Applied Psycholinguistics (2025), **46**, e43, 1–27
doi:10.1017/S0142716425100313

ORIGINAL ARTICLE

Semantic priming and ERP correlates of predictive processing in Chinese aMCI patients

Jingjing Yang[1,2], Jing Wang[1,2] and Lihe Huang[1,2]

[1]School of Foreign Studies, Tongji University, Shanghai, China and [2]Research Center for Ageing, Language and Care, Tongji University, Shanghai, China
Corresponding author: Lihe Huang; Email: cranehlh@tongji.edu.cn

(Received 15 May 2024; revised 30 September 2025; accepted 9 October 2025)

Abstract

Purpose: One of the typical symptoms of patients with aMCI is impaired semantic memory, but it remains unclear whether this impairment affects all types of semantic relationships equally. The primary goal of this study is to assess whether there are differences in the performance of aMCI patients and healthy older adults in tasks involving antonymic and categorical semantic relationships.
Method: A delayed congruency judgment task involving different types of semantic relationships (antonymic and categorical) was conducted on 13 normal aging adults and 13 aMCI patients. Participants were presented with word cues for antonyms or category exemplars, followed by targets that were either congruent or incongruent with the cues. Electrophysiological data were recorded simultaneously.
Results: The application of the delayed congruency judgment task across various semantic relationships led to the following main findings: 1) Different semantic relationships exhibit distinct semantic priming characteristics. Antonym relationships are highly restricted lexical-semantic relations, allowing participants to make precise predictions, while categorical relationships are less restricted, leading participants to engage in graded activation and activate related features; 2) This study suggests that aMCI patients may only be able to activate specific semantic features when processing antonym relationships and are unable to make precise predictions. In contrast, their impairment in categorical relationships primarily manifests as a narrower range of activation during graded activation.

Keywords: N400; predictive ability; semantic priming; aMCI

Introduction

Amnestic Mild Cognitive Impairment (aMCI) is a condition characterized by noticeable cognitive decline that exceeds what is expected with normal aging but does not reach the threshold for dementia (Azuma et al., 2013). A hallmark feature of aMCI is semantic memory impairment, which has been widely documented in numerous studies (Delage et al., 2024). Semantic memory refers to the cognitive system responsible for the storage and retrieval of general knowledge about the world, such as concepts, words, and facts, independent of the context in which they were learned (Heinze et al., 1998). Research has shown that aMCI patients experience difficulties in accessing and retrieving semantic knowledge, which leads to challenges in language comprehension, prediction, categorization, and word retrieval (Salmon & Bondi, 2009). Such impairments can significantly affect daily functioning, particularly in tasks that require the use of language and general knowledge.

To investigate semantic memory impairments in aMCI, researchers have employed a variety of assessment methods and found that aMCI patients may experience difficulty with tasks such as confrontational naming (Leyhe et al., 2010; Benoit et al., 2017), verbal fluency (Rinehardt et al., 2014; Nutter-Upham et al., 2008), and semantic priming (Brambati et al., 2012). Of the tasks mentioned, all except semantic priming involve intentional, non-automatic, and often effortful processing of semantic information. Some researchers have argued that this pattern of impairment reflects inefficient access to semantic knowledge rather than a loss of semantic knowledge per se (Nebes, Bardy, & Huff, 1989; Ober & Shenaut, 1995). In contrast, semantic priming, as an experimental paradigm that minimizes the effects of strategic confounds, allows for the assessment of semantic memory in a more automatic manner (Duong et al., 2006).

Semantic priming refers to the phenomenon in which the processing of a target word is faster and more accurate when it is preceded by the presentation of a semantically related prime word, as opposed to a semantically unrelated item. Semantic priming is most often explained by the network model of semantic memory, where activation of one node automatically spreads to adjacent nodes, leading to a reduction in reaction time across various tasks (Collins & Loftus, 1975).

Semantic priming has been widely used to assess the integrity of semantic memory in Alzheimer's disease and aMCI (Ober, 2002). In experimental studies, researchers typically use a variety of paradigms to examine how semantic relationships between prime and target words influence lexical processing. Guglielmi et al. (2020) evaluated semantic impairment in aMCI versus healthy individuals using a lexical decision task. They found that converters (those who later progressed to dementia) lost the priming effect, suggesting impairment in semantic knowledge rather than accessibility of semantic stores in aMCI individuals who progress to dementia. Duong et al. (2006) also conducted a lexical semantic priming test among healthy elderly individuals, aMCI patients, and AD patients, using word pairs that included associated, unassociated, and neutral pairs. Participants were required to make lexical decisions on both the primes and target words, i.e., to determine whether the presented strings were real words. The results showed that AD patients exhibited a significant hyper-priming effect, meaning their reaction time to associated word pairs was significantly shorter than to unassociated word pairs. In contrast, there were no significant differences in the priming effect between the healthy older adult group and the aMCI group, suggesting that automatic semantic processing in aMCI patients did not show significant impairment in this experiment.

From the studies on semantic priming in aMCI patients, it is clear that current research on whether aMCI patients exhibit semantic priming effects is inconsistent. Is the decline in semantic priming in aMCI patients masked by certain behavioral data, and would using event-related potential (ERP) techniques to study the cognitive processing of semantic effects yield different results? Moreover, while many researchers have found semantic memory impairments in aMCI patients, does this impairment affect all types of semantic relationships? Does the type of semantic relationship between "prime-target" word pairs in the experiment influence whether aMCI patients show a semantic priming effect? One of the few studies to address this question is Caputi et al. (2016), who analyzed performance on four distinct types of associative relations: part/whole, function, superordinate, and contiguity. Their findings revealed a differential pattern of impairment across these relationships, suggesting that some types of semantic relationships are more vulnerable to decline than others, even in the early stages of aMCI. Specifically, Caputi et al. (2016) observed that the function relationship, which involves understanding the use or purpose of an object, appeared to be less affected than other types of relationships in the verbal modality. In contrast, the superordinate relationship, which involves understanding hierarchical categories (e.g., the relationship between "dog" and "animal"), showed the most rapid decline as the disease progressed.

These results suggest that the degree of semantic impairment in aMCI patients may vary across different semantic relationships, highlighting the need for further research into which semantic relationships may be selectively impaired in aMCI. To more accurately analyze the timing and mechanisms involved in processing different types of semantic relationships, this study primarily aims to use event-related potentials (ERPs) and the semantic priming paradigm to examine aMCI patients' performance in processing antonymic and categorical semantic relationships.

ERPs are electrophysiological responses that reflect the brain's instantaneous reaction to specific stimuli, offering high temporal resolution that allows researchers to track the timing and dynamics of semantic processing in real time (Paitel et al., 2021). A large number of studies have used ERPs to examine the semantic priming effect, including its influence on lexical decision (Federmeier et al., 2010), orthographic priming (Beyersmann et al., 2014), repetition priming (Wagner et al., 2006), masked priming effects, and more. Among the various ERP components, the N400 stands out as a particularly sensitive marker for the processing of semantic information. This ERP component manifests as a negative deflection in the waveform, typically peaking around 400 ms after the presentation of a word, and is thought to reflect the integration of semantic information (Kutas & Federmeier, 2000). In semantic priming tasks, the N400 amplitude is typically reduced (i.e., more positive) when the prime and target words are semantically related, indicating that the prime words activate the semantic network of the target words, making the brain process the related meaning more efficiently (León-Cabrera et al., 2021). Conversely, when the prime and target words are semantically unrelated, the N400 is typically more pronounced, reflecting the additional cognitive effort required to process the target word in the absence of semantic facilitation. The amplitude disparities between expected and unexpected target words are known as the N400 effect, reflecting the additional costs associated with processing unexpected words or words that were not pre-activated based on prior processing steps, and these have been linked to a specific electrophysiological correlate (Roehm et al., 2007).

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

Therefore, this study suggests that the N400 amplitude and the N400 effect can be used to assess the degree of activation of target words, providing insight into whether their semantic memory has been impaired. For example, if, after seeing the stimulus "fruit," a participant shows the same N400 response for both the congruent target "apple" and the incongruent target "bed," it may indicate that the participant fails to pre-activate related information for the congruent target under the influence of the stimulus word "fruit," suggesting semantic impairment. The N400 effect can also be used to evaluate the degree of activation for a specific target. For instance, if a semantic priming effect occurs, the N400 amplitude will be greater for a target with low typicality (e.g., "strawberry") and the incongruent target "bed" than for a target with high typicality (e.g., "apple"). If the amplitude differences between the high-typicality target and the low-typicality target are smaller than the amplitude differences between the high-typicality target and the incongruent target, it suggests that the participant's activation of the low-typicality target is greater than that of the incongruent target. Conversely, if the amplitude differences between the low-typicality target and the incongruent target are the same, it may indicate damage to the semantic network.

Therefore, this study uses the ERP method to investigate the semantic priming performance of aMCI patients in both antonymic and categorical semantic relationships, aiming to assess the extent of impairment in these two types of semantic relationships within the aMCI population.

The current study

This study adopts the experimental paradigm from Federmeier et al. (2010), using a semantic categorization task to analyze the semantic priming performance of aMCI patients in both antonymic and categorical semantic relationships.

