Bilingualism: Language and Cognition

cambridge.org/bil

Time course of indirect reply processing in native and non-native Mandarin speakers: An ERP study

Xiuping Zhang[1], Xiaoxi Pan[1], Yizhu Wang[1], Maoyao Xu[1] and Adam John Privitera[2,3]

[1]School of Psychology, Beijing Language and Culture University, Beijing, China; [2]Science of Learning in Education Centre, National Institute of Education, Nanyang Technological University, Singapore, Singapore and [3]Centre for Research and Development in Learning, Nanyang Technological University, Singapore, Singapore

Research Article

Cite this article: Zhang, X., Pan, X., Wang, Y., Xu, M. and Privitera, A.J. (2025). Time course of indirect reply processing in native and non-native Mandarin speakers: An ERP study. *Bilingualism: Language and Cognition* **28**, 589–600. https://doi.org/10.1017/S1366728924000695

Received: 26 October 2023
Revised: 9 August 2024
Accepted: 29 August 2024
First published online: 13 December 2024

Keywords:
Second language processing; Indirect Replies; Mandarin Chinese; Event-related Potentials (ERP); N400/ LPC/delayed positivity

Corresponding author:
Xiuping Zhang;
Email: zhangxp@blcu.edu.cn

Abstract

To communicate successfully, listeners must decode both the literal and intended meanings of a speaker's message. This ability is especially crucial when processing indirect replies as intended meanings can differ significantly from what was said. How native and non-native speakers differ in this ability is an open question. The present study investigated differences in the time course of indirect reply processing in native and non-native Mandarin speakers. EEG signals were recorded while participants were presented with conversations that differed in their directness. For indirect replies, native speakers exhibited a larger left anterior N400 and posterior late positive component (LPC). Conversely, non-native speakers exhibited a larger left-distributed LPC and delayed LPC. Findings support that non-native speakers exhibit delayed processing of indirect replies, potentially because of cognitive resource limitations. Findings from the present study have implications for a broad range of investigations on human communication and second language processing.

Highlights

- Native speakers exhibited a larger left anterior N400 and posterior LPC.
- Non-native speakers exhibited a larger left-distributed LPC and delayed LPC.
- Non-native speaker's exhibit delayed processing of indirect replies.
- The processing delay may be because of cognitive resource limitations.

1. Introduction

To achieve successful communication, listeners must decode not only the literal meaning of a statement but also the intended meaning conveyed by the speaker, a process requiring pragmatic inference. This process is essential for communication as the literal meaning of a statement might differ significantly from its intended meaning. Considering the following example—A: Is the teacher in his office now? B: The teacher's car is outside the office. In this conversation, Speaker B's reply is not directly relevant to Speaker A's question. However, according to Grice's theory of conversational implicature (Grice, 1975), interlocutors adhere to cooperative principles during conversation, including the principle of relevance maximization. Thus, Speaker A might assume that Speaker B is trying to give a relevant reply (i.e., the teacher is in his office now), and infer Speaker B's intention beyond the literal meaning of the statement.

1.1. Processing of indirect replies

Relevance theory (Sperber & Wilson, 1986) posits that recognition of a speaker's intention is the core of human communication, and that inference plays a key role in understanding intention. To understand a speaker's intended meaning, listeners must first construct the explicature (what is explicitly said) and then derive further implicatures (additional meanings implied by what is said). Both the construction of explicatures and the derivation of implicatures are driven by relevance principles, that is, searching for an optimally relevant interpretation in a given context, with as little processing effort as possible. Processing effort is related to contextual cues, that is, the number and intensity of contextual signals available for interpretation. Interpreting the intended meaning of an indirect reply requires contextual disambiguation and deriving implicit meanings conveyed by the speaker in a given context. When people interpret a statement, given the lexical retrieval from memory and contextual information, many different assumptions from various sources come to mind. From these assumptions, listeners select the most immediately relevant interpretation with the greatest contextual relevance and the least processing effort.

CAMBRIDGE UNIVERSITY PRESS

https://doi.org/10.1017/S1366728924000695 Published online by Cambridge University Press

A range of diverse research designs have been used to study the processing of indirect replies. Using a priming paradigm, Holtgraves (1999) investigated how listeners make sense of indirect replies, demonstrating that both literal and intended meanings were activated in conversation and that pragmatic inference was involved during processing. Neuroimaging work utilizing functional magnetic resonance imaging (fMRI) extends this finding, supporting that detection of literal irrelevance and inferring a speaker's indirect meaning are also involved (Bašnáková et al., 2014; Feng et al., 2017; Feng et al., 2021; Jang et al., 2013; Shibata et al., 2011; van Ackeren et al., 2016; Zhang, Xu et al., 2023). Several neurophysiological studies using electroencephalography (EEG) to measure event-related potential (ERP) during indirect reply processing reported indirect effects both during early and late processing stages (Guo et al., 2023; Zhang et al., 2021; Zhang, Pan et al., 2023). In their investigation of the role of working memory (WM) capacity during indirect reply processing, Zhang et al. (2021) reported that, in high WM span readers, indirect replies elicited earlier ERP effects (i.e., P200 and P300), while in low span readers, only a delayed late positive component (LPC) effect was observed.

While the studies described above focused exclusively on processing indirect replies in a listener's native language, processing in a second language (L2) has also garnered considerable attention. In a series of studies, Taugchi investigated L2 indirect reply processing, focusing on the relationship between contextual cues, processing load, and conventionality (e.g., Taguchi, 2005, 2008a, 2008b, 2011; Taguchi et al., 2016). Taugchi's studies used offline comprehension judgment measures to investigate the comprehension of L2 indirect replies, which cannot provide insight into online processing. This limit can be addressed through the measurement of high temporal resolution ERPs, which can identify the time course of indirect reply comprehension during online processing. Additionally, while processing in L1 and L2 have both been investigated, little of this work has focused on how language influences the online processing of indirect replies.

