Assessment of Language Function in Older Mandarin-Speaking Adults with Mild Cognitive Impairment using Multifaceted Language Tests

Yun-Ting Tseng[a,b], Yu-Ling Chang[a,c,d,e,f,*] and Yen-Shiang Chiu[a]
[a]*Department of Psychology, College of Science, National Taiwan University, Taipei, Taiwan*
[b]*Department of Physical Medicine and Rehabilitation, Chi Mei Medical Center, Tainan, Taiwan*
[c]*Department of Neurology, College of Medicine, National Taiwan University, Taipei, Taiwan*
[d]*Graduate Institute of Brain and Mind Sciences, College of Medicine, National Taiwan University, Taipei, Taiwan*
[e]*Neurobiology and Cognitive Science Center, National Taiwan University, Taipei, Taiwan*
[f]*Center for Artificial Intelligence and Advanced Robotics, National Taiwan University, Taipei, Taiwan*

Abstract.
Background: Individuals with amnestic mild cognitive impairment (aMCI), especially for those with multidomain cognitive deficits, should be clinically examined for determining risk of developing Alzheimer's disease. English-speakers with aMCI exhibit language impairments mostly at the lexical–semantic level. Given that the language processing of Mandarin Chinese is different from that of alphabetic languages, whether previous findings for English-speakers with aMCI can be generalized to Mandarin Chinese speakers with aMCI remains unclear.
Objective: This study examined the multifaceted language functions of Mandarin Chinese speakers with aMCI and compared them with those without cognitive impairment by using a newly developed language test battery.
Methods: Twenty-three individuals with aMCI and 29 individuals without cognitive impairment were recruited. The new language test battery comprises five language domains (oral production, auditory and reading comprehension, reading aloud, repetition, and writing).
Results: Compared with the controls, the individuals with aMCI exhibited poorer performance in the oral production and auditory and reading comprehension domains, especially on tests involving effortful lexical and semantic processing. Moreover, the aMCI group made more semantic naming errors compared with their counterparts and tended to experience difficulty in processing items belonging to the categories of living objects.
Conclusions: The pattern identified in the present study is similar to that of English-speaking individuals with aMCI across multiple language domains. Incorporating language tests involving lexical and semantic processing into clinical practice is essential and can help identify early language dysfunction in Mandarin Chinese speakers with aMCI.

Keywords: Alzheimer's disease, language tests, Mandarin Chinese, mild cognitive impairment, neuropsychological tests

---

*Correspondence to: Yu-Ling Chang, PhD, Department of Psychology, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd, Taipei 10617, Taiwan. Tel.: +886 2 33663105; Fax: +886 2 23629909; E-mail: ychang@ntu.edu.tw; ORCID: 0000-0003-2851-3652.

INTRODUCTION

Mild cognitive impairment (MCI) represents a group of people with an increased risk of developing dementia [1, 2]. Because of heterogeneity in its clinical presentation, MCI is classified into four subtypes on the basis of two dimensions: primary episodic memory impairment (i.e., amnestic and nonamnestic) and impairment in single or multiple cognitive domains [3]. Amnestic MCI (aMCI) with impairment in single or multiple domains might be involved in the progression of Alzheimer's disease (AD) [4]. Compared with individuals with only memory impairment, those with aMCI in multiple domains, including language, are more likely to develop the AD type of dementia [5–8]. Thus, in individuals with aMCI, impairment in cognitive domains apart from episodic memory impairment should be clinically examined to determine their risk of dementia [9, 10].

Patients with dementia due to AD commonly exhibit language impairments in the aspects of verbal fluency [8, 11], semantic knowledge [12, 13], and connected speech processing [14, 15]. Similar to the pattern observed during the early stage of AD, language function deficits in aMCI are characterized at the lexical and semantic levels by impairments in verbal fluency and confrontation naming [15–19]. Moreover, individuals with aMCI make more semantic coordinate errors (e.g., naming a tiger as a bear) than other types of errors in naming tasks [13, 20]. The lexical-semantic deficits observed in aMCI are likely attributable to difficulties in lexical access and semantic processing, encompassing semantic representation and control [13, 21]. These difficulties in semantic processing are likely linked to compromised brain networks involving temporal and frontal regions and their connections [22, 23], as well as volumetric reductions in the perirhinal cortex, which is essential for integrating cross-modal object representations, especially in patients with aMCI [24, 25].

Although lexical–semantic impairments appear to be a crucial characteristic of language impairments in individuals with aMCI, conflicting results concerning individuals with aMCI have been reported [13, 18]. The verbal fluency task is among the most commonly used to assess lexical–semantic function, and it typically involves semantic and phonemic fluency tasks. In semantic verbal fluency tasks, participants are instructed to provide as many exemplars as possible for a given semantic category within a specific

time interval. In phonemic tasks, participants are asked to provide as many words as possible on the basis of a phonological feature (e.g., a specific letter) within a specific time interval. Compared with controls without cognitive impairment, individuals with aMCI exhibited poorer performance in both fluency tasks; however, the difference in performance was more prominent in the semantic fluency task than in the phonemic fluency task [19, 26–29]. Conversely, other studies have reported either no differential deficit between semantic and phonemic fluency in individuals with single-domain aMCI [30] or greater impairment in phonemic fluency than in semantic fluency [31] in individuals with aMCI compared with controls without cognitive impairment. These mixed results were likely due to the inclusion of heterogeneous MCI samples across studies (e.g., the inclusion of individuals with single-domain versus those with multidomain aMCI) and specific task methods (e.g., use of different semantic categories) in each study [18, 30].

Although many studies have examined the lexical-semantic aspect of language impairments in MCI, such as in naming and fluency tests, other language deficits have been observed. Comprehensive language test batteries, employed by Jokel et al. [32] and Tsantali et al. [33], have demonstrated that individuals with aMCI exhibit deficits in tasks involving language functions beyond lexical-semantic processing, such as sentence comprehension and auditory spelling. In addition, studies utilizing language tests that focus on aspects other than lexical and semantic processing, such as reading aloud, repetition, and writing abilities, have found more significant impairments in older adults with aMCI compared to healthy controls [34–38]. Poor performance in these tasks may be partially attributed to decreased executive control and working memory span, which can negatively impact processes involved in understanding various syntactic structures and temporally storing separate letters for different words [39]. These findings suggest that a multifaceted test battery can provide valuable information about the language function of individuals with aMCI.

The aforementioned studies have primarily enrolled English-speaking participants. Thus, whether their findings can be generalized to non-English-speaking participants, such as Mandarin Chinese speakers, remains unclear [40]. Mandarin Chinese, a logographic language, contains different phonological, semantic, and orthographical processing features from the English language, an

alphabetical language. For example, the phonological processing unit in Mandarin Chinese is the syllable rather than phoneme, as in English, and such syllable-based processing aligns with the fact that each Chinese character is an orthographical processing unit that also maps to a syllable [41]. In semantic processing, each Chinese character is a morpheme that directly represents meaning, whereas letters in alphabetical languages primarily represent sounds [42]. Moreover, lexical tone is a specific suprasegmental feature in Mandarin Chinese and reflects the categorical perception of pitch contours for native Mandarin Chinese speakers [43]. Studies examining neural networks for language processing have revealed common and distinct patterns between Mandarin and English [19, 44]. Thus, given the differences in linguistic features between the two languages, the extent to which the findings based on English speakers with aMCI are applicable to Mandarin Chinese speakers with aMCI should be explored.

Some studies enrolling Mandarin Chinese speakers have demonstrated that compared with healthy control participants (HCs), individuals with aMCI had deficits in lexical semantics [19, 45], tonal discrimination [46], sentence repetition [47], or increased semantic and phonemic errors in writing samples [48]. However, comparing the results of the aforementioned studies is difficult because they used diverse inclusion criteria for individuals with aMCI. Moreover, the aforementioned studies examined only a single aspect of language function without evaluating various language functions in the same individuals. Thus, the detailed and comprehensive language profile of Mandarin Chinese speakers with aMCI should be determined using a multifaceted language test battery.

Currently, there is a limited availability of test batteries to evaluate the multifaceted language functions of Mandarin Chinese-speaking populations. Translations of test batteries in other languages may prove insufficient due to the neglect of specific language features unique to Mandarin Chinese, such as challenges in tonal discrimination and dyslexia [40, 46, 49]. To the best of our knowledge, only the Concise Chinese Aphasia Test (CCAT) [50] has been developed and standardized to evaluate multifaceted language functions in Mandarin Chinese speakers. The CCAT consists of nine subtests that cover expressive and receptive language domains, including the tasks of answering simple questions regarding participants' personal information and life experiences,

auditory and reading comprehension, and writing. However, the CCAT was primarily developed for patients with focal brain lesions, and its test materials do not cover language aspects such as semantic knowledge and verbal fluency that are crucial for evaluating neurodegenerative disorders, such as dementia due to AD.

To address these research gaps, this study investigated and compared multifaceted language functions between older Mandarin Chinese speakers with aMCI and older adults without cognitive impairment by using a newly developed language test battery. This test battery covers five domains of language functions, namely oral production, auditory and reading comprehension, reading aloud, repetition, and writing, based on the input and output modalities during the language processing in each test [51]. Notably, incorporating writing tests into our battery development is imperative, especially when accounting for the distinctive features of Chinese in comparison to English. The Chinese writing system is distinguished by heightened character complexity, necessitating robust orthographic memory, a meticulous adherence to stroke order and structure, and heightened demands for lexical precision. This is particularly evident during a written dictation test, where participants are tasked with distinguishing between distinct characters that share the same phonetic representation (homophones) or exhibit similar phonetic variations (tonal variations). This challenge underscores the need for precise lexical discrimination in the context of the test.

Previous studies have indicated that individuals with aMCI primarily had lexical and semantic impairments. Accordingly, we hypothesized that compared with older adults without cognitive impairment, older adults with aMCI would exhibit impairments in the language function domains of oral production, and auditory and reading comprehension, thus requiring effortful lexical and semantic processing. In addition, we hypothesized that both the groups would demonstrate comparable performance in other language functions, including reading aloud, repetition, and writing, that require less effort in lexical and semantic processing. Further, we hypothesized that the aMCI group would primarily have language deficits in lexical and semantic processing and perform poorly in semantic verbal fluency and naming tests. Moreover, given the predominance of lexical–semantic deficits in aMCI, we hypothesized that compared with HCs, individuals with aMCI would make more semantic errors but comparable phonological errors in the

naming task. In light of the evidence suggesting executive dysfunction may contributed to difficulties in certain language tasks in aMCI [19, 32, 33, 39], we sought to investigate the associations between performances on executive function-related tests and those in each language domain.

MATERIALS AND METHODS

Participants

In this study, we included 23 individuals with aMCI (age: range = 61–82 years, $M \pm SD = 72.91 \pm 6.16$; 15 women and 8 men) and 29 HCs (age: range = 59–84 years, $M \pm SD = 71.45 \pm 7.16$; 17 women and 12 men). Twenty-eight of the 29 healthy controls underwent re-examination using the new language test battery to assess test-retest reliability. The revisit interval ranged from 21 to 39 days ($M = 27.68$, $SD = 4.66$). The individuals with aMCI were recruited from memory clinics at local hospitals, and the HCs were recruited from nearby residential communities. All participants were Asian, native Mandarin Chinese speakers, and had received a minimum of six years of education (with a range of schooling years between 6 to 19 years). This study is a part of an ongoing research project approved by the Research Ethics Committee of National Taiwan University Hospital. Written informed consent was obtained from all the participants before their participation in the study. Individuals were excluded if they had a history of, or were exhibiting, neurological diseases (e.g., epilepsy, stroke, and head injury), psychiatric illnesses, substance abuse (e.g., alcohol and illicit drug), or clinically significant sensory impairment (e.g., blindness or severe hearing deficiency). Moreover, individuals with developmental or acquired language impairments, such as developmental language disorder, learning disability, and aphasia after stroke, were excluded through thorough interviews assessing developmental and medical history.

