Journal of Neurolinguistics 78 (2026) 101315

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

Research paper

Cognitive reserve effects on discourse production processing in healthy aging

Andrea Marini [a,*], Francesco Petriglia [a], Silvia D'Ortenzio [a], Giulia Gasparotto [a], Ilaria Gabbatore [b,c]

[a] Cognitive Neuroscience Laboratory, University of Udine, Via Margreth, 33100, Udine, Italy
[b] Department of Humanities, University of Turin, Turin, Italy
[c] GIPSI - research Group on Inferential Processes in Social Interaction, University of Turin, Turin, Italy

ARTICLE INFO

Keywords:
Healthy aging
Discourse
Cognition
Cognitive reserve

ABSTRACT

Aging is marked by physiological changes that significantly affect discourse production skills. Cognitive reserve (CR), the brain's ability to adapt and cope with age-related changes, was examined in relation to narrative discourse production in 105 healthy adults aged 61-84. Verbal working memory, sustained attention, and inhibitory control were also assessed. CR was selectively associated with multiple cognitive and discourse measures, particularly those related to executive functioning, lexical accuracy, coherence, and informativeness. Based on CR level, participants were divided in three groups balanced for age and gender: 35 high CR, 35 moderate CR, and 35 medium CR. Participants with high CR consistently outperformed those with moderate and medium CR who did not differ from each other, indicating a non-linear pattern of CR effects. The differential contributions of education, working activity, and leisure activities to cognitive and discourse performance are also examined.

1. Introduction

The global population is aging rapidly (Medici, 2021; WHO, 2020). Aging is marked by physiological changes that significantly alter brain organization, ultimately leading to cellular atrophy in later life (Crivello et al., 2014; Vickery et al., 2024). It also involves progressive modifications in cognitive processing over time (Merlini et al., 2025; Salthouse, 2019). Recent evidence suggests that some changes follow a linear trajectory, while others are nonlinear, exhibiting periods of relative stability followed by accelerated decline at specific time points (Shen et al., 2024).

A crucial aspect to monitor is the potential influence of lifetime experiences on aging. In this regard, the concept of reserve has recently garnered significant attention. According to the Cognitive Reserve (CR) Theory (Stern, 2002; Stern et al., 2020), CR refers to an individual's resources that help protect against cognitive decline due to healthy aging, acquired brain lesions, or neurodegenerative diseases (Roberts et al., 2017). In the present study, CR is conceptualized primarily in functional terms, that is, as the capacity to flexibly recruit, adapt, and reorganize cognitive processes in response to age-related neural changes, rather than as a proxy of structural brain capacity. Specifically, it is hypothesized that individuals with greater CR can better compensate for neuronal degeneration, as their life experiences have strengthened the neural networks supporting those cognitive functions that decline with

* Corresponding author.
E-mail address: andrea.marini@uniud.it (A. Marini).

https://doi.org/10.1016/j.jneuroling.2026.101315
Received 6 October 2025; Received in revised form 21 January 2026; Accepted 6 February 2026
Available online 16 February 2026

aging (Montemurro et al., 2019; Roe et al., 2007). While closely related, cognitive reserve is conceptually distinct from brain reserve: the latter refers to structural characteristics of the brain (e.g., brain volume), whereas cognitive reserve emphasizes functional and compensatory mechanisms that allow individuals to maintain performance despite neural decline (Stern, 2017). The present study is concerned specifically with this functional dimension. From a functional point of view, it has been hypothesized that, through new learning and cognitive training, the brain adapts and reorganizes, enhancing its ability to create scaffolds and form new neural pathways to compensate for age-related decline (Scaffolding Theory of Aging; Borella et al., 2010; Park & Reuter-Lorenz, 2009).

CR is a dynamic concept, shaped by a complex interplay between genetic factors and lifetime experiences that jointly contribute to its development, shaping interindividual differences across the lifespan (Kelly et al., 2017; Pettigrew & Soldan, 2019). For this reason, CR cannot be assessed directly, and current research has focused on proxies that may serve as indirect measures. Some of the most used sociobehavioral indices include years of education (Sharp & Gatz, 2011; Stern et al., 1994), work experience (considering both duration and career achievements or professional status (Karp et al., 2009; Kröger et al., 2008; Rusmaully et al., 2017)), and engagement in leisure activities (Fancourt et al., 2018). However, due to its dynamic nature and susceptibility to various exposures during the lifespan, relying on a single measure of CR can be misleading. It is advisable to use multiple measures jointly to obtain a more accurate quantification of an individual's CR (Stern et al., 2020). To address these limitations in the assessment of CR, Nucci and colleagues (Nucci et al., 2012) developed the Cognitive Reserve Index questionnaire (CRIq), which is divided into four sections: a preliminary section collecting basic demographic information (e.g., age, gender, marital status) and three specific sections focusing on education (i.e., years of formal education plus any additional training courses completed throughout life), work experience and professional achievements, and leisure activities (categorized into weekly, monthly, annual, and repetitive activities, covering a broad range of cognitively, socially, and physically engaging behaviors). For each specific section, a subscore can be derived: CRI_Education, CRI_Working Activity, and CRI_Leisure Activities. These are ultimately combined to generate a comprehensive Cognitive Reserve Index (CRI). Originally developed in Italian, the CRIq has been validated across multiple languages and contexts (Cebi & Kulce, 2022; Farran & Darwish, 2023; Maiovis et al., 2016) and has proven reliable in adults with brain injuries or neurodegenerative diseases (Bertoni et al., 2022; Garba et al., 2020). As with all proxy-based measures, the CRIq does not capture reserve directly but provides an indirect estimate based on lifetime experiences, which may differentially contribute to functional compensation across individuals.

Given its undeniable importance in maintaining an adequate quality of life, verbal communication is one of the most essential cognitive skills to preserve during healthy aging. A crucial issue regards the ability to produce samples of narrative discourse, which is needed to convey stories, experiences, memories, and plays a major role in conversations. Narrative discourse constitutes a goal-oriented and context-sensitive form of communication, requiring the integration of linguistic, cognitive, and pragmatic processes to convey information effectively in everyday interactions. For this reason, it represents an ecologically valid proxy of real-life communicative functioning. Growing evidence suggests that this ability significantly declines in healthy aging (Marini, 2022; Marini & Andreetta, 2016; Marini et al., 2025; Petriglia et al., 2025). For example, Marini et al. (2025) showed that during discourse production microlinguistic skills (involved in lexical selection and grammatical structuring) decline gradually from midlife onward, whereas macrolinguistic abilities (including coherence and informativeness) remain relatively stable until a marked deterioration after age 75. Of note, among older individuals the interindividual variability is significantly higher than among younger adults (Hivliu et al., 2025; Juncos-Rabadán et al., 2005; Marini et al., 2005). This suggests that the interplay between those cognitive mechanisms that support discourse production (e.g., working memory, attention, inhibitory control, and theory of mind) and linguistic processing skills may be affected by different levels of CR.

To date the effect exerted by CR on discourse production skills in older individuals is still largely unexplored. The few studies focusing on such effects on healthy aging have mainly considered CR effects on lexical selection. Oosterhuis et al. (2023) investigated the relationship between CR and word-finding difficulties, which is one of the most prominent problems frequently observed in older adults (Burke & Shafto, 2004). Ninety participants were divided in young adults aged 18 to 30 years, middle-aged participants aged 40-55 years, and older adults aged 65-80 years. On a picture naming task older participants scored lower than middle-aged and young adults. Higher CR in middle-aged participants predicted greater accuracy in naming and suggested that high levels of CR may be beneficial not only in old age but also in midlife. Indirect evidence about the potential effects of CR comes also from few studies focusing on language abilities on patients with neurodegenerative diseases. For example, Montemurro et al. (2019) showed that measures of CR were associated with pragmatic comprehension abilities in patients with Parkinson's Disease. To the best of our knowledge, no studies have yet explicitly assessed how narrative discourse production is influenced by CR in healthy aging. This study aims to address this gap by examining the impact of CR on narrative discourse production in healthy older adults.

