study has shown that higher "cognitive reserve," operationalized as a higher score on the WAIS-III Information subtest, a measure of premorbid IQ, was associated with better phonemic fluency in older adults [32]. Consistent with the cognitive reserve hypothesis, the Lancet Commission on dementia has identified 14 modifiable risk factors for dementia, including early life education, hearing loss, depression, social participation and physical inactivity [33–35]. Together, the 14 factors identified in the 2024 version of the report amount to 45% of dementia risk. Other factors not included in the Lancet commission reports include musical activities and bilingualism. Previous work from our group and others suggest that singing is associated with cognitive benefits in working memory [36, 37]. A potential explanation is that singing involves the memorization of verbal content (lyrics), which may enhance working memory capabilities, a hypothesis that is consistent with the OPERA hypothesis [38, 39], which posits that musical training strengthens speech and verbal processing. Regarding bilingualism, while cognitive benefits are debated [40, 41], a longitudinal study showed that bilinguals outperformed monolinguals at the first testing session and across time in phonemic fluency, however, no interaction with age was found indicating that the rate of change across ages was similar for bilinguals and monolinguals [42]. Considering the importance of maintaining autonomy in older adults, understanding the role of potentially protective factors in preserving cognitive functioning and communication remains a critical priority, especially in the context of a rapidly aging global population.

The present study investigates verbal fluency performance in community-dwelling older adults in a comprehensive manner. The first specific objective was to examine, for the first time, the prevalence of nonreferential vocalizations, including self-talk and filled hesitation, during semantic and phonemic fluency tasks and determine whether such behaviors are associated with aging and with better performance (potentially acting as a compensation strategy). The second specific objective was to examine age differences in semantic and phonemic fluency in adults in terms of the nature of the errors that are committed. The third specific objective was to explore the potential moderating impact of eight factors that have been associated with the prevention of cognitive decline or with a higher level of cognitive performance: education, the practice of group singing, social participation, global cognitive level, self-reported health, positive outlook, multilingualism, and hearing on accuracy in semantic and phonemic fluency tasks. Our first hypothesis was that age would have a detrimental impact on lexical access, consistent with the literature. Our second hypothesis was that self-talk would be higher in older adults, and linked to verbal fluency performance positively, consistent with prior literature on other types of verbal and cognitive tasks. Our third hypothesis was that filled hesitation would be more frequent in older adults and negatively associated with performance, reflecting difficulties accessing the mental lexicon. Based on previous literature, we expected that the number of errors would be positively associated with age across categories. Finally, we expected accuracy to be moderated by higher values of eight potentially protective factors: education level, practice of group singing, social participation, cognitive level, self-reported health, positive outlook, multilingualism, and hearing.

2 | Methods

2.1 | Participants

The participants were 144 adults (mean 56 years, SD = 18, range: 20–87 years, 81 females; 63 males) selected from two prior studies from our group: 72 were selected from the PICCOLO project (from the French "Projet de recherche sur les effets de la **P**ratique d'un **I**nstrument ou du **C**hant sur la **CO**gnition, le **L**angage et l'**O**rganisation cérébrale") [36, 43], which is approved by the *Comité d'éthique de la recherche sectoriel en neurosciences et santé mentale, Institut Universitaire en Santé Mentale de Québec* (#2019-1733), and 72 were selected from the Quebec singing dataset (#192-2017) [44]. All participants provided informed consent. All participants from those studies who completed the fluency task were included. People who participated in both projects were included only once, with their first participation being retained. The PICCOLO project includes a group of instrument players who were not included in the sample. Both studies included singers, defined as those engaged in group singing at least once a week for at least 60 consecutive minutes.

In these prior studies, participants were recruited through emails, posters, Facebook, and flyers distributed to Université Laval's community, to the general community, and to choirs and music harmonies in the Quebec City area. The general inclusion criteria were to be right-handed according to the Edinburgh Handedness Inventory [45], speaker of Quebec French; to have normal or corrected-to-normal vision; no self-reported speech, voice, or respiratory disorder; no diagnosed language, hearing or psychological disorder and no neurological or neurodegenerative disorder. A description of the participants is provided in Table 1.

2.2 | Procedures

All participants visited the Speech and Hearing Neuroscience Laboratory in Quebec City, Canada as part of different projects which included verbal fluency tasks as well as a series of questionnaires and tests which allowed us to document eight potentially protective factors. These are detailed below.

2.2.1 | Assessment of Potentially Protective Factors

We assessed eight potentially protective factors: (1) engaging in group singing, (2) higher education level, (3) higher cognitive level, (4) positive outlook, (5) multilingualism, (6) self-reported general health, (7) active social participation, and (8) better hearing. These factors were defined as follows:

1. Group singing. Engagement in choir singing was measured using a custom questionnaire available online on Borealis (https://doi.org/10.5683/SP2/8IX6QZ). The participants were initially divided in two subgroups: nonsingers ($N = 74$) and singers ($N = 70$). Singers were defined as people who sang a minimum of 1 h a week (i.e., 60 consecutive minutes) for at least 2 years. However, because of the diversity of the singers' experiences, we divided the group into occasional singers ($N = 40$) and frequent singers ($N = 30$) based on the total number of hours they reported singing per week, which ranged from 0 to 14 h a week on average. In the absence of a clear empirical cutoff, we applied an arbitrary cutoff that allowed groups to have a relatively similar size. Participants who reported singing less than 5 h per week were classified as occasional singers, while those singing 5 h or more were classified as frequent singers. The occasional and frequent singers differed significantly in terms of how much singing they engaged in weekly (2.2 h per week on average for the occasional singers and 7.67 h per week on average for the frequent singers), combining group sessions and repetitions. They did not differ on age, sex distribution, or any of the other protective factors. Singers' characteristics are detailed in Table S2, and the singing groups are compared in Table S3.
2. Education. For the analyses, a four-level categorical Education variable was computed to reduce collinearity in the model and because some levels were infrequent (e.g., primary and secondary, PhD, medical fellowships): primary/secondary ($N = 16$), college ($N = 49$), undergraduate university degree ($N = 52$) and graduate level education (combining masters, doctorates and medical fellowships; $N = 24$).
3. Cognition. General cognitive functioning was assessed using the Montreal Cognitive Assessment (MoCA) [46].

TABLE 1 | Participants' characteristics (81 females; 63 males).

Characteristics	M	SD	Min	Max	SE	Skew	Kurtosis
Age	56.09	18	20	87	1.5	−0.24	−1
Hearing level[a]	9.64	9.18	−6.7	38.33	0.77	−1.04	0.82
Education (years)[b]	14.92	2.66	6	23	0.22	−0.39	1.99
Self-reported health (/7)[c]	5.19	0.91	3	7	0.08	−0.04	−0.08
Positive outlook (/30)[d]	2.81	2.8	0	15	0.23	−1.86	5.11
Cognition (MoCA; /30)[e]	27.58	1.82	21	30	0.15	−0.62	0.18
Social participation (/4)[f]	2.71	1.03	0	4	0.09	−0.47	−0.34

Note: M, mean; SD, standard deviation of the mean; SE, standard error of the mean. Skewness and kurtosis values assess normality. Most values fell within the commonly accepted range of −2 to +2, indicating that the distributions do not significantly deviate from normality.

[a]Hearing level, measured as the better ear pure tone average (PTA) thresholds at 0.5, 1, and 2 kHz, measured in decibels (dB HL).

[b]Number of years of education, standardized. Elementary = 6; High school = 11; CEGEP (general) = 13; CEGEP (technique) = 14; Undergraduate = 16; Master = 18 (includes medical doctors); PhD = 21; Medical doctors with specialization = 23.

[c]Self-reported health = self-reported physical health status on a scale of 0–7 (0 being lowest physical health level).

[d]Positive outlook = Geriatric Depression Scale (GDS). The GDS includes 30 yes/no questions. The normal score is between 0 and 9, a score between 10 and 19 suggests a mild depression, and a score between 20 and 30 indicates a severe depression. No participant scored above 15 (mild). This was used to create the positive outlook score by inverting the scale.

[e]Cognition was assessed via the Montreal Cognitive Assessment (MoCA). Higher scores indicate better cognitive functions.

[f]Social participation = A measure of social participation level, derived from responses to our custom questionnaire, ranging from 0 (no social participation) to 4 (high level of social participation).