1.2. Differences in pragmatic processing between native and non-native speakers

Both theoretical (Clahsen & Felser, 2006a, 2006b, 2018; Morishima, 2013) and empirical studies on the related pragmatic phenomena (e.g., Foucart, Moreno et al., 2015; Martin et al., 2013; Newman et al., 2012; Pérez et al., 2019) have indicated that pragmatic processing differs between L1 and L2. Two hypotheses posit that pragmatic processing differs between native speakers and non-native speakers. The shallow parsing hypothesis states that differences can be attributed to the use of different parsing mechanisms. Relative to natives, non-native speakers employ a "shallow" parsing approach and are more sensitive to semantic and pragmatic information compared with syntactic information, making efficient use of various non-grammatical information for complex language comprehension (Clahsen & Felser, 2006a, 2006b, 2018). The limited cognitive resources hypothesis assumes that L1/L2 processing is similar and that differences can be explained in terms of capacity-based limitations being exceeded during L2 processing (Morishima, 2013). For native speakers, lower-level language processes such as lexical access and parsing are automatized and are thought to consume few cognitive resources. Thus, native speakers are able to allocate their cognitive resources to high-level processing such as discourse comprehension (Harrington & Sawyer, 1992; Perfetti, 1985). In contrast, for non-native speakers, the construction of surface form and text base are cognitively demanding. As a consequence, relatively fewer resources are left for the construction of higher discourse-level representations, such as semantic and pragmatic representations (Rai et al., 2011).

Findings from previous neurophysiological studies also provide mixed support for these two hypotheses. Results from studies on sentence and discourse comprehension have found that, for non-native speakers, pragmatic processing elicited earlier ERP effects (i.e., N400) (Foucart, Garcia et al., 2015; Foucart et al., 2014; Foucart et al., 2016). This finding is thought to reflect non-native speakers' increased sensitivity for and reliance on contextual cues for pragmatic processing. However, other studies reported a delayed or absent N400 effect during pragmatic processing in non-native speakers (Foucart, Moreno et al., 2015; Martin et al., 2013; Newman et al., 2012; Pérez et al., 2019), attributed to the reduced cognitive resources available for non-native speakers during L2 processing, biasing them toward a reliance on the less demanding reactive control mode.

1.3. The present study

While a number of studies have investigated differences between native and non-native speakers in pragmatic and discourse processing, whether differences can be identified during online indirect reply comprehension is an open question. Using EEG to measure ERPs, the present study aimed to investigate this question in a sample of native and non-native Mandarin speakers. Direct and indirect reply conversations were written in Chinese, and native and non-native speakers were asked to read each conversation while EEG signals were recorded.

Several relevant ERP components, including P200, N400, LPC, and late negativity, have been reported in previous studies investigating figurative and non-literal language comprehension. The P200 component is thought to reflect initial semantic analysis during pragmatic processing (Regel & Gunter, 2017; Regel et al., 2011; Regel et al., 2014), while the N400 has been associated with the difficulty of integrating current information with prior contexts (Coulson & Van Petten, 2002; Pynte et al., 1996; Weiland et al., 2014). Finally, both the LPC and late negativity have been related to pragmatic reanalysis of mismatched information (Arzouan et al., 2007; Coulson & Van Petten, 2002; Goldstein et al., 2012; Regel & Gunter, 2017; Regel et al., 2011; Spotorno et al., 2013; Weiland et al., 2014; Zhang et al., 2021; Zhang, Pan et al., 2023). Based on previous ERP studies investigating indirect reply processing (Guo et al., 2023; Zhang et al., 2021; Zhang, Pan et al., 2023), we predict that native speakers may proactively process indirect replies, thus both earlier (e.g., P200/N400) and later components (e.g., LPC/late negativity) may be elicited for indirect replies compared with direct replies. For non-native speakers, we predict that heightened sensitivity to pragmatic information, as supported by the shallow parsing hypothesis (Clahsen & Felser, 2006a, 2006b, 2018), will lead to indirect reply effects in earlier ERP components (e.g., P200 and N400). Alternatively, based on the limited cognitive resources hypothesis (Morishima, 2013), indirect reply processing may be delayed because of the demands imposed by L2 use, presenting indirect reply effects in the LPC or late negativity.

2. Methods

2.1. Participants

Thirty-two native Mandarin speakers (16 females, M_{age} = 22.22, SD_{age} = 2.21), and 30 non-native Mandarin speakers (14 females, M_{age} = 24.37, SD_{age} = 3.47) participated in this experiment. Non-native speakers were proficient users of Mandarin who had begun studying the language as young adults (M_{age} = 19.55, SD_{age} = 2.50), continued studying the language for an average of 57.83 months (*SD* = 26.27), and passed either the HSK 5 or HSK 6 examination. Additionally, non-native speakers had spent an average of 36.58 months living in China (*SD* = 25.84). Non-native Mandarin speakers reported native languages of Korean (9), Urdu (6), Bengali (4), Russian (3), Samoan (2), Spanish (2), Indonesian (1), Kazakh (1), Nepali (1), and Swahili (1). All participants were undergraduate or graduate students, right-handed, with normal or corrected-to-normal vision, reporting no reading, neurological, or psychiatric disorders. Written informed consent was collected from all participants before the start of the study in accordance with the Declaration of Helsinki. Participants received 100 RMB (~$14 USD) for their participation. Approval for this study was granted by the ethics committee of the Beijing Language and Culture University.