According to standards proposed by the International Working Group [2], individuals were diagnosed as having MCI when they met all the following conditions: 1) being neither normal nor having dementia, 2) having preserved activities of daily living and either intact or minimally impaired complex instrumental functions, and 3) having evidence of objective cognitive impairment on neuropsychological tests. Objective cognitive impairment was determined if participants scored 1 SD below age-corrected norms on at least two neuropsychological

tests for one cognitive domain [52]. Individuals with memory impairment were defined as having aMCI [53]. Among the 23 participants with aMCI included in our study, 12 had single-domain aMCI and 11 had multiple-domain aMCI. The Clinical Dementia Rating (CDR) [54] was used to evaluate the daily functioning of the participants. All the participants with aMCI obtained a global score of 0.5 in the CDR. To stage the severity of cognitive impairment, the CDR Sum of Boxes score was calculated [55], and each participant completed the Chinese version of the Mini-Mental State Examination [56].

Neuropsychological evaluation

To evaluate baseline functions, the participants were administered a comprehensive neuropsychological test battery covering four cognitive domains: attention and processing speed, learning and memory, language, and executive function. Attention and processing speed were measured using the Digit Span Forward length of the Wechsler Adult Intelligence Scale–Third Edition (WAIS-III) [57] and the word condition of the Color–Word Interference Test (CWIT) of the Delis–Kaplan Executive Function System (D-KEFS) [58]. Learning and memory were evaluated using the immediate and delayed recall conditions of the Logical Memory (LM) subtest of the Wechsler Memory Scale–Third Edition (WMS-III) [59] and the immediate and delayed recall conditions of the Rey Complex Figure Test (RCFT) [60]. Language function was evaluated using the Vocabulary subtest of the WAIS-III and the 30-item Boston Naming Test [61]. Executive function was measured using the Part 2-minus-Part 1 reaction time of the Color Trails Test (CTT) [62] and the inhibition and switching condition of the CWIT. Participants' depressive status was assessed using the short form of the Geriatric Depression Scale (GDS) [63].

Mandarin Chinese language test: Construction, administration, and scoring

The Mandarin Chinese language test battery was constructed on the basis of the English language–processing framework from previous studies [64, 65]. Specific language features in Mandarin Chinese, such as lexical tones and Mandarin phonetic symbols (i.e., Zhuyin), were included to comprehensively investigate the processing of Mandarin Chinese. The test battery consisted of 20 tests belonging to five domains: 1) oral production, 2) auditory

and reading comprehension, 3) reading aloud, 4) repetition, and 5) writing ability. Table 1 summarizes the 20 tests. Each domain score was computed as a composite score by averaging z scores, which were based on the mean and SD of the HC group, across the tests within the corresponding language domain.

Picture stimuli were black-and-white line drawings developed on the basis of prior studies including sets of standardized pictures [51, 66, 67]. Auditory stimuli were prerecorded using a male voice for tests relying on auditory input processing to standardize test administration. Because the characteristics of language stimuli may affect test performance in clinical populations [51, 66], an independent group of participants without cognitive impairment ($n = 16$, age = 64.93 ± 11.66 years, 8 women and 8 men, years of education = 11.68 ± 3.53) was recruited to rate the familiarity and concreteness of the auditory and written word stimuli used in the test battery on a 5-point Likert scale. Specifically, the familiarity scale ranged from 1 (*not very familiar*) to 5 (*very familiar*). The concreteness scale ranged from 1 (*not very concrete*) to 5 (*very concrete*). All test materials in the present study were rated as having a medium-to-high level of familiarity and concreteness (ratings for all items were above 3.75). Practice items were included in all the tests prior to the formal testing phase to ensure that the participants understood instructions.

Statistical analysis

The independent t test and chi-square test were used to compare demographic, clinical, and neuropsychological data between the groups. If Levene's test for equality of variances yielded significant results, an independent t test with an adjusted degree of freedom was performed. For demographic and clinical data analyses, the α level of statistical significance was set at the conventional threshold ($p < 0.05$). To analyze performance on 10 neuropsychological tests, an α threshold of 0.005 was used with Bonferroni adjustment.

A 2 (group) $\times$ 5 (domain) mixed-design analysis of variance (ANOVA) was performed to analyze group differences among the five language domains. The group served as the between-subject factor, comprising the aMCI and HC groups, while the language domain acted as the within-subject factor, encompassing oral production, auditory and reading comprehension, reading aloud, repetition, and writing domains. Greenhouse–Geisser correction was employed when the assumption of sphericity was

violated in the ANOVA analysis. We analyzed the planned simple main effects of the groups on each language domain and the effects of the domains on each group. Whenever a simple main effect reached significance, *post hoc* analysis of pairwise comparisons was conducted. An α level of 0.05 was set for the main and interaction effects of group and language domains. For the simple main effects within each group on the five language domains, we used an α threshold of 0.01 based on the Bonferroni correction. Additionally, for the simple main effects of the language domains between the two groups, an α threshold of 0.025 was applied. The *post hoc* analysis of simple main effects used an α threshold of 0.005, following the Bonferroni correction for 10 pairwise comparisons.

When the main effect of the language domains or the interaction of group and language domains reached significance according to the ANOVA results, an independent t test was performed to further analyze group differences at the test level. If Levene's test for equality of variances yielded significant results, we conducted an independent t-test with adjusted degrees of freedom. We used the following α threshold values: 0.0125 (Bonferroni correction for four tests) for oral production and writing, 0.0071 (Bonferroni correction for seven tests) for auditory and reading comprehension, 0.0167 (Bonferroni correction for three tests) for reading aloud, and 0.025 (Bonferroni correction for two tests) for repetition domains.

Total numbers of semantic and phonological errors on the noun and verb naming test were calculated. The semantic errors were further classified into four distinct types, namely coordinate (e.g., "dog" for cat), superordinate (e.g., "pet" for cat), subordinate (e.g., "British shorthair" for cat), and associate (e.g., "scratchers" for cat). A more comprehensive definition of these four semantic errors can be found in the supplementary materials. Independent t tests were conducted for semantic and phonological error types between the groups with adjusted degrees of freedom if Levene's test for equality of variances yielded significant results. The chi-square tests were performed for the four semantic error types to compare the occurrence of errors between the groups. Furthermore, we calculated whether the percentages of participants in each group had errors with ≥ 1 or < 1 SD than the mean number of errors in the HC group. The chi-square test was performed to examine the group difference in the percentages of the participants based on the cutoff. An α threshold of 0.025, adjusted for

Table 1
Summary of 20 tests included in the Mandarin Chinese language test battery

Test	Assessed ability	Description
Oral production domain (i.e., phoneme, word, sentence, and discourse levels of oral expression)		
Zhuyin blending and deletion test	Pronouncing and using Zhuyin symbols	Two sections: 1. Reading eight monosyllabic Zhuyin symbols presented in written form, such as *ben1* ㄅㄣ (score range: 0–16 points) 2. Deleting the first Zhuyin symbols (e.g., *shi1* ㄕ) from eight auditory syllables (e.g., *shan1* ㄕㄢ) and pronouncing the remaining phonetic blends (e.g., *an1* ㄢ; score range: 0–16 points)
Sentence production and spontaneous speech test	Verbally producing sentences and discourses	Two sections: 1. Speaking complete sentences for five pictures depicting constrained subjects, objects, and verbs, such as "A groom is hugging a bride" (score range: 0–28 points) 2. Answering six questions on personal information and experiences (e.g., How was your day yesterday?) as well as describing one picture within 3 min (score range: 0–20 points)
Noun and verb naming test	Naming objects and actions on pictures	Producing the names of 32 pictures that depict an object (e.g., shoes; score range: 0–16 points) or an action (e.g., laugh; score range: 0–16 points)
Verbal fluency test	Generating exemplars on the basis of a rule within a time interval	Two sections: 1. Providing exemplars on the basis of a single semantic category (i.e., animals, fruits, and appliances) within 1 min 2. Providing exemplars with a rule to alternate between vegetables and vehicles within 1 min
Auditory and reading comprehension domain (i.e., phoneme, tone, word, and sentence levels of comprehension abilities)		
Phoneme and tone recognition test	Discriminating monosyllabic auditory pairs and identifying lexical tones	Two sections: 1. Orally responding to "the same" or "different" for the 19 auditory pairs of Zhuyin symbols or tones, such as *man4* ㄇㄢ vs. *fan4* ㄈㄢ, or *ba2* vs. *ba1* ㄅㄚ (score range: 0–38 points) 2. Choosing the best answer from four written tonal signs after listening to each auditory stimulus (e.g., *xi1* ㄒㄧ; eight stimuli in total; score range: 0–16 points)
Word–picture matching test	Matching the meaning of a word to its corresponding picture	Choosing one picture that best matches the meaning of 14 target nouns (e.g., rooster), which are presented either auditorily or visually shown above the four picture stimuli (score range: 0–56 points for the sum of auditory and written forms)
Attribute verification test	Processing semantic attributes of concrete nouns through a sentence verification task	For each concrete noun (e.g., watermelon; 10 nouns in total), providing a "yes" or "no" response to two statements related to the noun (e.g., "Watermelon can be eaten.") as well as choosing the best answer for a question (e.g., "Is the shape of watermelon round or diamond?"; score range: 0–120 points for the sum of auditory and written forms)
Relatedness judgement test for concrete and abstract words	Identifying semantic relations between concrete or abstract two-character words	Indicating which choice is semantically related to the target word (e.g., determining whether 天氣 *tian1 qi4* [weather] or 姿勢 *zi1 shi4* [posture] is semantically related to 氣象 *qi4 xiang4* [meteorology]; 20 words in total). The target word is read by the examiner in the auditory form before two choice words or is displayed above two choice words on paper in the written form (score range: 0–80 points for the sum of auditory and written forms)
Sentence and syntactic comprehension test	Matching the meaning of a sentence to its corresponding image	Choosing one picture that best matches the meaning of 17 sentences (e.g., "A cook is chasing a mail carrier."), which are presented auditorily by the examiner or visually shown above the four picture stimuli (score range: 0–68 points for the sum of auditory and written forms)
Idiom comprehension test	Explaining nonliteral meanings of idioms from semantic knowledge	Verbally elucidating the nonliteral or deeper levels of meaning for eight idioms presented auditorily (e.g., "A dog is barking at a train."; score range: 0–16 points)
Written lexical decision test	Identifying the valid written lexical forms of two-character words	Responding "yes" or "no" to 16 written words (e.g., 以前 *i3 qian2* [before] and 燃兒 *ran2 er2*) to indicate their validity as words (score range: 0–32 points)

(Continued)

Table 1
(Continued)