4. Positive outlook. This was measured using the 30-question Geriatric Depression Scale (GDS, /30) [47]. No participant exhibited signs of major depression, but three showed signs of mild depression. The scores were converted such that higher scores represented those with the least symptoms (i.e., those with a more positive outlook).
5. Multilingualism. The self-reported number of languages spoken by each participant was recorded, ranging from 1 to 5. Due to the uneven distribution of participants across levels (1 language = 12; 2 = 87; 3 = 42; 4 = 2; 5 = 1), a binary variable was created. Participants who spoke one or two languages were grouped as monolingual/bilingual (N = 99, M age = 56.30, SD = 17.96), while those who reported speaking three or more languages (i.e., 3–5) were categorized as multilingual (N = 45, M age = 55.62, SD = 18.28). The age distribution was similar between the two groups (t = 0.21, p = 0.84; Figure S1).
6. Self-reported health measured on a scale of 0–7 (0 being lowest physical health level) was assessed using a custom scale.
7. Social participation. Participants answered a series of questions on a custom questionnaire that were used to compute a score of social participation, on a scale of 0 (no social participation) to 4 (maximal social participation). Our calculation for this factor was based on a Japanese study that evaluated the impact of social participation on dementia [48]. In this study, a 5-point scoring system assessed five measures: marital status, exchange with family, contact with friends, participation in community groups, and remunerated work. A person with a score of 5 was found to have 46% less chance to develop dementia than someone with 0 or 1 point. Here, we used a 4-point scoring system based on four measures: living with other people (yes = 1; no = 0), having contacts with family and friends (yes = 1; no = 0), participating in a group activity regularly (yes = 1; no = 0) and having a regular (though not necessarily full-time) occupation (i.e., remunerated workers or full-time students; yes = 1; no = 0).
8. Hearing. Pure tone thresholds in dB HL were measured with a calibrated clinical audiometer (AC40, Interacoustics, Denmark) at the following frequencies: 0.5, 1, and 2 kHz. These measurements were used to compute a better ear (i.e., lowest thresholds between the two ears) pure tone average (PTA). Participants' PTAs are illustrated in Figure 1. These scores were converted (inverted) to examine the potentially

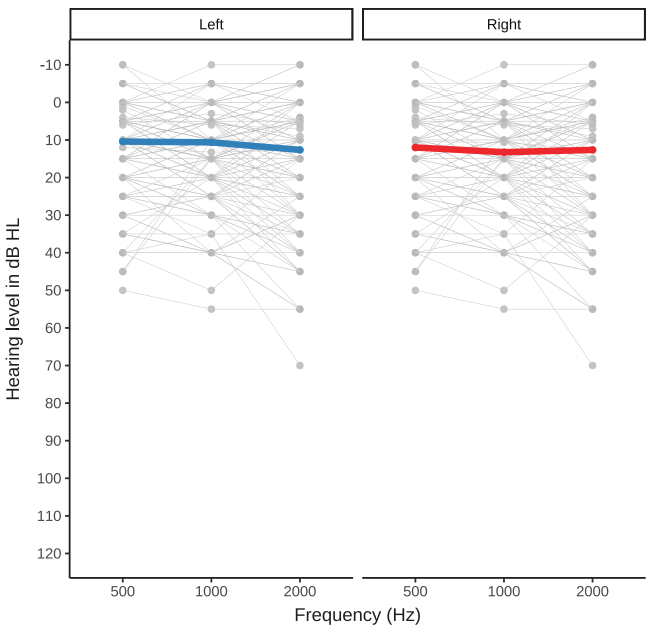

FIGURE 1 | Pure tone averages (PTAs) for each participant (thin gray lines) for the left and right ear. The thick lines represent the average across all participants.

protective effect of better hearing (higher score = better hearing).

2.2.2 | Verbal Fluency

To assess verbal fluency, we used the standard T–N–P tests [49], as well as the animal fluency test, always in that order, without a break in between. For the phonemic fluency, participants were given the following instructions: We are going to do a test where you will need to tell me as many words as possible in 1 min. I will give you a letter of the alphabet. At my signal, you will need to say as many words as possible that start with this letter, as quickly as you can. You cannot use proper nouns (names of people, places, companies, holidays, or planets). Additionally, you cannot give me two words referring to a similar concept by simply changing the ending of the word. For example, if you say "eve," you cannot also say "evening." Similarly, if you say "serious," you cannot also say, "seriousness." For the animal fluency, additional instructions were provided, "The next test you will take will be similar. However, this time, I would like you to name as many animals as you can, regardless of the letter they start with, as quickly as possible, at my signal. You cannot give me two animals by simply changing the ending, like 'Chicken' and 'Chick'. You have 60 s." All tests were recorded for offline analysis. All recordings were performed under identical conditions in a double-walled sound-attenuated room. Participants were seated in a comfortable armchair. Speech samples were recorded using a high-quality head-worn microphone (Microflex Beta 53) connected to a Quartet USB audio interface (Apogee Electronics, Santa Monica, USA) that fed into an iMac computer. The recordings were made using the Sound Studio 4 software (Felt Tip Inc., NYC, USA) at a sampling rate of 48 kHz and 24-bit quantization.

2.3 | Analysis of Fluency Data

A protocol was elaborated by our team at the beginning of the project to transcribe and annotate participants' recorded responses. The recordings were analyzed with the software Praat [50] on an iMac computer. Each recording was segmented, transcribed, and annotated manually. The segmentation was automatically computed using the "To Textgrid (silences)" function in Praat, which generates a TextGrid marking silent and sounding intervals in the selected recording, thereby identifying each utterance, including words, comments, and filled hesitations. Segmentation was manually adjusted by the annotators whenever the software lacked precision (e.g., if a filled hesitation was not segmented because of low voice intensity). Each recording was then transcribed and annotated manually by two different trained team members.

Six different codes were used: correct response, intrusion, repetition, nonwords, filled hesitation, and self-talk. There were too few instances of comments and play to analyze separately, leading to the decision to analyze them together as "self-talk."

The codes are described in Table S1. Two independent raters classified responses according to the codes. Discrepancies between raters were resolved through discussion. For each participant, we calculated the number of correct responses as well as the number of occurrences of self-talk, filled hesitations, intrusions, repetitions, and nonwords over the total number of productions. The full protocol is available in Borealis (https://doi.org/10.5683/SP3/U5ZF7O).

2.4 | Statistical Analyses

All data were analyzed with R Studio V 2023.12.1+ [51]. First, the data were inspected using density plots and by calculating kurtosis and skewness to ensure that the distributions were normally or relatively normally distributed (using the −1 and 1 interval as the cutoff). All variables were converted to z scores. A series of linear mixed model (LMM) analyses and multiple regression analyses was conducted. The models were fitted using the *buildmer* package version 1.9 [52] and the *lme4* package version 1.1.23 [53]. The *buildmer* package starts with the full model specified and determines the order of the fixed and random effects in the model that best explain the variance [54]. The effects are systematically reduced with backward stepwise elimination based on likelihood ratio tests to arrive at a final converging model with the best fit. The results were extracted using the *sjPlot* packages for R [55] for reporting and plotting model results (marginal means). The normality of the residuals of each model was inspected using Q-Q plots. Interactions were probed using the *Interactions* and *JTools* packages for R and the Johnson–Neyman interval approach.

The first set of analyses addresses our first objective, namely, to examine the prevalence of nonreferential vocalizations, including self-talk, during semantic and phonemic fluency tasks and determine whether such behaviors are associated with aging and with better performance. For this analysis, the dependent variable was the z-normalized number correct responses. In a first analysis, we compared the three letter fluency tasks to determine whether all three tasks should be included in the main analysis or whether they could be reduced to one phonemic fluency task. The full model examined the interaction between age, letter (T, N, P), and self-talk and filled hesitations. Sex was included as a covariate and was allowed to interact with Age and Letter. The full model was: Correct responses ∼ Age*Letter*Self-talk + Age* Letter*filled Hesitations + Age*Letter*Sex + (1|Participant). This analysis yielded no main effect of Letter and no interaction involving Letter. The full results are provided as Tables S4 and S5. As a result, we created an average phonemic fluency score across all three letters. The main analysis examined the interaction between age, Condition (phonemic, semantic), and self-talk and filled hesitations. Sex was included as a covariate and was allowed to interact with Age and Condition to capture potential sex differences. The full model was: Correct responses ∼ Age*Condition*Self + Age*Condition*Hesitation + Age*Condition*Sex + (1|Participant).

In addition, to gather more information about potential age differences in nonreferential vocalizations, despite their distributions being zero-inflated, we implemented a two-step generalized additive model (GAM) approach. In the first step, we examined whether age predicted the probability of making each type of

TABLE 2 | Descriptive statistics for raw (nontransformed), nonreferential behaviors, and errors separately for each condition.

	Phonemic						Semantic					
	Mean	**SD**	**Min**	**Max**	**Skew**	**Kurtosis**	**Mean**	**SD**	**Min**	**Max**	**Skew**	**Kurtosis**
Correct responses	13.6	3.02	6.7	21.7	0.17	−0.36	21.1	5.46	8	35	−0	−0.31
Number of words produced	14.2	3.06	7	21.7	0.05	−0.41	21.6	5.53	8	35	−0.1	−0.26
Self-talk ratio	0.14	0.15	0	0.9	2.3	6.71	0.05	0.08	0	0.45	2.31	6.61
Filled hesitation ratio	0.32	0.32	0	1.57	1.51	2.34	0.4	0.27	0	1.07	0.51	−0.46
Repetition error ratio	0.03	0.03	0	0.15	1.43	1.76	0.02	0.04	0	0.2	2.38	6.44
Nonword error ratio	0.01	0.01	0	0.06	2.09	3.72	0	0.01	0	0.05	5.81	32.49
Intrusion error ratio	0.01	0.02	0	0.1	2.14	4.81	0	0.01	0	0.06	3.53	10.96

Abbreviations: SD, standard deviation of the mean; Skew, skewness of the distribution.

nonreferential behavior (self-talk, filled hesitation) using a binomial GAM with a smooth term for age (i.e., a nonparametric spline function estimated using restricted maximum likelihood). In the second step, we assessed whether age predicted the proportion of each behavior among participants who made at least one such production again, using a binomial GAM with a smooth term for age.

