

Original Research

“See What I Mean?” Linguistic Digital Markers of Cognitive Decline in Nursing Home Residents

Sonja Cecchetti¹, Andrew T. Duchowski², Marco Cavallo^{1,*}¹Department of Theoretical and Applied Sciences, eCampus University, 22060 Novedrate, Italy²School of Computing, Clemson University, Clemson, SC 29634, USA*Correspondence: marco.cavallo@unicampus.it (Marco Cavallo)

Academic Editor: Bettina Platt

Submitted: 2 January 2026 Revised: 10 April 2026 Accepted: 14 April 2026 Published: 26 May 2026

Abstract

Background: Cognitive decline in nursing homes is often under-recognized. Access to specialized neuropsychological assessments, which involve detailed evaluations of how the brain influences behavior and thinking, is often limited. In this study, we examined the feasibility of administering an ecological reading task, involving reading activities that imitate real-life situations, in a real-world nursing home setting. We also explored whether reading-derived linguistic metrics and measures of language use during reading, along with an eye-tracking component that monitored participants' eye movements, were associated with cognitive impairment. **Methods:** This cross-sectional observational pilot study included 60 nursing home residents aged 65 years or older, classified as either cognitively impaired (CI, $n = 30$) or healthy control (HC, $n = 30$) based on neuropsychological profiles and clinical evaluations. All participants completed the Mini-Mental State Examination (MMSE), Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), Frontal Assessment Battery (FAB), Geriatric Depression Scale (GDS), Barthel Index, and Cognitive Reserve Index questionnaire (CRIq). The reading task involved reading aloud a 177-word Italian passage followed by eight comprehension questions. Primary outcomes were total reading errors (TRE), total comprehension errors (TCE), and words per minute (WPM). Eye-tracking data (Gazepoint GP3, 60 Hz) were available for a subsample with usable calibration and validation data (CI: $n = 26$; HC: $n = 24$). Analyses included descriptive comparisons, covariate-adjusted generalized linear models, receiver operating characteristic (ROC) analyses, and partial Spearman correlations. **Results:** All 60 participants completed the task. Usable audio recordings were obtained for everyone, and usable eye-tracking data were available for 50 participants. The CI group showed higher TRE and TCE, lower WPM, and longer reading times compared with the HC group. In models adjusted for gender, age, education, GDS, Barthel Index, and CRIq, TCE showed the strongest association with CI status (odds ratio [OR] = 11.108, 95% CI 2.662–246.186, $p = 0.024$), and TRE was also associated with CI status (OR = 1.238, 95% confidence interval [CI] 1.054–1.629, $p = 0.045$). ROC analyses showed high areas under the curve (AUCs) for TCE (0.982), TRE (≈ 0.95), and audio-based WPM (0.936). In the eye-tracking subsample, timing-related measures also showed good discrimination (AUC = 0.960 for eye tracking [ET]-WPM; 0.942 for ET-Total Reading Time), whereas conventional first-order oculomotor metrics did not differ significantly between groups. **Conclusions:** An ecological reading task was feasible within a nursing home setting. It generated reading-derived linguistic metrics linked to cognitive impairment and broader cognitive-functional status. Measures related to comprehension, overall reading error load, and reading speed were shown to be useful as digital linguistic markers.

Keywords: aged; cognitive dysfunction; feasibility studies; language; nursing homes; eye movements; eye-tracking technology

1. Introduction

1.1 Clinical Context and Significance

Population aging is a major demographic trend [1]. Italy ranks second, behind Japan, and is among Europe's longest-lived countries [2,3]. This shift has serious consequences for healthcare systems. Advancing age raises the risk of multimorbidity, including neurocognitive disorders [4]. Alzheimer's disease (AD), the leading cause of dementia, is marked by progressive cognitive, functional, and behavioral decline, and is a major cause of disability [5]. The latest World Alzheimer Report estimates that over 50 million people worldwide have dementia, a number expected to triple by 2050 [6]. Mild cognitive impairment (MCI), a transitional state between normal aging and dementia, is also common and may progress over time [7,8]. Early iden-

tification of cognitive decline is crucial, as it enables timely intervention, preserves abilities, and improves the lives of older adults and caregivers [9].

This issue is especially important in nursing homes, where residents are often frail and medically complex, and at increased risk of cognitive deterioration [10]. Earlier recognition of cognitive decline in these settings can support more timely, personalized care planning and facilitate non-pharmacological interventions, such as cognitive stimulation and computerized training [11]. Early detection may also reduce adverse events related to unrecognized decline, wandering, falls, and agitation, which increase staff burden and care costs in residential facilities [12,13]. Progression from MCI to dementia further raises care needs and supervision requirements, affecting staffing and resource alloca-



tion in nursing homes [14]. However, cognitive impairment in institutionalized older adults often goes unrecognized, and early detection can be harder than in traditional clinical contexts [15]. Thus, digital markers, including linguistic and eye-tracking metrics, have been proposed as promising, sensitive, non-invasive tools for detecting subtle cognitive changes in neurodegenerative disorders [16,17].

1.2 Existing Evidence and Research Gaps

Reading is a highly complex cognitive task that involves the coordinated integration of perceptual, linguistic, and executive processes [18]. Although reading paradigms are commonly used in eye-movement research, the existing evidence on aging and cognitive impairment remains varied and should be interpreted carefully [19,20].

Previous studies suggest that reading aloud may remain relatively intact in some individuals with mild AD, even when comprehension is impaired. Single-case observations by Ripamonti *et al.* [21] support this dissociation. Martínez-Nicolás *et al.* [22] also found that AD patients could accurately read words without necessarily understanding their meanings. The same study indicated that detailed temporal measures of reading fluency, such as speed and pausing, might more sensitively distinguish early AD from asymptomatic controls than broader global assessments, such as the Mini-Mental State Examination (MMSE). Education level and cognitive reserve could also influence performance [23].

Evidence from eye-tracking studies is similarly mixed. A meta-analysis by Moreno *et al.* [24] showed that healthy older readers tend to have longer fixation durations, more regressions, and longer overall reading times than young adults, suggesting decreased efficiency despite preserved basic reading ability. In clinical populations, Lueck *et al.* [25] observed slower reading and altered oculomotor behavior in patients with probable mild-to-moderate AD, including longer fixations, more forward saccades, and more regressions. However, saccade duration was similar to that of healthy controls. More recent multimodal studies further suggest that lexical difficulty may significantly influence oculomotor behavior. For example, Shah *et al.* [26] reported that high-difficulty words elicited more atypical gaze patterns in AD than in MCI or control participants, whereas low-difficulty words provided less information. Similarly, Groznik *et al.* [27] found that cognitively impaired participants read more slowly and with more irregularity. Nonetheless, they concluded that a short reading task was better suited for a broader assessment battery than as a standalone test.

Overall, these findings indicate that measures derived from reading can provide valuable insights into study outcomes. However, substantial methodological differences across reading tasks, clinical classifications, oculomotor metrics, eye trackers, and data collection protocols limit direct comparisons between studies and hinder practical ap-

plication. Evidence is particularly limited among nursing home populations, where ecological constraints, participant frailty, and concerns about data quality complicate feasibility and interpretation.

1.3 Rationale for the Proposed Approach

Reading aloud with comprehension is a complex and ecologically meaningful behavior. It depends on the integration of perceptual, lexical, phonological, semantic, attentional, and executive processes. Therefore, reading is a valuable task for detecting cognitive decline, as measured by accuracy, comprehension, and efficiency. Established cognitive models inform the present study of reading and comprehension. Dual-route word recognition models, such as the Dual Route Cascaded model [28], help interpret reading accuracy and error patterns through lexical and sublexical processing. Construction–integration accounts of comprehension, such as Kintsch’s model [29], support our understanding of comprehension performance in terms of semantic integration, working memory, and executive control. Models of eye movement control, such as E-Z Reader [30], provide a foundation for considering oculomotor indices as markers of online lexical and attentional processing. In this study, E-Z Reader is used solely as a theoretical reference for the eye-tracking component. A schematic appears in **Supplementary Fig. 1**. No model fitting or parameter estimation was performed. Instead, the study was conducted as part of the CogET project, which investigates language and cognition in nursing home residents. The approach combines neuropsychological assessments with a reading task supported by multimodal recordings, including audio and, when possible, eye-tracking. Reading-derived linguistic performance measures were selected as primary outcomes because of their low burden, ecological validity, and direct links to reading accuracy, comprehension, and fluency. Eye-tracking was added as an additional exploratory process measure.

1.4 Objectives and Hypotheses

The objectives of this cross-sectional pilot study were fourfold. First, we aimed to assess the feasibility of administering an ecological reading task in a nursing home, including task completion, audio recording acquisition, eye-tracking calibration and validation, and the proportion of usable recordings. Second, we aimed to compare the cognitively impaired (CI) and healthy control (HC) groups on primary reading outcomes, such as total reading errors (TRE), total comprehension errors (TCE), and words per minute (WPM), as well as secondary outcomes like Total Reading Time and specific error types. Third, we sought to evaluate how well reading-derived measures distinguish between CI and HC groups. Fourth, we aimed to examine the relationships between reading-derived measures and neuropsychological, functional, affective, and cognitive reserve measures. Given the feasibility-focused and exploratory nature

of this pilot study, we did not predefine directional hypotheses for all outcomes. Instead, we focused on estimating feasibility parameters, between-group differences, discriminative ability, and covariate-adjusted associations. The results will help guide the design of future confirmatory studies.

2. Materials and Methods

2.1 Design and Participants

The study followed STROBE checklist guidelines and is part of the CogET project, pre-registered on the Open Science Framework (OSF DOI: <https://doi.org/10.17605/OSF.IO/VG3TD>). Data collection occurred from June to December 2025. Recruitment used a convenience sampling strategy, considered appropriate for exploratory and pilot studies [31].

Participants were permanent residents of the Saint George nursing home in Cavallermaggiore, Cuneo, Italy. Inclusion criteria were: (a) age 65 years or older; (b) residence in the nursing home at the time of recruitment; and (c) ability to read a standard Italian text displayed on a computer screen. Exclusion criteria were: (a) a documented neurological or psychiatric disorder other than mild cognitive impairment or Alzheimer's disease; (b) a history of epilepsy; (c) uncorrected visual impairment or ophthalmological conditions interfering with eye tracking; and (d) inability to provide informed consent. All participants had normal or corrected-to-normal vision and were able to understand and complete the experimental tasks.