In Federmeier's study, the researchers aimed to examine age-related changes in predictive processing and explored older adults' tendency to predict words from simple cues based on lexical and world knowledge under conditions with minimal time pressure. They presented younger and older adults with phrasal cues for category exemplars (e.g., "an insect") or antonyms (e.g., "the opposite of closed"), followed by targets that were congruent ("ant" or "open") or incongruent ("gate" or "final") with the cues. To specifically probe predictive processing, they included congruent but low-typicality category targets (e.g., "hornet" following "an insect"). Their findings showed that condition-related effects on the N400 followed a similar pattern across groups: N400 amplitudes were more negative for incongruent targets (ANT-IN, CAT-IN) compared to expected ones (ANT-EX, CAT-EX). N400 responses to CAT-LO targets were intermediate between CAT-HI and CAT-IN. Young adults' N400 responses were similarly sensitive to both categorical and antonymic relationships. The mean N400 effects for these two types of semantic relations were indistinguishable in size, although the antonym effect peaked earlier. For the category cues, typicality modulated the N400 response, with low-typicality items eliciting N400 responses intermediate in amplitude between high-typicality and wholly incongruent targets. In addition to the N400, they also found that for younger adults (but not older adults), CAT-LO targets were characterized by a later-occurring positivity (500–900 ms) over frontal electrode sites.

The antonym relationships were characterized by significant feature differences and high levels of lexical association, while the categorical relationships were based on high levels of feature overlap but low levels of lexical association (Federmeier et al., 2010). This study adopts this paradigm to investigate semantic memory impairment in aMCI patients and aims to determine whether semantic impairment in aMCI patients is consistent across these two types of relationships. In the Antonym Experiment, participants are presented with cues (e.g.客观/ objective), followed by targets that could either be congruent (ANT-EX) (e.g. 主观/subjective), or entirely unrelated (ANT-IN) (e.g. 壮观/spectacular) to the cue. In the Category Experiment, cues (e.g.一种颜色/ a kind of color) are presented, followed by three types of targets: congruent targets with high typicality (CAT-HI) (e.g.红色/red), congruent targets with low typicality (CAT-LO) (e.g.棕色/brown), and completely unrelated targets (CAT-IN) (e.g.苹果/apple).

In both experiments, a delayed congruency judgment paradigm is utilized, in which participants are tasked with determining whether the target word is the antonym of the cue (in the Antonym Experiment) or whether it belongs to the category indicated by the prime (in the Category Experiment). Simultaneously, ERPs are recorded to provide insights into the neural processing dynamics accompanying these tasks.

In the study by Federmeier et al. (2010), only older adults with higher category fluency scores exhibited later-occurring positivity components similar to those of younger adults when processing low-typicality items. Therefore, considering the performance of aMCI patients and older adults in verbal fluency tests, this study hypothesizes that not all participants will show the frontal positivity component, and thus, this component will not be included in the analysis. The study will focus on analyzing the N400 component in the semantic categorization task.

Based on the results of Federmeier et al. (2010), this study hypothesizes that in normally aging older adults, there is a normal priming effect in both types of semantic relationships. Therefore, N400 amplitudes will be more negative for incongruent targets (ANT-IN, CAT-IN) compared to expected ones (ANT-EX, CAT-HI). The N400 has been previously shown to be sensitive to semantic structure, as reflected in typicality (Heinzi, Muente, & Kutas, 1998). Thus, we hypothesize that in normally aging older adults, N400 responses to CAT-LO targets will be intermediate between CAT-HI and CAT-IN. In the aMCI group, if there is no impairment in either semantic relationship, they will also exhibit semantic priming effects, showing N400 components similar to those of healthy control participants. However, if the aMCI group shows impairment in both types of semantic relationships, they may not exhibit semantic priming effects, or their performance on the N400 component may differ from that of healthy older adults.

Method

Participants

Individuals with aMCI were recruited from the Department of Neurology and the Department of Memory Clinic of the affiliated hospital of the Tongji University. Age, gender, and education-matched NCs were recruited from the community

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

through a free medical examination program held at the same hospital. All subjects are native Chinese speakers, right-handed with normal vision, or correct to normal vision. There were no significant differences in terms of educational background between the two groups. Initially, a total of 17 individuals with multi-domain amnestic mild impairment (aMCI) and 17 normal controls (NC) were recruited for this study based on inclusion and exclusion criteria (showing below). However, subsequent quality checks on the EEG data revealed that some participants' data were excluded due to excessive artifact rate (>15% of trials), significant channel loss (>30% of electrodes), and poor signal-to-noise ratio (SNR < 1), which did not meet the predefined quality standards for analysis. As a result, data from 4 aMCI patients and 4 normal controls were excluded from the final analysis, leaving a final sample of 13 aMCI patients and 13 normal controls. The research protocol was approved by the ethics committee, and all participants provided written informed consent.

Exclusion Criteria: 1) age below 60 years old; 2) definite history of stoke; 3) definite history of other diseases of the central nervous system such as infection, demyelinating diseases, and Parkinson's disease; 4) definite history of mental illness such as schizophrenia, major depressive disorder; 5) severe impaired vision, hearing, aphasia disorder, and other physical disease that seriously affect neuropsychological testing; 6) alcohol or drug addiction.

For the NC group, the inclusion criteria were as follows: 1) had no complaint of memory/ cognitive decline; 2) MOCA-B (Montreal Cognitive Assessment-Basic)> 19 for primary school or above, MOCA-B > 22 for junior high schools or above (Huang et al., 2018; 3) all cognitive domains were within the normal range; 4) the ability to perform daily errands was not impaired; 5) no history of diabetes.

For aMCI group, the general criteria had been defined previously (Petersen 2004): 1) complaints of memory/ cognitive decline by the subject and his caregiver; 2) CDR (Clinical Dementia Rating Scale) = .5; 3) any of the following two criteria was met: i) at least two tests within the memory cognitive domain scored below the established cut-off; ii) at least two cognitive domains (memory and other cognitive domains) were impaired. Participant information is shown in Table 1.

Materials

Experiment 1, termed the Antonym Test, involved the presentation of 60 phrasal cues for antonyms (e.g.客观/objectivity). These cues were followed by target phrases that could be either congruent (主观/ subjectivity) or incongruent (壮观/ spectacular) with the cues. To ensure that potential confounding factors were eliminated, both the congruent and incongruent target phrases were meticulously matched for part of speech (all nouns), length, number of word stroke and work frequency. Half of the phrasal cues were followed by congruent phrases (ANT-EX) while the other half were followed by incongruent phrases (ANT-IN). In this Experiment, the congruent target phrases represented the expected responses, while the incongruent target phrase were considered unexpected. The antonym pairs in this experiment were sourced from the *Chinese Antonym Dictionary* (Zhu, 2014).

Experiment 2, referred to as the Category Test, involved the presentation of 90 category cues (e.g., "一种颜色"/a kind of color). These cues were followed by three types of targets: incongruent target (e.g., "苹果"/ apple), congruent targets with high

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

Table 1. Demographic data, neuropsychological performance for all subjects (Significant differences were indicated between the two groups)

	NC(n = 13)	aMCI(n = 13)	p
age	69.46(8.96)	68.2(11.01)	0.757
education	12.46(2.53) yrs	10.92(2.47) yrs	0.130
gender	5 M 8 F	3M 10F	
MMSE	27.46(1.33)	23.54(3.57)	0.002**
MoCA-B	24.46(3.69)	16.15(4.32)	0.000**
Boston Naming	23.62(3.66)	20.00(4.74)	0.040*
Verbal Fluency	17.07(2.49)	12.69(3.66)	0.002**

Note: Data are presented as the mean (SD). p value of gender was obtained by chi-square test; p values for comparison in other demographic data, neuropsychological performance and behavioral data were acquired by independent sample T-test; * denotes $p < 0.05$, ** denotes $p < 0.01$. SD, standard deviation.

typicality (e.g., “红色”/red), and congruent target with low typicality (e.g., “棕色”/brown). All the exemplars were carefully matched for part of speech (all nouns) and word length. Among the 90 category cues, 30 were followed by incongruent targets, 30 by congruent targets with low typicality, and 30 by congruent targets with high typicality. In this experiment, the congruent target with high typicality (CAT-HI) was anticipated to elicit the highest level of expectation, followed by the target with low typicality (CAT-LO), while the incongruent target (CAT-IN) was expected to evoke the lowest level of expectation.

The category cues for this experiment were sourced from Yoon et al. (2004), which selected 105 semantic categories and asked 100 Chinese older adults to generate five words per category, producing a set of high-typicality words for each category. From this database, we selected 90 semantic categories, with 30 high-typicality words as targets for the CAT-HI condition directly sourced from Yoon et al. (2004). These selections were made to reflect the typicality norms of the Chinese population, thus ensuring the stimuli were well-suited for our participants.

Procedure

A delayed congruency judgment task was employed in these two experiments to ensure that participants were actively engaged with the stimuli and could discern the relationship between the cues and targets. The entire experimental procedure took place within a sound-attenuated, electrically-shielded chamber. Participants were comfortably seated in a chair positioned in front of a 17-inch computer screen, situated at a distance of approximately 0.5 meters from their eyes. The stimulus presentation was administrated through E-prime 3.0 software. The subjects were offered instructions prior to the commencement of the formal experimental. Each experiment was preceded by a practice phase consisting of 15 trials runs, designed to ensure that every participant comprehended and felt at ease with the operational aspects of the task.