To address our second objective, namely, to examine whether the kinds of errors per participant varied as a function of age and Condition (Phonemic, Semantic), we examined the distribution of the different types of errors that were produced. First, each error type was converted into a ratio (errors/total responses). Because there were few errors of each type, the distributions were severely zero-inflated. To address this, we applied the same two-part GAM approach used to examine self-talk and filled hesitations.

Finally, to address our third objective, namely, to explore the potential moderating impact of a series of factors that have been associated with the prevention of cognitive decline or with a higher level of cognitive performance, a LMM analysis was run using a process similar to the one described for the first analyses. The full model was: Correct responses ∼Age*Group*Condition + Age*Outlook*Condition + Age*Education*Condition + Age*MoCA*Condition + Age*Health*Condition + Age*Languages*Condition + Age*Hearing*Condition + Age*SocialLife*Condition + Sex + (1|Participant). Protective factors were only allowed to interact with Age and Condition to prevent overfitting and to maintain model parsimony. Sex could not be allowed to interact with all the factors because this produced very high collinearity in the model (operationalized as VIF values). We therefore included Sex only as a covariate. Importantly, we compared the sex groups on all the protection factors as well as age and found no difference (S6).

3 | Results

3.1 | Analysis of Correct Responses

On average, participants produced 21 ± 5.5 words in the semantic task and 14 ± 3 words in the phonemic task. The details of the verbal productions are provided in Table 2.

The first statistical analysis addresses our first objective—to explore the prevalence of nonreferential vocalizations, including self-talk, during semantic, and phonemic fluency tasks and determine whether such behaviors are associated with aging and with better performance. There were too few instances of comments and play to analyze using linear models, leading to the decision to analyze them together as "self-talk" (but see additional analyses in Section 3.2). Results of the analyses are illustrated in Figures 2 and 3 and detailed in Table 3. The final model was: CorrectResponses ∼ 1 + Age + Self-Talk + filled hesitation + Sex + Age:Sex + Condition + Age:Condition + Sex:Condition + Age:Condition:Sex. The VIF values for the model were all below 5 (Table S7).

The analysis revealed an interaction between Age and Condition, between Age and Sex and between Age, Sex, and Condition. The interaction between Age and Condition indicate that semantic fluency performance was negatively affected by age ($p \leq 0.01$). This interaction was moderated by Sex. While men and women showed lower performance with age in the semantic condition, women, but not men, tended to have better phonemic fluency in older age (Figure 3).

As can be seen in the forest plot presented in Figure 2A, both filled hesitation and self-talk were associated with accuracy. While self-talk was associated with fewer correct responses, hesitations were associated with more correct responses overall (Figure 2B,C). The amount of filled hesitation and self-talk was not influenced by age, condition, or sex.

3.2 | Nonreferential Behavior

As shown in Table 2, self-talk and filled hesitations were relatively rare, with ratios ranging from 0 to 1.57. The distributions are illustrated as histograms in Figure S2. To account for the zero-inflated nature of the distributions, we used a two-step GAM approach. In the first step, we assessed whether age predicted the likelihood of producing at least one occurrence of self-talk or filled hesitation. For self-talk, the binomial GAM revealed a significant nonlinear effect of age (edf = 3.102, $F = 3.879$, $p = 0.007$), indicating that the likelihood of producing at least one instance of self-talk varied with age (see Figure 4A). Specifically, the pattern followed a modest U-shape, with a slight dip in midadulthood and an increase in later years. For filled hesitation

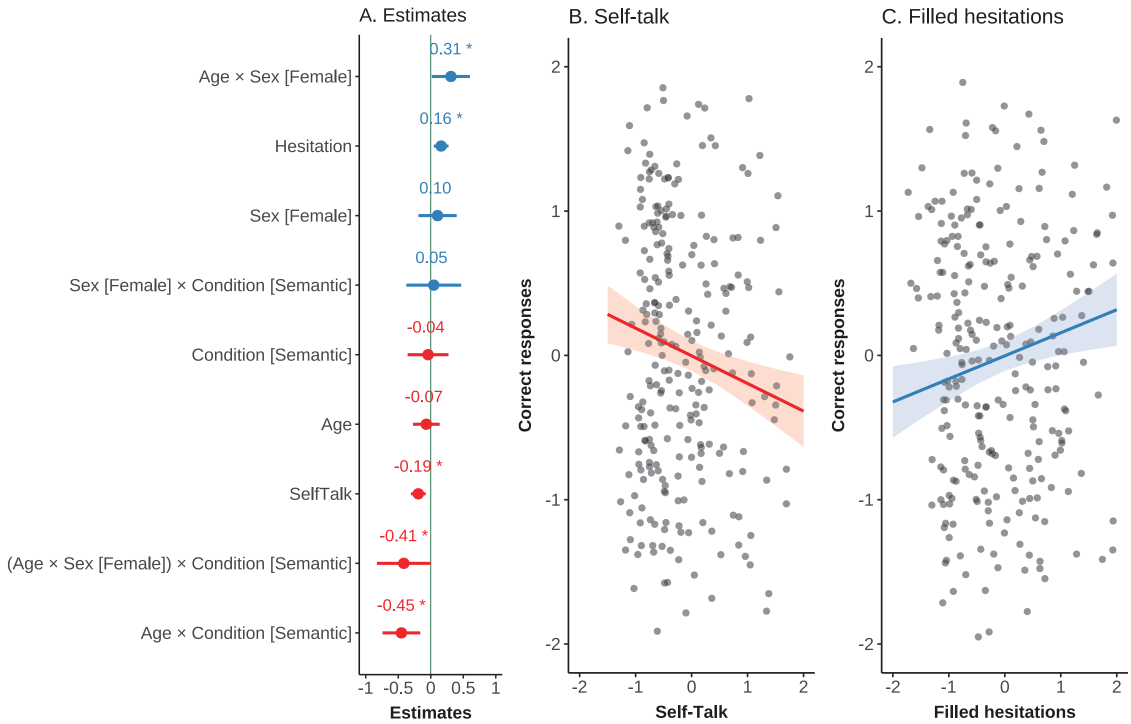

FIGURE 2 | (A) The forest plot illustrates the results (in the form of estimates) for the analysis of correct responses as a function of nonreferential behavior. Note that Hesitation refers to filled hesitations. The values above the lines represent the actual estimates. Positive estimates are plotted in blue, while negative estimates are plotted in red. The lines represent the error bars for the estimates. Significant estimates ($p \leq 0.05$) are followed by an asterisk. (B) The scatterplot illustrates the observed data from the analyses. Significant detrimental association between self-talk and correct responses. And (C) the scatterplot illustrates the significant positive association between filled hesitations and correct responses. In both scatterplots, each dot represents an individual participant. The lines represent predicted values from the model, and the shaded areas around the lines indicate the 95% confidence intervals of the estimates.

TABLE 3 | Results for the analysis of Correct Responses as a function of self-talk and filled hesitations.

Predictors	β	SE	CI	p
(Intercept)	−0.05	0.11	−0.27 to 0.17	0.681
Age	−0.07	0.1	−0.27 to 0.14	0.515
Self-talk	−0.19	0.06	−0.30 to −0.08	**0.001**
Filled hesitation	0.16	0.06	0.05–0.27	**0.006**
Sex [female]	0.1	0.15	−0.19 to 0.40	0.484
Condition [semantic]	−0.04	0.16	−0.36 to 0.27	0.795
Age × sex [female]	0.31	0.15	0.02–0.60	**0.038**
Age × condition [semantic]	−0.45	0.15	−0.74 to −0.16	**0.002**
Sex [female] × condition [semantic]	0.05	0.21	−0.38 to 0.47	0.832
(Age × sex [female]) × condition [semantic]	−0.41	0.21	−0.83 to −0.00	**0.05**
Observations	287			
R^2/R^2 adjusted	0.240/0.215			
AIC	755.7			

Abbreviations: CI, confidence interval of β; SE, standard error of β; β, standardized estimate. Bold values indicate a significant p-value.

the age effect was not significant (edf = 1.306, F = 1.552, $p = 0.33$).