Mandarin proficiency of non-native speakers was assessed before the start of the experiment using a literacy test based on Jiang and Liu (2004). During the test, 100 high-frequency Chinese characters were presented, and participants were asked to provide the correct *pinyin* for each character. Objective proficiency was operationalized as the total number of correct responses given out of 100. The average objective proficiency for non-native speakers was 68.64 (*SD* = 14.24). Participants were also asked to subjectively rate their Mandarin proficiency from 1 to 7 with 1 representing "not proficient at all" and 7 representing "very proficient." The average subjective proficiency for non-native speakers was 4.77 (SD = 0.73).

2.2. Design and materials

A 2 × 2 mixed experimental design was adopted with Group (Mandarin L1 and Mandarin L2) as the between-subjects factor and Reply (Direct and Indirect) as the within-subjects factor. Seventy-four written conversations were selected through pre-experimental screening sessions (see below). Each conversation included two speakers (Speaker A and Speaker B), one question (asked by Speaker A), and one reply (given by Speaker B). Speaker B's reply varied as either direct or indirect depending on Speaker A's question, resulting in two conditions (see Table 1). All conversations conveyed neutral information and represented everyday topics commonly discussed within Chinese culture.

2.2.1 Experimental material pretests

Experimental conversation stimuli were selected based on a series of pretests that assessed for the potentially confounding influence of indirectness, emotionality, cloze probability, and comprehensibility. Participants in each pretest did not take part in any other pretest or the formal experiment. Average scores or ratings for all measures of interest are reported in Table 2.

The first pretest was conducted to assess the level of indirectness for each written conversation to select the most appropriate experimental stimuli. Eighteen participants took part in the pretest. We generated 94 written conversation stimuli and divided them into two sets. Conversations in each set were rated by nine participants. During this pretest, participants were asked to read carefully and to rate how directly the given reply answered the question on a 7-point scale (1: most direct; 7: most indirect). If the reply was indirect (scores between 5 and 7), participants were asked to provide their interpretation of the speakers' intended meaning. We coded participants' interpretations as either 0 (did not identify what the speaker meant) or 1 (correctly identified what the speaker meant). Experimental conversations used in the formal study were selected according to the following criteria: ratings for direct replies ≤2.5; ratings for indirect replies >4; and, comprehension accuracy >77%. A total of 74 conversations were selected for the formal experiment. Indirect replies were rated as significantly more indirect than direct replies as revealed by a paired-samples *t*-test, $t(73) = 46.8$, $p < .001$.

To test for differences in stimulus emotionality, another pretest was conducted. Twenty participants were asked to rate the valence and arousal conveyed by each written conversation. Valence was rated on a scale from −3 (most negative) to 3 (most positive), and arousal was rated on a scale from 1 (least arousing) to 7 (most arousing). Paired *t*-tests revealed no significant differences between direct replies and indirect replies for both valence, $t(73) = -0.248$, $p = .805$, and arousal ratings, $t(73) = 1.14$, $p = .261$.

Next, we tested for differences in cloze probability in a sample of 18 participants. For all conversations, the third part of the target reply (underlined in Table 1) was removed, and participants were asked to complete the reply by providing the first word or phrase that came to their mind. No significant difference in cloze probability was identified between direct and indirect reply conditions, $t(73) = -0.93$, $p = .356$.

Finally, we tested for differences in the comprehensibility of Speaker A's question. Questions were presented randomly and 16 participants were asked to rate the degree of comprehensibility on a 7-point scale (1: most incomprehensible; 7: most comprehensible). No significant difference in question comprehensibility was found between direct and indirect reply conditions, $t(73) = 0.455$, $p = .650$.

2.3. Procedure

Data collection occurred in a sound-attenuating, shielded room at Beijing Language and Culture University. Participants were seated in a comfortable chair 60 cm in front of a computer screen (43.18 cm) on which all experimental stimuli were presented. Experimental trials were presented using E-Prime 2.0 (Psychology Software Tools, Pittsburg, PA, USA). Each trial began with the presentation of a central fixation cross (+) for 1000 ms.

Table 1. Example of experimental stimuli

Condition	Speaker A	Speaker B	Comprehension judgment
Direct reply	李明平时都做什么呢? What does Li Ming usually do?	他/每天/<u>都在工作</u>/。 **He/every day/<u>keeps working</u>/.** He keeps working every day.	Li Ming works every day.
Indirect reply	李明是个有意思的人吗? Is Li Ming an interesting person?		Li Ming is not an interesting person.

Note. Critical sentences (bolded) were divided into three parts with slashes, with the third part as the target segment (underlined). All materials were originally in Chinese. English versions shown above were translated word-by-word for illustrative purposes only.

https://doi.org/10.1017/S1366728924000695 Published online by Cambridge University Press

Table 2. Pretests results of the stimuli

Condition	Indirectness	Comp. Accuracy	Valence	Arousal	Cloze Prob.
Direct Reply	1.55(0.42)	—	0.16(0.75)	3.91(0.65)	24.8% (33.2%)
Indirect Reply	5.20(0.55)	91%(11%)	0.14(0.70)	3.99(0.62)	22.2% (30.2%)

Note. Data are expressed as mean (standard deviation). Comp: comprehension, Prob: probability

Immediately after, Speaker A's question was presented. Participants were asked to press the space bar once they had finished reading the question to advance to the next screen. Next, Speaker B's reply (critical sentence) was presented as three individual segments (see Table 1), each on a separate screen. The first two segments were presented for 500 ms each, followed by a blank screen for 200 ms. The third segment (target segment) was presented for 600 ms, followed by a period (500 m), and then a blank screen for 200 ms. The third segment in the critical sentence was crucial for deriving the intended meaning of the indirect replies. Accurate comprehension required not only processing the new information contained in the third segment but also contextualizing it based on information contained in the two previous segments and Speaker A's initial question. To enhance comprehension of the meanings conveyed in experimental conversations, the duration of the third segment was extended longer than the preceding two segments in accordance with practices reported in previous studies (Zhang et al., 2021, 2023; Guo et al., 2023).