A total of 60 participants were enrolled, comprising 30 healthy controls and 30 older adults with CI of varying severity. The group assignment was determined based on the overall neuropsychological profile and the clinical evaluation conducted by two qualified neuropsychologists. Healthy controls were required to have a MMSE score above 26, interpreted according to Italian normative values corrected for age and education, and no clinical evidence of cognitive impairment. Participants in the CI group showed clinical evidence of cognitive decline supported by neuropsychological assessment. The MMSE was used as a global cognitive screening measure, whereas the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) and Frontal Assessment Battery (FAB) contributed to the broader characterization of the cognitive profile and supported clinical classification.

2.2 Neuropsychological Assessment

All participants underwent a neuropsychological, affective, and functional assessment.

Global cognitive functioning was screened using the MMSE [32]. The MMSE includes 30 items evaluating orientation, short-term memory, attention and calculation, learning, language, and constructive praxis. Scores were interpreted using the Italian standardization by Magni *et al.* [33], which provides normative adjustments for age and ed-

ucation. In this study, the MMSE served as a general cognitive screening tool, and healthy controls were required to score above 26.

A broader cognitive profile was obtained using the RBANS [34], a brief standardized battery comprising 12 subtests that yields a Total Scale score and five Index scores: Immediate Memory, Visuospatial/Constructional abilities, Language, Attention, and Delayed Memory. The RBANS provides age-adjusted standard scores, aiding in characterizing cognitive performance across domains and supporting comparisons between groups with varying levels of cognitive decline [35,36]. In this study, the RBANS was used to support the clinical determination of the cognitive profile.

Executive functioning was further evaluated with the FAB [37], a six-subtest screening tool for frontal executive functions, including conceptualization, phonemic fluency, motor programming, sensitivity to interference, inhibitory control, and environmental autonomy. FAB scores were interpreted using the Italian norms by Appollonio *et al.* [38], adjusted for age and educational level.

Affective state was assessed using the Geriatric Depression Scale (GDS) [39], administered in its Italian version [40]. Functional independence was measured with the Barthel Index [41,42]. Lastly, cognitive reserve was estimated using the Cognitive Reserve Index questionnaire (CRIq), a semi-structured interview that quantifies lifetime cognitive enrichment through education, occupational attainment, and cognitively stimulating leisure activities [43,44]. The GDS, Barthel Index, and CRIq were used to characterize affective, functional, and reserve-related aspects of the sample, but were not criteria for group assignment.

2.3 Equipment and Setup

Eye movements were recorded in a dedicated room within the nursing home, easily accessible to residents and free of architectural barriers. Recordings were obtained using a Gazepoint GP3 video-based eye tracker sampling at 60 Hz (Gazepoint Research Inc., Vancouver, BC, Canada). According to the manufacturer, the device has a spatial accuracy of approximately 0.5–1.0° of visual angle under optimal conditions and a spatial resolution of approximately 0.1°. Binocular gaze data were streamed through the Gazepoint Application Programming Interface (v7.1.0, Gazepoint Research Inc.).

The experiment was run on a laptop computer (Intel Core i5-1235U processor, 16 GB RAM, Windows 11) (Dell Inspiron 15, Dell Inc., Round Rock, TX, USA) using a custom PsychoPy script (Release 2023.1.3, Open Science Tools Ltd., Nottingham, UK) [45]. Stimuli were presented on a 24-inch Dell Full HD monitor (Dell Inc.) with a resolution of 1920 × 1080 pixels. The viewing distance was maintained at approximately 60–65 cm (mean = 62.45 cm) from the participants' eyes, in accordance with the device

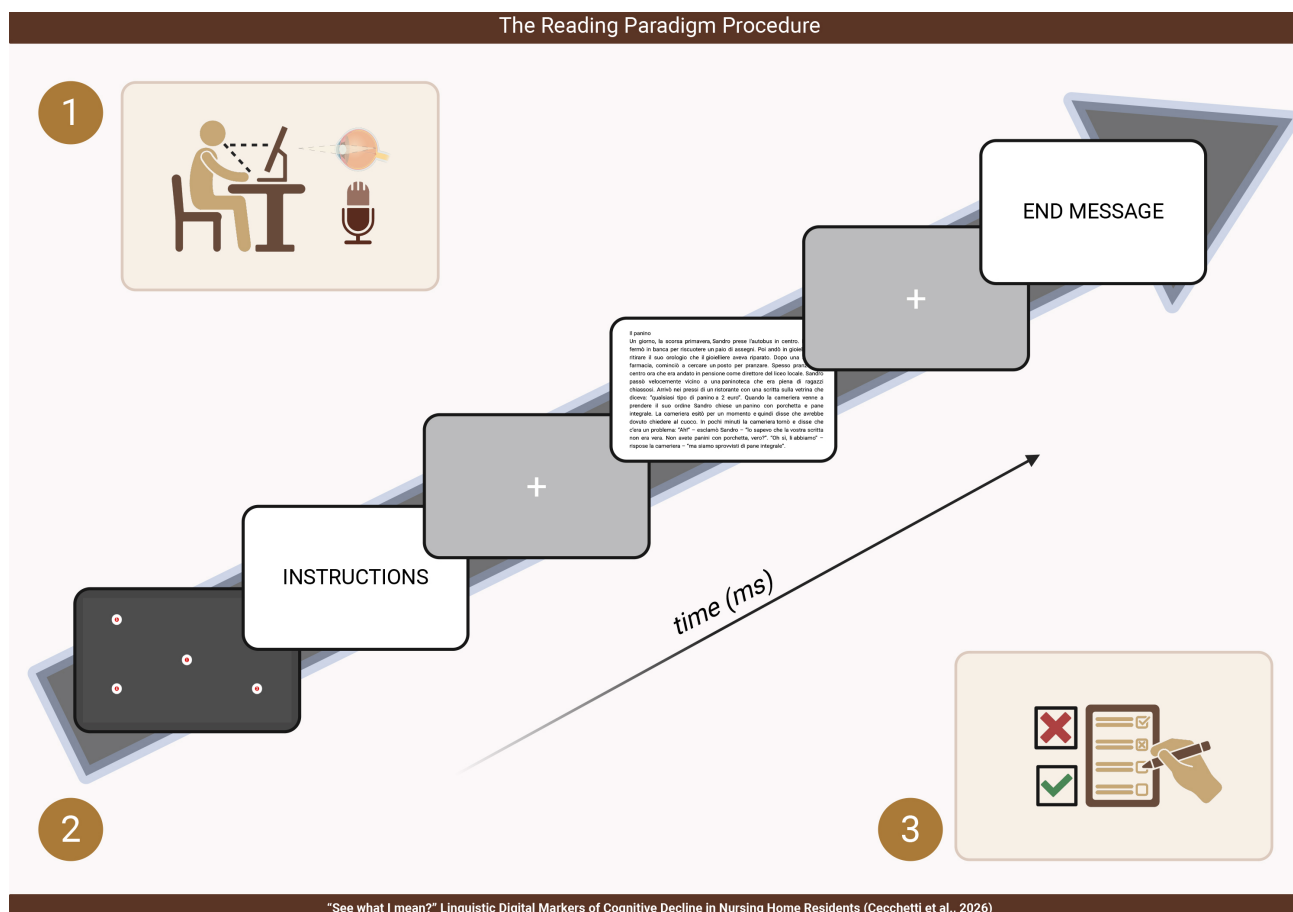


Fig. 1. The reading paradigm procedure. (1) Experimental Setup and extracted features. (2) Description of the procedure developed in PsychoPy. (3) Questionnaire for the evaluation of comprehension. Created in BioRender (<https://www.biorender.com/>). Cecchetti, S. (2026) <https://BioRender.com/2m4mv47>.

recommendations. The Gazeport monitoring window and pre-task positioning checks were used to optimize setup. Participants sat comfortably in front of the screen in a quiet, well-lit room. No chin rest was used, but participants were asked to minimize head movements during recording. The reading task setup and procedure are summarised in Fig. 1.

2.4 Reading Task Procedure

The reading stimulus was the 'sandwich' story, a 177-word passage spread over 15 lines, selected from the Discourse Comprehension Test (DCT) included in the Narrative Comprehension and Production Test battery (NCPT) [46]. This text was chosen for multiple reasons. First, it offers a brief, meaningfully realistic reading task suitable for older adults in a nursing home setting. Second, it was presented in the participants' native language. Third, it is part of a test battery with prior clinical use, including work with adults with AD [47]. Finally, following recommendations by Brookshire and Nicholas [48], the passage contains a mildly humorous element intended to enhance engagement and lower tension during testing. Participants were instructed to read the passage aloud as quickly and accurately

as possible while also trying to understand its content, as comprehension questions would follow. Instructions first appeared on the screen and were then repeated orally by the examiner to ensure understanding of the task. Just before the text was presented, a fixation cross was shown to focus gaze on the screen. The maximum time allowed for the reading task was set at 180 seconds.

Comprehension was assessed immediately afterward using eight structured questions related to the passage (see **Supplementary Table 1**). These questions targeted different levels of processing, including explicitly stated information (MIS), implicit main ideas (MII), and peripheral details presented either explicitly (DTS) or implicitly (DTI). This structure allowed for characterizing performance not only in terms of literal information retrieval but also in relation to inferential processing and the integration of implicit and secondary story elements. The questions were administered in a paper-and-pencil format. The overall reading paradigm is illustrated in Fig. 1.

2.5 Pre-Processing Data

Eye-tracking pre-processing and analysis followed the general logic of the gaze analysis pipeline described by Duchowski [20], including raw data inspection, event extraction, area-of-interest assignment, and aggregation of summary measures for statistical analysis. Raw gaze data consisted of time-stamped horizontal and vertical gaze coordinates. From these data, fixation- and saccade-related measures were derived and summarised using custom Python scripts (Python 3.10, Python Software Foundation, Beaverton, OR, USA).

Specifically, the raw data stream (the raw gaze signal) output by the eye tracker consists of gaze points $g_i = (x_i, y_i, t_i)$, where x_i and y_i are the 2D gaze point coordinates on the screen and t_i is the timestamp of every sample. A third-order ($s = 3$) Savitzky-Golay (SG) filter $h_i^{t,s}$ of width $p = 5$ is used to differentiate the raw positional gaze signal into its velocity estimate:

$$\dot{x}_n^s(t) = 1/(\Delta t^s) \left(\sum_{i=-p}^p h_i^{t,s} x_{n-i} \right)$$

applied independently to the x_i and y_i 2D gaze point coordinates. The resulting velocity signal (\dot{x}_i, \dot{y}_i) is thresholded at 36 degrees per second. The signal where velocity exceeds this threshold is labeled as part of a saccade; elsewhere, the signal is labeled a fixation. See Duchowski [20] for further details.