In the second step, we examined whether age predicted the proportion of self-talk and filled hesitation (relative to total responses) among participants who produced at least one occurrence of self-talk and/or filled hesitation. For self-talk, the continuous GAM revealed a modest but significant effect of age (edf = 5.813, F = 6.944, p = 0.016), again showing a nonlinear increase with age (see Figure 4B). For filled hesitations, no

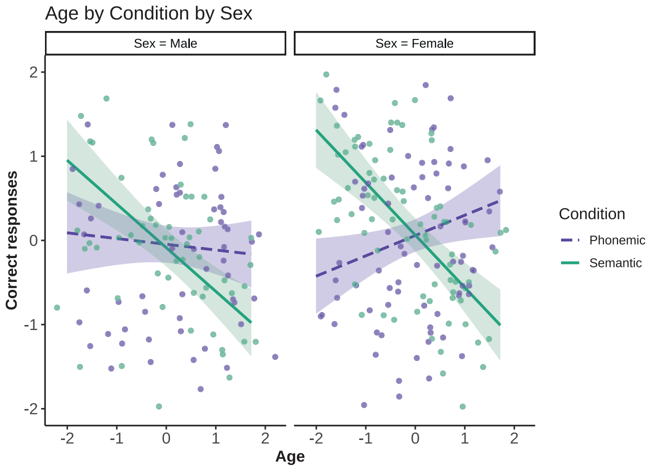

FIGURE 3 | The scatterplots illustrate the observed data for the three-way interaction between age, condition, and sex on correct responses. Each dot represents an individual participant. The lines represent predicted values from the model, and the shaded areas around the lines indicate the 95% confidence intervals of the estimates.

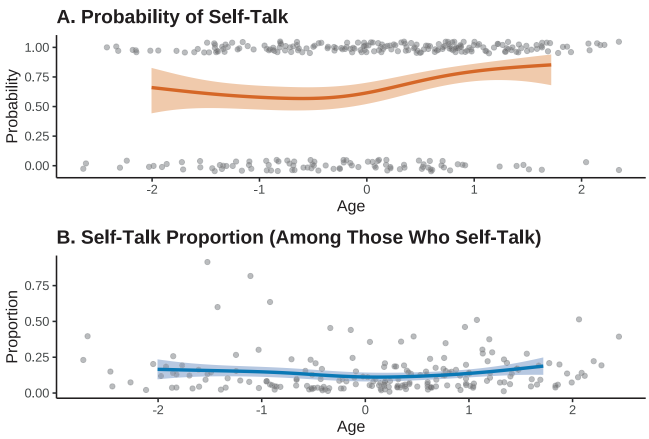

FIGURE 4 | (A) The probability of producing at least one instance of self-talk as a function of age, modeled using a binomial generalized additive model (GAM), and (B) the proportion of self-talk (relative to the total number of words produced) among participants who produced at least one instance of self-talk, modeled using a Gaussian GAM. Shaded areas represent 95% confidence intervals around the smooth terms.

significant effect was observed (edf = 1.597, $F = 1.988$, $p = 0.202$). Together, these findings suggest that both the probability and proportion of self-talk production varies nonlinearly across the adult lifespan, following a modest U-shaped trajectory.

3.3 | Error Type Analysis

As shown in Table 2, while all error types were rare, repetitions were the most frequent (see the histograms in Figure S3). To account for the zero-inflated nature of the distribution of errors, we used the same two-step (GAM) approach to model nonlinear relationships between errors and age. In the first step, we assessed whether age predicted the likelihood of producing at least one error. For repetition errors, the binomial GAM

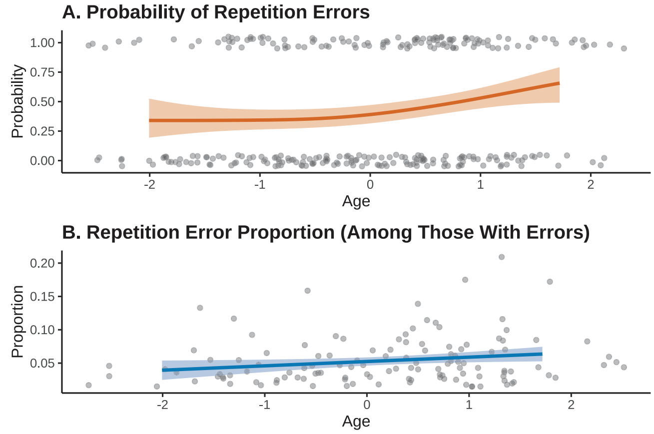

FIGURE 5 | (A) The probability of producing at least one repetition error as a function of age, modeled using a binomial generalized additive model (GAM), and (B) the proportion of repetition errors (relative to the total number of words produced) among participants who made at least one repetition error, modeled using a Gaussian GAM. Shaded areas represent 95% confidence intervals around the smooth terms.

revealed a significant nonlinear effect of age (edf = 1.772, $F = 2.224$, $p = 0.008$), indicating that older participants were more likely to produce repetition errors (see Figure 5A). For intrusion (edf = 1, $F = 1$, $p = 0.211$) and nonword errors (edf = 2.291, $F = 2.897$, $p = 0.093$), no such effect was found. In the second step, we examined whether age predicted the proportion of errors (relative to total responses) among participants who made at least one error. For repetition errors, the model showed a significant effect of age (edf = 1.82, $F = 7.52$, $p < 0.001$), revealing that error proportion increased with age almost linearly (Figure 5B). For the intrusion (edf = 4.391, $F = 1.348$, $p = 0.268$) and nonword errors (edf = 1.791, $F = 1.204$, $p = 0.283$), no effect of age was found. These results indicate a modest but significant nonlinear increase in both the probability of producing repetition errors and their relative frequency with advancing age.

3.4 | Mediators of Verbal Fluency Performance

This final analysis addresses our third specific objective, namely, to examine potentially moderating impact of eight factors on verbal fluency (education, the practice of group singing, social participation, global cognitive level, self-reported health, positive outlook, multilingualism, and hearing). The final model was: CorrectResponses ~ 1 + Cognition + Age + Education + Singing Group + Multilingualism + Age:Multilingualism + Hearing + Age:Hearing + Condition + Age:Condition. The VIF of the model were all below 5 (Table S8). The results are illustrated as the forest plot presented in Figure 6 and detailed in Table 4 and Figure 7. Of the eight different protection factors that were examined, five influenced response accuracy in the fluency tasks: education, cognitive level, singing group, multilingualism, and hearing. Two of these interacted with age (multilingualism and hearing).

The main effect of education is shown in Figure 7A. Performance in those who completed postsecondary education (college,

TABLE 4 | Results for the analysis of the mediators of accuracy.

Predictors	β	SE	CI	p
(Intercept)	−0.45	0.16	−0.77 to −0.13	**0.006**
Age	0.19	0.1	−0.01 to 0.38	0.058
Hearing	−0.11	0.08	−0.25 to 0.04	0.161
Age × hearing	0.15	0.06	0.04–0.26	**0.009**
Cognition	0.16	0.06	0.05–0.28	**0.006**
Education [college]	0.32	0.17	−0.01 to 0.65	0.058
Education [undergrad.]	0.53	0.16	0.21–0.86	**0.001**
Education [graduate]	0.47	0.2	0.09–0.86	**0.016**
Group [occasional]	0.35	0.12	0.11–0.59	**0.004**
Group [frequent]	0.11	0.14	−0.16 to 0.39	0.422
Multilingualism [multi]	0.13	0.11	−0.09 to 0.36	0.248
Condition [semantic]	0	0.1	−0.20 to 0.20	0.992
Age × multilingualism [multi]	−0.41	0.11	−0.63 to −0.19	**< 0.001**
Age × condition [semantic]	−0.66	0.1	−0.86 to −0.46	**< 0.001**
Observations	287			
R^2/R^2 adjusted	0.310/0.277			
AIC	735.849			

Note: For the group factor, the reference group is the Control group.
Abbreviations: CI, confidence interval of β; SE, standard error of β; β, standardized estimate. Bold values indicate a significant p-value.

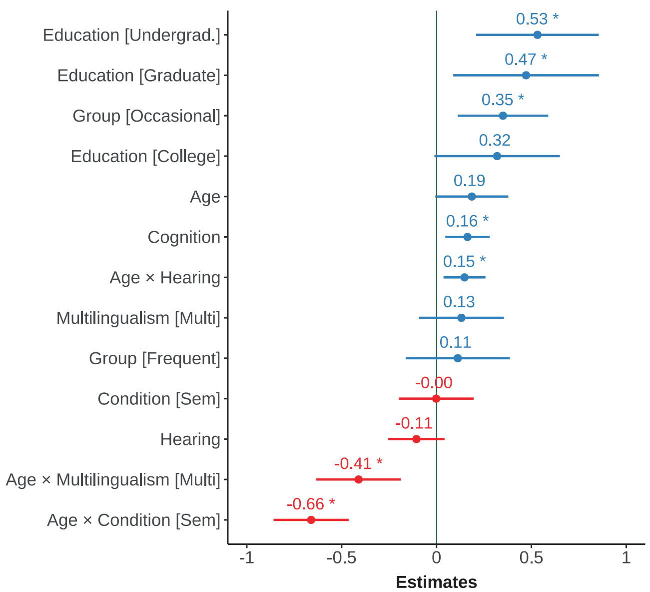

FIGURE 6 | The forest plot illustrates the results (in the form of estimates) for the analysis of the mediators of accuracy. Positive estimates are plotted in blue, while negative estimates are plotted in red. The values above the lines represent the actual estimates. The lines represent the error bars for the estimates. Significant estimates ($p \leq 0.05$) are followed by an asterisk.

undergraduate, and graduate) was associated with more correct responses compared to performance in those with primary or secondary education (Figure 7A). The main effect of singing is shown in Figure 7B. Performance in those who occasionally sing was associated with more correct responses compared to performance in those who did not sing, with no other significant difference between groups (Figure 7B). The main effect of cognitive level (operationalized as the MoCA score) is shown in Figure 7C. As can be seen in the figure, higher MoCA score was associated with a higher number of correct responses. An interaction between age and multilingualism revealed a slightly positive relationship between age and performance in mono-/bilingual participants ($b = 0.19$, $p = 0.06$). In contrast, among multilingual participants, there was a slightly negative ($b = -0.22$, $p = 0.06$) relationship between age and performance (Figure 7D). Finally, for Hearing, the interaction with Age revealed a positive association between age and performance in those with a better than average hearing ($p = 0.02$), while in those with average or below average hearing there was no association between age and performance (Figure 7E).