The 74 experimental conversations consisted of 37 different conversations written using both a direct or indirect reply, resulting in 74 conversations in total. Experimental conversations were divided into two sets ensuring that each conversation was only presented once per set, and sets were counterbalanced across participants and groups. Both sets contained 37 conversations for each condition. In the direct reply condition, Speaker A first provided a what/why/where question, while a yes/no question was given in the indirect condition. To balance the question type and to prevent participants from generating predictions of particular reply types based on Speaker A's question, we added 60 filler conversations using the opposite question type, that is, yes/no questions followed by direct replies and what/why/where questions followed by indirect replies. This resulted in 134 total conversations for each set of stimuli. After reading each conversation, participants were asked to complete a comprehension judgment task (yes or no judgment based on whether they understood the written conversation). Participants were instructed to press "*J*" for "yes" and "*F*" for "no." Trials on which participants indicated they did not understand the conversation were removed before analysis.

Experimental stimuli were divided into two blocks, each containing 67 conversations. Each block took around 10 minutes to complete, and participants were given a short break between blocks. Conversations were presented in a pseudo-random order to ensure that stimuli from a given condition were not presented more than three times in succession. Before the formal experiment, participants completed six trials as a practice to familiarize themselves with the procedure. Practice stimuli were different from those used in the formal experiment, but were identical in structure.

Accuracy and reaction time data from behavioral trials were analyzed using R software (version 4.2.0; R Core Team, 2023). To test for main effects and interactions, repeated measures ANOVA with Group (Mandarin L1 and Mandarin L2) as the between-subjects factor, Reply (Direct and Indirect) as the within-subjects factor, and the interaction between Group and Reply were used. Significant interactions were followed up with appropriate *t*-tests for pairwise comparison.

2.4. Electrophysiological recording and analysis

Electrophysiological data were recorded with 64 Ag/AgCl electrodes mounted on an elastic cap (10–20 system), using a SynAmps amplifier and Curry 7 Acquisition Software (Neuroscan Inc, USA). Data were collected at a sampling rate of 500 Hz with a bandpass filter of 0.05–100 Hz. Electrodes were referenced online to the left mastoid and re-referenced offline to the average of both left and right mastoids. An electrode between Fz and FPz electrodes served as the ground. Horizontal electrooculograms (HEOG) were monitored through two electrodes placed at the outer canthus of each eye. Vertical electrooculograms (VEOG) were recorded via two electrodes above and below the left eye. Electrode impedances were maintained below 5 kΩ during the whole experiment.

EEG data were preprocessed using EEGLAB (EEGLAB 14.1.1b, http://www.sccn.ucsd.edu/eeglab), an open-source MATLAB toolbox (Delorme & Makeig, 2004). Continuous EEG data were bandpass filtered (0.05–30 Hz) and segmented from 200 ms (−200 ms) before the onset of the third segment of the critical sentence to 1300 ms after the onset of the third segment of the critical sentence. The mean amplitude from −200 to 0 ms served as a baseline for each trial. Independent component analysis (ICA) was applied to epoched data to isolate and remove components representing ocular or other artifacts. Trials contaminated with artifacts (excessive amplitude exceeding ±80 μV) were removed. For each participant, average ERPs were calculated for each trial condition at each electrode. Grand average ERPs for each group and condition were calculated using all included trials.

For ERP statistical analysis, cluster-based random permutation tests were conducted (implemented in FieldTrip, Maris & Oostenveld, 2007). This nonparametric statistical procedure uses a cluster method to correct for multiple comparisons, effectively controlling for the overall Type I error rate when no prior information is available to guide the choice of latency windows and electrode sites (Luck & Gaspelin, 2017). The permutation test was performed within 0–1300 ms post-the third segment onset (in 2-ms steps) over 60 electrodes (FP1/FP2, FPZ, AF3/AF4; F7/F8, F5/F6, F3/F4, F1/F2, FZ; FT7/FT8, FC5/FC6, FC3/FC4, FC1/FC2, FCZ; T7/T8, C5/C6, C3/C4, C2/C1, CZ; TP7/TP8, CP5/CP6, CP3/CP4, CP2/CP1, CPZ; P7/P8, P5/P6, P3/P4, P1/P2, PZ; PO7/PO8, PO5/PO6, PO3/PO4, POZ; O1/O2, OZ). A simple dependent *t*-test was performed on each data point ("electrode by time") comparing two conditions (e.g., Direct vs. Indirect reply). All adjacent data points exceeding a preset significance level (p = .05) were grouped into clusters. Cluster-level statistics were calculated by taking the sum of the *t*-values for every cluster. The significance probability of the clusters was calculated using the Monte Carlo method with 1000 random draws.

3. Results

3.1. Behavioral results

Average comprehension rates and response times for direct and indirect replies are reported in Table 3. For the Mandarin L1 group, average comprehension was 96.45% (*SD* = 3.08%), indicating that experimental conversations were well understood. Average comprehension for the Mandarin L2 group was 84.36% (*SD* = 4.05%), indicating that participants understood most of the conversations used in the present study. Statistical analysis of comprehension data revealed significant main effects of Group [$F(1, 56) = 103.14$, $p < .001$] and Reply [$F(1, 56) = 118.49$, $p < .001$], as well as a significant interaction between Group and Reply [$F(1, 56) = 58.70$, $p < .001$]. Pairwise comparisons revealed that, while comprehension of indirect replies was lower than that of direct replies for both the Mandarin L1, $t(31) = 7.72$, $p < .001$, and Mandarin L2 groups, $t(26) = 36.37$, $p < .001$, overall comprehension was lower in the Mandarin L2 group (Indirect: $t(57) = 10.44$, $p < .001$; Direct: $t(57) = 5.42$, $p < .001$).