Recent evidence suggests that during aging linguistic production weakenings parallel declines in cognitive skills necessary for producing structurally adequate and informative narrative discourse (Hilviu et al., 2025; Marini et al., 2025). We aimed to assess whether age-related declines in discourse production, as well as related cognitive skills, are influenced by CR. Importantly, although age-related changes provide the theoretical background for this work, the present study did not aim to model age effects directly. Rather, it examined how CR explains interindividual differences in cognitive and discourse performance among older adults of comparable age. To answer this **first research question**, we assessed the cognitive skills and narrative speech samples of 105 Italian-speaking healthy adults aged 61 to 84 years. All participants completed the CRIq (Nucci et al., 2012). We hypothesized that the Cognitive Reserve Index would significantly predict variance in all cognitive variables that contribute to efficiently produce a narrative discourse (i.e., the passive and active components of phonological working memory, sustained attention, and inhibitory control; **Hypothesis 1)**. We also predicted that the CRI would significantly affect variance in all those linguistic production skills that are prone to decline with healthy aging (i.e., fluency, phonological, semantic and grammatical accuracy, discourse organization, and ability to produce informative speech samples; **Hypothesis 2)**. Based on these predictions, participants were recruited and divided into groups with high, moderate, and medium CR. These groupings were adopted for analytical purposes and do not imply categorical distinctions, as both aging and CR are conceived as continuous dimensions. We further hypothesized that persons with high CR would outperform those with moderate CR, who, in turn, would perform better than those with average CR. We expected a graded advantage associated with higher levels of CR to be evident across all skills known to decline with aging and assessed in this study **(Hypothesis 3)**.

The **second research question** investigated whether different types of CR related to education levels, working activity and achievements, and leisure activities predict discourse production in healthy aging. We hypothesized that all three types of CRI would significantly affect discourse abilities (**Hypothesis 4**).

2. Methods

2.1. Participants

105 Italian-speaking healthy adults were included in the study (for a thorough description of the demographic, cognitive, and narrative characteristics of the whole sample please refer to Supplementary Tables S1, S2, S3, and S4). These participants were part of a larger cohort of healthy adult individuals recruited for a project aimed at standardizing the multilevel procedure of discourse analysis. Although cognitive reserve is conceptualized as a dynamic and continuous construct, participants were grouped according to CRIq (Nucci et al., 2012) thresholds to facilitate the examination of interindividual variability in cognitive and discourse measures. This group-based approach was adopted as an analytical strategy rather than to imply discrete categories of reserve, allowing for clearer comparison of performance patterns across different levels of CR. Specifically, they were divided into three groups with different levels of Cognitive Reserve Index (CRI).

Participants with a CRI>131 were included in the group with high CR, those ranging from 116 to 130 formed the cohort with moderate (defined as “medium-high” in Nucci et al., 2012) CR, and those between 85 and 115 were considered with medium (or average) CR. The group with high CR included 35 participants (mean age: 70 years [range: 61-84 years]; Females: 63%; right handedness: 40%). The cohort with moderate CR was formed by 35 adults (mean age: 72 years [range: 61-82 years]; Females: 49%; right handedness: 34%). The group with medium CR consisted of 35 individuals (mean age: 73 years [range: 61-82 years]; Females: 71%; right handedness: 46%).

Inclusion criteria were the absence of neurological or neuropsychiatric diseases, as well as above the cutoff performance on the Italian version of the Montreal Cognitive Assessment (MoCA; Conti et al., 2015) and the short version of the Token test (De Renzi & Vignolo, 1962). The three groups did not differ in age ($F(2, 102) = 2.855$, $p = .062$, $\eta^2 = .053$).

A Welch's ANOVA was conducted to compare the CRI across the three groups because Levene's test indicated a violation of homogeneity of variances ($p < .001$). The results showed a significant difference between groups ($F(2, 59.485) = 194.777$, $p_{Welch} < .001$; partial $\eta^2 = .832$). Post-hoc Games-Howell tests revealed that all groups differed from each other (all ps < .001). Significant differences were also observed for on each of the three subscales of the CRIq: Work activity ($F(2, 102) = 37.171$, $p < .001$, $\eta^2 = .422$); Leisure activities ($F(2, 102) = 26.938$, $p < .001$, $\eta^2 = .346$), and continuous Education ($F(2, 102) = 49.229$, $p < .001$, $\eta^2 = .491$). Tukey's post-hoc analyses confirmed that the three groups were different from each other on all subscales of cognitive reserve (all ps < .001 but for leisure activities where the contrast between participants with high CR and those with moderate CR was $p < .007$). Regarding their years of schooling, the participants with high CR had more years of education (M = 16.37, SD = 2.87) than those with moderate CR (M = 12.31, SD = 2.91) who had more years than adults with medium CR (M = 8.57, SD = 2.79).

The three groups did not differ in gender distribution with a large effect size measured as Cramérs V [X^2 (2, N = 105) = 3.921, p = .141, Cramérs V = .193]. The three groups produced a similar total number of words ($F(2, 102) = 1.038$, $p = .358$, $\eta^2 = .020$) (See Table 1).

This study was not preregistered. The Ethical Committee of the University of Udine approved the study (Protocol CGPER-2024-02-27-01). All research was conducted in accordance with relevant guidelines and regulations, in compliance with the Declaration of Helsinki. All participants signed an informed consent form and provided written permission to participate in this study.

Table 1

General information about the three participant groups. Data are presented as means with standard deviations in brackets. For gender, the number of females in each group is reported along with the corresponding percentage in brackets. The Asterisk (*) indicates significant group differences. The last column reports the results of post-hoc tests where significant effects were found. *Legend*: CRI: Cognitive Reserve Index; CRI_Education: Education level as measured using the Cognitive reserve index questionnaire; CRI_Working activity: Work experience as measured using the Cognitive reserve index questionnaire; CRI_Leisure activities: Leisure activities as measured using the Cognitive reserve index questionnaire; Mod: Moderate CR.

	High CR (N = 35)	Moderate CR (N = 35)	Medium CR (N = 35)	Post-hoc
Age	69.51 (5.60)	72.26 (7.04)	72.89 (6.11)	–
Sex	F = 22 (63%)	F = 17 (49%)	F = 25 (71%)	–
CRI *	145.36 (11.18)	121.73 (4.35)	100.91 (7.82)	High > Mod > Medium
CRI_Education*	127.14 (13.97)	110.54 (10.54)	99.95 (9.62)	High > Mod > Medium
CRI_Working activity*	127.78 (16.24)	107.74 (17.41)	93.26 (16.80)	High > Mod > Medium
CRI_Leisure activities*	147.33 (24.84)	130.48 (24.34)	107.86 (17.87)	High > Mod > Medium
Total production of words	512.23 (230.20)	535.66 (391.04)	440.63 (204.83)	–

2.2. Cognitive assessment

The cognitive assessment targeted key components of verbal working memory, sustained attention, and inhibitory control.

To evaluate phonological working memory, both its passive and active components were assessed using the Forward and Backward Digit Recall tasks (Monaco et al., 2013). The Forward Digit Recall task measures the phonological buffer (passive storage) by requiring participants to repeat sequences of digits spoken by the examiner. The sequences start at 3 digits and progressively increase to 9 digits, with each length presented twice (e.g., 7-9-2 and 6-1-8). Participants earn 1 point for each correctly repeated sequence, and their span corresponds to the longest sequence they can correctly recall at least once. The Backward Digit Recall task evaluates the active, "working" component of verbal working memory (Baddeley, 2012). Participants must repeat digits sequences in reverse order, with sequence lengths ranging from 3 to 8 digits. As in the Forward Digit Recall task, each length is presented twice (e.g., 8-4-9 and 1-5-3), and 1 point is awarded for each correct response. The span is determined by the longest sequence successfully recalled at least once. This task specifically assesses the ability to manipulate verbal information within working memory.

Sustained attention and inhibitory control were assessed using the Trail Making Test (TMT; Giovagnoli et al., 1996), which consists of two parts. In Part A, participants connect 25 numbered circles in ascending order (1 to 25) as quickly as possible. In Part B, the circles contain both numbers (1 to 13) and letters (A to N), requiring participants to alternate between them in ascending order (e.g., 1-A-2-B, etc.). The time taken to complete Part A reflects sustained attention, while Part B assesses both sustained attention and inhibitory control. The B-A difference score is commonly used as a specific measure of inhibitory control (Arbuthnott & Frank, 2000) and was adopted in the present study to emphasize inhibitory and set-shifting components while minimizing shared variance with basic processing speed and sustained attention.