4 | Discussion

This study examined age-related differences in lexical access during language production using semantic and phonemic fluency tasks. Beyond standard metrics, we analyzed in an innovative way whether response content revealed age-related patterns in error types and nonreferential vocal behaviors, operationalized as self-talk and filled hesitations. Importantly, we explored the moderating effects of factors linked to cognitive resilience and performance—including education, group singing, social participation, cognitive level, self-reported health, positive outlook, hearing, and multilingualism—on accuracy. Consistent with the literature, we hypothesized that aging would negatively impact lexical access. We also expected self-talk to be more prevalent in

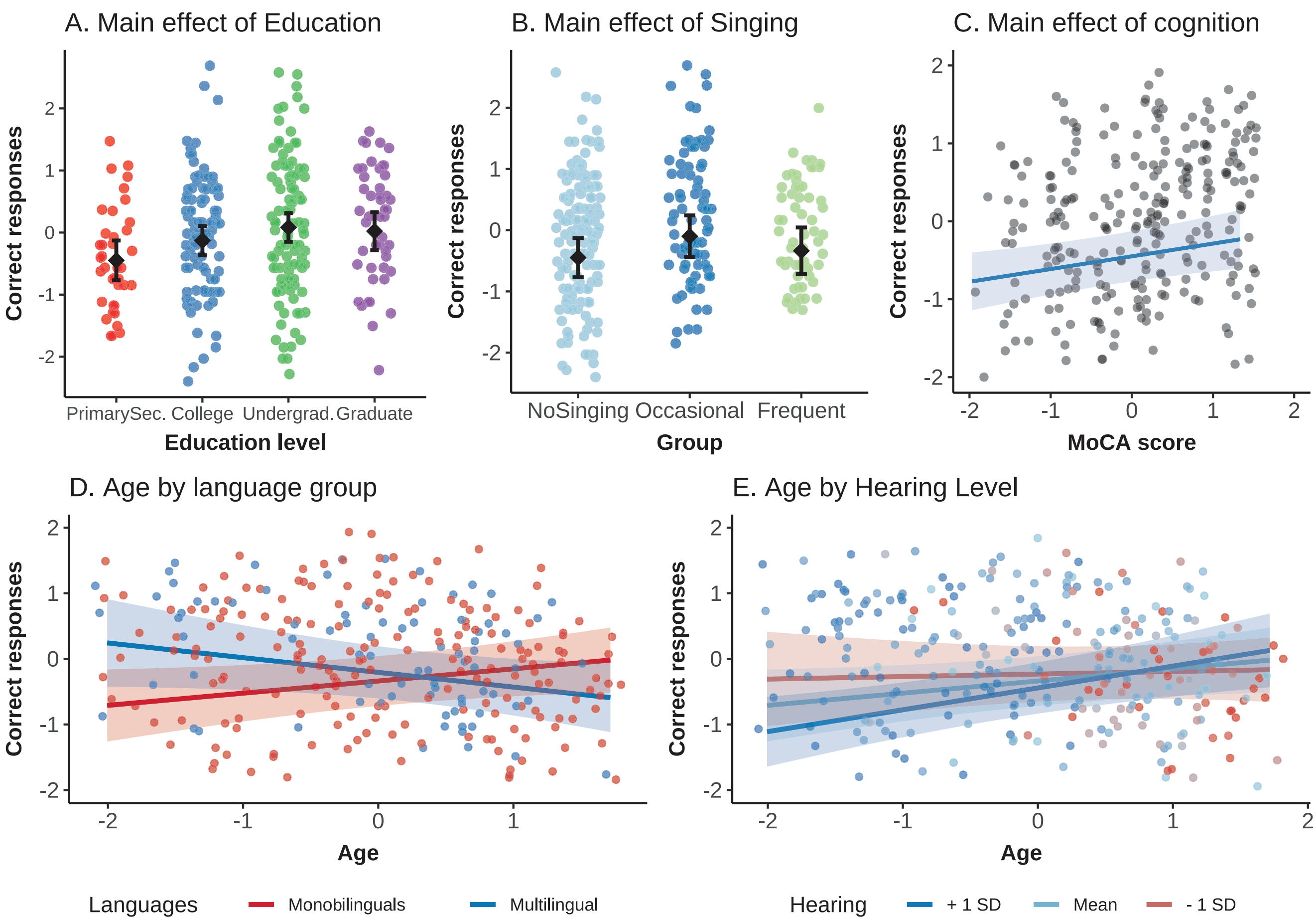

FIGURE 7 | Results for the analysis of potentially protective factors for verbal fluency, measured in terms of correct responses. All figures illustrate predicted values. (A) Relationship between education level and correct responses. In the Quebec system, primary and secondary (here abbreviated as PrimarySec) education amounts to 11 years. College is a 2- or 3-year degree that can be general (preuniversity) or technical. "Undergrad." refers to a baccalaureate degree, while graduate level refers to master's, PhDs, and medical fellowships. Each dot represents one participant. (B) Relationship between singing and correct responses, where each dot represents one participant. (C) The scatterplot illustrates the relationship between cognitive level (operationalized as the Montreal Cognitive Assessment [MoCA] score) and the number of correct responses. The shade around the regression line represents the 95% confidence of the regression line. (D) The scatterplot illustrates the interaction between age (z scores) and language groups (monobilinguals in red vs. multilinguals in blue) on the number of correct responses overall. (E) The scatterplot illustrates the interaction between age and hearing. Note that the scores here have been inverted such that a higher score reflects better hearing. The dark blue line illustrates the relationship between age and correct responses for those with better hearing than the average; the red line illustrates the relationship between age and correct responses for those with lesser hearing than the average; the pale blue line illustrates the relationship between age and correct responses for those with an average hearing level.

older adults compared to younger ones. Finally, we anticipated that accuracy in fluency tasks would be moderated by factors associated with cognitive resilience, including higher education, group singing, social participation, global cognitive level, positive outlook, hearing ability, good self-reported health, and multilingualism. Our results show that, in our sample, age had a negative impact on performance in the semantic task but not in the phonemic one in both men and women. Self-talk was associated with worse performance across all ages, but filled hesitations were associated with better accuracy. While the number of nonwords and intrusion errors did not vary with age, repetition errors were more frequent in older participants compared to younger ones. We found some evidence of a positive impact of cognitive reserve on verbal fluency performance. These findings are discussed in more detail below.

4.1 | Age Differences

Our results show that the semantic task (animal fluency) was overall easier, being associated with more correct responses (21 ± 5.5 words) compared to the phonemic tasks (14 ± 3 words). While fluency tasks are affected by the number of words in a language that are congruent with the criteria used, there is some evidence suggesting that the animal fluency tasks might be easier compared to other categories (e.g., tools, sports) as it is familiar and has higher semantic relatedness [56]. Despite being easier, the semantic fluency task was negatively associated with age in both men and women, consistent with prior studies, for example, refs. [10, 49, 57–59], while the harder phonemic fluency task was not, also consistent with prior studies [60, 61], though other studies have found an effect of age on phonemic fluency [62–64]. Our results suggest that access to semantic memory becomes less efficient over time. In contrast, phonemic fluency appears to remain stable across the lifespan, which may reflect differences in how these tasks engage cognitive and neural resources. Interestingly, in the present study, there were some biological sex differences in the effect of age on phonemic fluency. While overall phonemic fluency was not affected by age, our results show that women—but not men—exhibited better performance with increasing age. This contrasts with findings from St-Hilaire et al. [49], who found no significant effect of sex on phonemic fluency in a large sample of Quebec French speakers. However, our result is consistent with a recent meta-analysis reporting a small but reliable female advantage in phonemic fluency [65].

The analysis of nonreferential vocal behaviors during verbal fluency highlights the dynamic and multifaceted nature of speech production. Across age, filled hesitations and self-talk were both linked to task performance, but in contrasting ways: filled hesitations—sometimes referred to as disfluencies—were associated with better performance, whereas self-talk was linked to poorer performance. These findings suggest that filled hesitations may serve a functional role in lexical retrieval, acting as a marker of active search or temporary lexical blocking that allows individuals to maintain the flow of speech. In contrast, self-talk, at least in the context of a timed fluency task, may reflect greater lexical access difficulty or reduced task focus/engagement, potentially representing a verbal manifestation of cognitive strain.