Test	Assessed ability	Description
Reading aloud domain (i.e., word to sentence levels of reading aloud ability)		
Word reading test	Recognizing and pronouncing written characters	Reading aloud 10 regular characters (e.g., 試 *shi4*) and 10 irregular characters (e.g., 抽 *chou1*) presented visually (score range: 0–20 points)
Pseudoword reading test	Reading aloud pseudoncharacters	Reading aloud four pseudo-characters presented visually (e.g., 稿; score range: 0–8 points); predetermining correct answers based on pronunciations from pseudo-characters' orthographical neighbors (i.e., characters with the same phonetic radicals)
Sentence reading test	Recognizing and pronouncing written sentences	Reading aloud 10 sentences, each with lengths spanning from 3 to 12 characters and encompassing diverse sentence structures, such as active, passive, and interrogative forms (e.g., "A cat is catching a sparrow."; score range: 0–20 points)
Repetition domain (i.e., phoneme, word, and sentence levels of repetition abilities)		
Word and nonword repetition test	Repeating words and nonwords	Two sections: 1. Repeating 10 auditory words, including two monosyllabic Zhuyin symbol(e.g., *be1* ㄅ) and eight two-character words (e.g., 問題 *un4 ti2* [problem]; score range: 0–20 points) 2. Repeating five auditory nonwords, each comprising two to six syllables and formed from low-frequency characters (e.g., 乖扛 *guai1 kang2*; score range: 0–10 points)
Sentence repetition test	Repeating sentences	Repeating 10 sentences with lengths ranging from 3 to 12 characters and encompassing different sentence types, such as active, passive, and interrogative forms; for example, "The teacher bought three books" (score range: 0–20 points)
Writing domain (i.e., word to sentence levels of writing ability)		
Written naming test	Writing the names of objects on pictures	Writing the names of five objects ranging from one to three characters (e.g., ruler; score range: 0–10 points based on the correctness of each character)
Written to dictation test	Writing words that are presented auditorily	Writing five nouns read by the examiner and ranging from two to three characters (e.g., cat; score range: 0–11 points based on the correctness of each character)
Written sentence test	Writing basic sentences to describe pictures	Writing five complete sentences on the basis of constrained subjects, verbs, and objects shown on pictures, such as "A groom is hugging a bride" (score range: 0–28 points)
Copy of the written word test	Constructing written words and nonwords	Copying four real words (e.g., 下 *xia4* [down]) and one pseudo-character precisely (e.g., 嶚; score range: 0–101 points based on the correctness of each stroke)

See the Supplementary Material for detailed information on construction, administration, and scoring of each language test.

the Bonferroni correction to account for semantic and phonological error types, as well as an α threshold of 0.0125, reflecting the Bonferroni correction for the four semantic error types, were used in the analysis of naming errors. Moreover, for the tests that exhibited significant group effects, an analysis of the passing rates of items, with a cutoff of < 90%, by language characteristics (e.g., semantic categories or veracity of statements) was performed in the aMCI group.

To evaluate the associations between language domains and executive function, we generated a composite variable for executive function by standardizing individual scores on the two executive function measures using the means and standard deviations of the entire cohort. The z-scores from both measures were then averaged. Subsequently, Pearson correlations were calculated between each language

domain and the composite executive function score, using data from the full cohort. An α threshold of 0.01 (Bonferroni correction for five language domains) was used for the correlation analysis.

To ensure the reliability and validity of the language tasks utilized in this study, we conducted several psychometric evaluations, including test-retest reliability, inter-rater reliability, and convergent validity. Test-retest reliability was assessed by calculating the intraclass correlation coefficients (ICCs) and 95% confidence intervals (CIs) of each language domain's performance during the first and second administrations, utilizing a single-rating, absolute agreement, and 2-way mixed-effects model [68]. To evaluate inter-rater reliability, the spontaneous speech section and the idiom comprehension test were examined given their potentially subjective

Table 2
Demographic, clinical, and cognitive characteristics of individuals with amnestic mild cognitive impairment (aMCI) and healthy controls (HCs)

	aMCI ($n = 23$) $M \pm SD$	HC ($n = 29$) $M \pm SD$	Statistics	p
Demographic and clinical characteristics				
Age (y)	72.91 ± 6.16	71.45 ± 7.16	$t(50) = 0.78$	0.440
Education (y)	12.91 ± 2.82	14.14 ± 2.81	$t(50) = -1.56$	0.126
Sex (female/male)	15/8	17/12	$\chi^2(1, N = 52) = 0.24$	0.627
CDR-SB	0.61 ± 0.40	0.28 ± 0.29	$t(50) = 3.51$	**0.001**[*]
MMSE	27.35 ± 1.99	28.31 ± 1.51	$t(50) = -1.98$	0.053
GDS	2.35 ± 2.39	2.17 ± 2.12	$t(50) = 0.28$	0.780
Neuropsychological tests				
Attention/Processing speed				
WAIS-III Digit Span forward length	7.35 ± 1.47	7.52 ± 1.40	$t(50) = -0.42$	0.673
CWIT-word reading (s)	27.61 ± 7.70	25.38 ± 5.43	$t(50) = 1.22$	0.227
Learning and memory				
WMS-III LM immediate	29.13 ± 11.41	43.52 ± 10.30	$t(50) = -4.77$	**< 0.001**[#]
WMS-III LM delayed	15.74 ± 9.53	28.41 ± 7.98	$t(50) = -5.22$	**< 0.001**[#]
RCFT immediate	8.22 ± 6.25	17.50 ± 5.42	$t(50) = -5.73$	**< 0.001**[#]
RCFT delayed	7.78 ± 5.98	16.41 ± 5.60	$t(50) = -5.36$	**< 0.001**[#]
Language				
WAIS-III Vocabulary	39.57 ± 9.24	44.14 ± 10.65	$t(50) = -1.63$	0.110
30-item Boston Naming Test[†]	25.57 ± 3.45	27.45 ± 2.98	$t(50) = -2.11$	0.040
Executive function				
CTT Part 2-1 (s)	82.00 ± 30.48	53.79 ± 23.19	$t(49) = 3.76$	**< 0.001**[#]
CWIT-inhibition and switching (s)	107.96 ± 32.43	78.48 ± 16.06	$t(30.49) = 3.99$	**< 0.001**[#]

CDR-SB, Clinical Dementia Rating-Sum of Boxes; CTT, Color Trials Test; CWIT, Color–Word Interference Test; GDS, Geriatric Depression Scale; LM, Logical Memory subtest; MMSE, Mini-Mental State Examination; RCFT, Rey Complex Figure Test; WAIS-III, Wechsler Adult Intelligence Scale–Third Edition; WMS-III, Wechsler Memory Scale–Third Edition. [†]raw scores based on the spontaneous naming condition. [*]$p < 0.05$ for demographic and clinical characteristics. [#]$p < 0.005$, based on Bonferroni correction for 10 neuropsychological tests.

scoring involvement. To do so, we randomly selected five cases from the sample and had another rater, who was blind to the clinical diagnosis, rate them in addition to the original scores rated by the first author. The inter-rater reliability was determined using ICCs and 95% CIs based on a single-rating, absolute agreement, and 2-way random-effects model. To validate the new language tasks used in this study, we employed Pearson correlations to assess the relationship between the newly developed language battery and the CCAT battery across the entire cohort. Specifically, scores within the oral production domain underwent correlation analysis with the combined scores of the CCAT Simple Response, Expository Speech, and Naming subtests. Additionally, correlational analyses were conducted between the domain scores of auditory and reading comprehension in our developed test and the aggregate scores derived from the CCAT Auditory Comprehension and Reading Comprehension subtests. We also performed a correlational analysis on the reading aloud domain scores and the scores from the CCAT Reading Comprehension subtest. Furthermore, repetition domain scores were correlated with the CCAT Repetition subtest scores. Lastly, scores within the

writing domain underwent correlation analysis with the aggregated scores of the CCAT Copying and Spontaneous Writing subtests. The α threshold was set at 0.01 (Bonferroni correction) for the correlation analysis.

Effect sizes (Cohen's d, φ, and η_p^2) were calculated for significant results of the t test, χ^2 test, and ANOVA, respectively. All statistical analyses were conducted using SPSS (version 26.0; IBM Corp., Armonk, NY, USA) and JASP software (version 0.16.2) [69].

RESULTS

Differences in demographic, clinical, and cognitive characteristics between the groups

Table 2 presents the demographic, clinical, and cognitive data of the aMCI and HC groups. Age, years of education, sex distribution, and Mini-Mental State Examination (MMSE) and GDS scores did not significantly differ between the groups (all $p > 0.05$). The aMCI group had significantly higher CDR Sum of Boxes scores than did the HC group, $t(50) = 3.51$, $p < 0.001$, $d = 0.980$. In terms of raw

scores on neuropsychological tests, the aMCI group performed significantly poorer than did the HC group on the following cognitive tests: the immediate recall, $t(50) = -4.77$, $p < 0.001$, $d = -1.332$, and delayed recall, $t(50) = -5.22$, $p < 0.001$, $d = -1.458$, of the LM subtest of the WMS-III, the immediate recall, $t(50) = -5.73$, $p < 0.001$, $d = -1.600$, and delayed recall, $t(50) = -5.36$, $p < 0.001$, $d = -1.497$, of the RCFT, the Part 2-minus-Part 1 score of the CTT, $t(49) = 3.76$, $p < 0.001$, $d = 1.062$, and the inhibition and switching conditions of the CWIT, $t(30.49) = 3.99$, $p < 0.001$, $d = 1.196$. The two groups exhibited comparable performance in the Digit Span Forward length of the WAIS-III, the word condition of the CWIT of the D-KEFS, the Vocabulary subtest of the WAIS-III, and the 30-item Boston Naming Test (all $p > 0.005$).

Regarding missing data, one participant from the aMCI group did not complete the CTT due to color blindness. His data were excluded while comparing the CTT score between the groups. Another participant in the HC group did not complete the Zhuyin blending and deletion test due to a lack of experience with Zhuyin symbols. His data were excluded from the analyses concerning Zhuyin blending and deletion test scores.

Group differences in the Mandarin Chinese language test battery

The two-way ANOVA (group × domain) exhibited a significant main effect of group, $F(1, 49) = 6.59$, $p = 0.013$, $\eta_p^2 = 0.118$, and language domains, $F(3.42, 167.70) = 3.90$, $p = 0.007$, $\eta_p^2 = 0.074$. Furthermore, a two-way interaction of group and language domains reached significance, $F(3.42, 167.70) = 2.85$, $p = 0.033$, $\eta_p^2 = 0.055$. The aMCI group exhibited significantly poorer performance in the oral production domain, $F(1, 216.70) = 9.20$, $p = 0.003$, $\eta_p^2 = 0.040$, and auditory and reading comprehension domain, $F(1, 216.70) = 7.23$, $p = 0.008$, $\eta_p^2 = 0.032$, than did the HC group. The two groups demonstrated similar performance in reading aloud, repetition, and writing domains (all $p > 0.01$; Fig. 1). Moreover, within the aMCI group, language functions significantly differed among the five domains, $F(4, 167.70) = 5.21$, $p < 0.001$, $\eta_p^2 = 0.111$; however, no significant difference among the domains was observed within the HC group ($p > 0.025$). *Post hoc* pairwise analysis among the five language domains within the aMCI group revealed significant differences between the following pairwise domains: lower
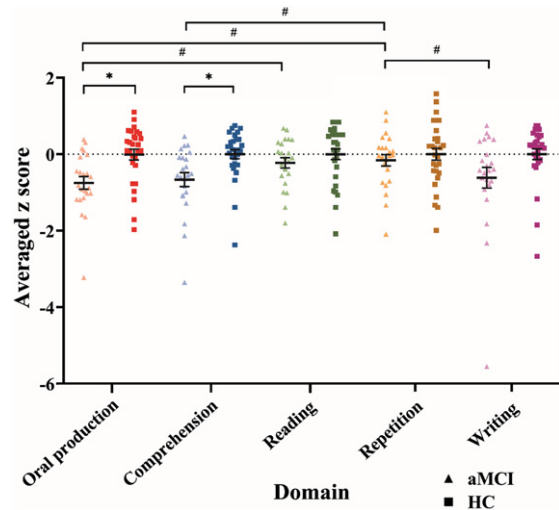

Fig. 1. Language performance of the two groups in the five language domains. Error bars denote the standard error. *significant group difference within each domain at $p < 0.01$. #significant pairwise comparison between the domains within each group at $p < 0.005$. aMCI, amnestic mild cognitive impairment; HC, healthy control.

scores were noted in the oral production domain than in the reading aloud domain, $t(167.70) = -3.37$, $p < 0.001$, $d = -0.728$, lower scores were observed in the oral production domain than in the repetition domain, $t(167.70) = -3.82$, $p < 0.001$, $d = -0.784$, lower scores were noted in the auditory and reading comprehension domain than in the repetition domain, $t(167.70) = -3.27$, $p = 0.001$, $d = -0.632$, and higher scores were observed in the repetition domain than in the writing domain, $t(167.70) = 2.95$, $p = 0.004$, $d = 0.433$. No significant difference was observed in the other six pairwise comparisons (all $p > 0.005$; Fig. 1).