2.3. Assessment of narrative discourse production

The narrative assessment was based on speech samples elicited using five picture-based stimuli: two single-picture scenes ("Picnic" by Kertesz, 1982; "Cookie Theft" by Goodglass & Kaplan, 1972) and three cartoon-picture sequences ("Flowerpot" by Huber & Gleber, 1982; "Quarrel" by Nicholas & Brookshire, 1993; "Nest Story" by Paradis & Libben, 1987). To mitigate the impact of short-term memory limitations, participants had continuous access to the images while narrating their stories.

The stimuli were presented in a randomized order on a laptop screen facing the participant, ensuring the examiner had no visual reference. Participants were informed that the examiner was unfamiliar with the images and were instructed to describe the scenes clearly, avoiding ambiguity.

All narratives were audio-recorded. Speech samples were transcribed using a semi-automated system developed by one of the authors (F.P.), which formatted the transcripts for direct integration into the database. Transcriptions were manually reviewed and corrected, if necessary, by comparing them with the original audio recordings. The final transcripts documented story duration (in seconds), false starts, phonological errors, and neologisms. Utterances were segmented following the four main criteria established in Marini et al. (2011). The transcripts were then analyzed according to a multilevel framework evaluating Productivity, Lexical and Grammatical Accuracy, and Discourse Organization (Marini et al., 2011).

Productivity was assessed by counting phonologically well-formed words, excluding false starts, phonological paraphasias, and neologisms. This count was also used to determine Speech Rate (words per minute).

Lexical accuracy was measured through percentages of Phonological errors and Semantic errors. Phonological errors are excluded from the count of phonologically well-formed words and include false starts, phonological paraphasias, and neologisms. False starts occur when a speaker initiates but does not complete a word. Phonological paraphasias involve phoneme substitutions, omissions, insertions, or inversions (e.g., *sable, *tabe, *tabiler, *battle instead of table). Neologisms are unintelligible, novel words. Of note, participants did not produce phonological paraphasias or neologisms. Therefore, the percentage of phonological errors was replaced by a percentage of false starts computed by dividing false starts by the total number of speech units that were not counted as phonologically well-formed words, then multiplying by 100.

Semantic errors include both semantic and verbal paraphasias. Semantic paraphasias occur when a word is replaced with another semantically related but incorrect word (e.g., "tree" instead of "flower"). Verbal paraphasias occur when a word is replaced with another semantically unrelated word (e.g., "tree" instead of "chair"). The percentage of Semantic Errors was calculated by dividing the number of these errors by the total number of words and multiplying the result by 100.

Grammatical skills were evaluated by calculating the percentage of Complete Sentences, which includes all grammatically correct sentences, excluding those with omissions or morphological errors. This was computed by dividing the number of complete sentences by the total number of utterances and multiplying this result by 100.

Discourse organization was assessed by calculating the percentages of Errors in Local and Global Coherence. Local coherence refers to the logical connection between consecutive utterances. Errors occur when utterances lack referential clarity or contain abrupt, incomplete shifts (e.g., "Children are trying to steal some …/The woman is washing the dishes"). The percentage of Local Coherence Errors was computed by dividing these errors by the utterances and multiplying by 100. Global coherence evaluates the overall thematic consistency of the narrative. Errors include tangential utterances that stray from the main narrative (e.g., "The tree is all green/Now in my garden the tree has no leaves"), conceptually incongruent utterances that introduce irrelevant or contradictory content (e.g., "The tree is all green/A plane is flying over it" in the Nest story), filler utterances that provide no new information (e.g., "The tree is all green/I like this tree"), and repeated utterances (e.g., "The tree is all green/It's all green"). The percentage of Errors in Global Coherence was computed by dividing the total of such errors by the total number of utterances and multiplying by 100.

Finally, a key measure of communicative effectiveness, percentage of Lexical Informativeness, was derived by dividing the number of informative words by the number of words, then multiplying the result by 100. Words were deemed informative if they were free from phonological, semantic, or morphological errors, were not fillers or repetitions, and were contextually relevant even if they were in utterances with cohesion or local coherence issues. This metric reflects an individual's ability to use precise and contextually appropriate words in narrative speech (Marini & Urgesi, 2012).

Discourse production measures were aggregated across the five stimuli to obtain stable estimates of participants' narrative abilities. Indeed, Brookshire and Nicholas (1994) recommend that discourse samples consist of at least 300 words to obtain reliable analyses. One way to address this issue is to combine multiple picture stimuli. This approach reduces stimulus-specific variability and is consistent with previous work adopting multistimulus discourse paradigm in aging research (e.g., Bryant et al., 2016; D'Ortenzio et al., 2025; Kong et al., 2025; Petriglia et al., 2025).

All assessments were conducted by two expert raters (F.P. and G.G.) under the supervision of A.M. Interrater reliability was evaluated using the Kappa statistic, with a reliability threshold of $k \geq .70$ considered acceptable (Carletta, 1996). Agreement between the two raters was almost perfect across all measures and substantial for lexical informativeness (see Marini et al., 2025).

2.4. Statistical analyses and sample size estimation

When planning this study, we aimed to assess the potential effects of cognitive reserve on cognitive and narrative skills in healthy adults. All analyses were conducted using JASP. Although a small numerical trend toward older age was observed in the moderate and medium CR groups, age differences did not reach statistical significance and were therefore not included as covariate in subsequent analyses, in line with previous work showing that individuals with a mean age of 69-72 years do not differ in their cognitive or narrative production skills (e.g., Hilviu et al., 2025; Marini et al., 2025).

To explore the **first research question** (i.e., the relation between Cognitive Reserve Index and cognitive and narrative measures and potential differences in cognitive and narrative production skills in the three groups of participants), a series of multiple linear regression analyses were preliminarily conducted with the CRI as predictor and the cognitive and linguistic measures as dependent variables. For those variables where the CRI was a significant predictor, potential differences between the three groups of participants were explored with ANOVAs with Group (high, moderate, and medium CR) as the independent variable and the cognitive and narrative measures as dependent variables. Before running the ANOVAs, preliminary controls were performed. When the Levene's test for equality of variances indicated a violation of homogeneity of variances, Welch's ANOVAs were conducted. Effect sizes were calculated in terms of partial eta squared. When the general effect of group was significant, specific group differences were further explored with paired comparisons using Tukey or Games-Howell tests where appropriate.

To explore the **second research question** (i.e., the effect potentially exerted by each of the three subscores derived by the CRIq on the cognitive and narrative variables) a series of multiple linear regressions were performed with the three subscores (CRI_Education, CRI_Work activity, and CRI_Leisure activites) as independent variables, and the cognitive and narrative variables of the whole group of participants as dependent variables.

2.5. Sample size estimation

To address **the first research question**, for the multiple linear regression analyses we followed the recommendations by Stevens (1996) who suggests that 15 subjects per predictor are needed for a reliable equation. Even using the more stringent methodology by Tabachnick and Fidell (2001) who suggest using formula $N > 50 + 8m$ (where m = number of independent variables), the minimum number of participants should be 58. Given that the study included 105 participants, the sample size requirements were met and even exceeded for this preliminary exploration. To further assess the potential impact of CRI on cognitive and narrative skills, an *a priori* power analysis using G*Power 3.1 software (Faul et al., 2009) was conducted. In this regard, previous research on narrative skills on a large cohort of healthy adults aged 20 through 89 (Marini et al., 2025) reported effect sizes for age-related group differences, ranging from .135 to .572 (Cohen's F) for the age ranges considered in the present study. Therefore, the power analysis was performed with 3 groups, assuming an average effect size of .354 and an alpha level of .05. The estimated sample size to achieve a statistical power of >.90 was 105 participants.

To explore **the second research question**, following the recommendations by Stevens (1996), as we expected to have three independent variables, the required minimum number of participants was 45 participants. Even using the more stringent methodology by Tabachnick and Fidell (2001) the minimum number of participants should have been 74. Given that the study included 105 participants, the sample size requirements were met and even exceeded also to explore the second research question.