Before the experiments, instructions will be given to participants. In the Antonym Experiment, we will inform the participants: “You will complete an

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

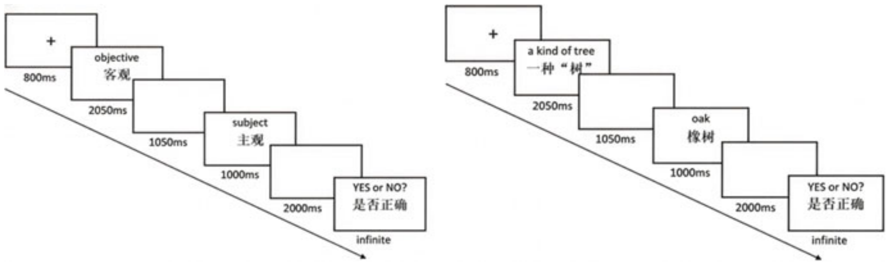

Figure 1. Procedure of Antonym Experiment (left) and Category Experiment (right) for one trial.

antonym judgment test. You will then see two words in sequence. The first word is the prime, and the second word may or may not be its antonym. After seeing the second word, please judge whether it is the antonym of the first word. If it is, press the right key; if it is not, press the left key.” In the Category Experiment, we will inform the participants: “You will complete a category judgment test. You will then see two words in sequence. The first word is a category, and the second word may or may not belong to the category of the first word. After seeing the second word, please judge whether it belongs to the category shown by the first word. If it does, press the right key; if it does not, press the left key.”

The complete cue phrase was presented in the center of the screen for 2050 ms, followed by a blank screen of 1050 ms. The target word was then presented for 1000 ms. At the offset of the target word, an interval of 2000 ms was set to capture ERP responses before the cue “Yes or No?” Participants were asked to respond to this cue with button presses indicating if the target was congruent or incongruent with the cue phrase. With an interval of 1500 ms, it proceeded to the next trial. The Antonym Experiment and Category Experiment were separated, and in each experiment the order of all trials was randomly presented. Participants were given a short break after every 30 trials. The paradigm is demonstrated in Figure 1.

EEG recording and data processing

EEG was recorded from 64 single Ag/AgCl scalp electrodes positioned based on the international extended 10–20 system using a NeuroScan SynAmps2 (SynAmps2TM Model 8050 EEG amplifier and a data acquisition system, Abbotsford, Victoria, Australia). Four facial electrodes were positioned adjacent to each eye’s left and right outer canthus and above and below the left orbit to measure eye movement. The data were processed in MATLAB (Version R2016a) using the EEGLAB toolbox (Version 14.1.1). Electrodes were grounded to Groud (GND) and referenced online to an electrode between the CZ and CPZ electrode positions. Impedances were kept below 10 KΩ. EEG was sampled at 1000 Hz with an online bandpass of 0.01–250Hz. ERPs were time-locked to the onset of the target word, and the ERP data for a total period of 1200 milliseconds (including 200 ms before the presence of the stimulus and 1000 ms after the presentation) were analyzed offline with filtering (30Hz), removing electrooculogram artifacts, being referenced to the average of the left and right mastoids and being baseline-corrected with respect to a 200ms pre-stimulus recording interval. Before averaging, trails with eye movements, blinks, and

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

excessive muscle activity were rejected. ERPs were derived by averaging correctly classified trials on each type of target word for each participant.

Statistical analysis

To analyze the reaction times (RTs), we standardized the reaction times using Z-score normalization and removed outliers with Z-scores greater than 3 or less than −3. We fitted a linear mixed model to standardized RT data using the lme4 package (Bates et al., 2015) in R. The model included fixed effects for group (aMCI and NC), material type (ANT-EX, ANT-IN in the Antonym Experiment; CAT-HI, CAT-LO, CAT-IN in the Category Experiment), and interaction item (group by material type). For the random effects, we first built the model with the maximal random effects structure. Based on our experiment, material types vary within-subjects but between-items. We thus needed to specify a by-subject random slope material type. Subject group varied between-subjects but within-items. The design thus called for a by-item random slope for group. The maximal random effect structure contained four random effects: a by-subject random intercept, a by-subject random slope for material type, a by-item random intercept, and a by-item random slope for group. If the model failed to converge or overfitted, we used the Principal Components Analysis (Meteyard et al., 2020; Schad et al., 2020; Barr et al., 2013) of the random effects structure to identify the variance components that could be removed without a loss in goodness of fit of the model. The parameters of the optimization model were reported. For each of the three fixed effects, its statistical significance was assessed using the Satterthwaite approximation for degrees of freedom from the lmerTest R package (Kuznetsova et al., 2016), and that we reported the corresponding F values and p values. For accuracy, we fitted a generalized linear mixed model with a binary distribution. For the fixed and random effects, the model was built in the same manner as in the analysis of RTs. In the statistical analyses, we reported the χ^2 values and corresponding p values for each of the fixed effects.

ERPs were isolated from the ongoing EEG by simple averaging, so the ERP data for each participant on each condition was the average of corresponding items. To analyze the ERP data of N400 component, we extracted the data of interested channels (in Figure 2) and fitted a linear mixed model with group, material and interaction item as fixed effects, while participant and channel as random effects. In the analysis of ERP, channel varied within subjects and within material types, thus the maximal random effect structure contained six random effects: a by-subject random intercept, a by-subject random slope for material type, a by-channel random intercept, a by-channel random slope for group, a by-channel random slope for material, and random slopes for the effect of interaction across channel. The optimization model was built in the same manner as in the analysis of behavior data.

To analyze the N400 effect (the amplitude difference between ANT-EX and ANT-IN targets or the difference between CAT-HI and CAT-IN targets), we fitted a linear mixed model with group, experiment type and interaction item as fixed effects, while participant and channel as random effects. The same with the analysis of N400 amplitude, channel varies within subjects and within experiment types, thus the maximal random effect structure contained six random effects: a by-subject

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

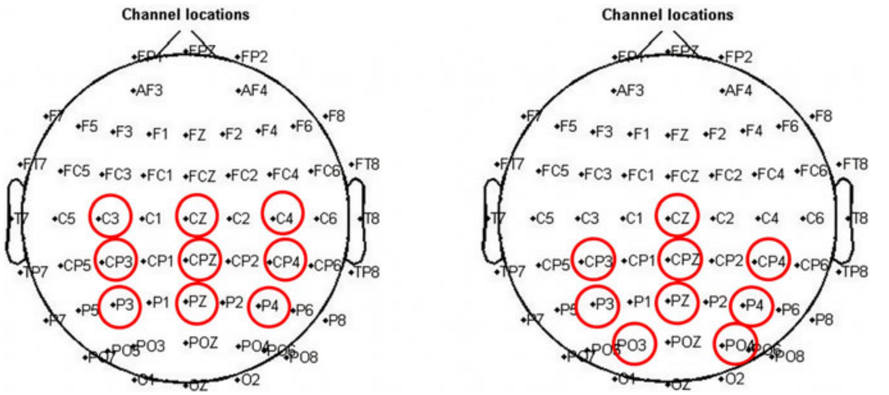

Figure 2. 9 selected channels marked by red circles in ANT Experiment (left) and CAT Experiment (right).

random intercept, a by-subject random slope for experiment type, a by-channel random intercept, a by-channel random slope for group, a by-channel random slope for experiment type, and random slopes for the effect of interaction across channel. The optimization model was built in the same manner as in the analysis of N400 amplitude.

Post hoc analysis adjusted by 'Bonferroni' was calculated using the emmeans package (Lenth, 2023) when the main effect of predictors or interaction effects were significant. Confidence intervals (95% CI) were calculated using the Wald method.

Regression analysis was adopted to explore the relationship between all variables. The conventional significance level of $\alpha = .05$ was used for the overall analysis.

Results

Antonym experiment

Behavioral results

The accuracy of the Antonym test varies as a function a between-group. Mean accuracy data are detailed in Table 2. The behavioral data underscore that accuracy levels were close to ceiling for normal aging adults, aligning with previous research indicating the preservation or augmentation of knowledge stores with increasing age. Subsequent to the PCA analysis, which led to the removal of the random slope for material across participants and random slope for group across items, an optimization model for accuracy in the Antonym Experiment was established. The results of this model can be found in Appendix A.

In the optimization model, we found a significant main effect of group ($\chi^2 = 7.948$, $df = 1$, $p = .0048$), wherein the NC group had higher accuracy than the aMCI group. The main effect of material type ($\chi^2 = .7313$, $df = 1$, $p = .3924$) and the interaction between group and material ($\chi^2 = 2.4393$, $df = 1$, $p = .1183$) were not significant.