Group differences at the test level

We analyzed group differences among the various tests for each language domain (Table 3). In the oral production domain, the aMCI group performed poorer than did the HC group on the noun and verb naming test, $t(50) = -3.19$, $p = 0.002$, $d = -0.890$, and verbal fluency test, $t(50) = -3.31$, $p = 0.002$, $d = -0.923$. By contrast, the two groups exhibited comparable performance on the Zhuyin blending and deletion test and sentence production and spontaneous speech test (both $p > 0.0125$).

In the auditory and reading comprehension domain, the aMCI group performed poorer than did the HC group on the attribute verification

Table 3
Domain composite scores and z-scores of tests in the Mandarin Chinese language test battery

	aMCI ($n = 23$) $M \pm SD$	HC ($n = 29$) $M \pm SD$	Statistics	p
Oral production[a]	−0.75 ± 0.79	−0.01 ± 0.75	$F(1, 216.70) = 9.20$	**0.003**[*]
Zhuyin Blending and Deletion test	−0.28 ± 1.28	0.0 ± 0.86	$t(49) = -0.93$	0.358
Sentence Production and Spontaneous Speech test	−0.82 ± 1.63	0.0 ± 0.88	$t(50) = -2.31$	0.025
Noun and Verb Naming test	−0.77 ± 0.90	0.0 ± 0.84	$t(50) = -3.19$	**0.002**[#]
Verbal Fluency test	−0.82 ± 0.73	0.0 ± 1.00	$t(50) = -3.31$	**0.002**[#]
Auditory and reading comprehension[a]	−0.66 ± 0.88	−0.01 ± 0.66	$F(1, 216.70) = 7.23$	**0.008**[*]
Phoneme and Tone Recognition test	−0.22 ± 1.01	0.0 ± 0.84	$t(50) = -0.87$	0.387
Word-picture Matching test	−0.49 ± 1.06	0.0 ± 0.84	$t(41.40) = -1.80$	0.079
Attribute-verification test	−1.56 ± 2.03	0.0 ± 0.87	$t(28.41) = -3.45$	**0.002**[†]
Relatedness Judgement test for Concrete and Abstract Words	−0.45 ± 1.46	0.0 ± 0.94	$t(50) = -1.34$	0.187
Sentence and Syntactic Comprehension test	−0.76 ± 1.05	0.0 ± 0.88	$t(50) = -2.84$	0.007
Idiom Comprehension test	−0.17 ± 1.26	0.0 ± 1.00	$t(50) = -0.56$	0.581
Written Lexical Decision test	−0.50 ± 1.39	0.0 ± 1.00	$t(50) = -1.49$	0.142
Reading aloud[a]	−0.23 ± 0.64	0.04 ± 0.74	$F(1, 216.70) = 1.15$	0.285
Word Reading test	−0.11 ± 0.66	0.0 ± 0.71	$t(50) = -0.57$	0.571
Pseudoword Reading test	−0.40 ± 1.48	0.0 ± 1.00	$t(50) = -1.15$	0.254
Sentence Reading test	−0.13 ± 1.07	0.0 ± 1.00	$t(50) = -0.44$	0.661
Repetition[a]	−0.16 ± 0.72	0.02 ± 0.85	$F(1, 216.70) = 0.54$	0.462
Word and Nonword Repetition test	0.03 ± 0.48	0.0 ± 0.77	$t(50) = 0.16$	0.877
Sentence Repetition test	−0.35 ± 1.00	0.0 ± 1.00	$t(50) = -1.26$	0.212
Writing[a]	−0.61 ± 1.31	−0.03 ± 0.75	$F(1, 216.70) = 5.80$	0.017
Written Naming test	−0.55 ± 1.24	0.0 ± 1.00	$t(50) = -1.79$	0.080
Written to Dictation test	−1.20 ± 2.06	0.0 ± 1.00	$t(30.20) = -2.57$	0.015
Written Sentences test	−0.59 ± 2.38	0.0 ± 1.00	$t(50) = -1.21$	0.231
Copy of the Written Word test	−0.11 ± 0.63	0.0 ± 1.00	$t(50) = -0.47$	0.642

aMCI, amnestic mild cognitive impairment; HC, healthy control. [a]The sample size of the HC group was 28 because one participant with missing data was excluded. [*]$p < 0.01$, based on Bonferroni correction for five domain composite scores. [#]$p < 0.0125$, based on Bonferroni correction for four test scores in the oral production domain. [†]$p < 0.0071$, based on Bonferroni correction for seven test scores in the auditory and reading comprehension domain.

test, $t(28.41) = -3.45$, $p = 0.002$, $d = -1.043$, and sentence and syntactic comprehension test, $t(50) = -2.84$, $p = 0.007$, $d = -0.793$. The two groups performed similarly on the phoneme and tone recognition test, word–picture matching test, relatedness judgement test for concrete and abstract words, idiom comprehension test, and written lexical decision test (all $p > 0.0071$). The performance of the two groups did not differ across the tests for the domains of reading aloud (all $p > 0.0167$), repetition (all $p > 0.025$), or writing (all $p > 0.0125$).

Group comparisons of naming error types and item analyses of passing rates

For the naming errors on the noun and verb naming test, the aMCI group made significantly more semantic errors compared with the HC group, $t(50) = -3.73$, $p < 0.001$, $d = -1.041$; Fig. 2A. A higher proportion of the participants in the aMCI group (61%) than in the HC group (17%) made semantic errors greater than or equal to 1 SD (i.e., 1.68 errors) on average, $\chi^2(1, N = 52) = 10.53$, $p = 0.001$; Fig. 2B. Phonological

errors were not analyzed because no participant in either group made a phonological error. Furthermore, in comparison to the HC group, the aMCI group exhibited a significantly greater occurrence of semantic coordinate errors, $\chi^2(1, N = 52) = 8.46$, $p = 0.004$. Both groups demonstrated similar rates of superordinate errors, $\chi^2(1, N = 52) = 0.55$, $p = 0.686$, subordinate errors, $\chi^2(1, N = 52) = 1.29$, $p = 0.442$, and associate errors, $\chi^2(1, N = 52) = 2.62$, $p = 0.191$.

In the item analysis, the passing rate cutoff was set at 90% for items on which the aMCI group performed poorly. For nouns in the noun and verb naming test, four items had low passing rates: 75% (three out of four items) of vegetable and fruit items (i.e., passing rates for the items "pepper," "onion," and "apple" were 30.4%, 78.3%, and 87.0%, respectively) and 33% (one out of three items) of tool items (i.e., the passing rate for the item "hammer" was 47.8%). No passing rate was below the cutoff for items belonging to the categories of animals and clothing. For verbs, 11 items had low passing rates: 67% (four out of six items) of intransitive verbs (i.e., passing rates for the items "pray," "cry," "jump," and "crawl" were 82.6%,
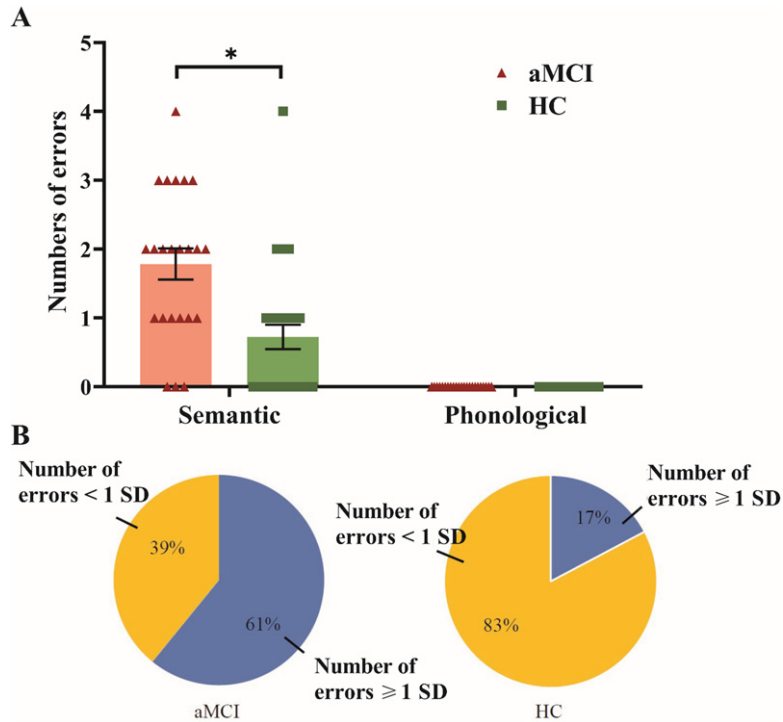
Fig. 2. Semantic and phonological naming errors between groups. Error bars denote the standard error. Panel A shows the total number of semantic and phonological errors by groups. *significant group difference in each type of naming errors at $p < 0.025$. Panel B displays the proportions of participants in each group who made semantic errors ≥ 1 or < 1 *SD* than the mean number of errors in the HC group. aMCI, amnestic mild cognitive impairment; HC, healthy control.

73.9%, 56.5%, and 8.7%, respectively) and 70% (7 out of 10 items) of transitive verbs (i.e., passing rates for the items "zip," "spilt," "throw," "write," "stir," "pull," and "climb" were 69.6%, 30.4%, 34.8%, 87.0%, 39.1%, 82.6%, and 0%, respectively).

Five items in the attribute verification test had low passing rates in the aMCI group on the basis of the average percentage of auditory and written forms. These items included 100% (all three items) of vegetable items (i.e., passing rates for the items "corn," "bitter gourd," and "sweet potato" were 78.3%, 89.2%, and 84.8%, respectively), 50% (one out of two items) of fruit items (i.e., passing rates for the item "watermelon" was 89.2%), and 50% (one out of two items) of animal items (i.e., the passing rate for the item "sparrow" was 82.7%). None of the items belonging to the tool category had a passing rate below the cutoff. Moreover, item analysis of the attribute verification test in terms of statement veracity revealed that four items had a low passing rate: 30% (3 out of 10 items) of false statements and 10% (1 out of 10 items) of factual statements.

Correlations between language domains and executive function scores

The correlations results between executive function scores and of language domains are illustrated in Fig. 3. The composite executive function scores exhibited significant positive correlations with the following language domains: oral production ($r = 0.58$, $p < 0.001$), auditory and reading comprehension ($r = 0.59$, $p < 0.001$), and writing ($r = 0.38$, $p = 0.006$). Furthermore, the composite executive function scores demonstrated a marginal positive correlation with the reading aloud domain ($r = 0.33$, $p = 0.019$) but were not found to be significantly correlated with the repetition domain ($r = 0.21$, $p = 0.145$).