3. Results

3.1. Preliminary inspection of the impact of CR on cognitive and narrative variables

Linear regression analyses were preliminarily conducted with CRI as predictor and the cognitive and narrative measures as dependent variables. Alpha level was adjusted were appropriate using Bonferroni corrections for multiple comparisons: .025 (.05/2 variables) for measures assessing phonological working memory (Forward and Backward Digit Recall), sustained attention and inhibitory control (TMT_A and TMT_B-A), productivity (Words_Mean and Speech rate), lexical accuracy (% False starts and % Semantic errors), and discourse organization (% Local coherence errors and % Global coherence errors). For % Complete sentences and % Lexical Informativeness alpha level was set at $p < .05$.

The regression model was not significant for Forward Digit Recall ($F(1, 104) = 2.557$, $p = .113$; $R^2 = .024$), which assesses the passive component of the phonological buffer. However, it was significant for Backward Digit Recall ($F(1, 104) = 6.795$, $p < .010$; $R^2 = .062$), assessing its active component, with CR explaining 6.2% of total variance. CR significantly predicted also 6.1% of the variance in TMT_A ($F(1, 104) = 6.676$, $p < .011$; $R^2 = .061$) and 7.2% of that observed in TMT_B-A ($F(1, 104) = 7.986$, $p < .006$; $R^2 = .072$) assessing sustained attention and inhibitory control, respectively.

Considering productivity, the model was not significant for Words_Mean ($F(1, 104) = 2.381$, $p = .126$; $R^2 = .023$) and Speech rate ($F(1, 100) = 1.963$, $p = .164$; $R^2 = .019$). However, CR significantly predicted 7.6 % the variance in % False starts ($F(1, 103) = 8.425$, $p < .005$; $R^2 = .076$) and 6.2% in % Semantic errors ($F(1, 104) = 6.791$, $p < .001$; $R^2 = .062$). The model assessing the influence of CR on the % of Complete Sentences was not significant ($F(1, 104) = 3.276$, $p = .073$; $R^2 = .031$). CR significantly predicted 5.4 % of total variance in % Local coherence errors ($F(1, 104) = 5.931$, $p < .017$; $R^2 = .054$) and 8% in % Global coherence errors ($F(1, 104) = 8.987$, $p < .003$; $R^2 = .080$). Finally, CR predicted 6.1% of variance in % Lexical informativeness ($F(1, 104) = 6.737$, $p < .011$; $R^2 = .061$).

Notably, CRI was not associated with forward digit recall, productivity measures (mean words and speech rate), or grammatical accuracy (% complete sentences), suggesting that CR effects were selective rather than uniform across cognitive and linguistic domains.

3.2. Cognitive skill differences associated with CRI

Group differences on those cognitive variables that were affected by CR were explored with one-way ANOVAs with Group as independent variable and the cognitive measures as dependent variables (Table 2). Alpha level was set at $p < .05$ for backward digit recall and at $p < .025$ after Bonferroni correction for multiple comparisons for scores from TMT_A and TMT_B-A.

The three groups differed on backward digit recall ($F(2, 102) = 3.522$, $p < .033$; partial $\eta^2 = .065$). Tukey's post-hoc analysis showed that participants with high CR had higher spans than those with medium CR ($p < .034$). No group-related differences were found for the other contrasts (high vs. moderate CR: $p = .837$; moderate vs. medium CR: $p = .127$).

The three groups differed on measures assessing sustained attention (TMT_A: $F(2, 102) = 4.770$, $p < .010$; partial $\eta^2 = .086$) and inhibitory control (TMT_B-A: $F(2, 57.621) = 5.227$, $p_{Welch} < .008$; partial $\eta^2 = .077$). Participants with high CR completed TMT_A earlier than persons with both moderate ($p < .047$) and medium CR ($p < .013$) who did not differ from each other ($p = .882$). Regarding TMT_B-A the participants with high CR had significantly lower inhibition costs than subjects with medium CR ($p < .016$). However, no significant differences were observed between participants with high and moderate CR ($p = .154$) nor between those with moderate and medium CR ($p = .441$).

Across cognitive measures, group effects were driven primarily by differences between the high CR group and the other groups, whereas moderate and medium CR participants did not reliably differ. This pattern is consistent with a threshold-like rather than graded relationship between CR level and cognitive performance.

3.3. Narrative skill differences associated with CRI

Group differences on narrative variables were explored with one-way ANOVAs with Group as independent variable and the narrative measures as dependent variables (Table 3). Alpha level was corrected using Bonferroni correction for multiple comparisons where appropriate. Specifically, p-value was set at $p < .025$ for lexical errors (% False starts and % Semantic errors), and macro-linguistic errors (% Local coherence errors, and % Global coherence errors). Alpha level was set at $p < .05$ for % Lexical Informativeness.

A significant difference was found for % false starts ($F(2, 63.362) = 6.965$, $p_{Welch} < .002$; partial $\eta^2 = .108$). Post-hoc Games-Howell tests revealed that participants with high CR produced fewer errors than those with medium CR ($p < .003$). Similarly, a significant difference was found for % semantic errors ($F(2, 64.825) = 4.028$, $p_{Welch} < .022$; partial $\eta^2 = .074$). Post-hoc Games-Howell tests revealed that participants with high CR produced significantly fewer errors than those with medium CR ($p < .014$). A group-related difference was found in % Local coherence errors ($F(2, 63.699) = 5.707$, $p_{Welch} < .005$; partial $\eta^2 = .077$). The Games-Howell post-hoc tests revealed that participants with high CR produced significantly fewer local coherence errors than those with moderate ($p < .022$) and medium CR ($p < .026$). A significant effect of CR was found on % of Global coherence errors ($F(2, 102) = 6.247$, $p < .003$; partial $\eta^2 = .109$). Tukey's post-hoc tests revealed that participants with high CR produced significantly fewer errors of global coherence than those with moderate ($p < .018$) and medium CR ($p < .004$). Considering that global coherence errors could be ascribed to four different types of errors, further analyses explored the possibility that differences on this measure could depend on differences in the production

Table 2

Results of cognitive assessment in the three groups. Data are expressed as means with standard deviations in brackets. The Asterisk (*) indicates significant group differences. In the last column the results of post-hoc tests are reported when significant contrasts were found. *Legend*: *Legend*: CR: Cognitive reserve; TMT_A = Trail-Making Test, Part A; TMT_B-A = Cost of inhibitory control, calculated by subtracting the time to complete TMT_Part A from the time to complete TMT_Part B; Mod: Moderate CR.

	High CR (N = 35)	Moderate CR (N = 35)	Medium CR (N = 35)	Post-hoc
Backward digit recall*	5.29 (1.76)	5.03 (1.69)	4.14 (2.19)	High > Medium
TMT_A (seconds) *	38.67 (14.50)	49.60 (21.74)	51.77 (19.99)	High < Mod = Medium
TMT_B-A (seconds) *	52.77 (23.80)	70.09 (48.83)	87.11 (65.87)	High < Medium

Table 3
Narrative assessment results. Data are presented as means with standard deviations in brackets. The Asterisk (*) indicates significant group differences. In the last column the results of post-hoc tests are reported when significant contrasts were found. *Legend*: CR: Cognitive reserve; Mod: Moderate CR.

	High CR (N = 35)	Moderate CR (N = 35)	Medium CR (N = 35)	Post-hoc
% False starts*	.79 (.59)	1.20 (.85)	1.49 (1.02)	High < Medium
% Semantic errors*	1.35 (1.51)	2.06 (1.72)	2.79 (2.87)	High < Medium
% Local coherence errors*	23.47 (11.40)	33.89 (19.45)	33.22 (18.32)	High < Mod = Medium
% Global coherence errors*	8.05 (4.91)	12.59 (7.38)	13.43 (7.87)	High < Mod = Medium
% Lexical informativeness*	83.90 (5.51)	78.17 (78.06)	78.06 (9.63)	High > Mod = Medium

of tangential, incongruent, repeated, or filler utterances (Table 4). Alpha level was set at $p < .013$ after Bonferroni correction for multiple comparisons (.05/4 variables). The three groups did not differ in the production of Tangential utterances ($F(2, 58.710) = .970$, $p_{Welch} = .385$; partial $\eta^2 = .025$), Filler utterances ($F(2, 102) = 2.510$, $p = .086$; partial $\eta^2 = .047$), and Repeated utterances ($F(2, 64.396) = 4.593$, $p_{Welch} = .014$; partial $\eta^2 = .062$). A group related difference was found for the production of Conceptually incongruent utterances ($F(2, 60.807) = 6.580$, $p_{Welch} < .003$; partial $\eta^2 = .075$). The Games-Howell post-hoc tests revealed that participants with high CR produced significantly fewer conceptually incongruent utterances than those with medium CR ($p < .003$). No significant differences were found between participants with high and moderate ($p = .146$) or moderate and medium CR ($p = .701$). Finally, the group-related difference in % Lexical informativeness was significant ($F(2, 62.447) = 7.361$, $p_{Welch} < .001$; partial $\eta^2 = .093$). Post-hoc Games-Howell tests showed that participants with high CR produced significantly more informative words than those with moderate ($p < .018$) and medium CR ($p < .016$). No significant differences were found between participants with moderate and medium CR in any linguistic measure, mirroring the cognitive results and further supporting a threshold-like pattern in which high CR is associated with more preserved discourse organization and informativeness.