Mean response time are demonstrated in Table 2. The optimization model for RTs removed the random slope for group across item and random slope for material across participant to avoid overfitting, and the summary of the model was presented

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

Table 2. Mean accuracy (Percentages of Correct Judgments), mean response time (in Milliseconds) and N400 amplitude of the Antonym Test for the NC and aMCI group

	NC group	aMCI group
Accuracy Test	99.36(1.08)	94.36(7.47)
accuracy on ANT-EX	98.97(2.10)	95.38(4.63)
accuracy on ANT-IN	99.74(0.92)	93.33(10.72)
Response Time all	675.01(192.23)	1635.48(1654.03)
response time on ANT-EX	724.11(249.99)	1308.84(842.09)
response time on ANT-IN	625.92(188.65)	1959.39(2508.83)
amplitude of N400 on ANT-EX	10.102 (4.451)	8.436 (7.039)
amplitude of N400 on ANT-IN	4.659 (2.846)	5.309 (5.982)

in Appendix B. We found that the main effect of group ($F = 4.960$, $df = 24$, $p = .0357$), the main effect of material type ($F = 7.7822$, $df = 58$, $p = .00713$), and the interaction between group and material ($F = 29.7362$, $df = 1402$, $p = 5.843e{-}08$) were significant. A post hoc analysis found that in the aMCI group, the response time of ANT-IN was much longer than that of ANT-EX ($F\text{-}ratio = 28.041$, $df1 = 1$, $df2 = 129$, $p < .0001$), and for the ANT-IN item, the aMCI group had much longer response time than the NC group ($F\text{-}ratio = 9.913$, $df1 = 1$, $df2 = 25.6$, $p = .0041$).

ERP results

Previous work using ERPs to explore the comprehension of targets following cues like these found significant N400 priming for congruous (compared to incongruous) targets. The differential waveforms between congruent targets and incongruent targets were used to determine the time window of N400 component.

N400 latency and amplitude were assessed at the 9 selected channels where such responses are most prominent, sites were marked with circles in Figure 2. The ERP responses of the NC group were shown in Figure 3, and those of the aMCI group were shown in Figure 4. Peak latency of the N400 effect was determined by the difference of the incongruent target (ANT-IN) and its highly expected counterpart (ANT-EX), which was 380ms for NC group and 400ms for aMCI group (Figure 5).

Amplitude was measured on the waveforms that were bandpass filtered from 0.2 to 5 Hz. N400 mean amplitudes were assessed in a 60ms window centered on the effect peak for each group (i.e., 350–410ms for NC group and 370–430ms for aMCI group). Table 2 demonstrated the amplitude of N400 component of two groups in the Antonym experiment. The model with maximal random effects structure contained six random effects, but presented the issue of overfitting. Appendix C provided the summary of the optimization model for N400 amplitude. We found a significant main effect of material type ($F = 63.698$, $df = 24$, $p = 3.29 * e{-}08$), and the interaction effect between material type and group ($F = 4.6531$, $df = 24$, $p = .0412$), while the main effect group was not significant ($F = .068$, $df = 24.939$,

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

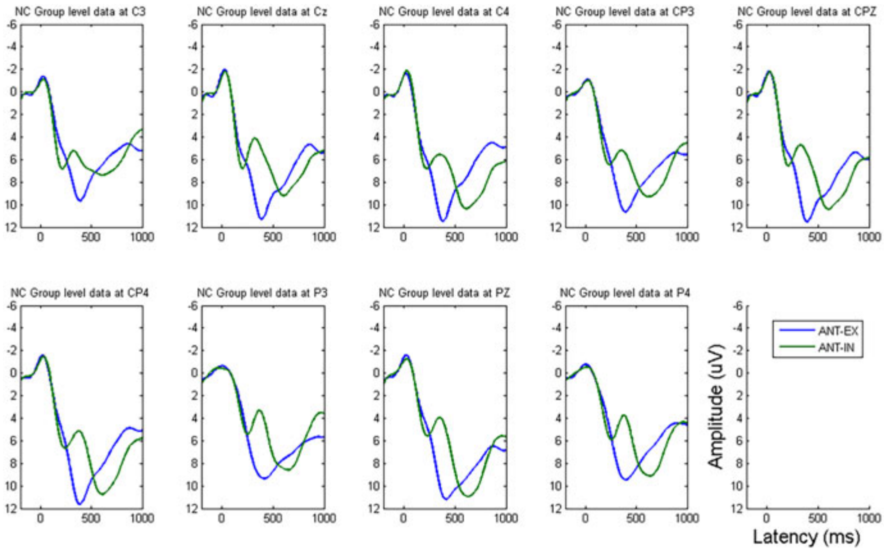

Figure 3. Normal aging adults' ERP responses to antonym targets of the selected electrodes. *Note:* Negative voltage is plotted up in this and all subsequent figures. N400 responses elicited by expected words are presented in blue line, while incongruent targets in green line. For the figures presented, the data were filtered using a 5 Hz low-pass filter.

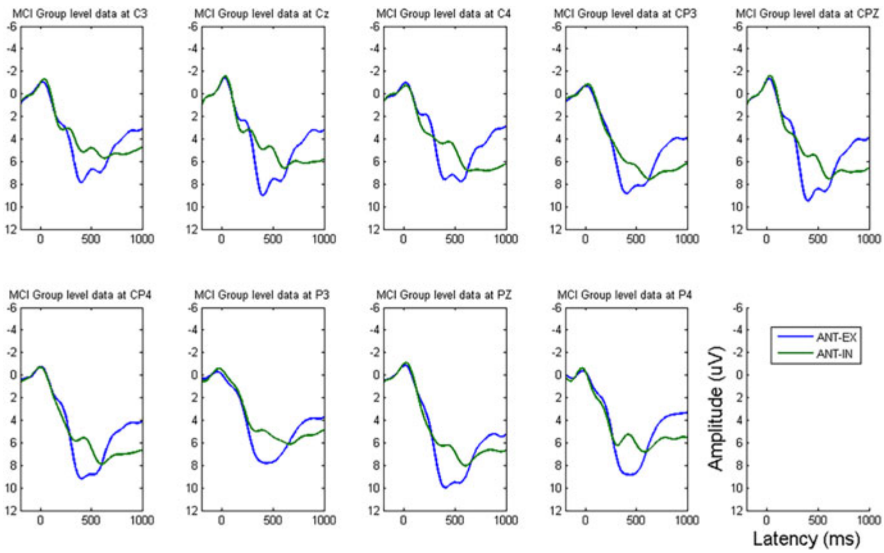

Figure 4. aMCI patients' ERP responses to antonym targets of the selected electrodes. *Note:* N400 responses elicited by expected words are presented in blue line, while incongruent targets in green line. For the figures presented, the data were filtered using a 5 Hz low-pass filter.

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

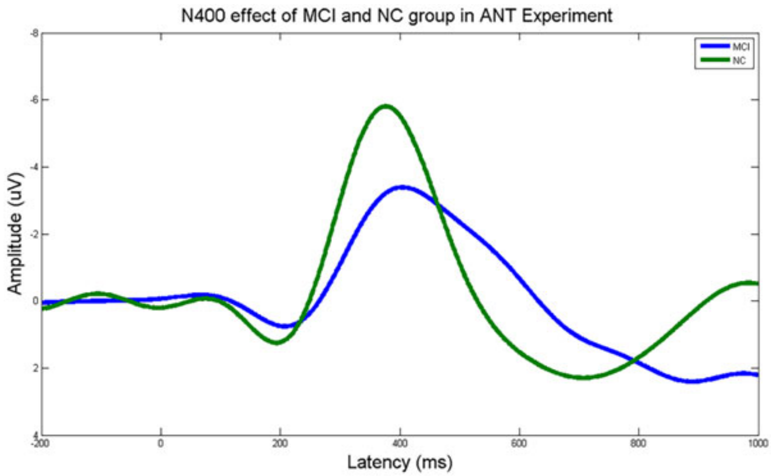

Figure 5. The N400 effect of aMCI and NC group in ANT Experiment. *Note:* The N400 effect was the amplitude difference between ANT-EX and ANT-IN targets. Peak latency of the N400 effect was 380 ms in NC group, and 400 ms in aMCI group. For the figures presented, the data were filtered using a 5 Hz low-pass filter.

$p = .796$). The post hoc analysis revealed that no matter in the aMCI ($t = 4.118$, $df = 24$, $p = .0004$) or the NC group ($t = 7.169$, $df = 24$, $p < .0001$), the N400 amplitude of the ANT-IN was larger than that of the ANT-EX.

Category experiment

Behavioral results

The accuracy of Category Test was demonstrated in Table 3. The optimization model of accuracy in the Category Experiment was presented in Appendix E.

In the optimization model (Appendix E), we found a significant main effect of group ($\chi^2 = 53.129$, $df = 1$, $p = 3.124 * \text{e}{-13}$), wherein the NC group had higher accuracy than the aMCI group, and a significant main effect of material type ($\chi^2 = 14.683$, $df = 2$, $p = .0006$). The interactive effect was not significant ($\chi^2 = 1.391$, $df = 2$, $p = .499$). The post hoc analysis revealed that in the aMCI group, the accuracy of CAT-IN was larger than that of the CAT-LO ($z = 3.558$, $p = .0011$), and the accuracy of CAT-HI was larger than that of the CAT-LO ($z = 3.201$, $p = .0041$).