Before analysis, reaction time data underwent a logarithmic transformation to address issues with non-normality. Analyses revealed significant main effects of Reply [$F(1, 56) = 39.66$, $p < .001$] and Group [$F(1, 56) = 121.55$, $p < .001$], indicating that participants' responses to questions following indirect replies were faster compared with direct replies, and that the Mandarin L1 group faster than the Mandarin L2 group. The interaction between Reply and Group was not significant [$F(1, 56) = 0.038$, $p = .846$].

3.2. ERP results

Because of issues with EEG quality, data for one participant from the Mandarin L1 and three participants from the Mandarin L2 group were removed before analysis. This resulted in the inclusion of 31 participants for the Mandarin L1 group and 27 participants for the Mandarin L2 group. Pre-analysis trial removal for comprehension and EEG artifact contamination issues resulted in the inclusion of an average of 35 direct reply and 34 indirect reply condition trials for the Mandarin L1 group, and 30 direct reply and 24 indirect reply condition trials for the Mandarin L2 group. Figures 1 and 2 display grand average ERP waveforms generated by the critical segment across different conditions and groups. Topographic distributions of the indirect reply effect in both groups are shown in Figure 3.

For the Mandarin L1 group, compared with direct replies, indirect replies elicited a larger negativity from 314 to 500 ms after critical segment onset (negcluster, $p = .007$), and a larger positivity in the 692 to 1300 ms time window (poscluster, $p = .047$). The negativity effect was over anterior regions (electrodes: FP1, FPz, FP2, AF3, AF4, F7, F5, F3, F1, Fz, FT7, FC5, FC3, and FC1), with a left hemisphere dominance, whereas the positivity effect was over posterior regions (electrodes: P7, P5, P3, P1, Pz, P4, P6, P8, PO7, PO5, PO3, POz, PO4, PO6, PO8, O1, Oz, and O2).

Table 3. Mean (SD) comprehension and reaction time (RT) for replies

	Comprehension		RT	
	Direct	Indirect	Direct	Indirect
Native	98.46%	94.23%	1290.73	1132.77
	(2.87%)	(4.97%)	(725.02)	(746.58)
Non-native	92.04%	73.27%	3183.39	2824.12
	(6.06%)	(9.02%)	(1525.86)	(1665.21)

For the Mandarin L2 group, indirect replies, relative to direct replies, elicited a larger positivity in the 522–966 ms time window (poscluster 1, $p = .029$), and in the time window 984–1176 ms (poscluster 2, $p = .043$). For the time window from 522 to 966 ms, the positivity effect was mainly distributed in central and posterior regions in the left hemisphere and the midline (electrodes: C1, C3, C5, Cz, T7, TP7, CP5, CP3, CPz, P7, P5, P3, P1, Pz, PO7, PO5, PO3, POz, O1, and Oz). For the time window from 984 to 1176 ms, the positivity effect was distributed across the whole cortex (electrodes: FP1, FPz, FP2, F4, FC3, FC5, FCz, C5, C3, C1, Cz, TP7, CP5, CP3, CP1, CPz, CP2, CP4, P7, P5, P3, P1, Pz, PO7, PO5, PO3, POz, and O1).

Given that the present study aimed to examine differences between native and non-native speakers during indirect reply comprehension, we conducted cluster-based permutation tests to examine whether Reply and Group interacted in the significant time windows identified during our initial ERP analysis. Accordingly, amplitudes at significant electrodes for each time window were averaged and 2 × 2 repeated-measure ANOVAs were performed, followed by pairwise comparisons to probe significant interactions and main effects.

For the time window from 314 to 500 ms (N400), a significant interaction between Group and Reply was identified [$F(1, 56) = 5.86$, $p = .023$, $\eta_p 2 = .184$]. Pairwise comparisons revealed that, for the Mandarin L1 group, N400 amplitude of indirect replies was significantly higher than direct replies, $t(30) = 3.64$, $p = .001$. However, this finding was not observed in the Mandarin L2 group, $t(26) = 1.09$, $p = .286$.

For the time window from 522 to 966 ms (LPC 1), no significant interaction between Group and Reply was identified [$F(1, 56) = 1.00$, $p = .327$, $\eta_p 2 = .037$]. Main effects of Group [$F(1, 56) = 45.14$, $p < .001$, $\eta_p 2 = .635$] and Reply [$F(1, 56) = 5.00$, $p = .034$, $\eta_p 2 = .161$] were significant. Pairwise comparisons showed that for the Mandarin L1 group, LPC amplitude for indirect and direct replies did not differ, $t(30) = 0.49$, $p = .627$. Conversely, LPC amplitude of indirect replies was higher than that of direct replies in the Mandarin L2 group, $t(26) = 2.43$, $p = .022$.

For the time window from 692 to 1300 ms (LPC 2), no significant interaction between Group and Reply was identified [$F(1, 56) = 0.360$, $p = .554$, $\eta_p 2 = .014$]. However, main effects of Group [$F(1, 56) = 4.51$, $p = .043$, $\eta_p 2 = .148$] and Reply [$F(1, 56) = 7.67$, $p = .010$, $\eta_p 2 = .228$] were significant. Pairwise comparisons revealed that, for the Mandarin L1 group, LPC amplitude of indirect replies was higher than that of direct replies, $t(30) = 2.10$, $p = .044$. However, this finding was not observed in the Mandarin L2 group, $t(26) = 1.50$, $p = .146$.