Importantly, the absence of CR-related differences in productivity and grammatical accuracy indicates that CR effects were specific to lexical accuracy and discourse-level organization and effectiveness, rather than reflecting a generalized increase in output or syntactic well-formedness. Furthermore, the finding that group differences were driven specifically by conceptually incongruent utterances, rather than tangential, repeated, or filler utterances, provides a fine-grained insight into the nature of discourse disruptions associated with lower CR.

3.4. Do different types of CR differentially predict cognitive and narrative processing in healthy aging?

To explore whether specific types of CR affected those cognitive and linguistic variables that have been found different across the three groups, multiple linear regression analyses were conducted with the three indices derived by the CRIq (i.e., CR_Education, CR_Working activity, and CR_Leisure activities) as predictors and the cognitive and narrative measures as dependent variables. Also in this case, p-value was set at .025 for measures of sustained attention and inhibitory control (TMT_A and TMT_B-A), Lexical errors (% False starts and % Semantic errors), and macrolinguistic errors (% Local coherence errors and % Conceptually incongruent utterances). Alpha level was set at .050 for % Lexical Informativeness and Backward digit recall.

Considering the participants' cognitive performance, the model predicted 13.3 % of the variance in Backward Digit Recall ($F(3, 104) = 5.158$, $p < .002$; $R^2 = .133$). Only CR_Working activity was significantly associated with this variable ($\beta = .397$; $p < .001$; part correlation coefficient: .312): this predictor uniquely explained 9.7% of the variance in Backward Digit Recall. The model was not significant for TMT-A ($F(3, 104) = 2.761$, $p < .046$; $R^2 = .076$). None of the standardized coefficients reached significance. Therefore, collectively the predictors explained a significant portion of the variance in the dependent variable, but, individually, none of them had a uniquely significant contribution when controlling for the others. The three factors explained 15.6% of the variance in TMT_B-A ($F(3, 104) = 6.218$, $p < .001$; $R^2 = .156$). Only CR_Education was significantly associated with TMT_B-A ($\beta = .376$; $p < .026$; part correlation coefficient: −.207), accounting for 4.3% of total variance.

Regarding lexical accuracy, the regression model was not significant for the production of False starts ($F(3, 103) = 3.165$, $p < .028$; $R^2 = .087$). The regression model explained 9.8% of the variance in the production of semantic errors ($F(3, 104) = 3.642$, $p < .015$; $R^2 = .098$). Only CR_Working activity was significantly associated with this measure ($\beta = -.252$; $p < .038$; part correlation coefficient: −.198) and this predictor uniquely explained 3.9% of the variance in such errors.

Considering macrolinguistic skills, the three predictors explained 20.6% of the variance in the production of local coherence errors

Table 4
Analytic assessment of the four types of global coherence errors. Data are presented as means with standard deviations in brackets. The Asterisk (*) indicates significant group differences. In the last column the results of post-hoc tests are reported when significant contrasts were found. *Legend*: CR: Cognitive reserve.

	High CR (N = 35)	Moderate CR (N = 35)	Medium CR (N = 35)	Post-hoc
Tangential utterances	.049 (.206)	.091 (.303)	.308 (1.180)	–
Conceptually incongruent utterances*	.646 (1.130)	1.535 (2.503)	1.969 (1.984)	High < Medium
Repeated utterances	2.471 (2.078)	4.107 (3.748)	4.029 (2.865)	–
Filler utterances	4.887 (3.608)	6.854 (4.857)	7.122 (5.069)	–

($F(3, 104) = 8.717$ $p < .001$; $R^2 = .206$), 9.4% of the variance in the production of conceptually incongruent utterances ($F(3, 104) = 3.501$, $p < .018$; $R^2 = .094$), and 7,5 % of the variance in the production of informative words ($F(3, 104) = 2.719$, $p < .049$; $R^2 = .075$). All three subscores of the CRI were significantly associated with % Local coherence errors (CR_Education $\beta = -.239$; $p < .040$; part correlation coefficient: −.185; CR_Working activity $\beta = -.243$; $p < .034$; part correlation coefficient: −.191; CR_Leisure activities $\beta = .184$; $p < .048$; part correlation coefficient: .178). Specifically, CR_Education uniquely predicted 3.4%, CR_Work 3.6%, and CR_Leisure activities 3.2% of total variance, respectively. Regarding the production of Conceptually incongruent utterances and Informative words, none of the standardized coefficients reached significance suggesting that, individually, none of the factors had a uniquely significant contribution when controlling for the others. This suggests shared variance among CR proxies for Conceptually incongruent utterances and Informative words.

4. Discussion

This study explored two key research questions. First, we aimed to determine whether cognitive reserve plays a role in mitigating the decline in discourse production and related cognitive skills in healthy aging. Second, we investigated whether different life experiences (specifically, in terms of lifelong education, work, and leisure activities) contribute differently to shaping discourse production skills and cognitive abilities in later life. We hypothesized that the Cognitive Reserve Index would account for significant variance across all tasks assessing cognitive abilities (**Hypothesis 1)**. Additionally, we expected that CR would influence a range of discourse measures, including speech rate, measures of lexical and grammatical accuracy as well as local and global coherence errors and lexical informativeness (**Hypothesis 2)**. Therefore, hypotheses 1 and 2 addressed complementary levels of analysis: Hypothesis 1 focused on cognitive mechanisms supporting discourse production, whereas Hypothesis 2 examined their linguistic and discourse-level outcomes. We also anticipated that the impact of CR would vary depending on its level: individuals with higher CR were expected to outperform those with moderate CR who, in turn, would perform better than those with medium CR **(Hypothesis 3)**. Finally, we proposed that all three subcomponents of CR would serve as predictors for the full set of dependent variables (**Hypothesis 4**). The findings of this study largely confirmed some of these expectations while also revealing some interesting nuances. As the present study examined how CR explains interindividual differences in cognitive and discourse performance among older adults of comparable age, the present findings should be interpreted as reflecting CR-related variability within later adulthood, rather than as evidence of CR moderating age-related decline per se.

Our results supported the utility of the CRIq (Nucci et al., 2012) in capturing the three major forms of CR and demonstrated a selective protective role of CR on specific cognitive and discourse-level processes in healthy aging. This aligns with previous hypotheses regarding the role of CR in protecting cognitive functioning (Stern et al., 2020) and provides the first evidence of its role in preserving discourse production skills.

CR significantly accounted for variance in nearly all cognitive tasks, except for the forward digit recall task, thereby lending partial support to Hypothesis 1 and indicating a selective rather than uniform influence. Similarly, while CR was found to influence most narrative measures, speech rate and the percentage of complete sentences did not follow the expected pattern, leading to partial support for Hypothesis 2 and reinforcing the idea that CR operates through specific cognitive mechanisms rather than as a global buffer against decline. When considering the impact of CR levels, our findings suggest that individuals with high CR benefited from a consistent advantage across cognitive and narrative measures. In contrast, those with average to moderate CR performed similarly with no clear differentiation. This pattern only partially aligned with our initial prediction in Hypothesis 3. Finally, when analyzing the effects of CR subtypes, we observed selective influences on specific measures: CR_Working primarily influenced phonological working memory skills and the production of semantic errors, while CR_Education significantly affected inhibitory control. Interestingly, all three CR subtypes played a role in predicting local coherence errors. However, for most other measures, only a general effect of CR emerged rather than specific contributions from its subcomponents, leading to partial support for **Hypothesis 4**. Taken together, these findings provide valuable insights into the relationship between cognitive reserve and discourse production in healthy aging. They reinforce the idea that cognitive reserve serves as a protective factor, though its influence is not uniform across all measures.