Mean response time are demonstrated in Table 3. The result of optimization model was presented in Appendix D. In this model, we found a significant main effect of group ($F = 8.3101$, $df = 24$, $p = .0082$) but no significant main effect of material ($F = 2.7046$, $df = 2083$, $p = .067$). The interactive effect between group and material also meets the significant level ($F = 3.392$, $df = 2083$, $p = .0338$).The post hoc analysis revealed that the NC group had significantly faster response times on all three target types compared to the aMCI patients (CAT-HI: $t = 3.440$, $df = 28$, $p = .0018$; CAT-LO: $t = 2.663$, $df = 28$, $p = .0126$; CAT-IN: $t = 2.194$, $df = 28.1$, $p = .0367$). Within the aMCI group, there were no significant differences in reaction times between the three target types. In the NC group, the reaction time for

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

Table 3. Mean accuracy (Percentages of Correct Judgments), mean response time (in Milliseconds) and N400 amplitude of the Category Test for the NC and aMCI group

	NC group	aMCI group
Mean Accuracy	96.33(2.95)	87.77(8.51)
ACC on CAT-IN	97.18(7.31)	91.79(9.59)
ACC on CAT-LO	95.12(5.37)	80.51(10.96)
ACC on CAT-HI	96.67(3.33)	91.28(9.48)
Mean Response Time	675.01(192.23)	1635.48(1654.03)
RT on CAT-IN	822.65(442.96)	2558.43(3188.47)
RT on CAT-LO	946.21(411.17)	2049.81(1436.28)
RT on CAT-HI	926.56(491.95)	1641.62(995.02)
N400 amplitude on CAT-IN	2.516 (3.320)	2.088 (5.171)
N400 amplitude on CAT-LO	3.147 (3.423)	3.141 (5.274)
N400 amplitude on CAT-HI	4.526 (3.076)	5.078 (5.551)

CAT-IN was significantly faster than that for CAT-LO ($t = 2.641$, $df = 299$, $p = .0261$).

ERP results

The N400 was a sensitive component to semantic memory structure as reflected in typicality (Heinze et al., 1998). The same with Antonym Experiment, the differential waveforms between congruent targets and incongruent targets were used to determine the peak latency of N400 in Category Experiment, which was 420ms for the NC group (Figure 6) and 450ms for the aMCI group (Figure 7).

N400 mean amplitudes were assessed in a 60ms window centered on the effect peak for each group (i.e., 390–450ms for the NC group and 420–480ms for the aMCI group) and was demonstrated in Table 3. The results of optimization model were represented in Appendix F, in which we found a significant main effect of material type ($F = 14.982$, $df = 2$, $p = 5.9887$*e05). The post hoc analysis found that in the aMCI group, the N400 amplitude of CAT-IN and CAT-LO were much larger than that of the CAT-HI (CAT-IN vs. CAT-HI $t = 4.625$, $df = 24$, $p = .0003$; CAT-LO vs. CAT-HI $t = 2.898$, $df = 24$, $p = .0237$), while in the NC group, only the difference between CAT-IN and CAT-HI met the significant level ($t = 3.110$, $df = 24$, $p = .0143$).

The result of N400 effect analysis

The N400 effect of aMCI and NC group in Category Experiment were shown in Figure 8. Table 4 showed the averaged N400 effect across selected channels of the Antonym and Category experiment of the two groups of participants, and the results of the optimization model were presented in Appendix G. In this model, we found a significant main effect of experiment type ($F = 7.14$, $df = 1$, $p = .0123$)

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

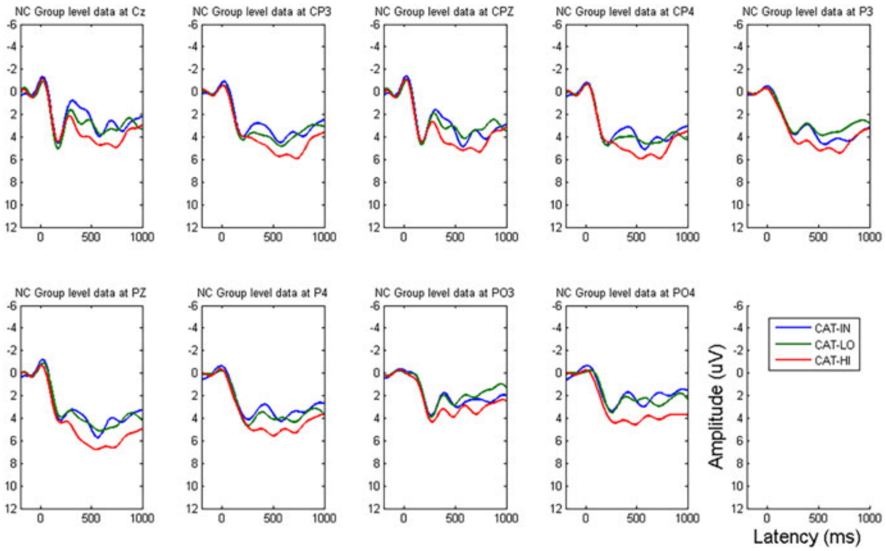

Figure 6. Normal aging adults' ERP responses to category targets of the selected electrodes. *Note:* N400 responses elicited by incongruent targets are presented in blue line, while low typicality targets in green and high typicality targets in red. For the figure presented, the data were filtered using a 5 Hz low-pass filter.

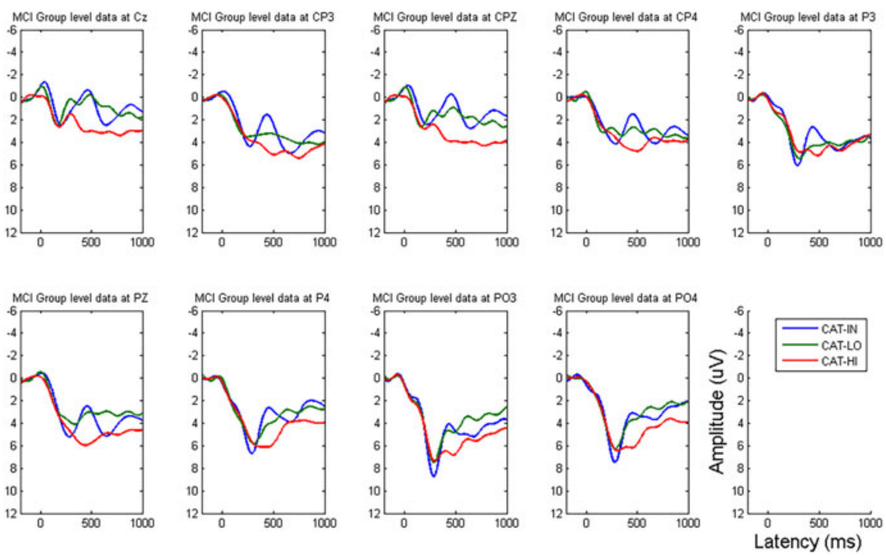

Figure 7. aMCI patients' ERP responses to category targets of the selected electrodes. *Note:* N400 responses elicited by incongruent targets are presented in blue line, while low typicality targets in green and high typicality targets in red. For the figure presented, the data were filtered using a 5 Hz low-pass filter.

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

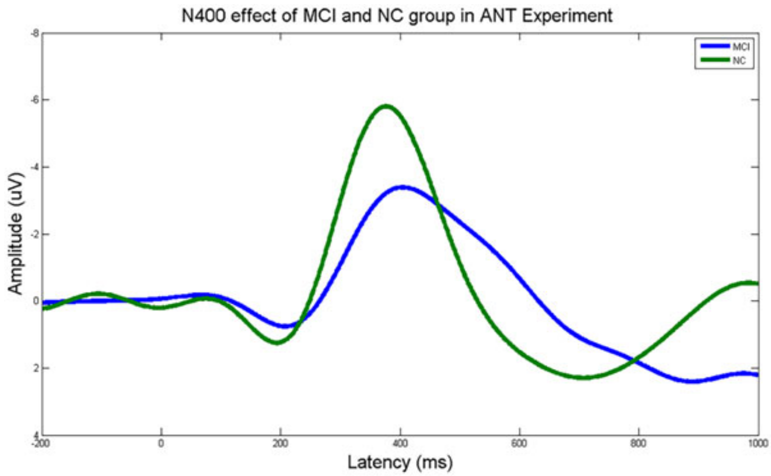

Figure 8. The N400 effect of aMCI and NC group in CAT Experiment. *Note:* The N400 effect was the amplitude difference between CAT-HI and CAT-IN targets. Peak latency of the N400 effect was 450ms in the NC group, and 480ms in the aMCI group. For the figures presented, the data were filtered using a 5 Hz low-pass filter.

Table 4. N400 effect of Antonym and Category Test for the NC and aMCI group

	NC group	aMCI group
ANT	5.640(2.992)	3.317(2.919)
CAT	2.049(2.573)	3.176(2.910)

Table 5. Pearson correlation coefficient r (*p*-value) between cognitive abilities and accuracy of different materials

	Boston Naming	Verbal Fluency
ANT-ACC	0.317 (0.114)	0.321 (0.110)
ANT-RT	−0.421* (0.032)	−0.369 (0.064)
CAT-ACC	0.538 **(0.005)	0.480 *(0.013)
CAT-RT	−0.375 (0.059)	−0.605** (0.001)

Note. *$p < .05$. **$p < .01$. ***$p < .001$

and a significant interactive effect between experiment type and group ($F = 7.244$, $df = 1$, $p = .0128$). The post hoc analysis found that in the Antonym experiment, NC group's N400 effect was much larger than that of the aMCI group ($t = 2.475$, $df = 24$, $p = .0208$), and for the NC group, the N400 effect of Antonym experiment was larger than that of the Category experiment ($t = 3.789$, $df = 27.4$, $p = .0008$).

The results of Pearson correlation analysis

Pearson r correlations were employed to explore the relationship between cognitive abilities and subjects' performance of these two experiments. As shown in Table 5, the results revealed that there existed a significant relationship between Boston Naming and the reaction time of the Antonym experiment, ($r = -.421$, $p = .032$), and the accuracy of the Category experiment ($r = .538$, $p = .005$). The scores of Verbal Fluency test were positively related with the accuracy of the Category experiment ($r = .480$, $p = .013$), and negatively with the reaction time ($r = -.605$, $p = .001$).