For text-based eye-tracking analysis, Areas of Interest (AOIs) were defined at the word level, with one AOI assigned to each word of the passage, following a word-based approach similar to that described by Busjahn and Tamm [49]. AOIs were created in document editing software (Scribus, v. 1.7.0, The Scribus Team, open-source, <https://www.scribus.net/>) based on the exact visual layout of the displayed text. Gaze events were then assigned to the predefined AOIs, and events falling outside these word-level AOIs were excluded from AOI-based analyses. A more detailed description of AOI construction and the custom processing workflow is provided in the **Supplementary Fig. 2** to facilitate reproducibility.

The resulting eye-tracking metrics included conventional aggregate oculomotor measures, such as fixation count, fixation duration, and saccade amplitude, as well as AOI-based aggregate indices. Because the study was conducted in an ecological nursing home setting, without head stabilization and using a 60 Hz device, the analyses focused on robust summary measures rather than highly fine-grained temporal parsing.

Audio recordings were processed separately. Each file was manually reviewed in ELAN (version 7.0, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands) [50,

51], where reading errors, including mispronounced words, repetitions, excess words, and omissions, were identified and annotated.

2.6 Study Outcomes

The study outcomes were predefined as feasibility, primary, secondary, and exploratory outcomes. Feasibility outcomes included completion of the reading task, successful acquisition of audio recordings, successful eye-tracking calibration and validation when attempted, and the proportion of usable audio and eye-tracking recordings available for analysis.

The primary outcomes focused on key reading metrics: TRE, TCE, and WPM. TRE was defined as the total number of reading errors made during aloud reading, including substitutions, omissions, repetitions, excess responses, and mispronunciations. TCE was defined as the total number of incorrect responses on the comprehension questions administered after the reading passage. WPM was defined as the number of correctly read words per minute.

Secondary outcomes included Total Reading Time, individual components of reading and comprehension performance, and timing-based eye-tracking metrics. Reading error types comprised substitutions, omissions, repetitions, excess responses, and mispronounced words. Comprehension error types included MII, MIS, DTI, and DTS errors. Timing-based eye-tracking measures included eye-tracking-based WPM and Total Reading Time.

Exploratory outcomes encompassed traditional first-order eye-tracking metrics, such as fixation- and saccade-based indices. Since usable eye-tracking data were available only for a smaller subsample, analyses involving these measures were considered exploratory.

2.7 Statistical Analyses

Descriptive statistics are reported for socio-demographic variables, neuropsychological measures, and reading-derived outcomes, stratified by group (HC vs CI). For between-group comparisons of continuous variables, normality was assessed within each group using the Shapiro-Wilk test. When normality was supported in both groups ($p > 0.05$), group differences were evaluated using independent-samples t -tests. Homogeneity of variances was assessed using Levene's test, and, when violated, Welch's correction was applied. When at least one group deviated from normality, the Wilcoxon rank-sum test was used. Categorical variables were compared using Pearson's chi-squared test.

To evaluate the association between reading-derived measures and cognitive impairment status, multivariable generalized linear models (GLMs) with a binomial distribution and a logit link function were fitted with CI vs HC status as the dependent variable. Separate models were estimated for each reading-derived predictor, including TRE,

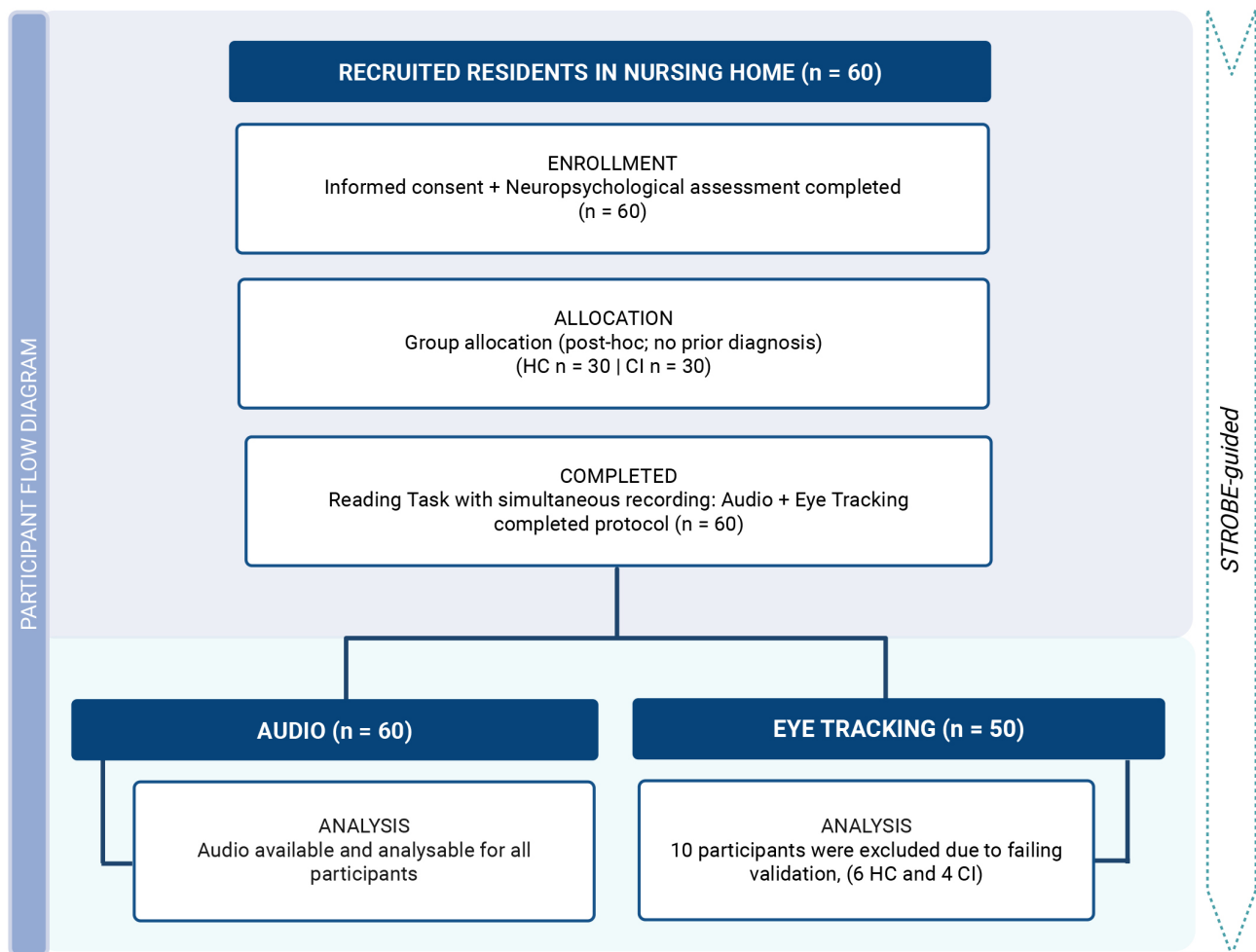


Fig. 2. Participant flow diagram. Created in BioRender (<https://www.biorender.com/>). Cecchetti, S. (2026) <https://BioRender.com/48bg3ju>. HC, healthy control; CI, cognitively impaired.

TCE, WPM, Total Reading Time, and the specific reading and comprehension error-type measures, while adjusting for gender, age, education, GDS score, Barthel Index, and CRIq. Model results are reported as odds ratios (ORs) with 95% confidence intervals (CIs). To account for multiple testing across these models, *p*-values were adjusted using the false discovery rate (FDR) procedure.

For each GLM, discriminative performance was further assessed by receiver operating characteristic (ROC) analysis, reporting the area under the curve (AUC) together with sensitivity, specificity, and accuracy at the selected probability threshold.

Finally, partial Spearman correlation analyses were conducted to explore relationships between reading-related measures and neuropsychological, functional, affective, and cognitive reserve variables, adjusting for age, gender, and education. A non-parametric approach was chosen because several variables were not normally distributed. The correlation analysis included audio-based and eye-tracking-based reading measures; eye-tracking variables were available only in a smaller subsample (*n* = 50). To account for

multiple testing, *p*-values were adjusted using the FDR procedure, and only correlations that remained significant after FDR correction ($q < 0.05$) were considered meaningful.

Exploratory eye-tracking analyses were also performed on the subsample with usable oculometric data following calibration and validation. In addition to eye-tracking timing measures, traditional first-order oculomotor metrics were examined. Statistical tests were selected based on assessments of normality and homogeneity of variance. Due to the smaller subsample size and the exploratory nature of these analyses, findings related to conventional eye-tracking measures were interpreted with caution.

All analyses were conducted in R (version 4.5.2; R Core Team, R Foundation for Statistical Computing, Vienna, Austria) [52]. Tables were generated using the *gtsummary* package (version 2.5.0) [53], and figures were produced using *ggplot2* (version 4.0.2) [54]. Statistical significance was set at $\alpha = 0.05$ (two-sided).

Table 1. Sociodemographic characteristics.

| Characteristic | HC (n = 30) | CI (n = 30) | <i>p</i> -value |
|----------------------|----------------------|----------------------|-----------------|
| Age, years | 74.50 (69.00; 81.00) | 88.00 (82.00; 92.00) | <0.001 |
| Education, years | 8.00 (5.00; 13.00) | 5.00 (5.00; 8.00) | <0.001 |
| Female gender, n (%) | 15/30 (50.00%) | 23/30 (76.67%) | 0.032 |

Values are presented as median (Q1; Q3) for continuous variables and n (%) for categorical variables. *p*-values were calculated using the Wilcoxon rank-sum test for continuous variables and Pearson's chi-squared test for categorical variables.

Table 2. Neuropsychological assessment.

| Characteristic | HC (n = 30) | CI (n = 30) | <i>p</i> -value |
|----------------------------|------------------------|----------------------|-----------------|
| MMSE (adjusted) | 28.00 (27.20; 28.40) | 21.85 (18.40; 24.20) | <0.001 |
| RBANS total (age-adjusted) | 88.00 (76.00; 105.00) | 60.00 (53.00; 69.00) | <0.001 |
| FAB (adjusted) | 15.51 (1.68) | 11.01 (2.84) | <0.001 |
| GDS | 4.00 (1.00; 6.00) | 8.50 (6.00; 14.00) | <0.001 |
| Barthel Index | 100.00 (98.00; 100.00) | 54.00 (28.00; 85.00) | <0.001 |
| CRIq | 102.97 (19.57) | 80.17 (12.83) | <0.001 |

Values are presented as median (Q1; Q3) for non-normally distributed variables and mean (SD) for normally distributed variables. *p*-values were calculated using the Wilcoxon rank-sum test or Welch's two-sample *t*-test, as appropriate.