From a cognitive perspective, the results provided partial support for **Hypothesis 1**. Considering verbal working memory, while CR did not significantly predict variance in the passive phonological buffer (forward digit recall), it accounted for 6.2% of the variance in the active component (backward digit recall). This provides indirect support for current working memory models, which propose that phonological working memory consists of two distinct yet interacting systems: a phonological buffer for temporary information storage and an active mechanism for maintaining and manipulating this information (Baddeley, 2012). Growing evidence suggests that verbal working memory declines with healthy aging (Monov et al., 2024) and that this decline may differentially impact its two components (Marini et al., 2025). Specifically, Marini et al. (2025) found that the storage capacity of the phonological buffer, as measured by the forward digit recall test, remained largely intact across the adult lifespan. This likely explains why CR did not influence the passive component of phonological working memory, which remained stable in our sample. Its active component, which is responsible for manipulating stored information, is more susceptible to aging and in this case the effect of CR was significant reinforcing prior evidence on its role in mitigating cognitive aging.

Sustained attention (measured by TMT_A) and inhibitory control (assessed via TMT_B-A) have been shown to decline with aging (Harada et al., 2013; Wecker et al., 2000). In this case CR significantly predicted variance in sustained attention (6.1% in TMT_A) and inhibitory control (7.2% in TMT_B-A). Therefore, individuals with higher CR levels exhibit better cognitive control and attentional capacities, supporting the protective role of CR in maintaining executive functions despite aging. This aligns with previous studies showing that CR contributes to the maintenance of executive functions in older adults, likely allowing them to recruit alternative neural networks to compensate for cognitive decline (Tucker & Stern, 2011).

From a discourse perspective, the results provided partial support for **Hypothesis 2**. CR did not significantly predict variance in productivity measures or grammatical accuracy. However, it significantly predicted lexical accuracy, as evidenced by its contribution to variance in false starts (7.6%) and semantic errors (6.2%). Furthermore, CR significantly predicted local (5.4%) and global coherence errors (8%), and lexical informativeness (6.1%). These findings align with previous evidence suggesting that it helps preserve linguistic abilities in older adults (Oosterhuis et al., 2023), particularly in tasks with lower cognitive demands. Previous studies (Hilviu et al., 2025; Marini et al., 2005) suggest that individuals with higher CR are better able to maintain lexical access and coherence in speech. Our results also align with Marini et al. (2005), who found that individuals with higher education levels produced more coherent and informative discourse, supporting the notion that CR plays a crucial role especially in mitigating declines in macrolinguistic aspects of discourse production. The absence of significant effects on speech rate and sentence completeness further supports the idea that CR may exert its primary influence on higher-order discourse organization rather than basic language production mechanisms. Group comparisons further elucidated these patterns. Participants with high CR produced significantly fewer false starts and semantic errors than those with average CR, although no significant differences were observed between groups with moderate and medium CR. Similarly, individuals with high CR produced fewer local coherence errors compared to both moderate and medium CR groups. A similar pattern was observed for global coherence errors, particularly in the production of conceptually incongruent utterances, where individuals with high CR outperformed those with medium CR. Additionally, high CR participants produced significantly more informative words than those with moderate and medium CR, underscoring the role of CR in maintaining discourse efficiency. Importantly, in line with the Scaffolding Theory of Aging (Borella et al., 2010; Park & Reuter-Lorenz, 2009) this apparent preservation may reflect adaptive compensatory strategies rather than mere resistance to decline. Individuals with higher CR may flexibly recruit executive and monitoring resources to maintain discourse coherence and informativeness despite age-related vulnerabilities in underlying linguistic processes.

Contrary to our expectations, we did not find a gradual benefit of CR across groups. We had hypothesized that CR would exert a graded effect, with high-CR participants outperforming those with moderate CR, who, in turn, would outperform those with medium CR across all measured variables (**Hypothesis 3**). However, our results only partially supported this hypothesis. For microlinguistic variables assessing lexical selection (semantic errors) and lexical production accuracy (false starts), no significant differences emerged between participants with high and moderate CR. However, those with high CR outperformed participants with medium CR. A similar pattern was observed in two cognitive measures crucial for efficient lexical selection and production (i.e., phonological working memory and inhibitory control). Together, these findings suggest a threshold-like effect of CR, whereby only relatively high levels of reserve confer consistent advantages, rather than a graded benefit across the entire continuum. This interpretation reconciles the group-based findings with the conceptualization of CR as a continuous construct, while acknowledging that its functional impact may become most evident beyond a critical level.

Hypothesis 4 was also not fully supported, as different types of CR influenced cognitive and linguistic variables in complex ways. Working activity significantly predicted variance in backward digit recall (9.7%) and semantic error rates (3.9%), while education uniquely predicted inhibitory control (TMT_B-A, 4.3%). All three CR components significantly contributed to local coherence errors, with education, working activity, and leisure activities explaining 3.4%, 3.6%, and 3.2% of total variance, respectively. However, no single CR component significantly predicted lexical informativeness or the production of conceptually incongruent utterances when controlling for the other factors. Importantly, the absence of unique effects for individual CR components does not imply a lack of relevance; rather, it likely reflects the substantial shared variance among life-experience proxies that jointly contribute to discourse-level resilience. This highlights the differential contributions of CR components to discourse skills. While **formal education** may enhance inhibitory control, **working experience** appears to strengthen the active component of the phonological loop and lexical selection. **Leisure activities**, often associated with social engagement and cognitive stimulation, may play a more complex role in maintaining discourse coherence. This confirms that CR is a multidimensional construct, with different life experiences shaping distinct aspects of cognitive-linguistic resilience in aging. These findings underscore the need for future studies to include multiple measures of CR, as relying on a single indicator (e.g., education) may provide only a limited perspective on its effects.

This study has some limitations. First, the cross-sectional design prevents from drawing causal conclusions about the relationship between CR and discourse abilities. Longitudinal studies are necessary to determine whether CR actively protects discourse production skills against decline over time or merely reflects accumulated advantages. Second, the sample size, though sufficient for detecting group differences, may limit the generalizability of our findings to broader populations. Future research should aim to replicate these results with larger cohorts. Third, while the CRIq provides a comprehensive measure of CR, it may not capture all aspects of cognitive reserve, such as social engagement or bilingualism (Bialystok, 2021), which have been shown to influence cognitive aging. Finally, the study was not preregistered; therefore, and some of the analyses should therefore be considered exploratory. Future preregistered and longitudinal studies will be necessary to confirm and extend the present findings.

In conclusion, this study highlights the impact of CR on cognitive and discourse abilities in older adults. While CR does not uniformly affect all linguistic skills, its influence on executive functions and discourse coherence underscores its role as a critical factor in maintaining effective discourse-level communication in healthy older adults. Specifically, CR was associated with lexical accuracy, coherence, and informativeness, rather than with productivity or grammatical completeness. These findings emphasize the need for interventions that support lifelong learning, occupational engagement, and social participation to enhance higher-order communicative efficiency rather than all aspects of language production in aging populations. Finally, the results of the present study highlight the selective nature of CR effects, underscoring that reserve supports higher-order communicative efficiency rather than all aspects of language production.

CRediT authorship contribution statement

Andrea Marini: Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization. **Francesco Petriglia:** Writing – review & editing, Software, Methodology, Formal analysis, Data curation. **Silvia D'Ortenzio:** Methodology, Investigation, Formal analysis, Data curation. **Giulia Gasparotto:** Methodology, Investigation, Formal analysis, Data curation. **Ilaria Gabbatore:** Writing – review & editing, Methodology, Formal analysis, Data curation, Conceptualization.

Data availability statement

The supporting data for this research and the JASP file with the analyses are openly available for download at the following OSF link: https://osf.io/7jszp/?view_only=fdb4a7e5e3d2427e832f14b7b6cfe61f.