Finally, for the time window from 984 to 1176 ms (delayed LPC), no significant interaction between Group and Reply was identified [$F(1, 56) = 0.76$, $p = .391$, $\eta_p 2 = .028$]. The main effect of Group was significant [$F(1, 56) = 29.39$, $p < .001$, $\eta_p 2 = .531$], but the main effect of Reply was not [$F(1, 56) = 3.10$, $p = .090$, $\eta_p 2 = .106$]. Pairwise comparisons revealed that, for the Mandarin L1 group, the delayed LPC amplitude of indirect replies was not different from that of direct replies, $t(30) = 0.53$, $p = .599$. However, the delayed LPC amplitude of indirect replies was larger than that of direct replies in the Mandarin L2 group, $t(26) = 2.12$, $p = .043$.

In summary, we found that the interaction between Group and Reply was significant at the N400 time window, but not in the time windows of LPCs. Complete statistical test results can be found in Table 4.

https://doi.org/10.1017/S1366728924000695 Published online by Cambridge University Press

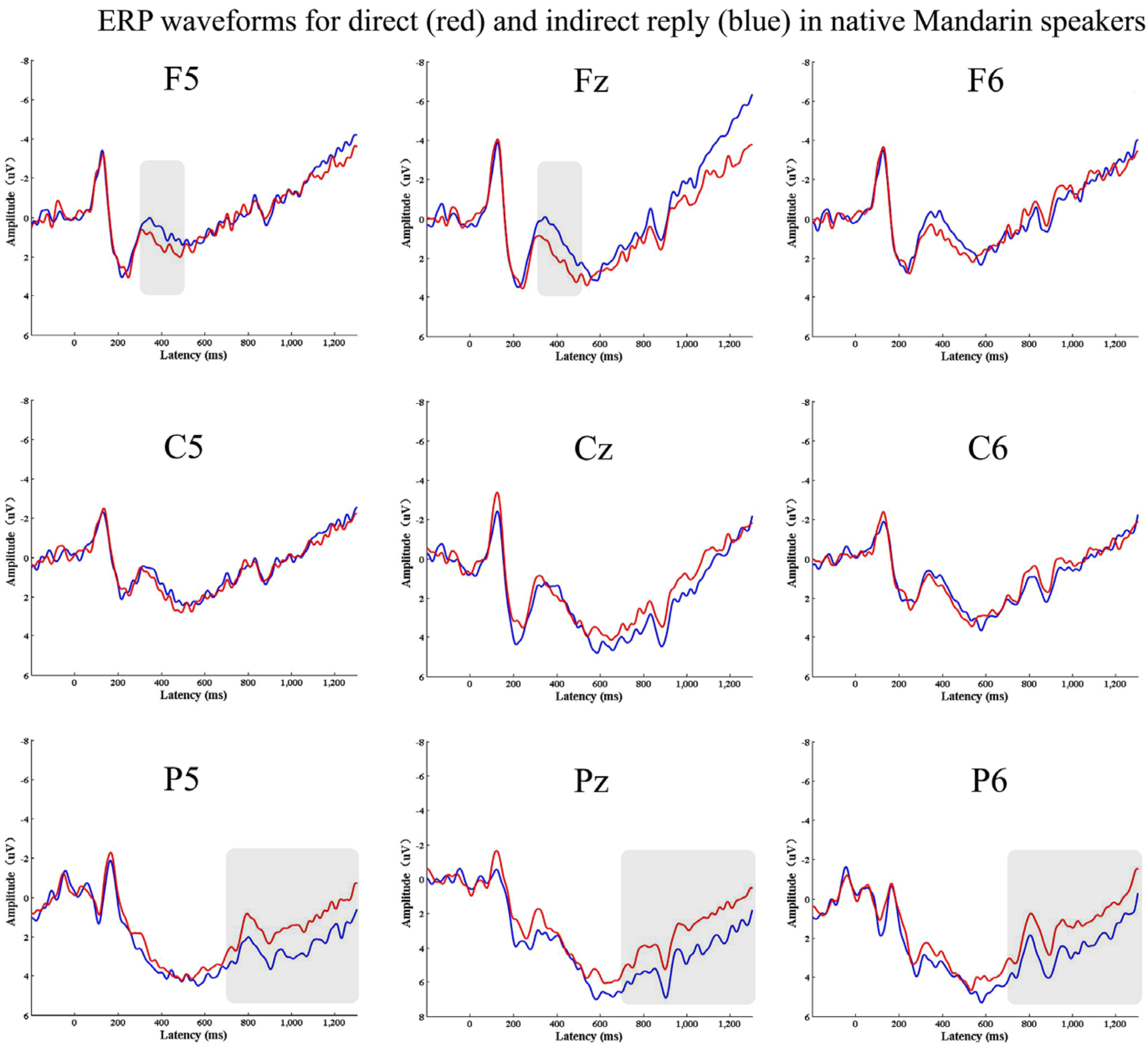

Figure 1. Grand average waveforms elicited by direct (red line) and indirect reply (blue line) conditions in native Mandarin speakers. Nine electrodes are presented for nine regions: left-anterior (F5), left-central (C5), left-posterior (P5), middle-anterior (Fz), middle-central (Cz), middle-posterior (Pz), right-anterior (F6), right-central (C6), and right-posterior (P6).