Discussion

The central focus of this study is to assess whether aMCI patients exhibit the same decline across different types of semantic relationships. The behavioral data presented above indicate that the aMCI group showed slower response times and lower accuracy levels compared to the NC group in both experiments. These findings suggest that semantic spreading activation efficiency may decline in aMCI patients, regardless of the type of semantic relationship. In the Category Experiment, both the NC group and the aMCI group demonstrated faster response times for both CAT-HI and CAT-IN targets. However, the accuracy of CAT-LO items remained higher in the NC group than in the aMCI group.

In terms of ERP data, the aMCI group exhibited general delays and amplitude reductions in the N400 component. Similar to the NC group, the aMCI group displayed significant N400 effects for both category-based and antonym-based relationships, indicating that the semantic priming effect still exists within the aMCI group. It is worth noting that the N400 effect in the aMCI group peaked later, and the magnitude of the N400 effect was significantly lower than that of the NC group in the Antonym Experiment. In contrast, in the Category Experiment, there was no notable difference in the size of the N400 effect between the two groups. Consistent with the results of Federmeier et al. (2010), in both groups, the N400 amplitude in the Antonym Experiment peaked earlier than in the Category Experiment.

The above summarizes the results of these two experiments, raising two key questions: one regarding the differences caused by semantic relationships, and the other regarding the differences caused by group characteristics.

Semantic relationship differences

In our study, we found that in both the ANT and CAT experiments, incongruent target words elicited an increase in N400 amplitude, indicating that semantic effects were activated in both antonym and categorical relationships. Regarding the results from healthy older adults, in the ANT experiment, the N400 amplitude for ANT-IN was significantly larger than that for ANT-EX. In the CAT experiment, a linear relationship was observed between target word typicality and N400 amplitude. Specifically, the N400 amplitude for CAT-LO was intermediate between that of CAT-HI and CAT-IN. This observation aligns with our initial hypothesis and is consistent with the findings of Federmeier et al. (2010) regarding both young and older adults.

Although both types of semantic relationships showed the presence of the N400 component, they exhibited different patterns in the N400 effect (defined as the amplitude difference between congruent and incongruent targets). The present

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

study found that within the NC group, the N400 effect for the antonymic relationship was significantly larger than that for the categorical relationship, suggesting that when an incongruent target appears, more cognitive resources are required to process the incongruent target in the antonym relationship. This study proposes that this may be related to the characteristics of the two types of semantic relationships. Federmeier et al. (2010) suggested that antonym relationships are characterized by important feature differences and high levels of lexical association, while categorical relationships are based on high levels of feature overlap but low levels of lexical association. Antonym relationships are considered highly restricted lexical-semantic relations (Roehm et al., 2007), where participants focus precisely on specific words. Therefore, in the ANT experiment, when participants see the prime, they are likely to make precise predictions, such as "black-white." In contrast, categorical relationships involve lower levels of lexical association. In the CAT experiment, participants may not make specific predictions but instead engage in passive spreading activation, where related words and semantic features are activated to varying degrees.

As mentioned in Lau et al. (2013), when processing pairs of words from the two semantic relationships, after the prime is encountered, a strongly associated target word is predictively added to a working memory representation of the prime-target pair. In contrast, for categorical relations, lexical facilitation for related targets arises primarily due to the passive priming of representations stored within long-term semantic memory. Since activation involves updating representations in semantic memory in advance of the input, incongruent words may increase lexical selection difficulty. This is because the lexical representation activated by the bottom-up input has to compete with the highly activated representation (Lau et al., 2013). The N400 effect is at least partially driven by the degree to which the context predicts the target (Federmeier, 2007).

From the results of our study, in the ANT experiment, the context has a stronger predictive effect on the target, as the N400 effect for healthy older adults in the ANT experiment was larger than that in the CAT experiment. These results align with those of Lau et al. (2013), who found that when the experimental context encourages participants to make more specific predictions, a greater N400 reduction occurs, corresponding to a greater N400 effect.

The differences in semantic priming observed above are attributable to the two types of semantic relationships. In summary, antonym relationships are highly restricted lexical-semantic relations, enabling participants to make precise predictions. In contrast, categorical relationships are less restricted lexical-semantic relations, where participants typically engage in graded activation, activating related features. Now, let's examine the semantic priming performance in both the aMCI group and the healthy older adult group for these two types of semantic relationships.

The performance of the aMCI group in antonym relationships

The decline in performance among aMCI patients in antonym relationships may primarily reflect their inability to make precise predictions based on the prime.

Analysis of the experimental results revealed that, in both the aMCI and NC groups, the amplitudes were more negative for incongruent targets (ANT-IN, CAT-IN) compared to expected targets (ANT-EX, CAT-HI), suggesting that both aMCI patients and healthy controls exhibit semantic priming effects. However, based on the performance of the N400 component, there appear to be some differences between the two groups.

First, regarding the size of the N400 effect, this study found that the NC group exhibited a difference in the N400 effect across the two experiments. The N400 effect (the difference between ANT-IN and ANT-EX) in the ANT experiment was significantly larger than the difference between CAT-IN and CAT-HI in the CAT experiment. Since the N400 effect reflects the degree to which the context predicts the target, this suggests that in the NC group, the prime had a stronger predictive effect on the target, enabling them to make precise predictions. In contrast, in the CAT experiment, the NC group seemed to engage in "graded activation." However, in the aMCI group, no such difference was observed. The aMCI group showed no significant difference in the N400 effect between the ANT and CAT experiments, suggesting that the activation levels for both highly restricted lexical-semantic relations and less restricted ones were the same in this group. So, did the aMCI group engage in graded activation or precise activation? Given that there were no differences in the N400 effect between the NC and aMCI groups in the CAT experiment, this study suggests that the aMCI group engaged in graded activation in the CAT experiment. Similarly, in the ANT experiment, the antonym prime only activated the semantic features of the target for the aMCI patients, without enabling the precise prediction seen in the NC group. As a result, compared to the NC group, the aMCI group showed a smaller N400 effect in the ANT experiment. This finding suggests that aMCI patients may experience reduced facilitation from constraining contexts.

Statistical analyses further revealed delays in the peak of the N400 effect among aMCI patients in both experiments. The prolonged and diminished N400 congruency effect suggests that aMCI patients exhibit slower and less effective use of available information to influence their word processing, especially in more predictive contexts. This finding highlights the difficulties aMCI patients face in effectively utilizing contextual cues to facilitate language processing.

The performance of the aMCI group in categorical relationships

The difference in the performance of aMCI patients on categorical relationships is primarily reflected in the fact that the range of activation in graded activation is relatively smaller compared to the NC group. This is evident in two main points. First, in the EEG data, the study found that in the aMCI group, the N400 amplitude for CAT-IN and CAT-LO was significantly larger than for CAT-HI, but there was no significant difference between CAT-LO and CAT-IN. This suggests that aMCI patients processed CAT-LO and CAT-IN similarly. Second, in the behavioral data, we observed differential performance between the two groups regarding items with varying typicality and predictability. In the NC group, although accuracy for CAT-LO items was lower than for CAT-IN and CAT-HI items, the difference did not reach statistical significance. In contrast, the aMCI patients showed a notable and statistically significant difference, indicating a tendency to treat low-typicality

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

targets as incongruent. This suggests that low-typicality targets may fall outside the expectancy set activated by the prime for aMCI patients, indicating that even within the framework of "graded activation," the aMCI group exhibits weaker activation and a more limited scope of pre-activation compared to the NC group.

The observed deficits in semantic memory among aMCI patients in this study are consistent with previous research, which suggests that aMCI patients may experience category-general semantic memory impairments (Chang et al., 2022). Semantic memory serves as a vast repository of lexical meaning information (Kutas & Federmeier, 2000). As highlighted by Brothers et al. (2017), semantic retrieval and graded activation form the foundation of predictive processing. When encountering a prime stimulus, individuals must access the semantic features of the prime and activate corresponding lexical representations in semantic memory. In highly constraining contexts with rich linguistic cues, activation accumulates for a specific lexical item, leading to precise word prediction. However, when the structure or contents of semantic memory are altered, it can manifest in different cognitive performances. The narrower activation range and extended reaction times observed in the aMCI group suggest that the spread of activation within semantic memory is impeded by the condition. The diminished capacity to fully leverage contextual information may stem from difficulties in retrieving semantic features, which in turn impacts the predictive process.

The rapid activation of semantic information and the ability to make accurate predictions are crucial for language processing. Numerous eye-movement studies have suggested that predictions generated in highly constraining contexts can significantly enhance the comprehension of both auditory and visual words (DeLong et al., 2014). Therefore, if aMCI patients have impaired abilities to use contextual information to make precise predictions, they may face difficulties in reading and experience slower sentence processing speeds.