The nonlinear analyses further revealed that self-talk and repetition errors followed nonlinear trajectories, with increases emerging primarily in later adulthood, whereas filled hesitations and other error types (nonwords and intrusions) remained relatively stable across the adult lifespan. This dissociation suggests that not all forms of nonreferential vocal behavior are equally sensitive to aging. Whereas filled hesitations may reflect transient lexical search processes that remain preserved with age, the increase in repetition errors and self-talk may point to more systematic changes in cognitive control and speech monitoring mechanisms. The age-related increase in self-talk echoes findings from less structured language production tasks—such as autobiographical narratives and conversation—in which older adults tend to produce more verbose, tangential, or self-referential speech [20, 66, 67]. These shifts are often attributed to age-related declines in inhibitory control, which may make it harder to suppress irrelevant or self-directed utterances. However, another possibility is that older adults deliberately or unconsciously engage in self-regulatory strategies to support task performance—for example, by using self-talk to maintain focus, structure their thoughts, or compensate for increased lexical retrieval difficulty. In a similar vein, repetition errors (or perseveration) may reflect limited monitoring or working memory capacity [68]. However, it is also conceivable that, in some cases, such repetitions serve as an implicit or even conscious strategy—for instance, to reactivate the lexical search process by repeating prior responses. While potentially redundant, this behavior could represent an attempt to facilitate continued lexical access under increased cognitive demand. Overall, these age-related differences may not be uniformly maladaptive. Rather, they may represent a reorganization of communicative strategies with age, consistent with theories of compensatory cognitive aging, for example, refs. [69–71]. From this perspective, increases in self-talk and repetitions may reflect adaptive efforts to preserve task engagement and maintain fluency in the face of cognitive or linguistic challenges. Taken together, these findings underscore the importance of considering not only the quantity of correct responses in verbal fluency tasks, but also the qualitative features of speech, which can provide valuable insights into the cognitive and linguistic changes that accompany aging. Further research should examine the functional role of self-directed speech in different task contexts to clarify when it supports or hinders performance in older adults.

4.2 | Cognitive Reserve and Verbal Fluency

This study examined several potentially moderating factors of cognitive resilience and performance—specifically, education, group singing, social participation, cognitive level, self-reported health, positive outlook, hearing, and knowledge of languages—on verbal fluency abilities. No significant relationships emerged for social participation, self-reported health, or positive outlook, despite adequate variability in these measures (Table 1). However, MoCA scores, singing, multilingualism, education, and hearing were associated with verbal fluency performance. Each of these factors is explored in the following paragraphs.

Higher MoCA scores were associated with better performance, irrespective of age, suggesting that general cognitive function plays a crucial role in word retrieval capacity. The MoCA effect supports previous research showing that global cognitive function is a key predictor of verbal fluency performance [72, 73]. The MoCA measures general cognitive level and is used to detect cognitive decline and mild cognitive impairments [46, 74]. The fact that MoCA scores influenced performance regardless of age suggests that individuals with stronger overall cognitive abilities maintain better lexical access throughout aging compared to those with lower scores, but it also means that it does not moderate the effects of aging, which would have translated into a MoCA by Age interaction effect. Furthermore, given that the MoCA includes a verbal fluency test, the relation between this test and verbal fluency scores might be expected.

Higher education, specifically higher-level education, including undergraduate and graduate-level education, appeared protective in our sample, with more educated participants showing a better performance compared to those with only primary or secondary education, irrespective of age or condition, aligning with a previous study showing that formal education contributes to cognitive reserve, specifically to better verbal fluency, for example, ref. [49]. This finding is also consistent with the *Lancet* standing commission on dementia, who identified early life education as one of 14 modifiable risk factors for dementia [33–35].

Occasional singing was linked to better fluency compared to nonsingers, with no difference between occasional and frequent singers. This aligns with the OPERA hypothesis [38, 39], which posits that musical training strengthens speech and verbal processing. Our findings are also consistent with previous singing training studies [37, 75] and prior research on the same sample, which found better working memory in singers compared to active controls [36]. However, some cross-sectional and longitudinal studies have not found similar advantage in fluency or verbal working memory [44, 76, 77], highlighting inconsistencies in the literature. These discrepancies suggest that sample characteristics or variations in singing practice may influence results. The lack of a difference between occasional and frequent singers also points to a nonlinear relationship, consistent with previous findings [44]. It has been proposed that prolonged training enhances domain-specific expertise, improving skills directly related to the practiced activity rather than broadly transferring to other cognitive domains [79].

Knowledge of more languages was associated with better verbal fluency performance in younger adults, but this advantage diminished with increasing age. The effect of age on performance varied by language background. While there was a nonsignificant trend for a lower performance with age in individuals who spoke more than two languages, those who spoke one or two languages showed a modest (nonsignificant) improvement with age, reducing the performance gap between the two subgroups (Figure 7D). This pattern suggests that although multilingualism may confer an early advantage in verbal fluency, it does not appear to provide additional protection against age-related change in this domain. The notion that bilinguals outperform monolinguals in cognitive control has gained traction in popular media and education, largely due to influential studies reporting a bilingual advantage, for example, refs. [79–81]. This so-called bilingual advantage is typically attributed to the cognitive control demands imposed by managing multiple language systems. However, recent larger-scale studies, for example, refs. [41, 82, 83], systematic reviews and meta-analyses, for example, refs. [84, 85] suggest that any such advantage is small, inconsistent, and may be task- or population-specific. Notably, a recent meta-analysis focusing on healthy adults reported a small bilingual *disadvantage* in verbal fluency tasks [85], possibly reflecting reduced lexical access within each language due to divided language use. Our findings are consistent with this literature, showing that multilingualism is not associated with better verbal fluency performance in older adults, nor with reduced age-related change in word retrieval.

Finally, the last potentially protective factor that showed a relationship to accuracy was hearing. Our results show that verbal fluency performance in those with good hearing performed improved with age, while accuracy in those with lower hearing was unaffected by age. However, accuracy in older age was similar across hearing level. Hearing loss has been identified by the *Lancet* standing commission on dementia prevention, intervention, and care as a potentially modifiable risk factor for dementia [33–35]. Our finding that better hearing mitigates age-related decline aligns with the *Lancet* standing commission as well as with the ACHIEVE study, which showed that hearing loss accelerates cognitive decline and that wearing hearing aids can provide mitigation at least in vulnerable people with more comorbidities and lower education [87]. These findings suggest that preserving auditory function is important for maintaining cognitive level in aging, but hearing alone is not sufficient to enhance performance in cognitively healthy adults.

To summarize, we identified several protective factors that were significantly associated with better verbal fluency performance, including postsecondary education, occasional singing, higher cognitive functioning, multilingualism, and better hearing. These findings support the view that individual differences in lifestyle and cognitive profile contribute meaningfully to verbal fluency outcomes.

However, our goal was not only to identify static associations but to examine whether these factors moderated age-related decline—that is, whether they contributed to differential preservation [87, 88]. From this perspective, only protective factors that interact with age—by moderating or slowing age-related decline—can be considered evidence in support of the mental exercise hypothesis. Such interactions were rare in our data. This does not undermine the relevance of the observed associations; rather, it highlights the theoretical and methodological challenges in demonstrating genuine protective effects on cognitive aging. As Salthouse [90] has argued, aging is inherently a process of change, and the strongest evidence for cognitive protection comes from factors that modify the slope of decline over time. While this view has been debated [91], it provides a useful framework for interpreting our findings. It remains possible that only specific combinations of protective factors, acting together, are strong enough to slow cognitive aging—a hypothesis that requires large samples and, ideally, longitudinal data to answer robustly. Our findings thus contribute to this growing literature by identifying relevant candidate factors, while also underscoring the need for further work to determine their long-term impact on cognitive trajectories.

5 | Conclusions

This study contributes to the growing literature on cognitive aging and cognitive reserve by examining error patterns, nonreferential vocal behavior, and the influence of various protective factors on verbal fluency in healthy adults. To our knowledge, this is the first study to document that community-dwelling older adults produce more self-talk and repetition errors during verbal fluency tasks, while other error types—such as intrusions and nonwords—remain stable across the adult lifespan. Several protective factors—including occasional singing, higher education, cognitive level, multilingualism, and hearing—were associated with better verbal fluency performance. However, we found limited evidence that these factors mitigate age-related decline in this domain. Longitudinal or intervention-based studies are needed to address this more directly. Importantly, future interventions that target multiple protective domains simultaneously may hold the most promise for limiting cognitive decline with age.

Author Contributions

Pascale Tremblay: conceptualization, funding acquisition, methodology, supervision, resources, project administration, formal analysis, visualization, data Curation, writing – original draft preparation. **Lydia Gagnon**: investigation, formal analysis. **Edith Durand**: conceptualization, methodology, supervision, writing – reviewing and editing. **Joël Macoir**: conceptualization, methodology, writing – reviewing and editing.

Acknowledgments

We thank all the participants.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The raw datasets generated during the current study is not publicly available because participants did not consent to public data sharing. However, the group data will be available on Borealis upon publication (https://doi.org/10.5683/SP3/U5ZF7O).

References

1. T. A. Salthouse, "Trajectories of Normal Cognitive Aging," *Psychology and Aging* 34 (2019): 17–24.

2. P. K. Parikh, A. K. Troyer, A. M. Maione, and K. J. Murphy, "The Impact of Memory Change on Daily Life in Normal Aging and Mild Cognitive Impairment," *Gerontologist* 56 (2016): 877–885.