Reliability analyses of the new language test battery

Reliability analyses were conducted on the new language test battery, and the results showed that the test-retest reliability ICCs were 0.917
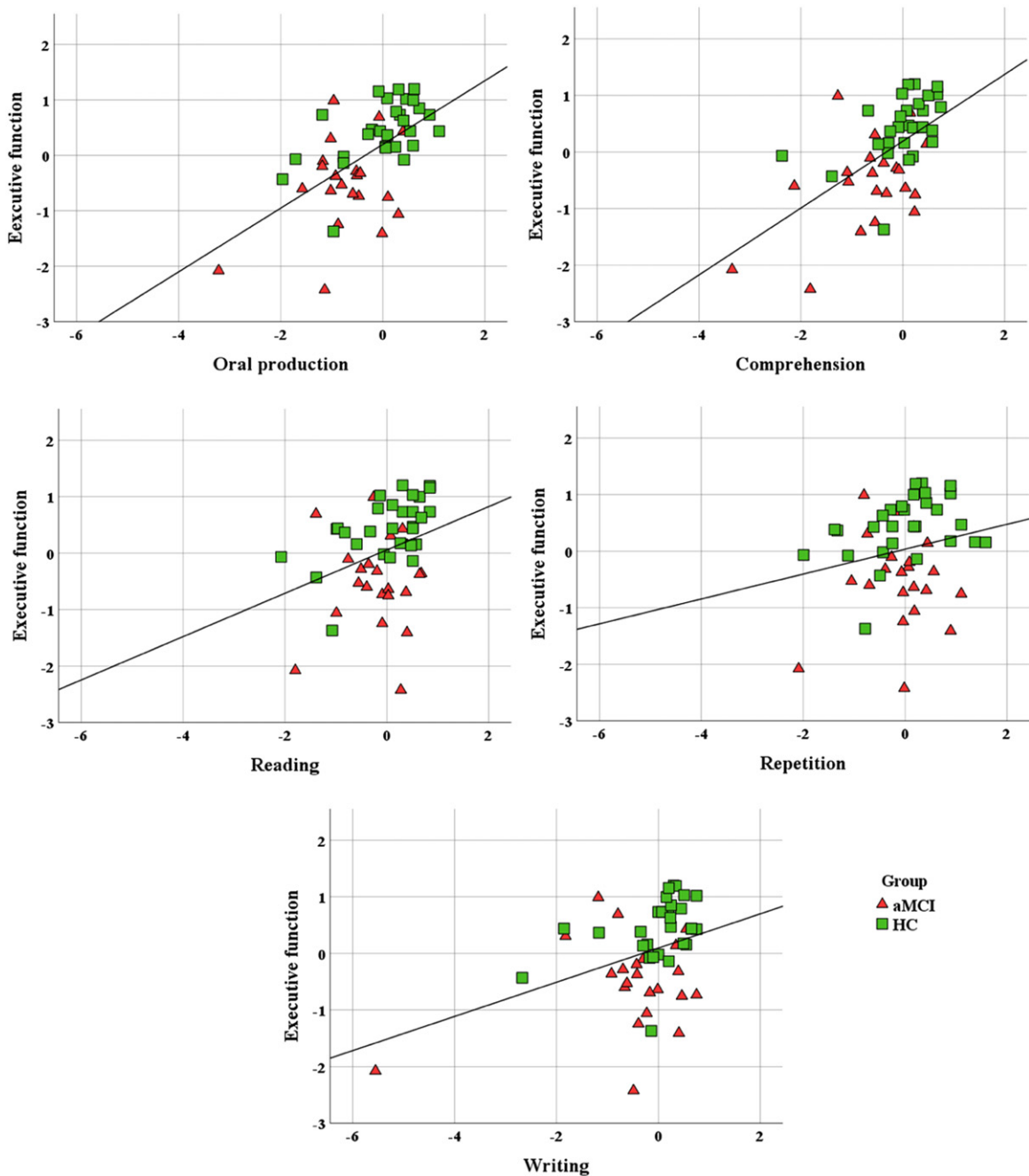
Fig. 3. Association between composite scores of executive function and of five language domains. aMCI, amnestic mild cognitive impairment; HC, healthy control.

(95% CI = 0.826–0.961) for the oral production domain, 0.894 (95% CI = 0.783–0.949) for the auditory and reading comprehension domain, 0.769 (95% CI = 0.557–0.886) for the reading domain, 0.793 (95% CI = 0.600–0.899) for the repetition domain, and 0.835 (95% CI = 0.673–0.920) for the writing domain. Additionally, inter-rater reliability ICCs were 0.778 (95% CI = –0.243–0.975) for the spontaneous speech section and 0.755 (95% CI = –0.042–0.985) for the idiom comprehension test.

Correlations between scores on the new language test battery and CCAT battery

Correlation analyses between scores on the new language test battery and those on the CCAT battery revealed significantly positive correlations between the following pairs of language test scores: the oral production composite scores and the combined scores of CCAT Simple Response, Expository Speech, and Naming subtests ($r = 0.64$, $p < 0.001$), the auditory and reading comprehension composite scores and the combined scores of CCAT Auditory and Reading Comprehension subtests ($r = 0.49$, $p < 0.001$), the reading aloud composite scores and CCAT Reading Comprehension subtest scores ($r = 0.39$, $p = 0.004$), the repetition composite scores and CCAT Repetition subtest scores ($r = 0.64$, $p < 0.001$), and the writing composite scores and the combined scores of CCAT Copying and Spontaneous Writing subtest ($r = 0.76$, $p < 0.001$).

DISCUSSION

The current study characterized a multifaceted language profile of the Mandarin Chinese speakers with aMCI and compared it with that of the HCs. Four main findings were obtained using a newly developed language test battery. First, compared with the HC group, the aMCI group exhibited poorer performance in the oral production and auditory and reading comprehension domains; however, the performance of both the groups was comparable in the reading aloud, repetition, and writing domains. Within the aMCI group, the participants' language function in the oral production domain was poorer than in the other domains. Second, the aMCI group exhibited poorer performance on tests requiring effortful lexical and semantic processing and made significantly more semantic errors compared with the HC group. Third, item analysis findings demonstrated that more items concerning living objects had low passing rates compared with those related to nonliving items in the aMCI group. Fourth, the performance of executive function exhibited positive correlations with nearly all language domains, with the exception of the repetition domain.

In line with our hypothesis, compared with the HCs, the older adults with aMCI exhibited poorer performance in the oral production and auditory and reading comprehension domains but comparable performance in the reading aloud, repetition, and writing domains. Furthermore, the individuals with aMCI

experienced more difficulty in the verbal fluency test, noun and verb naming test, and attribute verification test compared with the other tests. This pattern of impairment suggests a common underlying difficulty in effortful lexical and semantic processing in individuals with aMCI [18, 21]. This finding is consistent with those of studies either employing a comprehensive language test battery [32, 33] or focusing on a single aspect of language function (e.g., [17, 19, 39, 70, 71]).

In particular, in the oral production domain, both the verbal fluency test and noun and verb naming test were involved in the semantic processing (e.g., searching for the semantic associations of a superordinate target or a picture with attributes) and lexical processing of retrieving specific words from semantic networks [13, 19]. We observed that the individuals with aMCI made many sematic naming errors but no phonological errors on the noun and verb naming test; this finding supports the notion that the noun naming problem of individuals with aMCI may have a combined lexical and semantic origin [18, 20, 72, 73]. Moreover, we determined that compared with the HC group, the aMCI group made more semantic coordinate errors, which indicate responses belonging to the same category as the target word (e.g., naming dogs as cats). This finding is in line with evidence that among semantic errors, individuals with aMCI make more semantic coordinate errors compared with other semantic error subtypes, such as superordinate errors (i.e., replacing the target word with a general category), suggesting specific difficulty in accessing detailed information on the lexical form of an object [13, 20]. Similar to the error pattern observed on the noun naming test, the individuals with aMCI made more semantic errors on the verb naming test, indicating a similar lexical and semantic origin of these errors; this finding accords that of a study revealing that patients with AD exhibited a similar semantic error pattern on noun and verb naming tests [74].

In the oral production domain, the aMCI and HC groups exhibited comparable language functions in the Zhuyin blending and deletion test and sentence production and spontaneous speech test. Our findings suggest that individuals with aMCI have a relatively preserved ability to process orally produced Zhuyin symbols, syntactically simple sentences, and connected speech; this finding is consistent with those of studies revealing the involvement of intact brain regions in speech production and articulation even during the early stage of dementia due to AD [75, 76]. Recent studies have demonstrated differences in

detectable linguistic changes between older adults with aMCI and their counterparts by performing detailed and automatic linguistic analysis of connected speech data (e.g., [47, 77]). Although future studies are warranted, the inconsistent findings indicate that novel behavioral analysis techniques may detect subtle changes that may not be visible in behavioral data analyzed at the coarse-grained level, which is typical of clinical behavioral evaluations.

In the auditory and reading comprehension domain, the individuals with aMCI performed poorly on the attribute verification test compared with the controls; such a finding has been reported in previous studies [39, 71]. The results may be biased because items in the four categories included in the attribute verification test differed in concreteness ratings, with the animal category having a lower concreteness rating than items in the other three categories. However, the group differences on this task remained after items in the animal category were excluded from the analysis, $t(35.54) = 3.47$, $p = 0.001$, $d = 1.017$. The finding suggests that individuals with aMCI had difficulty in accessing attributes from semantic networks. This finding is consistent with evidence indicating that individuals with aMCI exhibit a volumetric reduction in the anterior temporal lobes and inferior prefrontal cortex, which are key regions involved in semantic storing and appropriate utilization of semantic representations while performing a semantic task [71]. Moreover, a study reported that performance on this task could be affected by the veracity of statements among patients with AD because task demands may differ between false and true statements [78]. In particular, more sets of possible attributes need to be searched for false statements than for factual statements, causing difficulty for patients with AD to verify false statements [78]. Consistent with this view, we observed a veracity effect in the item analysis and determined that the tendency to have low passing rates was higher for false statements than factual statements for the individuals with aMCI.

The individuals with aMCI exhibited poorer performance, with marginal significance, compared with their counterparts on the sentence and syntactic comprehension test. Although some studies [79–81] have observed preserved syntactic comprehension performance in individuals with aMCI, other studies [38, 82] have indicated that when performing tasks requiring effortfully decoding syntactic structures and taxing cognitive resources, the aMCI group may demonstrate poorer performance than the control group. The aforementioned evidence is consistent

with the finding of the item analysis of the sentence and syntactic comprehension test in our aMCI sample. We observed that the individuals with aMCI failed to comprehend more syntactically complex sentences (i.e., a sentence with an embedded clause) compared with syntactically simple sentences within reversible sentence items in the test. This pattern may be attributable to their decreased set-switching and inhibition abilities required for integrating various syntactic structures. For example, the individuals with aMCI appeared to experience more difficulty with sentences having an embedded clause for the subject in a sentence with a subject complement (e.g., "在叉子下的 <1> 杯子 <2> 是藍色的 <3> The cup <2> which is under the fork <1> is blue <3>"). This sentence required the participants to switch to an unfamiliar structure and inhibit the false organization of phrases. Notably, the sentences that the individuals with aMCI had a problem with were not always the longest, suggesting that working memory capacity did not account for their difficulty with syntactically complex sentences. The involvement of inhibition and switching abilities was evident in our finding that compared with the HC group, the aMCI group had impairments on standardized neuropsychological tests that involved the use of switching and inhibition abilities, namely the CTT and CWIT.