Abbreviations: MMSE, Mini-Mental State Examination; RBANS, Repeatable Battery for the Assessment of Neuropsychological Status; FAB, Frontal Assessment Battery; GDS, Geriatric Depression Scale; CRIq, Cognitive Reserve Index questionnaire.

3. Results

Regarding the first objective, feasibility results showed that 60 out of 105 residents screened were enrolled and completed the study protocol (see Fig. 2). Usable audio recordings were obtained from all included participants. Eye tracking was attempted with all 60 participants, producing usable data for 50 of them. Ten participants were excluded from the eye-tracking analysis due to failed calibration or validation, including four in the CI group and six in the HC group.

A total of 60 nursing-home residents participated in the study (HC, *n* = 30; CI, *n* = 30). As shown in Table 1, the CI group was significantly older than the HC group (median [Q1; Q3]: 88.00 [82.00; 92.00] vs 74.50 [69.00; 81.00] years, *p* < 0.001) and had fewer years of education (5.00 [5.00; 8.00] vs 8.00 [5.00; 13.00], *p* < 0.001). The gender distribution also differed between groups, with a higher proportion of females in the CI group compared to the HC group (23/30, 76.67% vs 15/30, 50.00%; *p* = 0.032). Assumption checks for continuous variables are detailed in **Supplementary Tables 2,3**.

Neuropsychological and functional assessment results are presented in Table 2. Relative to the HC group, the CI group showed lower performance on the adjusted MMSE (median [Q1; Q3]: 21.85 [18.40; 24.20] vs 28.00 [27.20; 28.40], *p* < 0.001) and lower age-adjusted RBANS total scores (60.00 [53.00; 69.00] vs 88.00 [76.00; 105.00], *p* < 0.001). The CI group also showed lower executive functioning on the adjusted FAB (11.01 ± 2.84 vs 15.51 ± 1.68,

p < 0.001), greater depressive symptomatology on the GDS (8.50 [6.00; 14.00] vs 4.00 [1.00; 6.00], *p* < 0.001), lower functional independence on the Barthel scale (54.00 [28.00; 85.00] vs 100.00 [98.00; 100.00], *p* < 0.001), and lower cognitive reserve on the CRIq (80.17 ± 12.83 vs 102.97 ± 19.57, *p* < 0.001).

Regarding the second objective, descriptive statistics for reading-related measures are provided in the **Supplementary Materials**. **Supplementary Table 4** reveals that the CI group had a higher TRE load than the HC group, with the most notable differences observed in omissions and, to a lesser degree, in substitutions and mispronunciations.

Supplementary Table 5 indicates that the CI group also exhibited more TCE than the HC group, with higher values across all types of comprehension errors (DTS, DTI, MIS, and MII).

Supplementary Table 6 further demonstrates that the CI group read more slowly than the HC group, with fewer WPM and longer overall reading times. Collectively, these descriptive findings provide the distributional context for the subsequent covariate-adjusted GLM and ROC analyses.

Consistent with these patterns, Table 3 shows the results of covariate-adjusted multivariable GLMs examining the relationship between each reading-derived measure and CI versus HC status, controlling for gender, age, education, GDS, Barthel Index, and CRIq. Among comprehension-related variables, higher scores of MII errors (OR = 7.595, 95% CI 1.213–66.313, *p* = 0.041), DTI errors (OR = 6.397, 95% CI 1.539–48.273, *p* = 0.027), DTS errors (OR = 278.096, 95% CI 6.343–139,396.507, *p* = 0.025), and TCE

Table 3. Multivariable generalized linear models (GLMs).

| Variable | n | OR | 95% CI | <i>p</i> -value ¹ | q-value ² |
|----------------------------|----|---------|--------------------|------------------------------|----------------------|
| MII errors | 60 | 7.595 | 1.213, 66.313 | 0.041 | 0.134 |
| MIS errors | 60 | 3.464 | 0.788, 24.396 | 0.143 | 0.238 |
| DTI errors | 60 | 6.397 | 1.539, 48.273 | 0.027 | 0.134 |
| DTS errors | 60 | 278.096 | 6.343, 139,396.507 | 0.025 | 0.134 |
| Total Comprehension Errors | 60 | 11.108 | 2.662, 246.186 | 0.024 | 0.134 |
| Substitutions | 60 | 1.130 | 0.866, 2.982 | 0.741 | 0.741 |
| Omissions | 60 | 2.014 | 1.154, 5.690 | 0.081 | 0.173 |
| Repetitions | 60 | 1.492 | 0.899, 3.547 | 0.306 | 0.417 |
| Excess | 60 | 1.320 | 0.710, 4.956 | 0.625 | 0.670 |
| Mispronounced | 60 | 1.052 | 0.946, 1.395 | 0.515 | 0.595 |
| Total Reading Errors | 60 | 1.238 | 1.054, 1.629 | 0.045 | 0.134 |
| WPM (audio) | 60 | 0.978 | 0.945, 1.008 | 0.171 | 0.256 |
| Total Reading Time (audio) | 60 | 1.013 | 0.987, 1.041 | 0.341 | 0.426 |
| WPM (ET) | 50 | 0.968 | 0.932, 0.999 | 0.057 | 0.142 |
| Total Reading Time (ET) | 50 | 1.021 | 0.995, 1.052 | 0.126 | 0.237 |

¹*p*-values come from a logistic model using gender, age, education, GDS, Barthel Index, and CRIq as covariates.

²False discovery rate correction for multiple testing.

Abbreviations: CI, confidence interval; OR, odds ratio; MII, implicit main ideas; MIS, explicitly stated information; DTI, peripheral details presented either implicitly; DTS, peripheral details presented either explicitly; WPM, words per minute; ET, eye tracking.

(OR = 11.108, 95% CI 2.662–246.186, *p* = 0.024) were linked to higher odds of being in the CI group. MIS errors showed a similar trend, but the association was not statistically significant (OR = 3.464, 95% CI 0.788–24.396, *p* = 0.143). Among reading-related variables, TRE was also positively associated with CI status (OR = 1.238, 95% CI 1.054–1.629, *p* = 0.045), whereas individual reading error types were not significantly associated with group status. No significant links emerged for reading speed or reading time measures. After FDR correction, none of the models stayed significant at the usual threshold of *q* < 0.05. However, several variables, including MII errors, DTI errors, DTS errors, TCE, TRE, and WPM (eye tracking, ET), showed borderline adjusted values. Overall, these results suggest that comprehension-related measures and the overall number of reading errors provide the strongest adjusted signals for differentiating the CI and HC groups in this pilot sample.

Regarding the third objective, ROC analyses based on the fitted models showed good discrimination between CI and HC status (Fig. 3). Among reading error metrics, the model including TRE (Fig. 3A) demonstrated the best performance (AUC ≈ 0.95; threshold = 0.571; sensitivity = 0.867; specificity = 1.00; accuracy = 0.933). For individual reading error categories (Fig. 3B), discrimination remained strong, with AUC values ranging from 0.931 to 0.961. Specifically, substitutions had an AUC of 0.931, omissions 0.961, repetitions 0.937, mispronounced words 0.931, and excess errors 0.931. Overall, these findings suggest that total reading errors offer the strongest discrimination of cog-

native impairment status, while individual reading error categories also provide meaningful, though slightly weaker, classification performance.

A similarly strong pattern emerged for comprehension-related metrics (Fig. 4). The model, including TCE (Fig. 4A), showed excellent discrimination between CI and HC status (AUC = 0.982), with an optimal probability threshold of 0.492, yielding sensitivity = 0.933, specificity = 0.967, and accuracy = 0.950. For the individual comprehension error categories (Fig. 4B), discriminatory performance was also good to excellent. MII errors showed an AUC of 0.938 (threshold = 0.381; sensitivity = 0.900; specificity = 0.833; accuracy = 0.867), MIS errors an AUC of 0.930 (threshold = 0.768; sensitivity = 0.767; specificity = 0.967; accuracy = 0.867), DTI errors an AUC of 0.949 (threshold = 0.671; sensitivity = 0.833; specificity = 0.967; accuracy = 0.900), and DTS errors an AUC of 0.973 (threshold = 0.534; sensitivity = 0.933; specificity = 0.967; accuracy = 0.950). Overall, these models showed consistently high specificity and good sensitivity, suggesting that comprehension-related measures, particularly TCE and DTS errors, may be especially informative for distinguishing CI from HC participants in this exploratory pilot sample. Taken together with the covariate-adjusted odds ratios reported in Table 3, these findings indicate that both aggregate and selected domain-specific comprehension error measures provided meaningful discriminatory information even after adjustment for gender, age, education, GDS, Barthel Index, and CRIq.

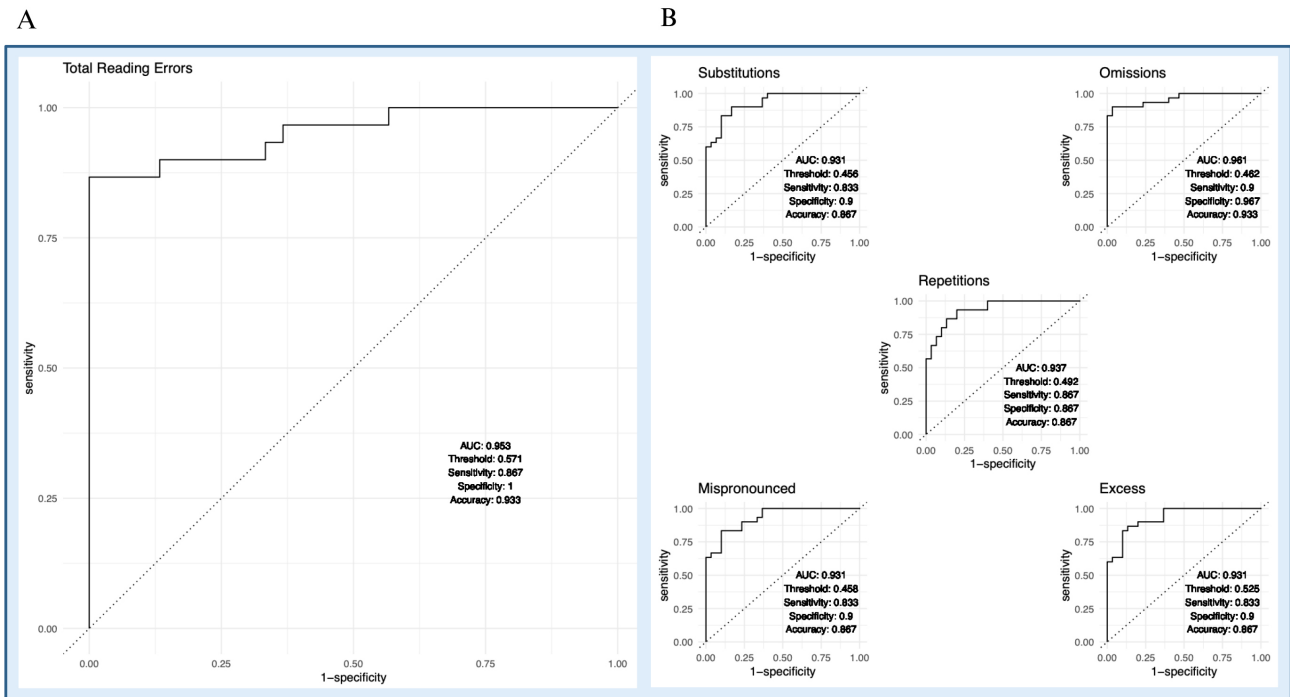


Fig. 3. Receiver operating characteristic (ROC) curves for reading error metrics. (A) Total reading errors. (B) Individual reading error categories, including substitutions, omissions, repetitions, mispronounced, and excess. Curves are derived from covariate-adjusted models that include age, gender, education, GDS, Barthel Index, and CRiQ. AUC, area under the curve.