Conclusion

Through the application of delayed congruency judgment tasks across various semantic relationships, we made the following key findings: 1) Different semantic relationships exhibit distinct semantic priming characteristics. Antonym relationships are highly restricted lexical-semantic relations, allowing participants to make precise predictions, while categorical relationships are less restricted, where participants generally engage in graded activation, activating related features; 2) aMCI patients show different patterns of impairment when processing different semantic relationships. This study suggests that aMCI patients may only be able to activate specific semantic features when processing antonym relationships, preventing them from making precise predictions. As a result, the difference between ANT-IN and ANT-EX is smaller in the aMCI group compared to the NC group. In contrast, aMCI patients' impairment in categorical relationships primarily manifests as a smaller range of activation during graded activation, making them more likely to identify low-typicality targets as incongruent.

The results of this study suggest that different semantic relationships may lead to distinct patterns of deterioration. When analyzing semantic memory in aMCI

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

patients, it is important to consider the influence of semantic relationships. However, these findings should be viewed in light of several limitations. First, despite efforts to minimize the impact of head movements and muscle artifacts through data processing, some residual artifacts remain unavoidable, which requires cautious interpretation of the results. Second, the sample size in this study was relatively small, which may influence the statistical power and limit the generalizability of the findings. As this study is exploratory in nature, it aims to provide insights and hypotheses for future research with larger sample sizes. Third, while the focus has been primarily on alterations in semantic memory and their impact on semantic aspects of word comprehension, it is clear that aMCI is associated with changes across multiple cognitive processing levels. Therefore, future research should prioritize understanding the underlying differences in predictive ability and their consequences for language processing, as a key avenue for investigating language changes associated with cognitive decline.

Replication package. All analysis, data, and code are available at: https://osf.io/5kc93/.

Financial support. This research was supported by the National Social Science Fund of China (Grant: 24BYY120).

Competing interests. The authors have no conflicts of interest to declare.

Ethical standards. This research received approval from the Tongji University School of Foreign Studies Review Ethics Committee, with the approval number tjsflrec202306.

References

Altmann, G., & Mirkovic´, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, **33**(4), 583–609.

Azuma, T., Sabbagh, M. N., & Connor, D. J. (2013). The effect of healthy aging and mild cognitive impairment on semantic ambiguity detection. *Journal of Neurolinguistics*, **26**(2), 271–282. https://doi.org/10.1016/j.jneuroling.2012.09.003

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of memory and language*, **68**(3), 255–278.

Bates, D., Maechler, M., Bolker, B., & Walker, S.(2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Benoit, S., Rouleau, I., Langlois, R., Dostie, V., Kergoat, M. J., & Joubert, S. (2017). The impact of time and repeated exposure on famous person knowledge in amnestic mild cognitive impairment and Alzheimer's disease. *Neuropsychology*, **31**(7), 697.

Beyersmann, E., Iakimova, G., Ziegler, J. C., & Colé, P. (2014). Semantic processing during morphological priming: an ERP study. *Brain Research*, **1579**, 45–55. https://doi.org/10.1016/j.brainres.2014.07.010

Brambati, S. M., Peters, F., Belleville, S., & Joubert, S. (2012). Lack of semantic priming effects in famous person recognition in mild cognitive impairment. *Cortex*, **48**(4), 414–420. https://doi.org/10.1016/j.cortex.2011.04.001

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of memory and language*, **93**, 203–216.

Caputi, N., Di Giacomo, D., Aloisio, F., & Passafiume, D. (2016). Deterioration of semantic associative relationships in mild cognitive impairment and alzheimer disease. *Applied Neuropsychology-Adult*, **23**(3), 186–195. https://doi.org/10.1080/23279095.2015.1030020

Chang, H. T., Chiu, M. J., Chen, T. F., Liu, M. Y., Fan, W. C., Cheng, T. W., … & Hua, M. S. (2022). Deterioration and predictive values of semantic networks in mild cognitive impairment. *Journal of Neurolinguistics*, **61**, 101025.

Collins, A.M.; Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, **82**, 407–428.

Delage, E., Rouleau, I., Akzam-Ouellette, M.-A., Roy-Cote, F., & Joubert, S. (2024). An examination of semantic performance in mild cognitive impairment progressors and nonprogressors. *Neuropsychology*, **38**(4), 309–321. https://doi.org/10.1037/neu0000947

DeLong, K. A., Quante, L., & Kutas, M. (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia*, **61**, 150–162.

Duong, A., Whitehead, V., Hanratty, K., & Chertkow, H. (2006). The nature of lexico-semantic processing deficits in mild cognitive impairment. *Neuropsychologia*, **44**(10), 1928–1935. https://doi.org/10.1016/j.neuropsychologia.2006.01.034

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and language*, **115**(3), 149–161. https://doi.org/10.1016/j.bandl.2010.07.006

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain research*, **1146**, 75–84. https://doi.org/10.1016/j.brainres.2006.06.101

Guglielmi, V., Quaranta, D., Mega, I., Costantini, E. M., Carrarini, C., Innocenti, A., & Marra, C. (2020). Semantic priming in mild cognitive impairment and healthy subjects: effect of different time of presentation of word-pairs. *Journal Of Personalized Medicine*, **10**(3), 57. https://doi.org/10.3390/jpm10030057

Heinze, H.- J., Muente, T. F., & Kutas, M. (1998). Context effects in a category verification task as assessed by event- related brain potential (ERP) measures. *Biological Psychology*, **47**, 121–135. https://doi.org/10.1016/S0301-0511(97)00024-0

Huang, L., Chen, K.L., Lin, B.Y., Tang, L., Zhao, Q. H., Lv, Y.R. & Guo, Q.H. (2018). Chinese version of montreal cognitive assessment basic for discrimination among different severities of Alzheimer's disease. *Neuropsychiatric disease and treatment*, 2133–2140.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, **4**(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B. (2016). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-30. http://CRAN.Rproject.org/package = lmerTest

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of cognitive neuroscience*, **25**(3), 484–502. https://doi.org/10.1162/jocn_a_00328

Lenth, R.(2023). emmeans: Estimated Marginal Means, aka Least-Squares Means. R Package Version 1.8.4-1. https://CRAN.R-project.org/package = emmeans

León-Cabrera, P., Pagonabarraga, J., Morís, J., Martínez-Horta, S., Marín-Lahoz, J., Horta-Barba, A., . . . & Rodríguez-Fornells, A. (2021). Neural signatures of predictive language processing in Parkinson's disease with and without mild cognitive impairment. *Cortex*, **141**, 112–127.

Leyhe, T., Muller, S., Eschweiler, G. W., & Saur, R. (2010). Deterioration of the memory for historic events in patients with mild cognitive impairment and early Alzheimer's disease. *Neuropsychologia*, **48**(14), 4093–4101, https://doi.org/10.1016/j.neuropsychologia.2010.10.011.

Meteyard, L., & Davies, R. A. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, **112**, 104092.

Nebes, R. D., Brady, C. B., & Huff, F. J. (1989). Automatic and attentional mechanisms of semantic priming in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, **11**, 219–230.

Nutter-Upham, K. E., Saykin, A. J., Rabin, L. A., Roth, R. M., Wishart, H. A., Pare, N., & Flashman, L. A. (2008). Verbal fluency performance in amnestic MCI and older adults with cognitive complaints. *Archives Of Clinical Neuropsychology*, **23**(3), 2

Ober, B. A. (2002). RT and non-RT methodology for semantic priming research with Alzheimer's disease patients: a critical review. *Journal of Clinical and Experimental Neuropsychology*, **24**, 883911.

Ober, B. A., & Shenaut, G. K. (1995). Semantic Priming in Alzheimer's Disease: Meta-Analysis and Theoretical Evaluation. In P. A. Allen & T. R. Bashore (Eds.), *Age differences in word and language processing* (pp. 247–271). Amsterdam: Elsevier/North-Holland.

Paitel, E. R., Samii, M. R., & Nielson, K. A. (2021). A systematic review of cognitive event-related potentials in mild cognitive impairment and alzheimer's disease. *Behavioural Brain Research*, **396**, 112904. https://doi.org/10.1016/j.bbr.2020.112904

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of internal medicine*, **256**(3), 183–194.

Rinehardt, E., Eichstaedt, K., Schinka, J. A., Loewenstein, D. A., Mattingly, M., Fils, J., Duara, R., & Schoenberg, M. R. (2014). Verbal fluency patterns in mild cognitive impairment and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, **38**(1–2), 1–9. https://doi.org/10.1159/000355558

Roehm, D., Bornkessel-Schlesewsky, I., Roesler, F., & Schlesewsky, M. (2007). To predict or not to predict: influences of task and strategy on the processing of semantic relations. *Journal of Cognitive Neuroscience*, **19**(8), 1259–1274. https://doi.org/10.1162/jocn.2007.19.8.1259

Salmon, D. P., & Bondi, M. W. (2009). Neuropsychological assessment of dementia. *Annual review of psychology*, **60**(1), 257–282.

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: a tutorial. *Journal of memory and language*, **110**, 104038.29–241. https://doi.org/10.1016/j.acn.2008.01.005

Wagner, M., Baving, L., Berg, P., Cohen, R., & Rockstroh, B. (2006). An ERP investigation of semantic priming, repetition priming, and negative priming in schizophrenic patients. *Journal of Psychophysiology*, **20**(3), 195–211.

Yoon, C., Feinberg, F., Hu, P., Gutchess, A.H., Hedden, T., Chen, H.Y.M., ... & Park, D.C. (2004). Category norms as a function of culture and age: comparisons of item responses to 105 categories by American and Chinese adults. *Psychology and aging*, **19**(3), 379.

Zhu, J. S. (2014). *Chinese Antonym Dictionary*. Language & Culture Press.