4. Discussion

In the present study, we investigated differences between native and non-native Mandarin speakers in the processing of indirect replies in written Mandarin conversations. Results showed that for native Mandarin speakers, indirect replies elicited both higher left anterior N400s and higher posterior LPCs than direct replies, whereas non-native speakers presented with higher globally distributed LPCs and higher delayed LPCs. These results support that native and non-native Mandarin speakers differ in the time course of indirect reply comprehension.

4.1. Processing of indirect replies by native Mandarin speakers

For native speakers, a left anterior-distributed N400 effect was found during indirect reply comprehension. In previous studies on figurative language comprehension, N400 effects have been associated with contextual integration of pragmatic meanings (Bambini et al., 2019; Coulson & Van Petten, 2002; Lai & Curran, 2013; Lai et al., 2009; Pynte et al., 1996; Weiland et al., 2014), the retrieval of semantic information and construction of action meaning (Amoruso et al., 2013; van Elk et al., 2010), and the degree to which WM is taxed (for reviews, see Coulson, 2004). In the present study, the anterior N400 effect for indirect replies could be associated with retrieving semantic information from memory and contextualizing it to immediately construct meaning. Thus, the anterior N400 effect of indirect replies could be linked to the immediate integration of the statement with context and world knowledge in a non-literal way for relevant interpretations.

The anterior-distributed negativity effect has also been established as an indicator of prediction generation and maintenance

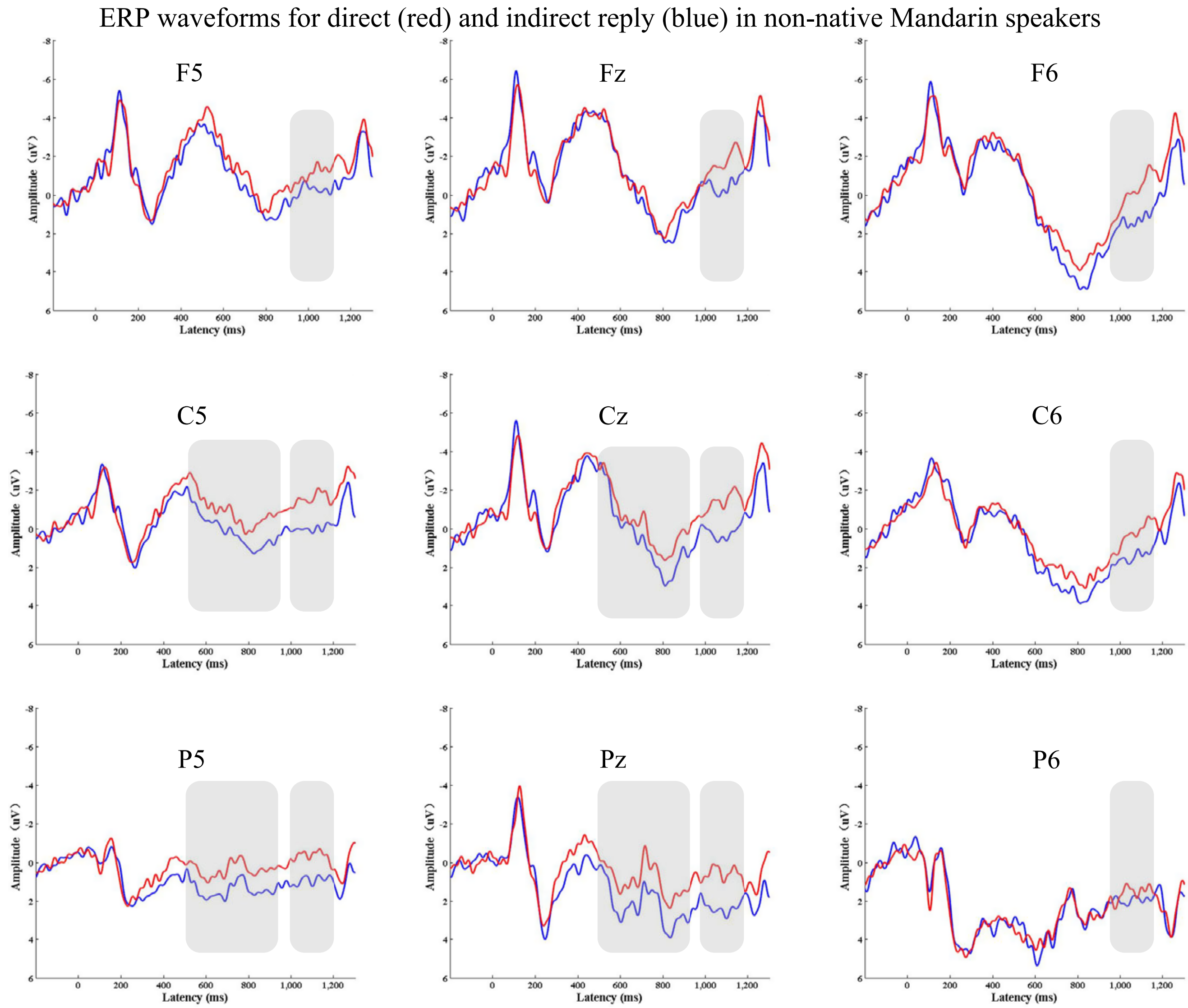

Figure 2. Grand average waveforms elicited by direct (red line) and indirect reply (blue line) conditions in non-native Mandarin speakers. Nine electrodes are presented for nine regions: left-anterior (F5), left-central (C5), left-posterior (P5), middle-anterior (Fz), middle-central (Cz), middle-posterior (Pz), right-anterior (F6), right-central (C6), and right-posterior (P6).