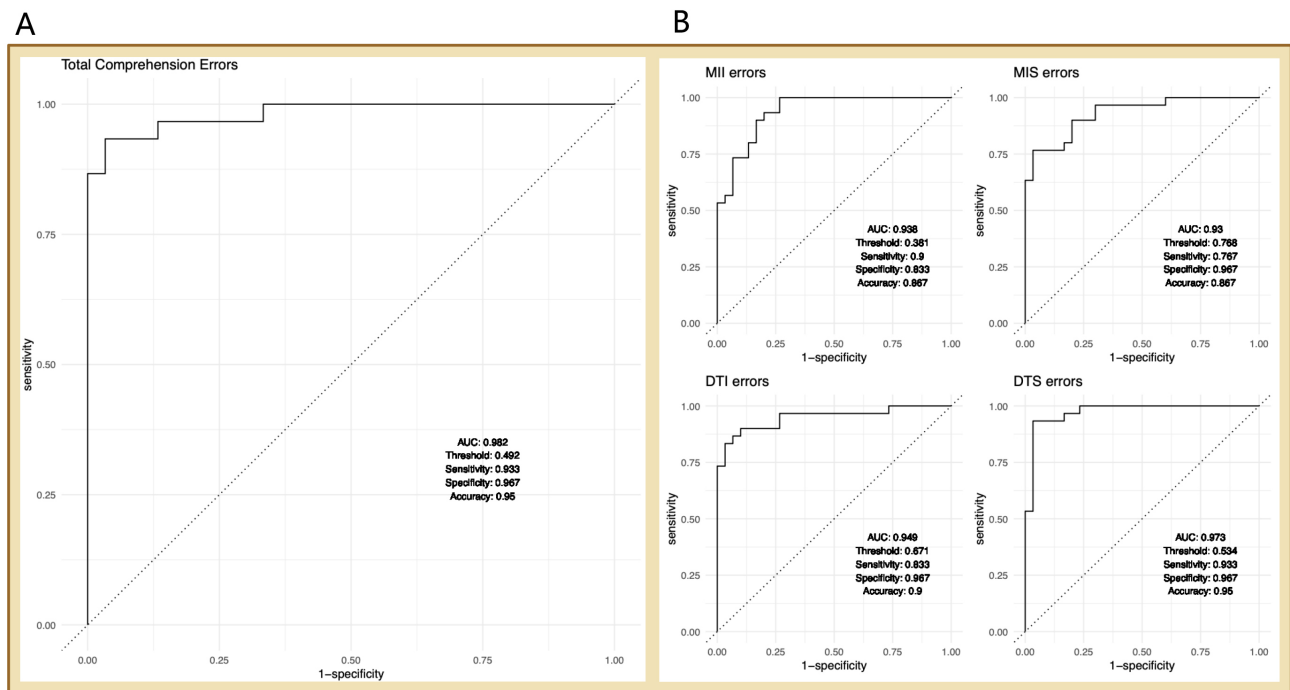


Fig. 4. ROC curves for comprehension error metrics. (A) Total comprehension errors. (B) Comprehension error type metrics (MII, MIS, DTI, and DTS). Models adjusted for age, gender, education, GDS, Barthel Index, and CRiQ.

For the audio timing-based reading metrics (Fig. 5), WPM (Fig. 5A) demonstrated effective discrimination between CI and HC status. WPM in the audio condition achieved an AUC of 0.936, with an optimal probability

threshold of 0.323, sensitivity of 0.900, specificity of 0.800, and accuracy of 0.850. Total Reading Time in the audio condition (Fig. 5B) showed similar performance, with an AUC of 0.930, an optimal probability threshold of 0.540,

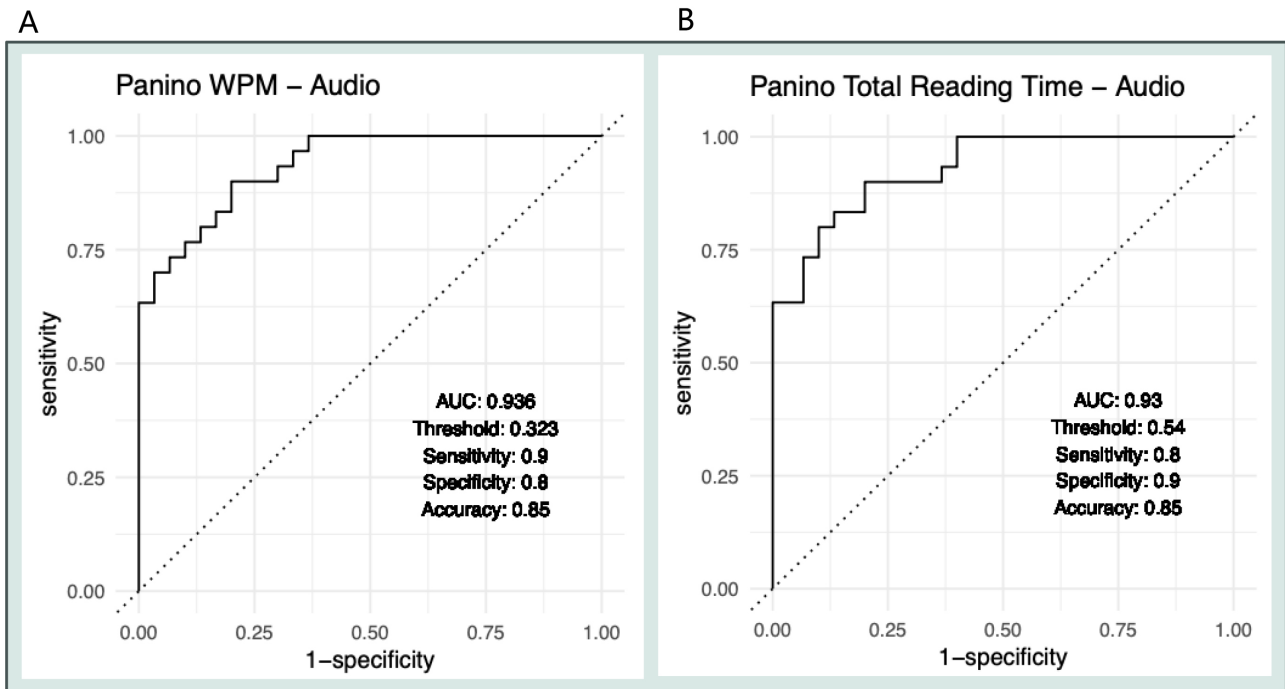


Fig. 5. ROC curves for timing-based reading metrics (audio). (A) WPM. (B) Total Reading Time: models adjusted for age, gender, education, GDS, Barthel Index, and CRIq.