Declaration of generative AI and AI-assisted technologies in the writing process

This manuscript was proofread and edited for language clarity using CHATGPT. All scientific content, interpretations, and conclusions remain the responsibility of the authors.

Funding

This research was supported by PRIN 2022 PNRR, Prot. n. P2022M9JCM, project title: "Standardization of the Multilevel prOcedure for discOurse analysis and Training program for narrative production in Healthy adults - SMOOTH". Avviso pubblico n. 1409 del September 14, 2022 – PRIN 2022 PNRR M4C2 Inv. 1.1. Ministero dell'Università e della Ricerca (Financed by EU, NextGenerationEU) – CUP G53D23007250001.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to thank SLTs Cinzia Baldanzi and Chiara Vitali (IRCCS Don Gnocchi), Giulia Fusari (University of Pavia), and Professors Francesco Ferretti and Ines Adornetti (Roma Tre University) for their precious help during the recruitment phase.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jneuroling.2026.101315.

Data availability

The link to the OSF repository with the analyses has been provided.

References

Arbuthnott, K., & Frank, J. (2000). Trail making test, part B as a measure of executive control: Validation using a set-switching paradigm. *J Clin EXP Neuropsychol, 22*(4), 518–528. https://doi.org/10.1076/1380-3395(200008)22:4;1-0;FT518

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Bertoni, D., Petraglia, F., Basagni, B., Pedrazzi, G., De Gaetano, K., Costantino, C., & De Tanti, A. (2022). Cognitive reserve index and functional and cognitive outcomes in severe acquired brain injury: A pilot study. *Appl Neuropsychol Adult, 29*, 684–694. https://doi.org/10.1080/23279095.2020.1804910

Bialystok, E. (2021). Bilingualism: Pathway to cognitive reserve. *Trends Cogn Sci., 25*(5), 355–364. https://doi.org/10.1016/j.tics.2021.02.003

Borella, E., Carretti, B., Riboldi, F., & De Beni, R. (2010). Working memory training in older adults: Evidence of transfer and maintenance effects. *Psychology and Aging, 25*(4), 767–778. https://doi.org/10.1037/a0020683

Brookshire, R. H., & Nicholas, L. E. (1994). Speech sample-size and test-retest stability of connected speech measures for adults with aphasia. *Journal of Speech & Hearing Research, 37*(2), 399–407. https://doi.org/10.1044/jshr.3702.399

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics and Phonetics, 30*(7), 489–518. https://doi.org/10.3109/02699206.2016.1145740

Burke, D. M., & Shafto, M. A. (2004). Aging and language production. *Current Directions in Psychological Science, 13*(1), 21–24. https://doi.org/10.1111/j.0963-7214.2004.01301006.x

Carletta, J. (1996). Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics, 22*, 249–254.

Cebi, M., & Kulce, S. N. (2022). The Turkish translation study of the Cognitive Reserve index questionnaire (CRIq). *Appl Neuropsychol Adult., 29*(6), 1536–1542. https://doi.org/10.1080/23279095.2021.1896519

Conti, S., Bonazzi, S., Laiacona, M., Masina, M., & Coralli, M. V. (2015). Montreal Cognitive Assessment (MoCA) - Italian version: Regression based norms and equivalent scores. *Neurological Sciences, 36*(2), 209–214. https://doi.org/10.1007/s10072-014-1921-3

Crivello, F., Tzourio-Mazoyer, N., Tzourio, C., & Mazoyer, B. (2014). Longitudinal assessment of global and regional rate of Grey Matter atrophy in 1,1172 helathy older adults: Modulation by sex and Age. *PLoS One, 9*(12), Article e114478. https://doi.org/10.1371/journal.pone.0114478

De Renzi, E., & Vignolo, L. A. (1962). The token test: A sensitive test to detect receptive disturbances in aphasics. *Brain, 85*, 665–678. https://doi.org/10.1093/brain/85.4.665

D'Ortenzio, S., Petriglia, F., Gasparotto, G., Andreetta, S., Gobbo, M., & Marini, A. (2025). Aging, cognitive efficiency, and lifelong learning: Impacts on simple and complex sentence production during storytelling. *Brain Sciences, 15*, 1120. https://doi.org/10.3390/brainsci15101120

Fancourt, D., Steptoe, A., & Cadar, D. (2018). Cultural engagement and cognitive reserve: Museum attendance and dementia incidence over a 10-year period. *British Journal of Psychiatry, 213*(5), 661–663. https://doi.org/10.1192/bjp.2018.129

Farran, N., & Darwish, H. (2023). Validation of the Cognitive Reserve index questionnaire (CRIq) in Arabic. *Behav Sci, 13*(12), 1006. https://doi.org/10.3390/bs13121006

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Garba, A. E., Grossberg, G. T., Enard, K. R., Jano, F. J., Roberts, E. N., Marx, C. A., & Buchanan, P. M. (2020). Testing the cognitive reserve index questionnaire in an alzheimer's disease population. *J Alzheimers Dis Rep, 4*(1), 513–524. https://doi.org/10.3233/ADR-200244

Giovagnoli, A. R., Del pesce, M., Mascheroni, S., Simoncelli, M., Laiacona, M., & Capitani, E. (1996). Trail making Test: Normative values from 287 normal adult controls. *The Italian Journal of Neurological Sciences, 17*(4), 305–309. https://doi.org/10.1007/BF01997792

Goodglass, H., & Kaplan, E. (1972). *The Boston diagnostic Aphasia examination*. Lea & Febiger.

Harada, C. N., Natelson Love, M. C., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in Geriatric Medicine, 29*, 737–752. https://doi.org/10.1016/j.cger.2013.07.002

Hilviu, D., Parola, A., Bosco, F. M., Marini, A., & Gabbatore, I. (2025). Grandpa, tell me a story! narrative ability in healthy aging and its relationship with cognitive functions and Theory of Mind. *Lang Cogn Neurosci., 40*(1), 103–121. https://doi.org/10.1080/23273798.2024.2401027

Huber, W., & Gleber, J. (1982). Linguistic and nonlinguistic processing of narratives in aphasia. *Brain and Language, 16*(1), 1–18. https://doi.org/10.1016/0093-934X(82)90069-4

Juncos-Rabadán, O., Pereiro, A. X., & Rodríguez, M. S. (2005). Narrative speech in aging: Quantity, information content, and cohesion. *Brain and Language, 95*(3), 423–434. https://doi.org/10.1016/j.bandl.2005.04.001

Karp, A., Andel, R., Parker, M. G., Wang, H.-X., Winblad, B., & Fratiglioni, L. (2009). Mentally stimulating activities at work during midlife and dementia risk after age 75: Follow-up study from the Kungsholmen Project. *Am J Geriatr Psychiatry, 17*(3), 227–236. https://doi.org/10.1097/JGP.0b013e318190b691

Kelly, M., Duff, H., Kelly, S., McHugh Power, J. E., Brennan, S., Lawlor, B. A., & Loughrey, D. G. (2017). The impact of social activities, social networks, social support and social relationships on the cognitive functioning of healthy older adults: A systematic review. *Systematic Reviews, 6*(1), 259. https://doi.org/10.1186/s13643-017-0632-2

Kertesz, A. (1982). *Western Aphasia battery*. Grune & Stratton.

Kong, A. P.-H., Cheung, C. Y.-N., & Wong, C. W.-Y. (2025). Establishing norm of connected speech measures for descriptive discourses in Cantonese-speaking adults. *International Journal of Language & Communication Disorders, 60*, Article e70055. https://doi.org/10.1111/1460-6984.70055

Kröger, E., Andel, R., Lindsay, J., Benounissa, Z., Verreault, R., & Laurin, D. (2008). Is complexity of work associated with risk of dementia? The Canadian Study of Health and Aging. *American Journal of Epidemiology, 167*(7), 820–830. https://doi.org/10.1093/aje/kwm382

Maiovis, P., Ioannidis, P., Nucci, M., Gotzamani-Psarrakou, A., & Karacostas, D. (2016). Adaptation of the Cognitive Reserve index questionnaire (CRIq) for the Greek population. *Neurological Sciences, 37*(4), 633–636. https://doi.org/10.1007/s10072-015-2457-x

Marini, A. (2022). Cognitive and linguistic characteristics of narrative discourse production in healthy aging. In C. Coelho, L. R. Cherney, & B. Shadden (Eds.), *Discourse Analysis in Adults with and Without Communication Disorders: A Source for Clinicians and Researchers* (pp. 15–31). Plural Publishing, Inc. ISBN: 978-1-63550-375-3.