Appendix A. The optimalization model for accuracy in ANT experiment

R model equation: ACC_ANT.opt = glmer(data = ACC,ACC~group*Material+(1|ID) +(1|item), family = 'binomial',control = glmerControl(optimizer = 'bobyqa',boundary.tol = 1e–04))

Fixed effect	Estimate	Standard error	95% CI	*Z* value	p
intercept	5.2563	0.526	4.235,6.276	10. 096	< 2e–16***
group1	2.5316	0.79	0.983,4.080	3.204	0.00135***
Material1	0.3653	0.6941	–0.995, 1.726	0.526	0.59864
group1:Material1	1.8586	1.19	–0.474,4.191	1.562	0.11833
Random effects					
Groups	Name	variance	SD		
item	intercept	1.383	1.176		
ID	intercept	1.350	1.162		
Analysis of Deviance Table (Type II Wald chisquare tests)					
	χ^2	df	p		
group	7.9478	1	0004815 **		
Material	0.7313	1	0.392449		
group:Material	2.4393	1	0.118331		
Model fit	Marginal		Conditional		
	0.235		0.582		

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

Appendix B. The optimization model for response time in ANT experiment

R model equation: RT_ANT.opt = lmer(data = RT, RT~group*Material+(1|item)+(1|ID), control = lmerControl(optimizer = 'bobyqa',boundary.tol = 1e–04))

Fixed effect	Estimate	Standard error	95% CI	*t* value	p
intercept	−0.07899	0.05085	−0.1786, 0.0206	−1.554	0.13276
group1	−0.22241	0.09986	−0.4181, − 0.0267	−2.227	0.03579*
Material1	0.07312	0.02621	0.02174, 0.1245	2.790	0.00714**
group1:Material1	−0.19425	0.03562	−0.26406, − 0.1244	−5.453	5.84e–08***
Random effects					
Groups	Name	variance	SD		
item	(Intercept)	0.005543	0.07445		
ID	(Intercept)	0.062743	0.25048		
Residual		0.117116	0.34222		
lmerTest *F*-values (Type III Analysis of Variance Table with Satterthwaite's method)					
	F value	df	p		
group	4.9601	24	0.035793*		
Material	7.7822	58	0.007136**		
group:Material	29.7362	1402	5.843e–08***		
Model fit	Marginal		Conditional		
	0.078		0.418		

Appendix C. The optimization model for N400 amplitude in ANT experiment

R model equation: N400_ANT.opt = lmer(data = N400, n400amp~group*material+(1+material|ID)+(1+group1|channel), control = lmerControl(optimizer = 'bobyqa'))

Fixed effect	Estimate	Standard error	95% CI	*t* value	p
intercept	1155.25	231.52	701.47, 1609.02	4.990	4.16e–05***
group1	−960.47	461.83	−1865.65, − 55.29	−2.080	0.048406*
Material1	277.54	108.74	64.43, 490.66	2.552	0.013352*
group1:Material1	−751.46	206.94	−1157.05, − 345.87	−3.631	0.000292***
Random effects					
Groups	Name	variance	SD		
ID	(Intercept)	16762	129.5		
channel	(Intercept)	1316802	1147.5		
Residual		4175330	2043.4		
lmerTest *F*-values (Type III Analysis of Variance Table with Satterthwaite's method)					
	F value	df	p		
group	4.3251	24	0.0484056*		
Material	6.5151	58	0.0133517*		
group:Material	13.1864	1474	0.0002916***		
Model fit	Marginal		Conditional		
	0.049		0.279		

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

Appendix D. The optimization model for response time in CAT experiment

R model equation: RT.opt3 = lmer(data = RT, zscore~group*Material+(1|ID)+(−1+group1|item), control = lmerControl(optimizer = ‘bobyqa’,boundary.tol = 1e−04))

Fixed effect	Estimate	Standard error	95% CI	*t* value	p
intercept	−9.903e−02	4.163e−02	−0.18061, − 0.01744	−2.379	0.02406***
group1	−2.958e−01	8.325e−02	−0.45895, − 0.13260	−3.553	0.00131**
Material1	4.145e−02	5.159e−02	−0.05965, 0.142568	0.804	0.42175
Material2	6.669e−04	5.835e−02	−0.11370, 0.11503	0.011	0.99088
group1:Material1	2.621e−01	1.032e−01	0.05983, 0.46428	2.540	0.01116*
group1:Material2	3.016e−01	1.167e−01	0.07288, 0.53035	2.584	0.00982**
Random effects					
Groups	Name	variance	SD		
item	group1	9.469e−16	3.077e−08		
ID	(Intercept)	3.874e−02	1.968e−01		
Residual		1.408e−01	3.752e−01		
lmerTest *F*-values (Type III Analysis of Variance Table with Satterthwaite’s method)					
	F value	df	p		
group	8.3101	24	0.00823**		
Material	2.7046	2083.36	0.06713		
group:Material	3.3920	2083.36	0.03383*		
Model fit	Marginal		Conditional		
	0.088				

Appendix E. The optimalization model for accuracy in CAT experiment

R model equation: ACC_CAT.opt = ACC.opt2 = glmer(data = ACC, ACC~group*Material+(−1+ material2|ID) +(1|item), family = ‘binomial’,control = glmerControl(optimizer = ‘bobyqa’, boundary.tol = 1e−04))

Fixed effect	Estimate	Standard error	95% CI	*Z* value	p
intercept	2.8734	0.2304	2.422, 3.325	12.471	< 2e−16***
group1	1.3110	0.3871	0.552, 2.069	3.387	0.000707***
Material1	−0.9178	0.7551	−2.397, 0.562	−1.215	0.224183
Material2	−0.1566	0.8740	−1.869, 1.556	−0.179	0.857830
group1:Material1	0.3480	1.2755	−2.152, 2.848	0.273	0.785002
group1:Material2	−0.1632	1.4714	−3.047, 2.721	−0.111	0.911694
Random effects					
Groups	Name	variance	SD		
item	intercept	0.3183	0.5642		
ID	material2	0.3232	0.5685		
Analysis of Deviance Table (Type II Wald chisquare tests)					
	χ^2	df	p		
group	53.1286	1	3.124e−13 ***		
Material	14.6832	2	0.000648 ***		
group:Material	1.3911	2	0.498802		
Model fit	Marginal		Conditional		
	0.134		0.210		

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

Appendix F. The optimization model for N400 amplitude in CAT experiment

R model equation: N400_CAT.opt = lmer(data = N400,n400amp~group*material+(1+material|ID)+(1+group1|channel), control = lmerControl(optimizer = 'bobyqa'))

Fixed effect	Estimate	Standard error	95% CI	*t* value	p
intercept	1.7547	0.7671	0.251, 3.258	2.288	0.0299*
group1	0.6115	1.5373	−2.402, 3.624	0.398	0.6938
Material1	5.6960	1.1217	3.497, 7.894	5.078	3.41e−05***
Materiasl2	7.3538	1.3446	4.719, 9.989	5.469	1.27e−05***
group1:Material1	−2.3229	2.2434	−6.719, 2.074	−1.035	0.3108
group1:material2	−2.8806	2.6892	−8.151, 2.390	−1.071	0.2947
Random effects					
Groups	Name	variance	SD	Corr	
ID	(Intercept)	13.0325	3.6101		
		23.3690	4.8342	−0.29	
		34.9571	5.9124	−0.27	0.94
channel	(Intercept)	0.4909	0.7007		
	Group1	2.0502	1.4318	−0.64	
Residual		6.2676	2.5035		
lmerTest *F*-values (Type III Analysis of Variance Table with Satterthwaite's method)					
	F value	df	p		
group	0.0007	28.424	0.9787		
Material	14.9824	24.000	5.987e−05***		
group:Material	0.5767	24.000	0.5693		
Model fit	Marginal		Conditional		
	0.051		0.714		

https://doi.org/10.1017/S0142716425100313 Published online by Cambridge University Press

Appendix G. The optimization model for N400 effect

R model equation: N400_diff.opt = lmer(data = diff, diff~group*experiment+(1+experiment|ID)+(−1+experiment1|channel), control = lmerControl(optimizer = ‘bobyqa’,boundary.tol = 1e–04))

Fixed effect	Estimate	Standard error	95% CI	*t* value	p
intercept	−3.5457	0.3570	−4.245, −2.846	−9.932	5.62e−10***
group1	−0.5974	0.7140	−1.996, 0.802	−0.837	0.4110
experiment1	1.8662	0.6982	0.498, 3.235	2.673	0.0123*
group1:experiment1	3.4508	1.2821	0.938, 5.964	2.691	0.0128*
Random effects					
Groups	Name	variance	SD	Corr	
ID	(Intercept)	3.1599	1.7776		
	Experiment1	10.0696	3.1733	0.05	
channel	Experiment1	0.6891	0.8301		
Residual		2.7697	1.6642		
lmerTest *F*-values (Type III Analysis of Variance Table with Satterthwaite’s method)					
	F value	df	p		
group	0.7001	24.000	0.41100		
experiment	7.1439	28.656	0.01228*		
group:experiment	7.2438	24.000	0.01275*		
Model fit	Marginal		Conditional		
	0.165		0.732		

Cite this article: Yang, J., Wang, J., & Huang, L. (2025). Semantic priming and ERP correlates of predictive processing in Chinese aMCI patients. *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716425100313