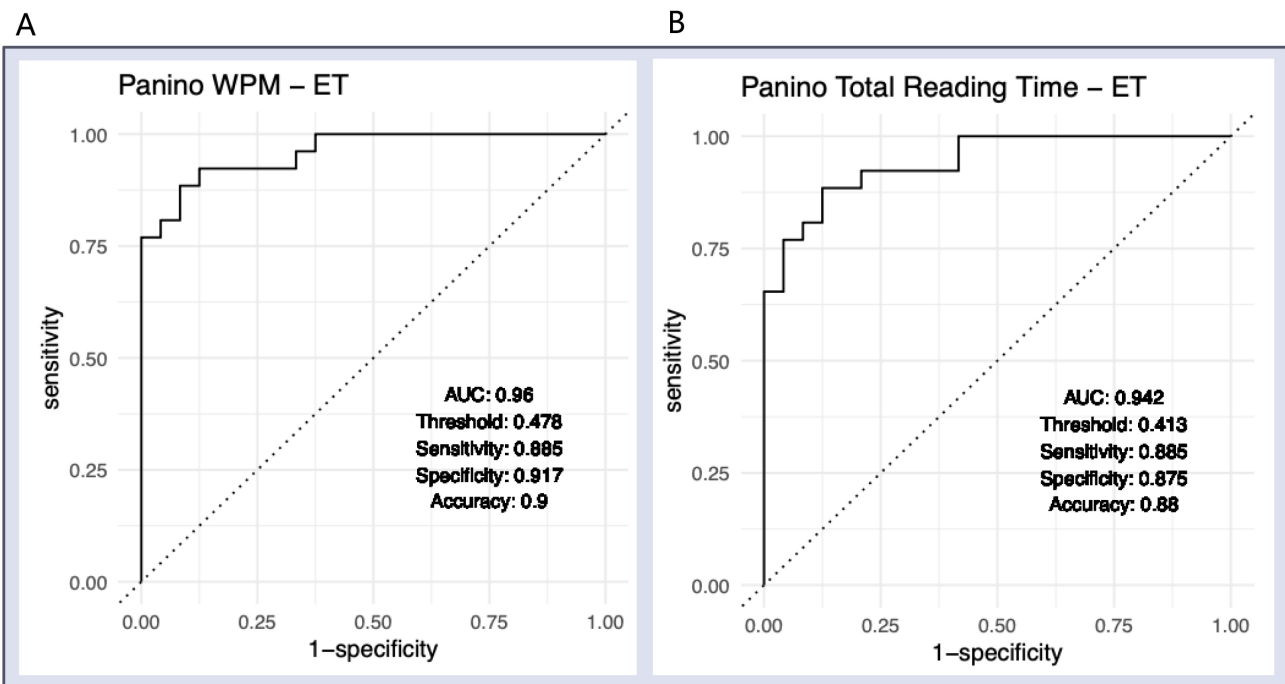
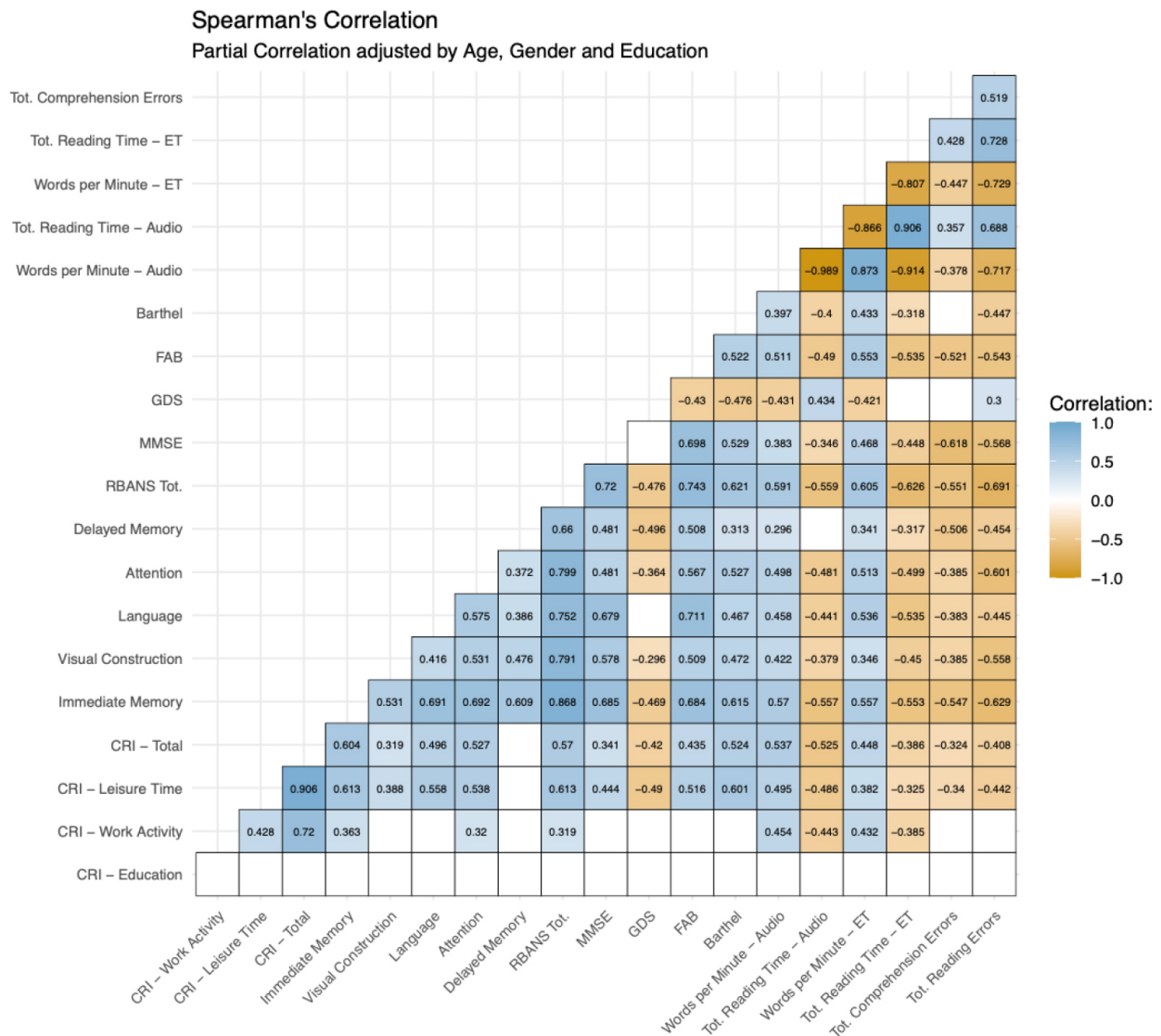


Fig. 6. ROC curves for timing-based reading metrics (Eye Tracking). (A) WPM. (B) Total Reading Time: models adjusted for age, gender, education, GDS, Barthel Index, and CRIq.

sensitivity of 0.800, specificity of 0.900, and accuracy of 0.850. Overall, these results indicate that both reading speed and total reading duration can effectively distinguish cognitive impairment status in this pilot sample, although neither metric surpassed the error-based measures.

For the eye-tracking timing-based reading metrics (Fig. 6), both WPM and Total Reading Time showed a strong ability to distinguish between the CI and HC groups. WPM in the eye-tracking condition (Fig. 6A) achieved an AUC of 0.960, with an optimal probability threshold of



All colored squares represent a significant correlation at level 0.05 after fdr correction

Fig. 7. Partial Spearman correlation matrix adjusted for age, gender, and education. Only correlations that remained significant after false discovery rate (FDR) correction for multiple comparisons are displayed and color-coded; blank cells indicate either redundant cells in the symmetric matrix or associations that did not survive FDR correction. Blue indicates positive correlations and orange indicates negative correlations, with color intensity reflecting the magnitude of the association. ET metrics were available for a reduced subsample of 50 participants. Abbreviations: CRI, Cognitive Reserve Index; RBANS Tot., Repeatable Battery for the Assessment of Neuropsychological Status total score; Tot. Reading Time, Total Reading Time; Tot. Comprehension Errors, total comprehension errors; Tot. Reading Errors, total reading errors.

0.478, sensitivity of 0.885, specificity of 0.917, and accuracy of 0.900. Total Reading Time in the eye-tracking condition (Fig. 6B) also performed well, with an AUC of 0.942, an optimal probability threshold of 0.413, sensitivity of 0.885, specificity of 0.875, and accuracy of 0.880. Overall, these results suggest that eye-tracking-derived timing metrics effectively differentiate CI from HC participants in this pilot sample and show slightly better discriminatory performance than the corresponding audio-based timing measures.

Regarding the fourth objective, partial Spearman correlations, adjusted for age, gender, and education, are shown in Fig. 7. Only correlations that remained significant after FDR correction are shown. Overall, the reading-derived measures were consistently associated with cognitive performance, executive functioning, depressive symptoms, functional status, and cognitive reserve. TRE was significantly linked to poorer neuropsychological performance, showing negative correlations with RBANS total score ($\rho = -0.69$), MMSE ($\rho = -0.57$), and FAB ($\rho = -0.54$), and a positive correlation with GDS ($\rho = 0.30$). TCE

exhibited a similar pattern, with negative correlations with the RBANS total score ($\rho = -0.55$), the MMSE ($\rho = -0.62$), and the FAB ($\rho = -0.52$).

Timing-based measures were also strongly interconnected. In the audio condition, WPM was highly inversely correlated with Total Reading Time ($\rho = -0.99$), positively correlated with eye-tracking WPM ($\rho = 0.87$), and negatively correlated with eye-tracking Total Reading Time ($\rho = -0.91$). Audio WPM was also related to fewer TRE ($\rho = -0.72$) and fewer TCE ($\rho = -0.38$), while audio Total Reading Time was positively associated with TRE ($\rho = 0.69$) and TCE ($\rho = 0.36$). A similar pattern appeared for eye-tracking metrics: eye-tracking WPM was inversely related to eye-tracking Total Reading Time ($\rho = -0.81$), TRE ($\rho = -0.73$), and TCE ($\rho = -0.45$), whereas eye-tracking Total Reading Time was positively linked to both TRE ($\rho = 0.73$) and TCE ($\rho = 0.43$).

Cognitive reserve was also meaningfully related to reading performance. Higher Cognitive Reserve Index (CRI)-Total scores were linked to faster reading speeds in both the audio ($\rho = 0.54$) and eye-tracking ($\rho = 0.45$) conditions, shorter reading times in the audio condition ($\rho = -0.53$), and fewer reading and comprehension errors (TRE: $\rho = -0.41$; TCE: $\rho = -0.32$). Similar, but slightly weaker, patterns were observed for CRI-Leisure Time and CRI-Work Activity. Overall, these adjusted and FDR-corrected correlations support the convergent validity of the reading-derived measures with established neuropsychological, affective, and functional indices, and indicate that cognitive reserve influences individual differences in reading efficiency and error rates.

Finally, exploratory analyses of conventional first-order eye-tracking metrics are reported in the **Supplementary Materials**. Unlike the eye-tracking-derived timing measures presented earlier in the main Results, traditional oculomotor variables did not distinguish between groups. Specifically, no significant between-group differences were observed in fixation durations, saccadic amplitudes, or saccadic durations (all $p > 0.05$; see **Supplementary Fig. 3**). Overall, these findings suggest that, under the present acquisition conditions, conventional first-order eye-movement measures were less sensitive to cognitive status than reading-derived timing and error-based indices.

4. Discussion

The present study investigated the feasibility and potential diagnostic value of an ecological reading task administered in a nursing home setting to identify digital linguistic markers associated with cognitive impairment in older adults. Specifically, the study aimed to determine whether reading-derived linguistic metrics, i.e., total reading errors, total comprehension errors, and WPM, as well as selected eye-tracking indicators, may distinguish cognitively impaired residents from healthy controls.

The results support the primary hypothesis that performance on an ecologically valid reading task is associated with cognitive status in institutionalized older adults. Overall, the findings demonstrate that the task was feasible in a real-world nursing home environment and that several reading-based measures showed strong discriminative capacity between cognitively impaired and cognitively healthy individuals. Consistent with the study hypotheses, participants in the cognitively impaired group performed significantly worse than healthy controls across multiple linguistic reading metrics. They exhibited higher rates of reading errors and comprehension errors, slower reading speeds, and longer overall reading times. These differences remained significant even after adjusting for relevant covariates, including age, gender, education, depressive symptoms, and functional status. Among the tested indicators, comprehension errors showed the strongest association with cognitive impairment, suggesting that comprehension-based measures may be particularly sensitive to cognitive decline in this population. This finding aligns with previous research indicating that individuals with neurodegenerative conditions may retain the ability to decode written words while experiencing deficits in semantic integration and comprehension processes. The strong predictive value of comprehension errors, therefore, supports the idea that higher-order language processing tasks may reveal early cognitive disruptions more effectively than simple reading accuracy alone.