Marini, A., & Andreetta, S. (2016). Age-related effects on language production: A combined psycholinguistic and neurolinguistic perspective. In H. H. Wright (Ed.), *Cognition, Language and Aging* (pp. 55–79). John Benjamins Publishing Company.

Marini, A., Andreetta, S., Del Tin, S., & Carlomagno, S. (2011). A multi-level approach to the analysis of narrative language in Aphasia. *Aphasiology, 25*(11), 1372–1392. https://doi.org/10.1080/02687038.2011.584690

Marini, A., Boewe, A., Caltagirone, C., & Carlomagno, S. (2005). Age-related differences in the production of textual descriptions. *Journal of Psycholinguistic Research, 34*(5), 439–463. https://doi.org/10.1007/s10936-005-6203-z

Marini, A., Petriglia, F., D'Ortenzio, S., Bosco, F. M., & Gasparotto, G. (2025). Unveiling the dynamics of discourse production in healthy aging and its connection to cognitive skills. *Discourse Processes, 62*(6-7), 479–501. https://doi.org/10.1080/0163853X.2025.2507548

Marini, A., & Urgesi, C. (2012). Please get to the point! A cortical correlate of linguistic informativeness. *J Cogn Neurosci, 24*, 2211–2222. https://doi.org/10.1162/jocn_a_00283

Medici, A. C. (2021). Health sector challenges and policies in the context of ageing populations, Technical Paper No. 2021/3 | Population Division. https://www.un.org/development/desa/pd/content/health-sector-challenges-and-policies-context-ageing-populations-technical-paper-no-20213.

Merlini, A., Gabbatore, I., Marini, A., & Bosco, F. M. (2025). Affective and cognitive theory of mind in aging: what's the role of executive functions and cognitive reserve on the decline? *Experimental Aging Research*, 1–29. https://doi.org/10.1080/0361073X.2025.2586970

Monaco, M., Costa, A., Caltagirone, C., & Carlesimo, G. A. (2013). Forward and backward span for verbal and visuo-spatial data: Standardization and normative data from an Italian adult population. *Neurological Sciences, 34*, 749–754. https://doi.org/10.1007/s10072-012-1130-x

Monov, G., Stein, H., Klock, L., Gallinat, J., Kühn, S., Lincoln, T., Krkovic, K., Murphy, P. R., & Donner, T. H. (2024). Linking cognitive integrity to working memory dynamics in the aging human brain. *Journal of Neuroscience, 44*(26), Article e1883232024. https://doi.org/10.1523/JNEUROSCI.1883-23.2024

Montemurro, S., Mondini, S., Crovace, C., & Jarema, G. (2019). Cognitive reserve and its effect in older adults on retrieval of proper names, logo names and common nouns. *Frontiers in Communication, 4*, 14.

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *J Speech and Hearing Res, 36*, 338–350. https://doi.org/10.1044/jshr.3602.338

Nucci, M., Mapelli, D., & Mondini, S. (2012). Cognitive Reserve Index questionnaire (CRIq): A new instrument for measuring cognitive reserve. *Aging Clin Exp Res, 24*, 218–226. https://doi.org/10.3275/7800

Oosterhuis, E. J., Slade, K., Smith, E., May, P. J. C., & Nuttall, H. E. (2023). Getting the brain into gear: An online study investigating cognitive reserve and word-finding abilities in healthy ageing. *PLoS One, 18*, Article e0280566. https://doi.org/10.1371/journal.pone.0280566

Paradis, M., & Libben, G. (1987). *Bilingual aphasia test*. Lawrence Erlbaum Associates.

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology, 60*, 173–196. https://doi.org/10.1146/annurev.psych.59.103006.093656

Petriglia, F., Gasparotto, G., D'Ortenzio, S., Gabbatore, I., & Marini, A. (2025). Assessing lexical diversity and informativeness across the adult lifespan: A comprehensive investigation. *Research Methods in Applied Linguistics, 4*, Article 100276. https://doi.org/10.1016/j.rmal.2025.100276

Petriglia, F., Marini, A., Gabbatore, I., & Bosco, F. M. (2025). Pragmatic and narrative decline in healthy aging: a multilevel assessment of their interplay. *Journal of Communication Disorders, 118*, 106593. https://doi.org/10.1016/j.comdis.2025.106593

Pettigrew, C., & Soldan, A. (2019). Defining cognitive reserve and implications for cognitive aging. *Current Neurology and Neuroscience Reports, 19*, 1. https://doi.org/10.1007/s11910-019-0917-z

Roberts, A., Savundranayagam, M., & Orange, J. B. (2017). Non-Alzheimer dementias. In L. Cummings (Ed.), *Perspectives in Pragmatics, philosophy & Psychology: Vol. 11. Research in clinical pragmatics* (pp. 347–377). Cham: Springer. https://doi.org/10.1007/978-3-319-47489-2_14.

Roe, C. M., Xiong, C., Miller, J. P., & Morris, J. C. (2007). Education and Alzheimer disease without dementia: Support for the cognitive reserve hypothesis. *Neurology, 68*(3), 223–228. https://doi.org/10.1212/01.wnl.0000251303.50459.8a

Rusmaully, J., Dugravot, A., Moatti, J. P., Marmot, M. G., Elbaz, A., Kivimaki, M., Sabia, S., & Singh-Manoux, A. (2017). Contribution of cognitive performance and cognitive decline to associations between socioeconomic factors and dementia: A cohort study. *PLoS Medicine, 14*(6), Article e1002334. https://doi.org/10.1371/journal.pmed.1002334

Salthouse, T. A. (2019). Trajectories of normal cognitive aging. *Psychology and Aging, 34*(1), 17–24. https://doi.org/10.1037/pag0000288

Sharp, E. S., & Gatz, M. (2011). Relationship between education and dementia: An updated systematic review. *Alzheimer Disease and Associated Disorders, 25*(4), 289–304. https://doi.org/10.1097/WAD.0b013e318211c83c

Shen, X., Wang, C., Zhou, X., Zhou, W., Hornburg, D., Wu, S., & Snyder, M. P. (2024). Nonlinear dynamics of multi-omics profiles during human aging. *Nature Aging, 4*(11), 1619–1634. https://doi.org/10.1038/s43587-024-00692-2

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society, 8*(3), 448–460. PMID: 11939702.

Stern, Y. (2017). An approach to studying the neural correlates of reserve. *Brain Imaging Behav, 11*(2), 410–416. https://doi.org/10.1007/s11682-016-9566-x

Stern, Y., Arenaza-Urquijo, E. M., Bartrés-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., et al. (2020). Whitepaper: Defining and investigating cognitive reserve. *Brain Reserve, and Brain Maintenance, 16*(9), 1305–1311. https://doi.org/10.1016/j.jalz.2018.07.219

Stern, Y., Gurland, B., Tatemichi, T. K., Wilder, D., & Mayeaux, R. (1994). Influence of education and occupation on the incidence of Alzheimer's disease. *JAMA, 271*(13), 1004–1010. PMID: 8139057.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahway, NJ: Lawrence Erlbaum Associates, Inc.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Allyn and Bacon.

Tucker, A. M., & Stern, Y. (2011). Cognitive reserve in aging. *Current Alzheimer Research, 8*(4), 354–360. https://doi.org/10.2174/156720511795745320

Vickery, S., Patil, K. R., Dahnke, R., Hopkins, W. D., Sherwood, C. C., Caspers, S., Eickhoff, S. B., & Hoffstaedter, F. (2024). The uniqueness of human vulnerability to brain aging in great ape evolution. *Science Advances, 10*(35), eado2733. https://doi.org/10.1126/sciadv.ado2733

Wecker, N. S., Kramer, J. H., Wisniewski, A., Delis, D. C., & Kaplan, E. (2000). Age effects on executive ability. *Neuropsychology, 14*(3), 409–414. https://doi.org/10.1037//0894-4105.14.3.409

World Health Organization. (2020). Decade of healthy ageing: Plan of action. https://www.who.int/publications/m/item/decade-of-healthy-ageing-plan-of-action.