3. J. Austin, K. Klein, N. Mattek, and J. Kaye, "Variability in Medication Taking Is Associated With Cognitive Performance in Nondemented Older Adults," *Alzheimer's & Dementia (Amsterdam)* 6 (2017): 210–213.

4. WHO, *World Report on Ageing and Health* (World Health Organization) (2015).

5. D. Villalobos, L. Torres-Simon, J. Pacios, et al., "A Systematic Review of Normative Data for Verbal Fluency Test in Different Languages," *Neuropsychology Review* 33 (2023): 733–764.

6. S. L. Aita, J. D. Beach, S. E. Taylor, et al., "Executive, Language, or Both? An Examination of the Construct Validity of Verbal Fluency Measures," *Applied Neuropsychology-Adult* 26 (2019): 441–451.

7. D. Kempler, E. L. Teng, M. Dick, et al., "The Effects of Age, Education, and Ethnicity on Verbal Fluency," *Journal of the International Neuropsychological Society* 4 (1998): 531–538.

8. E. H. Strauss, E. M. Sherman, and O. Spreen, *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary* (Oxford University Press, 2006).

9. R. Tomer and B. E. Levin, "Differential Effects of Aging on Two Verbal Fluency Tasks," *Perceptual and Motor Skills* 76 (1993): 465–466.

10. A. K. Troyer, M. Moscovitch, and G. Winocur, "Clustering and Switching as Two Components of Verbal Fluency: Evidence From Younger and Older Healthy Adults," *Neuropsychology* 11 (1997): 138–146.

11. J. K. Gordon, M. Young, and C. Garcia, "Why Do Older Adults Have Difficulty With Semantic Fluency?," *Neuropsychology, Development, and Cognition Section B Aging, Neuropsychology and Cognition* 25 (2018): 803–828.

12. T. N. Tombaugh, J. Kozak, and L. Rees, "Normative Data Stratified by Age and Education for Two Measures of Verbal Fluency: FAS and Animal Naming," *Archives of Clinical Neuropsychology* 14 (1999): 167–177.

13. G. Lubrini, J. A. Perianez, G. Laseca-Zaballa, et al., "Verbal Fluency Tasks: Influence of Age, Gender, and Education and Normative Data for the Spanish Native Adult Population," *Archives of Clinical Neuropsychology* 37 (2022): 365–375.

14. L. S. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes* (Harvard University Press, 1978).

15. L. S. Vygotsky, "Thinking and Speech," in *The Collected Works of L. S. Vygotsky*, ed. R. W. Rieber and A. S. Carton (Springer, 1987), 39–285.

16. A. R. Luriia, *The Role of Speech in the Regulation of Normal and Abnormal Behavior* (Liveright, 1961).

17. M. Gade and M. Paelecke, "Talking Matters—Evaluative and Motivational Inner Speech Use Predicts Performance in Conflict Tasks," *Scientific Reports* 9 (2019): 9531.

18. J. S. K. Nedergaard and G. Lupyan, "Not Everybody Has an Inner Voice: Behavioral Consequences of Anendophasia," *Psychological Science* 35 (2024): 780–797.

19. X. Guo and K. Dobkins, "Private Speech Amount Positively Predicts Memory Performance in Young Adults," *Consciousness and Cognition* 113 (2023): 103534.

20. L. E. James, D. M. Burke, A. Austin, and E. Hulme, "Production and Perception of "Verbosity" in Younger and Older Adults," *Psychology and Aging* 13 (1998): 355–367.

21. T. Ruffman, J. Murray, J. Halberstadt, and M. Taumoepeau, "Verbosity and Emotion Recognition in Older Adults," *Psychology and Aging* 25 (2010): 492–497.

22. P. V. Cooper, "Discourse Production and Normal Aging: Performance on Oral Picture Description Tasks," *Journal of Gerontology* 45 (1990): P210–214.

23. H. Bortfeld, S. D. Leon, J. E. Bloom, et al., "Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender," *Language and Speech* 44 (2001): 123–147.

24. W. S. Horton, D. H. Spieler, and E. Shriberg, "A Corpus Analysis of Patterns of Age-Related Change in Conversational Speech," *Psychology and Aging* 25 (2010): 708–713.

25. K. Ledoux, T. D. Vannorsdall, E. J. Pickett, et al., "Capturing Additional Information About the Organization of Entries in the Lexicon From Verbal Fluency Productions," *Journal of Clinical and Experimental Neuropsychology* 36 (2014): 205–220.

26. J. D. Henry and L. H. Phillips, "Covariates of Production and Perseveration on Tests of Phonemic, Semantic and Alternating Fluency in Normal Aging," *Neuropsychology, Development, and Cognition Section B Aging, Neuropsychology and Cognition* 13 (2006): 529–551.

27. J. McDowd, L. Hoffman, E. Rozek, et al., "Understanding Verbal Fluency in Healthy Aging, Alzheimer's Disease, and Parkinson's Disease," *Neuropsychology* 25 (2011): 210–225.

28. E. Kozora and C. Munro Cullum, "Generative Naming in Normal Aging: Total Output and Qualitative Changes Using Phonemic and Semantic Constraints," *Clinical Neuropsychologist* 9 (1995): 313–320.

29. Y. Stern, "What Is Cognitive Reserve? Theory and Research Application of the Reserve Concept," *Journal of the International Neuropsychological Society* 8 (2002): 448–460.

30. Y. Stern, "The Concept of Cognitive Reserve: A Catalyst for Research," *Journal of Clinical and Experimental Neuropsychology* 25 (2003): 589–593.

31. Y. Stern, "Cognitive Reserve," *Neuropsychologia* 47 (2009): 2015–2028.

32. L. Gonzalez-Burgos, J. Barroso, and D. Ferreira, "Cognitive Reserve and Network Efficiency as Compensatory Mechanisms of the Effect of Aging on Phonemic Fluency," *Aging (Albany New York)* 12 (2020): 23351–23378.

33. G. Livingston, J. Huntley, K. Y. Liu, et al., "Dementia Prevention, Intervention, and Care: 2024 Report of the Lancet Standing Commission," *Lancet* 404 (2024): 572–628.

34. G. Livingston, J. Huntley, A. Sommerlad, et al., "Dementia Prevention, Intervention, and Care: 2020 Report of the Lancet Commission," *Lancet* 396 (2020): 413–446.

35. G. Livingston, A. Sommerlad, V. Orgeta, et al., "Dementia Prevention, Intervention, and Care," *Lancet* 390 (2017): 2673–2734.

36. M. Joyal, A. Sicard, V. Penhune, et al., "Attention, Working Memory, and Inhibitory Control in Aging: Comparing Amateur Singers, Instrumentalists, and Active Controls," *Annals of the New York Academy of Sciences* 1541 (2024): 163–180.

37. M. C. Fu, B. Belza, H. Nguyen, et al., "Impact of Group-Singing on Older Adult Health in Senior Living Communities: A Pilot Study," *Archives of Gerontology and Geriatrics* 76 (2018): 138–146.

38. A. D. Patel, "Can Nonlinguistic Musical Training Change the Way the Brain Processes Speech? The Expanded OPERA Hypothesis," *Hearing Research* 308 (2014): 98–108.

39. A. D. Patel, "The OPERA Hypothesis: Assumptions and Clarifications," *Annals of the New York Academy of Sciences* 1252 (2012): 124–128.

40. A. de Bruin, A. S. Dick, and M. Carreiras, "Clear Theories Are Needed to Interpret Differences: Perspectives on the Bilingual Advantage Debate," *Neurobiology of Language (Cambridge)* 2 (2021): 433–451.

41. A. S. Dick, N. L. Garcia, S. M. Pruden, et al., "No Evidence for a Bilingual Executive Function Advantage in the Nationally Representative ABCD Study," *Nature Human Behaviour* 3 (2019): 692–701.

42. J. K. Ljungberg, P. Hansson, P. Andres, et al., "A Longitudinal Study of Memory Advantages in Bilinguals," *PLoS ONE* 8 (2013): e73029.

43. X. Zhang and P. Tremblay, "Aging of Resting-State Functional Connectivity in Amateur Singers, Instrumentalists and Controls," *Aperture Neuro* 5, (2025).

44. P. Tremblay and M. Perron, "Auditory Cognitive Aging in Amateur Singers and Non-Singers," *Cognition* 230 (2023): 105311.

45. R. C. Oldfield, "The Assessment and Analysis of Handedness: The Edinburgh Inventory," *Neuropsychologia* 9 (1971): 97–113.

46. Z. S. Nasreddine, N. A. Phillips, V. Bedirian, et al., "The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool for Mild Cognitive Impairment," *Journal of the American Geriatrics Society* 53 (2005): 695–699.

47. J. A. Yesavage, T. L. Brink, T. L. Rose, et al., "Development and Validation of a Geriatric Depression Screening Scale: A Preliminary Report," *Journal of Psychiatric Research* 17 (1982): 37–49.