The aMCI group exhibited comparable performance to the HC group on other tests involving a lower load of intentionally lexical and semantic processing in the auditory and reading comprehension domain, such as in the written lexical decision test, relatedness judgement test, and word–picture matching test. Our finding of similar performance on these tests aligns with the findings of prior studies. In these studies, individuals with MCI demonstrated comparable performance to controls in tasks involving the judgment of simple semantic associations for vocabularies or pictures, without demanding extensive cognitive resources [12, 32, 39]. Notably, despite both the attribute verification test and the relatedness judgment test evaluating semantic memory, the greater difficulty observed in individuals with aMCI in the former test compared to the latter, when compared with controls, may be attributed to several potential reasons. First, this pattern aligns with the "bottom-up" or "attribute-first" theory, suggesting that attributes are affected earlier in the course of AD compared to higher-level information such as coordinate or superordinate connections [83, 84]. In support of this notion, Rogers and colleagues found that patients with AD exhibited significant superordinate

and coordinate category member priming effects but showed no attribute priming effect [83]. Second, in contrast to the attribute verification test, the relatedness judgment test may place lower cognitive demands, as participants were tasked with evaluating semantic associations among three vocabularies without the need to process syntactic information at the sentence level. Additionally, the relatedness judgment test, comprising six Chinese characters in total for each trial, featured shorter mean character lengths compared to the attribute verification test ($M = 7.8$, range = 4–12 characters). Collectively, these factors may contribute to the observed disparate findings between the attribute verification test and the relatedness judgment test.

It is worth noting that we did not observe group difference in our word–picture matching test. Conversely, a study reported significant impairment in the multidomain aMCI group on the Peabody Picture Vocabulary Test, which is also a word–picture matching task [32]. However, it is crucial to consider that the aforementioned result may be confounded by the inclusion of individuals with aMCI at a more advanced stage of the disorder (mean MMSE score = 25.50) than our aMCI sample (mean MMSE score = 27.35). This is because our operational criterion for objective cognitive impairment was –1 instead of –1.5 SD from the norm. The inclusion of individuals with milder cognitive impairment in the memory and executive function domains may have resulted in our finding of comparable performance between the groups on the word–picture matching test. Moreover, our word–picture matching test has a relatively low task demand due to the inclusion of low-difficulty items (e.g., those with high levels of word frequency and familiarity). Collectively, consistent with the results of other studies [21, 39, 71], our findings suggested that individuals with aMCI exhibit deficits primarily in facets associated with effortful semantic processing.

The current study demonstrated comparable language functions between the older adults with aMCI and their counterparts in the reading aloud, repetition, and writing domains. This result is consistent with those of other studies [8, 32, 48, 85]. Our finding of comparable performance may be attributable to the lower difficulty of the test items, such as the use of a shorter word length and higher word frequency for repetition tasks, and lower task demand for assessing the writing ability at a simple sentence level. Other studies demonstrating significant impairment in the aMCI group [37, 38, 47] have included stimuli with

expanded word lengths for repetition tasks or used narrative stories as writing stimuli, which presumably required greater cognitive resources to perform the tasks than did those required in the tasks used in our study. Alternatively, we did not identify group differences in writing and reading aloud domains, possibly because we measured only behavioral accuracy at the coarse-grained level, similar to the clinical assessment context, unlike other studies that have used finer-grained measurements of written and reading latencies [34, 35] and analyzed data by using machine learning techniques [47]. Although semantic processing may be involved in the reading aloud, repetition, and writing domains, low task demands may enable older adults with aMCI to complete language tasks in an automatic manner by leveraging frequently learned connections between input and output lexicons in different modalities [32, 49].

Nevertheless, it is crucial to underscore that our writing tests incorporate a diverse range of stimuli intentionally designed to elicit various error types. We systematically manipulated character regularity, distinguishing between regular and irregular words. This deliberate design enables a comprehensive analysis, both quantitative or qualitative analysis, of group disparities in error types or the accuracy of regular versus irregular words within our assessment across diverse clinical populations. While we did not observe significant group differences in accuracy within the writing domain, qualitative analysis reveals a higher incidence of errors in the aMCI group compared to the control group. Particularly noteworthy are phonologically plausible substitution errors, instances of transcribing only Zhuyin, and leaving answers blank—a phenomenon indicative of challenges in orthographic memory for specific words [86]. These preliminary observations align with existing literature [48, 87], suggesting that individuals with aMCI are more prone to homophone errors, phonetically similar errors, and radicals misplacements. These observed tendencies resonate with prior research and may be attributed to the heightened reliance on non-semantic pathways by individuals with aMCI, stemming from the progressive deterioration of their lexical-semantic pathways [88].

The present study demonstrated differences between domains within the aMCI group, with relatively poor language function in the oral production domain. This result suggests a primary deficit of lexical and semantic processing in individuals with aMCI; this finding is supported by the results of passing rate analyses. Overall, the aMCI group failed in more

items in the categories of living objects (i.e., vegetables, fruits, and animals) than in the categories of nonliving objects (i.e., tools and clothing) in the noun and verb naming test as well as in the attribute verification test. Our finding of the individuals with aMCI experiencing relative difficulty processing living items is consistent with the findings of other studies [16, 39], suggesting that individuals with aMCI have difficulty processing items in semantic categories that contain more shared and fewer distinctive attributes, such as vegetables and animals [89]. A study has revealed a significant association between a semantic deficit in the MCI group and a volumetric reduction in the perirhinal areas [25]. These perirhinal areas have been identified as central regions that play a crucial role in integrating semantic representations [24]. Notably, in our study of individuals with MCI, we did not observe a syntactic deficit in the argument structures of the verb naming test. This was evident from their comparable performance between transitive and intransitive verbs in the noun and verb naming tests. These results align with a previous finding concerning patients with AD [90].

In addition to core language ability, our result suggest that executive function plays a role in various language tasks used in this study, consistent with previous studies [37–39, 47]. Notably, the verbal fluency test heavily relies on executive functions, such as self-initiation, inhibiting previously produced responses, and efficiently organizing verbal retrieval [11, 91, 92]. Moreover, the fluency test used in our study includes a switching condition that is presumed to heighten the demand for executive function, particularly in the context of switching ability [92, 93]. Studies have linked impairments in switching ability during the semantic verbal fluency task to the volumes of the superior frontal gyrus and inferior frontal gyrus in individuals with aMCI [94]. Therefore, it is plausible that both executive function and language ability contribute to impairments observed in the verbal fluency test among older adults with aMCI. However, we did not find any significant correlation between the repetition domain and the composite executive function score. A qualitative examination of this finding suggests that it may be attributed to the constrained variation in performance, possibly owing to a floor effect observed in the word and nonword repetition test. Participants, regardless of their MCI status, might have encountered challenges while repeating items within the nonword repetition section. This difficulty can be attributed to the relatively unfamiliar nature of two-to-six-syllable nonword items

derived from low-frequency characters, which Mandarin Chinese-speaking older adults find challenging to recognize and repeat.

The psychometric properties of the Mandarin Chinese language test battery employed in this study were preliminarily established. Notably, all indicators of test-retest reliability and inter-rater reliability exhibited high levels of reliability, ranging from good to excellent [68]. Moreover, our Mandarin Chinese language test battery demonstrated correlations with diverse subtest scores within the CCAT battery, offering considerable validation for the efficacy of our language assessments. While there may be room for discussion regarding the use of CCAT Reading Comprehension subtest scores as the index for correlating with our reading aloud composite scores, it emerges as the most suitable CCAT subtest for comparison. This choice is based on their shared involvement in the intricate process of reading Chinese characters. To mitigate potential concerns, a thoroughgoing validity analysis of the reading aloud composite scores is deemed necessary in future studies. In summary, the robust correlations observed across multiple tests between our test battery and the CCAT collectively affirm the validity of our assessment in evaluating the language function of the Mandarin Chinese speakers.

This study has some limitations that should be addressed. First, the sample size was relatively small. Despite the sample size, our findings provide preliminary evidence that Mandarin Chinese speakers with aMCI exhibit significantly poor performance on tests requiring effortful lexical and semantic processing. The effect sizes of Cohen's d derived from this study were comparable to those based on English-speaking participants [33]. Second, we included both individuals with single-domain and multidomain subtypes of aMCI. Because some language tests may interact with other cognitive functions, such as executive function (e.g., the verbal fluency test) or visuospatial ability (e.g., writing tests), the small sample size of our aMCI group precluded us from further examining this problem. Third, with the goal of applying the language test battery to clinical settings, we chose to use test stimuli with relatively low task demands (i.e., relatively high word frequency, familiarity, and concreteness rating scores); this might have resulted in between-group null findings in reading aloud, repetition, and writing domains. Nevertheless, we still identified a pattern of prominent deficits in lexical and semantic processing in the individuals with aMCI by performing a multifaceted language test battery. Fourth, our study was cross-sectional in design. Thus,

the predictive power of certain language impairments for identifying individuals with aMCI who may eventually develop dementia is yet to determined. Future studies should investigate the longitudinal trajectory of various language impairments in individuals with aMCI to evaluate the clinical utility of our language tests.

Fifth, although our primary focus did not center on investigating the relationship between education and language performance, previous studies have shed light on a notable association between language test performance and the level of education (e.g., [95, 96]). Moreover, high educational achievement has been linked to the postponement of cognitive decline associated with neurodegeneration (e.g., [97, 98]). In the present study, we identified a positive correlation between education levels and the five domains of language function: oral production ($r = 0.577$, $p < 0.001$), auditory and reading comprehension ($r = 0.533$, $p < 0.001$), reading aloud ($r = 0.452$, $p < 0.001$), repetition ($r = 0.421$, $p = 0.002$), and writing ($r = 0.554$, $p < 0.001$). These findings align with previous research, which has outlined varying degrees of influence of education level on different language domains [99]. Notably, in specific test like the Zhuyin blending and deletion test, we discover a positive association with the duration of education ($r = 0.303$, $p = 0.031$), suggesting that phonological awareness, particularly in complex tasks such as blending, may undergo continued development with increasing years of education [100, 101]. The development and contributing factors to phonological awareness in the Chinese language have rarely been systematically studied, particularly beyond populations of preschoolers or elementary school children [102]. Thus, our findings hold potential implications for future studies emphasizing issues such as the mutuality and assessment of phonological awareness. Future investigations should aim to increase sample sizes, encompassing diverse educational levels, to comprehensively explore the effects of demographic variables, such as education level, on language performance in clinical populations with cognitive impairment.

Overall, the language function profile of the Mandarin Chinese speakers with aMCI exhibited a behavioral pattern similar to that determined for English-speaking participants by using a multifaceted language test battery [32, 33]. This study extended the results to demonstrate a primary deficit of effortful lexical and semantic processing in Mandarin-speaking individuals with aMCI.

Clinicians can incorporate language tests examining functions in oral production and comprehension domains to evaluate individuals suspected of having MCI. Compared to the CCAT, our Mandarin Chinese language test battery has several strengths. It covers a more comprehensive range of language aspects and utilizes a simpler scoring system, making it a potentially valuable clinical tool for evaluating the multifaceted language abilities of Mandarin Chinese speakers. Future studies should validate the construct of the five language domains by conducting factor analyses to confirm the latent structure of the language test battery. Other psychometric properties, including sensitivity, specificity, and the cutoff for each test, of the Mandarin Chinese language test battery should also be explored with a larger sample size. The comprehensive language test battery developed in the present study can enhance our understanding of how and to what extent language function may change over time in different domains, such as reading, repetition, and writing, during disease progression from MCI to dementia in Mandarin speakers [35, 48, 103]. This battery also provides opportunities to investigate multifaceted language functions in different clinical populations, such as patients with primary progressive aphasia [104, 105] or amyotrophic lateral sclerosis [106], and further elucidate the complex brain–behavioral relationship.

ACKNOWLEDGMENTS

FUNDING

CONFLICT OF INTEREST

Yu-Ling Chang is an Editorial Board Member of this journal but was not involved in the peer-review process nor had access to any information regarding

its peer-review. Other authors have no conflicts of interests to report.