Reading speed also emerged as a meaningful marker. The lower WPM observed in the cognitively impaired group is consistent with earlier findings showing that neurodegenerative disorders often affect the temporal dynamics of reading. Slower reading may reflect reduced processing efficiency, increased cognitive load, or compensatory strategies to maintain comprehension. Importantly, the receiver operating characteristic analyses revealed high discriminative performance for all three linguistic metrics, with particularly high area under the curve values for comprehension errors and reading errors. These results suggest that even a brief reading task can generate quantifiable linguistic indicators with substantial diagnostic potential.

The eye-tracking component of the study provided additional insight into the mechanisms underlying these performance differences. Timing-related eye-movement measures appeared broadly consistent with the linguistic results, suggesting that altered reading behavior in cognitively impaired participants may partly reflect changes in processing dynamics during text engagement. However, first-order oculomotor metrics did not significantly differ between groups. This result suggests that the observed reading impairments are unlikely to stem primarily from basic oculomotor dysfunction. Instead, they may reflect higher-level cognitive processes such as lexical access, semantic integration, attention regulation, or executive control. In other words, the findings reinforce the interpretation that

cognitive impairment affects the cognitive architecture of reading rather than the mechanical eye-movement system itself. These results are consistent with prior studies suggesting that reading difficulties in neurodegenerative disorders often arise from disruptions in linguistic and cognitive processing rather than from fundamental oculomotor deficits.

Beyond the specific performance findings, an important contribution of this study pertains to the practical feasibility of it in a nursing home environment. All participants were able to complete the reading task, and audio recordings were obtained for the entire sample. Furthermore, eye-tracking data were successfully collected for the majority of participants despite the challenges typically associated with real-world data acquisition in older populations. These results indicate that ecological digital assessment approaches can be integrated into residential care settings without being too demanding for older adults. Given that cognitive decline is frequently under-recognized in institutionalized older adults, such accessible and scalable tools may play a valuable role in early detection and monitoring.

Despite these promising findings, several limitations should be considered when interpreting these preliminary results. First, the study employed a cross-sectional design, which prevents conclusions about causal relationships or the longitudinal progression of cognitive decline. Future research should examine whether the identified reading-based markers are sensitive to changes over time and whether they can predict transitions from mild cognitive impairment to dementia. Second, the sample size was relatively small and based on a single institutional context, an issue that may limit the generalizability of the findings. Replication in larger and more diverse populations, including residents from multiple facilities and individuals with different clinical profiles, would probably strengthen the external validity of these interesting preliminary results. Another limitation concerns the heterogeneity of cognitive impairment among participants. Although individuals were categorized based on clinical and neuropsychological evaluation, cognitive impairment in older adults can arise from multiple etiologies and may manifest differently across individuals. Future studies might benefit from examining specific diagnostic subgroups or incorporating biomarker information to better characterize underlying disease processes. Additionally, although eye-tracking data were available for most participants, calibration difficulties reduced the usable sample for this component of the analysis. While this is common in studies involving older adults, it highlights the need for continued refinement of eye-tracking methodologies for use in naturalistic clinical environments.

Along with its limitations, the study also has notable strengths. It combines ecological task design, linguistic analysis, and objective digital measurement within a population that is often underrepresented in cognitive research. The use of a realistic reading passage mirrors everyday ac-

tivities more closely than many traditional neuropsychological tests, thereby enhancing ecological validity. Moreover, the integration of linguistic performance metrics with eye-tracking data provides a multimodal perspective on reading behavior and cognitive function. The study also controlled for several potentially confounding variables, including education, depressive symptoms, and functional status, which strengthens confidence in the observed associations.

Future research in the field should further explore the potential of reading-based digital markers as screening or monitoring tools for cognitive decline. Longitudinal designs could clarify whether these metrics can detect early cognitive changes before they become clinically apparent. In addition, expanding the range of linguistic and eye-movement features, such as pause patterns, regressions, lexical difficulty responses, and variability in reading dynamics, may reveal more nuanced signatures of cognitive impairment. Advances in automated speech analysis and machine learning could also enable the development of scalable digital assessment platforms capable of extracting complex linguistic features from routine reading tasks. Finally, integrating these digital markers with other modalities, such as speech analysis, wearable sensor data, or neuroimaging, may contribute to more comprehensive models of cognitive health in aging populations.

5. Conclusions

In conclusion, the present findings suggest that an ecological reading task can be successfully implemented in nursing home settings and can generate meaningful linguistic indicators associated with cognitive impairment. Measures of comprehension accuracy, reading errors, and reading speed demonstrated strong discriminatory capacity between cognitively impaired residents and healthy controls, while eye-tracking results pointed toward cognitive rather than purely oculomotor mechanisms underlying these differences. Together, these findings support the potential of reading-derived digital markers as practical and informative tools for detecting cognitive decline in older adults living in residential care environments.

Declarations

Some of the preliminary findings from the study were presented as an oral communication at the Healthy Aging Week Conference 2025.

Abbreviations

AD, Alzheimer's disease; CI, cognitively impaired; DTI, peripheral details presented either implicitly; DTS, peripheral details presented either explicitly; FAB, Frontal Assessment Battery; GDS, Geriatric Depression Scale; MCI, mild cognitive impairment; MII, implicit main ideas; MIS, explicitly stated information; MMSE, Mini-Mental State Examination; TCE, total comprehension errors; TRE, total reading errors; WPM, words per minute.

Availability of Data and Materials

Raw data will be made available upon request. Researchers interested in participating in a multicenter study are asked to contact the corresponding author.

Author Contributions

SC designed the research study. SC conducted the research. MC and ATD collected the data. ATD and SC analyzed the data. All authors contributed to editorial changes to the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and have agreed to be accountable for all aspects of it.

Ethics Approval and Consent to Participate

The study was conducted in accordance with the Declaration of Helsinki. Approval of the research protocol was obtained from the Ethics Commission of eCampus University (Protocol No. 07/2024, 5 December 2024). All participants provided signed informed consent, and all caregivers were informed about the study and gave their agreement for their loved ones to participate.

Acknowledgment

We would like to gratefully acknowledge the management and all the residents of the Saint George nursing home in Cavallermaggiore (Cuneo), Italy. In particular, we would like to thank Lorenzo Signore, a resident who worked closely with the research team in a spirit of collaborative research, sharing his perspective and offering valuable insights that helped shape a person-centered study.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflicts of interest. Marco Cavallo is serving as one of the Editorial Board members and a Guest Editor of this journal. We declare that Marco Cavallo had no involvement in the peer review of this article and has no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to Bettina Platt.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at <https://doi.org/10.31083/JIN49652>.

References

- [1] Chen M, Hanewald K, Si Y, Gu Y, Beard JR. Intrinsic Capacity Across 15 Countries in the Survey of Health, Aging, and Retirement in Europe. *JAMA Network Open*. 2025; 8: e259792. <https://doi.org/10.1001/jamanetworkopen.2025.9792>.
- [2] Kitani T, Matoba S, Naito Y, Horvath S, Yoshimori T, Chan FKL, *et al.* Towards global healthy longevity: report from the 1st World Longevity Summit in Kyotango, Japan. *Npj Aging*. 2025; 11: 89. <https://doi.org/10.1038/s41514-025-00279-0>.
- [3] ISTAT INdS. Rapporto annuale 2025. L'invecchiamento attivo nel nostro Paese. 2025. Available at: <https://www.istat.it/new-s-dati-alla-mano/invecchiamo-bene/> (Accessed: 27 December 2025).
- [4] Livingston G, Huntley J, Liu KY, Costafreda SG, Selbæk G, Alladi S, *et al.* Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *Lancet* (London, England). 2024; 404: 572–628. [https://doi.org/10.1016/S0140-6736\(24\)01296-0](https://doi.org/10.1016/S0140-6736(24)01296-0).
- [5] Frisoni GB, Hansson O, Nichols E, Garibotto V, Schindler SE, van der Flier WM, *et al.* New landscape of the diagnosis of Alzheimer's disease. *Lancet* (London, England). 2025; 406: 1389–1407. [https://doi.org/10.1016/S0140-6736\(25\)01294-2](https://doi.org/10.1016/S0140-6736(25)01294-2).
- [6] ADI. World Alzheimer Report 2025: Reimagining life with dementia – the power of rehabilitation. 2025. Available at: <https://www.alzint.org/resource/world-alzheimer-report-2025/> (Accessed: 27 December 2025).
- [7] Salari N, Lotfi F, Abdolmaleki A, Heidarian P, Rasoulpoor S, Fazeli J, *et al.* The global prevalence of mild cognitive impairment in geriatric population with emphasis on influential factors: a systematic review and meta-analysis. *BMC Geriatrics*. 2025; 25: 313. <https://doi.org/10.1186/s12877-025-05967-w>.
- [8] Haraldsen IH, Hatlestad-Hall C, Marra C, Renvall H, Maestú F, Acosta-Hernández J, *et al.* Intelligent digital tools for screening of brain connectivity and dementia risk estimation in people affected by mild cognitive impairment: the AI-Mind clinical study protocol. *Frontiers in Neuroinformatics*. 2023; 17: 1289406. <https://doi.org/10.3389/fnbot.2023.1289406>.
- [9] Campbell J, Lavoie L, Farraia M, Huelin R, Zhang Q, Tahami Monfared AA. Quality of Life in Mild Cognitive Impairment and Mild Dementia Associated with Alzheimer's Disease: A Systematic Review. *Neurology and Therapy*. 2025; 14: 7–26. <https://doi.org/10.1007/s40120-024-00676-9>.
- [10] Malara A, Zucchelli A, Trevisan C, Borselli G, Incalzi RA, Onder G, *et al.* Polypharmacy appropriateness in Italian Long-Term Care Facilities: the nationwide prescription day point survey. *Aging Clinical and Experimental Research*. 2025; 37: 291. <https://doi.org/10.1007/s40520-025-03183-5>.
- [11] Cavallo M, Cecchetti S. "High-tech" cognitive assessment and intervention in neurocognitive disorders: Acceptable, feasible, effective? XXV Congresso Nazionale AIP, Sezione di Psicologia Clinica e Dinamica; 2025 Sep 17-20; Perugia, Italy. *Mediterranean Journal of Clinical Psychology*, 13 (2, Suppl.), 546-542. <https://doi.org/10.13129/2282-1619/mjcp-5106> Available at: <https://iris.uniecampus.it/retrieve/fa7ea230-d562-4a27-a33a-9f4d459e05a9/AIP%20Clinica%202025.pdf> Accessed: 22 March 2026.
- [12] Iaboni A, Spasojevic S, Newman K, Schindel Martin L, Wang A, Ye B, *et al.* Wearable multimodal sensors for the detection of behavioral and psychological symptoms of dementia using personalized machine learning models. *Alzheimer's & Dementia* (Amsterdam, Netherlands). 2022; 14: e12305. <https://doi.org/10.1002/dad2.12305>.
- [13] Badawi A, Elmoghazy S, Choudhury S, Elgazzar S, Elgazzar K, Burhan AM. Multimodal Detection of Agitation in People With Dementia in Clinical Settings: Observational Pilot Study. *JMIR Aging*. 2025; 8: e68156. <https://doi.org/10.2196/68156>.
- [14] Shi C, Wong GHY, Choy JCP, Tang JYM, Luo H, Lu S, *et al.* Association of dementia diagnosis, cognitive impairment levels, and their combination with care costs among publicly funded long-term care recipients. *Innovation in Aging*. 2025; 9: igaf088. <https://doi.org/10.1093/geroni/igaf088>.