48. T. Saito, C. Murata, M. Saito, et al., "Influence of Social Relationship Domains and Their Combinations on Incident Dementia: A Prospective Cohort Study," *Journal of Epidemiology and Community Health* 72 (2018): 7–12.

49. A. St-Hilaire, C. Hudon, G. T. Vallet, et al., "Normative Data for Phonemic and Semantic Verbal Fluency Test in the Adult French-Quebec Population and Validation Study in Alzheimer's Disease and Depression," *Clinical Neuropsychologist* 30 (2016): 1126–1150.

50. P. Boersma and D. Weenink. 2011. Praat: Doing Phonetics by Computer. http://www.praat.org/

51. R_Core_Team, *R: A Language and Environment for Statistical Computing*, (R Foundation for Statistical Computing, 2019). https://www.R-project.org/.

52. C. Voeten, "buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression," in *R Package* (2020).

53. D. Bates, M. Maechler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," *Journal of Statistical Software* 67 (2015): 1–48.

54. D. J. Barr, "Random Effects Structure for Testing Interactions in Linear Mixed-Effects Models," *Frontiers in Psychology* 4 (2013): 328.

55. D. Lüdecke, "sjPlot: Data Visualization for Statistics in Social Science," in *R Package* (2021).

56. F. Jebahi, R. Abou Jaoude, and C. Ellis, "Semantic Verbal Fluency Task: The Effects of Age, Educational Level, and Sex in Lebanese-Speaking Adults," *Applied Neuropsychology-Adult* 29 (2022): 936–940.

57. Z. Zhu, J. Deng, M. Li, et al., "Processing Speed Mediates the Relationship Between Brain Structure and Semantic Fluency in Aging," *Neuroscience Letters* 788 (2022): 136838.

58. C. Rodriguez-Aranda and M. Martinussen, "Age-Related Differences in Performance of Phonemic Verbal Fluency Measured by Controlled Oral Word Association Task (COWAT): A Meta-Analytic Study," *Developmental Neuropsychology* 30 (2006): 697–717.

59. D. H. Murphy and A. D. Castel, "Age-Related Similarities and Differences in the Components of Semantic Fluency: Analyzing the Originality and Organization of Retrieval From Long-Term Memory," *Neuropsychology, Development, and Cognition Section B Aging, Neuropsychology and Cognition* 28 (2021): 748–761.

60. K. I. Bolla, K. N. Lindgren, C. Bonaccorsy, and M. L. Bleecker, "Predictors of Verbal Fluency (FAS) in the Healthy Elderly," *Journal of Clinical Psychology* 46 (1990): 623–628.

61. A. J. Parkin and R. I. Java, "Deterioration of Frontal Lobe Function in Normal Aging: Influences of Fluid Intelligence Versus Perceptual Speed," *Neuropsychology* 13 (1999): 539–545.

62. A. M. Brickman, R. H. Paul, R. A. Cohen, et al., "Category and Letter Verbal Fluency Across the Adult Lifespan: Relationship to EEG Theta Power," *Archives of Clinical Neuropsychology* 20 (2005): 561–573.

63. A. S. Loonstra, A. R. Tarlow, and A. H. Sellers, "COWAT Metanorms Across Age, Education, and Gender," *Applied Neuropsychology* 8 (2001): 161–166.

64. S. Auriacombe, C. Fabrigoule, S. Lafont, et al., "Letter and Category Fluency in Normal Elderly Participants: A Population-Based Study," *Aging, Neuropsychology, and Cognition* 8 (2001): 98–108.

65. M. Hirnstein, J. Stuebs, A. Moe, and M. Hausmann, "Sex/Gender Differences in Verbal Fluency and Verbal-Episodic Memory: A Meta-Analysis," *Perspectives on Psychological Science* 18 (2023): 67–90.

66. M. Ceccaldi, Y. Joanette, F. Tikhomirof, et al., "The Effects of Age-induced Changes in Communicative Abilities on the Type of Aphasia," *Brain and Language* 54 (1996): 75–85.

67. l. Mortensen, A. S. Meyer, and G. W. Humphreys, "Age-Related Effects on Speech Production: A Review," *Language and Cognitive Processes* 21 (2006): 238–290.

68. V. M. Rosen and R. W. Engle, "The Role of Working Memory Capacity in Retrieval," *Journal of Experimental Psychology General* 126 (1997): 211–227.

69. R. Cabeza, "Hemispheric Asymmetry Reduction in Older Adults: The HAROLD Model," *Psychology and Aging* 17 (2002): 85–100.

70. R. Cabeza, N. D. Anderson, J. K. Locantore, and A. R. McIntosh, "Aging Gracefully: Compensatory Brain Activity in High-Performing Older Adults," *Neuroimage* 17 (2002): 1394–1402.

71. A. M. Freund and P. B. Baltes, "Selection, Optimization, and Compensation as Strategies of Life Management: Correlations With Subjective Indicators of Successful Aging," *Psychology and Aging* 13 (1998): 531–543.

72. M. Souza, F. Rodrigues Bernardes, C. Machado, et al., "Relationship Between Cognitive and Sociodemographic Aspects and Verbal Fluency of Active Elderly," *Revista CEFAC* 20 (2018): 493–502.

73. M. Bialek and A. Borkowska, "Predictors of Verbal Fluency Decline in People With Cognitive Dysfunctions," *Gerontology & Geriatrics: Research* 10, no. 3 (2024): 1103.

74. Z. S. Nasreddine, H. Chertkow, N. Phillips, et al., "Sensitivity and Specificity of the Montreal Cognitive Assessment (MoCA) for Detection of Mild Cognitive Deficits," *Canadian Journal of Neurological Sciences* 30 (2003): 695–699.

75. E. Pongan, B. Tillmann, Y. Leveque, et al., "Can Musical or Painting Interventions Improve Chronic Pain, Mood, Quality of Life, and Cognition in Patients With Mild Alzheimer's Disease? Evidence From a Randomized Controlled Trial," *Journal of Alzheimer's Disease* 60 (2017): 663–677.

76. E. Dubinsky, E. A. Wood, G. Nespoli, and F. A. Russo, "Short-Term Choir Singing Supports Speech-in-Noise Perception and Neural Pitch Strength in Older Adults with Age-Related Hearing Loss," *Frontiers in Neuroscience* 13 (2019): 1153.

77. B. Hanna-Pladdy and A. MacKay, "The Relation Between Instrumental Musical Activity and Cognitive Aging," *Neuropsychology* 25 (2011): 378–386.

78. F. Dege, "Music Lessons and Cognitive Abilities in Children: How Far Transfer Could Be Possible," *Frontiers in Psychology* 11 (2020): 557807.

79. E. Bialystok, "The Bilingual Adaptation: How Minds Accommodate Experience," *Psychological Bulletin* 143 (2017): 233–262.

80. E. Bialystok, F. Craik, and G. Luk, “Cognitive Control and Lexical Access in Younger and Older Bilinguals,” *Journal of Experimental Psychology. Learning, Memory, and Cognition* 34 (2008): 859–873.

81. T. H. Bak, J. J. Nissan, M. M. Allerhand, and I. J. Deary, “Does Bilingualism Influence Cognitive Aging?,” *Annals of Neurology* 75 (2014): 959–963.

82. K. R. Paap, H. A. Johnson, and O. Sawi, “Are Bilingual Advantages Dependent Upon Specific Tasks or Specific Bilingual Experiences?” *Journal of Cognitive Psychology* 26 (2014): 615–639.

83. J. A. Dunabeitia, J. A. Hernandez, E. Anton, et al., “The Inhibitory Advantage in Bilingual Children Revisited: Myth or Reality?” *Experimental Psychology* 61 (2014): 234–251.

84. M. Lehtonen, A. Soveri, A. Laine, et al., “Is Bilingualism Associated With Enhanced Executive Functioning in Adults? A Meta-Analytic Review,” *Psychological Bulletin* 144 (2018): 394–425.

85. K. R. Paap, H. A. Johnson, and O. Sawi, “Should the Search for Bilingual Advantages in Executive Functioning Continue?,” *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior* 74 (2016): 305–314.

86. F. R. Lin, J. R. Pike, M. S. Albert, et al., “Hearing Intervention Versus Health Education Control to Reduce Cognitive Decline in Older Adults With Hearing Loss in the USA (ACHIEVE): A Multicentre, Randomised Controlled Trial,” *Lancet* 402 (2023): 786–797.

87. T. A. Salthouse, “Mental Exercise and Mental Aging: Evaluating the Validity of the “Use It or Lose It” Hypothesis,” *Perspectives on Psychological Science* 1 (2006): 68–87.

88. T. A. Salthouse, “Influence of Experience on Age Differences in Cognitive Functioning,” *Human Factors* 32 (1990): 551–569.

89. T. A. Salthouse, “Implications of Within-Person Variability in Cognitive and Neuropsychological Functioning for the Interpretation of Change,” *Neuropsychology* 21 (2007): 401–411.

90. C. Schooler, “Use It- and Keep It, Longer, Probably: A Reply to Salthouse (2006),” *Perspectives on Psychological Science* 2 (2007): 24–29.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting Information: nyas70064-sup-0001-SuppMatt.docx