DATA AVAILABILITY

The data used in this study is regulated due to privacy and ethical restrictions. The data are available upon request from the corresponding author.

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: https://dx.doi.org/10.3233/JAD-230871.

REFERENCES

[1] Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L, Winblad B (2001) Current concepts in mild cognitive impairment. *Arch Neurol* **58**, 1985-1992.

[2] Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund LO, Nordberg A, Bäckman L, Albert M, Almkvist O (2004) Mild cognitive impairment–beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment. *J Intern Med* **256**, 240-246.

[3] Petersen RC (2016) Mild cognitive impairment. *Continuum (Minneap Minn)* **22**, 404.

[4] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC (2011) The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 270-279.

[5] Bozoki A, Giordani B, Heidebrink JL, Berent S, Foster NL (2001) Mild cognitive impairments predict dementia in nondemented elderly patients with memory loss. *Arch Neurol* **58**, 411-416.

[6] Han JW, Kim TH, Lee SB, Park JH, Lee JJ, Huh Y, Park JE, Jhoo JH, Lee DY, Kim KW (2012) Predictive validity and diagnostic stability of mild cognitive impairment subtypes. *Alzheimers Dement* **8**, 553-559.

[7] Piccininni C, Quaranta D, Gainotti G, Lacidogna G, Guglielmi V, Giovannini S, Marra C (2021) The destiny of multiple domain amnesic mild cognitive impairment: Effect of alternative neuropsychological definitions and their adjunctive role in respect of memory impairment. *Arch Clin Neuropsychol* **36**, 702-710.

[8] Tabert MH, Manly JJ, Liu X, Pelton GH, Rosenblum S, Jacobs M, Zamora D, Goodkind M, Bell K, Stern Y (2006) Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Arch Gen Psychiatry* **63**, 916-924.

[9] Belleville S, Fouquet C, Hudon C, Zomahoun HTV, Croteau J (2017) Neuropsychological measures that predict progression from mild cognitive impairment to Alzheimer's type dementia in older adults: A systematic review and meta-analysis. *Neuropsychol Rev* **27**, 328-353.

[10] Gallucci M, Di Battista ME, Battistella G, Falcone C, Bisiacchi PS, Di Giorgi E (2018) Neuropsychological tools to predict conversion from amnestic mild cognitive impairment to dementia. The TREDEM Registry. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn* **25**, 550-560.

[11] Henry JD, Crawford JR, Phillips LH (2004) Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia* **42**, 1212-1222.

[12] Adlam A-LR, Bozeat S, Arnold R, Watson P, Hodges JR (2006) Semantic knowledge in mild cognitive impairment and mild Alzheimer's disease. *Cortex* **42**, 675-684.

[13] Balthazar MLF, Cendes F, Damasceno BP (2008) Semantic error patterns on the Boston Naming Test in normal aging, amnestic mild cognitive impairment, and mild Alzheimer's disease: Is there semantic disruption? *Neuropsychology* **22**, 703.

[14] Forbes-McKay K, Shanks MF, Venneri A (2013) Profiling spontaneous speech decline in Alzheimer's disease: A longitudinal study. *Acta Neuropsychiatr* **25**, 320-327.

[15] Mueller KD, Hermann B, Mecollari J, Turkstra LS (2018) Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. *J Clin Exp Neuropsychol* **40**, 917-939.

[16] Krumm S, Berres M, Kivisaari SL, Monsch AU, Reinhardt J, Blatow M, Kressig RW, Taylor KI (2021) Cats and apples: Semantic fluency performance for living things identifies patients with very early Alzheimer's disease. *Arch Clin Neuropsychol* **36**, 838-843.

[17] Marra C, Piccininni C, Masone Iacobucci G, Caprara A, Gainotti G, Costantini EM, Callea A, Venneri A, Quaranta D (2021) Semantic memory as an early cognitive marker of Alzheimer's disease: Role of category and phonological verbal fluency tasks. *J Alzheimers Dis* **81**, 619-627.

[18] Taler V, Phillips NA (2008) Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *J Clin Exp Neuropsychol* **30**, 501-556.

[19] Chang Y-L, Chen T-F, Tseng W-Y (2020) White matter pathways underlying Chinese semantic and phonological fluency in mild cognitive impairment. *Neuropsychologia* **149**, 107671.

[20] Gallant M, Lavoie M, Hudon C, Monetta L (2019) Analysis of naming errors in healthy aging, mild cognitive impairment, and Alzheimer's disease. *Can J Speech Lang Pathol Audiol* **43**, 95–108.

[21] Venneri A, Jahn-Carta C, De Marco M, Quaranta D, Marra C (2018) Diagnostic and prognostic role of semantic processing in preclinical Alzheimer's disease. *Biomark Med* **12**, 637-651.

[22] Pineault J, Jolicoeur P, Grimault S, Bermudez P, Brambati SM, Lacombe J, Villalpando JM, Kergoat M-J, Joubert S (2018) Functional changes in the cortical semantic network in amnestic mild cognitive impairment. *Neuropsychology* **32**, 417.

[23] Chang YL, Chen TF, Tseng WY (2020) White matter pathways underlying Chinese semantic and phonological fluency in mild cognitive impairment. *Neuropsychologia* **149**, 107671.

[24] Taylor KI, Moss HE, Stamatakis EA, Tyler LK (2006) Binding crossmodal object features in perirhinal cortex. *Proc Natl Acad Sci U S A* **103**, 8239-8244.

[25] Venneri A, Mitolo M, Beltrachini L, Varma S, Della Pieta C, Jahn-Carta C, Frangi AF, De Marco M (2019) Beyond episodic memory: Semantic processing as independent predictor of hippocampal/perirhinal volume in aging and

mild cognitive impairment due to Alzheimer's disease. *Neuropsychology* **33**, 523-533.

[26] Chasles M-J, Tremblay A, Escudier F, Lajeunesse A, Benoit S, Langlois R, Joubert S, Rouleau I (2020) An examination of semantic impairment in amnestic MCI and AD: What can we learn from verbal fluency? *Arch Clin Neuropsychol* **35**, 22-30.

[27] Lonie JA, Herrmann LL, Tierney KM, Donaghey C, O'Carroll R, Lee A, Ebmeier KP (2009) Lexical and semantic fluency discrepancy scores in aMCI and early Alzheimer's disease. *J Neuropsychol* **3**, 79-92.

[28] Murphy KJ, Rich JB, Troyer AK (2006) Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia. *J Int Neuropsychol Soc* **12**, 570-574.

[29] Teng E, Leone-Friedman J, Lee GJ, Woo S, Apostolova LG, Harrell S, Ringman JM, Lu PH (2013) Similar verbal fluency patterns in amnestic mild cognitive impairment and Alzheimer's disease. *Arch Clin Neuropsychol* **28**, 400-410.

[30] Brandt J, Manning KJ (2009) Patterns of word-list generation in mild cognitive impairment and Alzheimer's disease. *Clin Neuropsychol* **23**, 870-879.

[31] Rinehardt E, Eichstaedt K, Schinka JA, Loewenstein DA, Mattingly M, Fils J, Duara R, Schoenberg MR (2014) Verbal fluency patterns in mild cognitive impairment and Alzheimer's disease. *Dement Geriatr Cogn Disord* **38**, 1-9.

[32] Jokel R, Lima BS, Fernandez A, Murphy KJ (2019) Language in amnestic mild cognitive impairment and dementia of Alzheimer's type: Quantitatively or qualitatively different? *Dement Geriatr Cogn Dis Extra* **9**, 136-151.

[33] Tsantali E, Economidis D, Tsolaki M (2013) Could language deficits really differentiate mild cognitive impairment (MCI) from mild Alzheimer's disease? *Arch Gerontol Geriatr* **57**, 263-270.

[34] Afonso O, Álvarez CJ, Martínez C, Cuetos F (2019) Writing difficulties in Alzheimer's disease and mild cognitive impairment. *Read Writ* **32**, 217-233.

[35] Arsenault-Lapierre G, Bergman H, Chertkow H (2012) Word reading threshold and mild cognitive impairment: A validation study. *BMC Geriatr* **12**, 38.

[36] Gardini S, Venneri A, Sambataro F, Cuetos F, Fasano F, Marchi M, Crisi G, Caffarra P (2015) Increased functional connectivity in the default mode network in mild cognitive impairment: A maladaptive compensatory mechanism associated with poor semantic memory performance. *J Alzheimers Dis* **45**, 457-470.

[37] Hayashi A, Nomura H, Mochizuki R, Ohnuma A, Kimpara T, Suzuki K, Mori E (2015) Writing impairments in Japanese patients with mild cognitive impairment and with mild Alzheimer's disease. *Dement Geriatr Cogn Dis Extra* **5**, 309-319.

[38] Sherman JC, Henderson Jr CR, Flynn S, Gair JW, Lust B (2021) Language decline characterizes amnestic mild cognitive impairment independent of cognitive decline. *J Speech Lang Hear Res* **64**, 4287-4307.

[39] Duong A, Whitehead V, Hanratty K, Chertkow H (2006) The nature of lexico-semantic processing deficits in mild cognitive impairment. *Neuropsychologia* **44**, 1928-1935.

[40] Weekes BSH (2020) Aphasia in Alzheimer's disease and other dementias (ADOD): Evidence from Chinese. *Am J Alzheimers Dis Other Demen* **35**, 1533317520949708.

[41] Chen T-M, Chen J-Y (2013) The syllable as the proximate unit in Mandarin Chinese word production: An intrinsic

or accidental property of the production system? *Psychon Bull Rev* **20**, 154-162.

[42] Hoosain R (2013) *Psycholinguistic implications for linguistic relativity: A case study of Chinese*, Psychology Press, New York.

[43] Peng G, Zheng H-Y, Gong T, Yang R-X, Kong J-P, Wang WS-Y (2010) The influence of language experience on categorical perception of pitch contours. *J Phon* **38**, 616-624.

[44] Wu C-Y, Ho M-HR, Chen S-HA (2012) A meta-analysis of fMRI studies on Chinese orthographic, phonological, and semantic processing. *Neuroimage* **63**, 381-391.

[45] Wang T, Lian C, Pan J, Yan Q, Zhu F, Ng ML, Wang L, Yan N (2019) Towards the speech features of mild cognitive impairment: Universal evidence from structured and unstructured connected speech of Chinese. *Interspeech*, International Speech Communication Association, Graz, Austria, pp. 3880-3884.

[46] Feng Y, Meng Y, Peng G (2019) The categorical perception of Mandarin tones in normal aging seniors and seniors with mild cognitive impairment. *19th International Congress of Phonetic Sciences,* Australasian Speech Science and Technology Association, Melbourne, Australia, pp. 909-913.

[47] Wang T, Hong Y, Wang Q, Su R, Ng ML, Xu J, Wang L, Yan N (2021) Identification of mild cognitive impairment among Chinese based on multiple spoken tasks. *J Alzheimers Dis* **82**, 185-204.

[48] Zhou J, Jiang B, Huang X-H, Kong L-L, Li H-L (2016) Characteristics of agraphia in Chinese patients with Alzheimer's disease and amnestic mild cognitive impairment. *Chin Med J (Engl)* **129**, 1553-1557.

[49] Yin WG, Weekes BS (2003) Dyslexia in Chinese: Clues from cognitive neuropsychology. *Ann Dyslexia* **53**, 255-279.

[50] Chung Y-M, Lee S-E, Chang M-H, Hsu T-C (1998) The Concise Chinese Aphasia Test (CCAT) and its applications. *J Speech Lang Hearing Association Taiwan* **13**, 119-137.