- [15] Cahill S, Diaz-Ponce AM, Coen RF, Walsh C. The underdetection of cognitive impairment in nursing homes in the Dublin area. The need for on-going cognitive assessment. *Age and Ageing*. 2009; 39: 128–131. <https://doi.org/10.1093/ageing/afp198>.
- [16] Cecchetti S, Roy U, Cavallo M. From the Eyes to the Suffering Brain: Can Eye Tracking be Considered as a Significant Digital Biomarker for Neurological Diseases? *Journal of Integrative Neuroscience*. 2024; 23: 167. <https://doi.org/10.31083/jjin2309167>.
- [17] Cecchetti S, Duchowski AT, Cavallo M. Eye-Tracking Metrics as a Digital Biomarker for Neurocognitive Disorders in Multiple Sclerosis: A Scoping Review. *Brain Sciences*. 2025; 15: 149. <https://doi.org/10.3390/brainsci15020149>.
- [18] Rayner K, Pollatsek A, Ashby J, Clifton C Jr. *Psychology of Reading*. 2nd edn. Psychology Press: New York, NY. 2012.
- [19] Holmqvist K, Andersson R. *Eye-tracking: A comprehensive guide to methods, paradigms and measures*. Oxford University Press: Oxford. 2017.
- [20] Duchowski A. *Eye Tracking Methodology*. Springer: Cham. 2017. <https://doi.org/10.1007/978-3-319-57883-5>.
- [21] Ripamonti E, Lucchelli F, Lazzati G, Martini E, Luzzatti C. Reading impairment in neurodegenerative diseases: a multiple single-case study. *Aphasiology*. 2017; 31: 519–541. <https://doi.org/10.1080/02687038.2016.1208802>.
- [22] Martínez-Nicolás I, Martínez-Sánchez F, Ivanova O, Meilán JGG. Reading and lexical-semantic retrieval tasks outperforms single task speech analysis in the screening of mild cognitive impairment and Alzheimer’s disease. *Scientific Reports*. 2023; 13: 9728. <https://doi.org/10.1038/s41598-023-36804-y>.
- [23] Martínez-Sánchez F, Meilán JGG, García-Sevilla J, Carro J, Arana JM. Oral reading fluency analysis in patients with Alzheimer disease and asymptomatic control subjects. *Neurología (English Edition)*. 2013; 28: 325–331. <https://doi.org/10.1016/j.nrleng.2012.07.017>.
- [24] Moreno JD, León JA, Arnal LAM, Botella J. Age Differences in Eye Movements During Reading: Degenerative Problems or Compensatory Strategy? *European Psychologist*. 2019; 24: 297–311. <https://doi.org/10.1027/1016-9040/a000344>.
- [25] Lueck KL, Mendez MF, Perryman KM. Eye movement abnormalities during reading in patients with Alzheimer disease. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*. 2000; 13: 77–82.
- [26] Shah HA, Andberg S, Koivisto AM, Bednarik R. A Multimodal Approach for Early Identification of Mild Cognitive Impairment and Alzheimer’s Disease With Fusion Network Using Eye Movements and Speech. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: a Publication of the IEEE Engineering in Medicine and Biology Society*. 2025; 33: 1449–1459. <https://doi.org/10.1109/TNSRE.2025.3561043>.
- [27] Groznik V, Možina M, Lazar T, Georgiev D, Semeja A, Sadikov A. Validation of reading as a predictor of mild cognitive impairment. *Scientific Reports*. 2025; 15: 12834. <https://doi.org/10.1038/s41598-025-94583-0>.
- [28] Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J. DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*. 2001; 108: 204–256. <https://doi.org/10.1037/0033-295x.108.1.204>.
- [29] Kintsch W. The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review*. 1988; 95: 163–182. <https://doi.org/10.1037/0033-295x.95.2.163>.
- [30] Reichle ED, Rayner K, Pollatsek A. The E-Z reader model of eye-movement control in reading: comparisons to other models. *The Behavioral and Brain Sciences*. 2003; 26: 445–476; discussion 477–526. <https://doi.org/10.1017/s0140525x03000104>.
- [31] Hessels RS, Niehorster DC, Nyström M, Andersson R, Hollerman GA, Hooge ITC. The fundamentals of eye tracking part 5: The importance of piloting. *Behavior Research Methods*. 2025; 57: 216. <https://doi.org/10.3758/s13428-025-02737-9>.
- [32] Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975; 12: 189–198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6).
- [33] Magni E, Binetti G, Bianchetti A, Rozzini R, Trabucchi M. Mini-Mental State Examination: a normative study in Italian elderly population. *European Journal of Neurology*. 1996; 3: 198–202. <https://doi.org/10.1111/j.1468-1331.1996.tb00423.x>.
- [34] Randolph C, Tierney MC, Mohr E, Chase TN. The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): preliminary clinical validity. *Journal of Clinical and Experimental Neuropsychology*. 1998; 20: 310–319. <https://doi.org/10.1076/jcen.20.3.310.823>.
- [35] Duff K, Humphreys Clark JD, O’Byrant SE, Mold JW, Schiffler RB, Sutker PB. Utility of the RBANS in detecting cognitive impairment associated with Alzheimer’s disease: sensitivity, specificity, and positive and negative predictive powers. *Archives of Clinical Neuropsychology: the Official Journal of the National Academy of Neuropsychologists*. 2008; 23: 603–612. <https://doi.org/10.1016/j.acn.2008.06.004>.
- [36] Duff K, Suhrie KR, Hammers DB, Dixon AM, King JB, Koppelmans V, *et al*. Repeatable battery for the assessment of neuropsychological status and its relationship to biomarkers of Alzheimer’s disease. *The Clinical Neuropsychologist*. 2023; 37: 157–173. <https://doi.org/10.1080/13854046.2021.1995050>.
- [37] Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: a Frontal Assessment Battery at bedside. *Neurology*. 2000; 55: 1621–1626. <https://doi.org/10.1212/wnl.55.11.1621>.
- [38] Appollonio I, Leone M, Isella V, Piamarta F, Consoli T, Villa ML, *et al*. The Frontal Assessment Battery (FAB): normative values in an Italian population sample. *Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*. 2005; 26: 108–116. <https://doi.org/10.1007/s10072-005-0443-4>.
- [39] Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, *et al*. Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of Psychiatric Research*. 1982; 17: 37–49. [https://doi.org/10.1016/0022-3956\(82\)90033-4](https://doi.org/10.1016/0022-3956(82)90033-4).
- [40] Isella V, Villa ML, Appollonio IM. Screening and Quantification of Depression in Mild-to-Moderate Dementia Through the GDS Short Forms. *Clinical Gerontologist*. 2002; 24: 115–125. https://doi.org/10.1300/J018v24n03_10.
- [41] Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. *Maryland State Medical Journal*. 1965; 14: 61–65.
- [42] Galeoto G, Lauta A, Palumbo A, Castiglia SF, Mollica R, Santilli V, *et al*. The Barthel Index: Italian Translation, Adaptation and Validation. *International Journal of Neurology and Neurotherapy*. 2015; 2: 2–7. <https://doi.org/10.23937/2378-3001/211028>.
- [43] Nucci M, Mapelli D, Mondini S. Cognitive Reserve Index questionnaire (CRIq): a new instrument for measuring cognitive reserve. *Aging Clinical and Experimental Research*. 2012; 24: 218–226. <https://doi.org/10.1007/BF03654795>.
- [44] Stern Y. Cognitive reserve. *Neuropsychologia*. 2009; 47: 2015–2028. <https://doi.org/10.1016/j.neuropsychologia.2009.03.004>.
- [45] Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, *et al*. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*. 2019; 51: 195–203. <https://doi.org/10.3758/s13428-018-01193-y>.
- [46] Carlomagno S, Vorano L, Razzano C, Marini A. Un test per la valutazione della produzione e della comprensione di narrative nell’adulto cerebroleso con deficit di comunicazione. *Rivista di Psicologia Clinica*, 1, 76-96. 2013. Available

- at: <https://www.quadernidipsicologiaclinica.com/index.php/rp-c-archivio/it/article/view/1290/1395> (Accessed: 2 April 2026).
- [47] Welland RJ, Lubinski R, Higginbotham DJ. Discourse comprehension test performance of elders with dementia of the Alzheimer type. *Journal of Speech, Language, and Hearing Research: JSLHR*. 2002; 45: 1175–1187. [https://doi.org/10.1044/1092-4388\(2002/095\)](https://doi.org/10.1044/1092-4388(2002/095)).
- [48] Brookshire RH, Nicholas LE. The discourse comprehension test. Communication Skill Builder/The Psychological Corporation. 1993. Available at: <https://books.google.it/books?id=XrCZ0AEACAAJ> (Accessed: 2 April 2026).
- [49] Busjahn T, Tamm S. A deeper analysis of AOI coverage in code reading. In *ACM Symposium on Eye Tracking Research and Applications*. 2021. <https://doi.org/10.1145/3448018.3457422>.
- [50] ELAN. Version 7.0. Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics, The Language Archive. 2025.
- [51] Lausberg H, Sloetjes H. Coding gestural behavior with the NEUROGES–ELAN system. *Behavior Research Methods*. 2009; 41: 841–849. <https://doi.org/10.3758/BRM.41.3.841>.
- [52] R Core Team. *R: A Language and Environment for Statistical Computing*. Version 4.5.0. Vienna, Austria: R Foundation for Statistical Computing. 2025.
- [53] Sjoberg DD, Whiting K, Curry M, Lavery JA, Larmarange J. Reproducible Summary Tables with the gtsummary Package. *The R Journal*. 2021; 13: 570–580. <https://doi.org/10.32614/RJ-2021-053>.
- [54] Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. 2016. ISBN 978-3-319-24277-4, Available at: <https://ggplot2.tidyverse.org> (Accessed: 2 April 2026).