[51] Swinburn K, Porter G, Howard D (2004) *Comprehensive aphasia test*, Psychology Press, Hove.

[52] Jak AJ, Bondi MW, Delano-Wood L, Wierenga C, Corey-Bloom J, Salmon DP, Delis DC (2009) Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am J Geriatr Psychiatry* **17**, 368-375.

[53] Petersen RC (2004) Mild cognitive impairment as a diagnostic entity. *J Intern Med* **256**, 183-194.

[54] Morris JC (1997) Clinical dementia rating: A reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *Int Psychogeriatr* **9**, 173-176.

[55] O'Bryant SE, Lacritz LH, Hall J, Waring SC, Chan W, Khodr ZG, Massman PJ, Hobson V, Cullum CM (2010) Validation of the new interpretive guidelines for the clinical dementia rating scale sum of boxes score in the national Alzheimer's coordinating center database. *Arch Neurol* **67**, 746-749.

[56] Guo N-W, Liu H-C, Wong P-F, Liao K-K, Yan S-H, Lin K-P, Chang C-Y, Hsu T-C (1988) Chinese version and norms of the Mini-Mental State Examination. *Rehabilit Pract Sci* **16**, 52-59.

[57] Chen H-Y, Chen Y-H (2002) *Wechsler Adult Intelligence Scale-Third Edition (WAIS-III) Manual for Taiwan*, The Chinese Behavioral Science Corporation, Taipei, Taiwan.

[58] Delis DC, Kaplan E, Kramer JH (2001) *Delis-Kaplan executive function system*, Pearson.

[59] Hua M-S, Chang B-S, Lin K-N, Yang C-M, Lu LH, Chen H-Y (2005) *Wechsler Memory Scale-(WMS-III) Manual for Taiwan*, The Chinese Behavioral Science Corporation.

[60] Meyers JE, Meyers KR (1995) Rey complex figure test under four different administration procedures. *Clin Neuropsychol* **9**, 63-67.

[61] Goodglass H, Kaplan E, Weintraub S (1983) *Boston naming test*, Lea & Febiger.

[62] Kuo H-Y, Hua M-S (2015) *Color Trails Test-Chinese version manual*, Chinese Behavioral Science Corporation, Taipei.

[63] Burke WJ, Roccaforte WH, Wengel SP (1991) The short form of the Geriatric Depression Scale: A comparison with the 30-item form. *J Geriatr Psychiatry Neurol* **4**, 173-178.

[64] Caplan D (1992) *Language: Structure, processing, and disorders*, The MIT Press.

[65] Marshall CR, Hardy CJD, Volkmer A, Russell LL, Bond RL, Fletcher PD, Clark CN, Mummery CJ, Schott JM, Rossor MN, Fox NC, Crutch SJ, Rohrer JD, Warren JD (2018) Primary progressive aphasia: A clinical approach. *J Neurol* **265**, 1474-1490.

[66] Snodgrass JG, Vanderwart M (1980) A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *J Exp Psychol Hum Learn* **6**, 174.

[67] Staffaroni AM, Weintraub S, Rascovsky K, Rankin KP, Taylor J, Fields JA, Casaletto KB, Hillis AE, Lukic S, Gorno-Tempini ML (2021) Uniform data set language measures for bvFTD and PPA diagnosis and monitoring. *Alzheimers Dement (Amst)* **13**, e12148.

[68] Koo TK, Li MY (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* **15**, 155-163.

[69] JASP Team (2022) JASP (Version 0.16.2) Computer software. https://jasp-stats.org

[70] Hwang YK, Kim E, Kim YB, Kim YW, Nam CM, Cho S-H, Kim H (2017) Diagnostic value of time-constrained naming test in mild cognitive impairment. *Dement Geriatr Cogn Disord* **44**, 171-181.

[71] Joubert S, Brambati SM, Ansado J, Barbeau EJ, Felician O, Didic M, Lacombe J, Goldstein R, Chayer C, Kergoat M-J (2010) The cognitive and neural expression of semantic memory impairment in mild cognitive impairment and early Alzheimer's disease. *Neuropsychologia* **48**, 978-988.

[72] Ahmed S, Arnold R, Thompson SA, Graham KS, Hodges JR (2008) Naming of objects, faces and buildings in mild cognitive impairment. *Cortex* **44**, 746-752.

[73] Willers IF, Feldman ML, Allegri RF (2008) Subclinical naming errors in mild cognitive impairment: A semantic deficit? *Dement Neuropsychol* **2**, 217-222.

[74] Almor A, Aronoff JM, MacDonald MC, Gonnerman LM, Kempler D, Hintiryan H, Hayes UL, Arunachalam S, Andersen ES (2009) A common mechanism in verb and noun naming deficits in Alzheimer's patients. *Brain Lang* **111**, 8-19.

[75] Braak H, Braak E (1991) Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol* **82**, 239-259.

[76] Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* **8**, 393-402.

[77] Tóth L, Hoffmann I, Gosztolya G, Vincze V, Szatlóczki G, Bánréti Z, Pákáski M, Kálmán J (2018) A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech. *Curr Alzheimer Res* **15**, 130-138.

[78] Peraita H, Díaz C, Anllo-Vento L (2008) Processing of semantic relations in normal aging and Alzheimer's disease. *Arch Clin Neuropsychol* **23**, 33-46.

[79] Hodges JR, Erzinçlioğlu S, Patterson K (2006) Evolution of cognitive deficits and conversion to dementia in patients with mild cognitive impairment: A very-long-term follow-up study. *Dement Geriatr Cogn Disord* **21**, 380-391.

[80] Lambon Ralph MA, Patterson K, Graham N, Dawson K, Hodges JR (2003) Homogeneity and heterogeneity in mild cognitive impairment and Alzheimer's disease: A cross-sectional and longitudinal study of 55 cases. *Brain* **126**, 2350-2362.

[81] Pereira DN, da Cruz Souza W, Belan AFR, de Arruda Camargo MvZ, Forlenza OV, Radanovic M (2022) Sentence processing in mild cognitive impairment. *J Neurolinguistics* **63**, 101070.

[82] Sung JE, Choi S, Eom B, Yoo JK, Jeong JH (2020) Syntactic complexity as a linguistic marker to differentiate mild cognitive impairment from normal aging. *J Speech Lang Hear Res* **63**, 1416-1429.

[83] Rogers SL, Friedman RB (2008) The underlying mechanisms of semantic memory loss in Alzheimer's disease and semantic dementia. *Neuropsychologia* **46**, 12-21.

[84] Warrington EK (1975) The selective impairment of semantic memory. *Q J Exp Psychol* **27**, 635-657.

[85] Hughes JC, Graham N, Patterson K, Hodges JR (1997) Dysgraphia in mild dementia of Alzheimer's type. *Neuropsychologia* **35**, 533-545.

[86] De Tollis M, De Simone MS, Perri R, Fadda L, Caltagirone C, Carlesimo GA (2021) Verbal and spatial memory spans in mild cognitive impairment. *Acta Neurol Scand* **144**, 383-393.

[87] Evyapan Akkuş D, Güler A (2016) The Ege Agraphia Test Battery for identifying the writing disorders in cases with mild cognitive impairment and Alzheimer's Disease. *Turk Psikiyatri Dergisi* **27**, 185.

[88] Yin W, He S, Weekes BS (2005) Acquired dyslexia and dysgraphia in Chinese. *Behav Neurol* **16**, 159-167.

[89] Taylor KI, Moss HE, Tyler LK (2007) The conceptual structure account: A cognitive model of semantic memory and its neural instantiation. In *Neural basis of semantic memory*, Hart J, Kraut MA, eds. Cambridge University Press, pp. 265–301.

[90] Kim M, Thompson CK (2004) Verb deficits in Alzheimer's disease and agrammatism: Implications for lexical organization. *Brain Lang* **88**, 1-20.

[91] Oh SJ, Sung JE, Choi SJ, Jeong JH (2019) Clustering and switching patterns in semantic fluency and their relationship to working memory in mild cognitive impairment. *Dement Neurocogn Disord* **18**, 47-61.

[92] Weakley A, Schmitter-Edgecombe M, Anderson J (2013) Analysis of verbal fluency ability in amnestic and non-amnestic mild cognitive impairment. *Arch Clin Neuropsychol* **28**, 721-731.

[93] Troyer AK, Moscovitch M, Winocur G, Alexander MP, Stuss D (1998) Clustering and switching on verbal fluency: The effects of focal frontal-and temporal-lobe lesions. *Neuropsychologia* **36**, 499-504.

[94] Peter J, Kaiser J, Landerer V, Köstering L, Kaller CP, Heimbach B, Hüll M, Bormann T, Klöppel S (2016) Category and design fluency in mild cognitive impairment: Performance, strategy use, and neural correlates. *Neuropsychologia* **93**, 21-29.

[95] Folia V, Liampas I, Ntanasi E, Yannakoulia M, Sakka P, Hadjigeorgiou G, Scarmeas N, Dardiotis E, Kosmidis MH

(2022) Longitudinal trajectories and normative language standards in older adults with normal cognitive status. *Neuropsychology* **36**, 626-639.

[96] Zamarian L, Karner E, Bodner T, Djamshidian A, Delazer M (2021) Differential impact of education on cognitive performance in neurological patients with progressive cognitive decline. *J Alzheimers Dis* **80**, 1491-1501.

[97] Rosselli M, Uribe IV, Ahne E, Shihadeh L (2022) Culture, ethnicity, and level of education in Alzheimer's disease. *Neurotherapeutics* **19**, 26-54.

[98] Chen G, Zhao M, Yang K, Lin H, Han C, Wang X, Han Y (2021) Education exerts different effects on cognition in individuals with subjective cognitive decline and cognitive impairment: A population-based study. *J Alzheimers Dis* **79**, 653-661.

[99] Pineda DA, Rosselli M, Ardila A, Mejia SE, Romero MG, Perez C (2000) The Boston diagnostic aphasia examination–Spanish version: The influence of demographic variables. *J Int Neuropsychol Soc* **6**, 802-814.

[100] Lin Y, Lin Y-J, Wang F, Wu X, Kong J (2020) The development of phonological awareness and Pinyin knowledge in Mandarin-speaking school-aged children. *Int J Speech Lang Pathol* **22**, 660-668.

[101] Shu H, Peng H, McBride-Chang C (2008) Phonological awareness in young Chinese children. *Dev Sci* **11**, 171-181.

[102] Ronen M, Lifshitz-Ben-Basat A, Taitelbaum-Swead R, Fostick L (2018) Auditory temporal processing, reading, and phonological awareness among aging adults. *Acta Psychol (Amst)* **190**, 1-10.

[103] Beales A, Whitworth A, Cartwright J, Panegyres PK, Kane RT (2019) Profiling sentence repetition deficits in primary progressive aphasia and Alzheimer's disease: Error patterns and association with digit span. *Brain Lang* **194**, 1-11.

[104] Tee BL, Kwan-Chen LYL, Chen T-F, Yan CT, Tsoh J, Chan AL-T, Wong A, Lo RY, Lu CL, Wang P-N (2022) Dysgraphia phenotypes in native Chinese speakers with primary progressive aphasia. *Neurology* **98**, e2245-e2257.

[105] Ting SKS, Foo H, Chia PS, Hameed S, Ng KP, Ng A, Kandiah N (2018) Dyslexic characteristics of Chinese-speaking semantic variant of primary progressive aphasia. *J Neuropsychiatry Clin Neurosci* **30**, 31-37.

[106] Pinto-Grau M, Hardiman O, Pender N (2018) The study of language in the amyotrophic lateral sclerosis-frontotemporal spectrum disorder: A systematic review of findings and new perspectives. *Neuropsychol Rev* **28**, 251-268.