(Boux et al., 2021; Grisoni et al., 2019; Li et al., 2017; Otten & Van Berkum, 2008). In the present study, the observed negativity effect might be related to prediction processing for indirect replies. Native speakers may proactively process upcoming information during conversations, generating predictions about what a speaker is likely to say (Grüter & Rohde, 2013; Grüter et al., 2014; Pérez et al., 2019). During the processing of indirect replies, this tendency to make predictions may support more efficient detection of the literal irrelevance of a statement in a given context, consequently generating an early effect for indirect reply comprehension.

Following the N400 effect, an enhanced LPC was observed for indirect replies. Previous studies have linked the LPC effect to pragmatic analysis (Coulson & Van Petten, 2002; Regel et al., 2010; Regel & Gunter, 2017; Regel et al., 2011; Spotorno et al., 2013; Weiland et al., 2014) and efforts in constructing the situation model (Coulson & Lovett, 2010; Leuthold et al., 2012). For indirect replies, readers may combine the upcoming information with previous contexts and world knowledge to make pragmatic inferences and reanalyze the intended meaning conveyed by a speaker. Thus, the late positivity effect might reflect that a detailed and elaborate situation model was built for indirect replies.

The biphasic N400-LPC pattern observed in the present study has been reported in previous investigations of figurative language in the comprehension of metaphors (Coulson & Van Petten, 2002; Weiland et al., 2014), jokes (Coulson & Kutas, 2001), or unfamiliar irony (Filik et al., 2014). The N400 effect possibly indicates that non-literal language comprehension requires more semantic efforts than literal language comprehension. For non-literal language comprehension, it may be necessary to retrieve additional semantic information from contexts or long-term memory, leading to a larger N400 (Brouwer et al., 2012; Coulson & Van Petten, 2002). The LPC effect could also reflect pragmatic processing to derive non-literal meaning (Coulson & Van Petten, 2002; Filik et al., 2014; Weiland et al., 2014). For non-literal language comprehension, readers may expend greater efforts to integrate the retrieved

https://doi.org/10.1017/S1366728924000695 Published online by Cambridge University Press

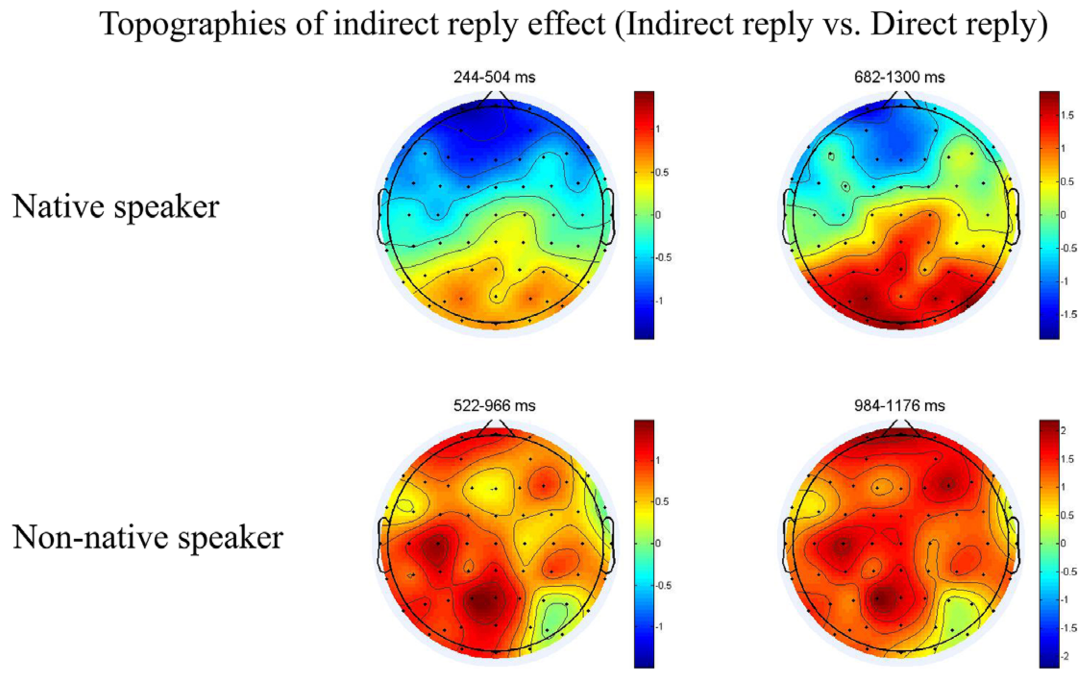

Figure 3. Topographies of indirect reply effect (Indirect reply vs. Direct reply) in native (upper) and non-native (lower) Mandarin speakers.

Table 4. Statistical analysis of ERP data at the reported time windows

	Group*Reply	MND L1 (IN-DI)	MND L2 (IN-DI)	
Time windows (ms)	*F* (1, 56)	*t* (30)	*t* (26)	Scalp distributions
N400 (314–500)	5.86*	−3.64**	1.09	Left anterior
LPC 1 (522–966)	1.00	0.491	2.43*	Left central/posterior and Midline
LPC 2 (692–1300)	0.36	2.10*	1.50	Posterior
Delayed LPC (984–1176)	0.76	0.531	2.12*	Whole-brain distributed

Note. "IN-DI" represents the difference between indirect replies (IN) and direct replies (DI). MND: Mandarin ** $p < 0.01$; * $p < 0.05$.

information into mental representations than during literal language comprehension.

4.2. Processing of indirect replies by non-native Mandarin speakers

For non-native speakers, an enhanced LPC mainly distributed in left-central and left-posterior regions, and an enhanced delayed positivity were observed during indirect reply processing relative to direct reply processing. These findings support that indirect meanings were accessed by non-native speakers at a later stage. This finding aligns with previous studies indicating that non-native speakers have more difficulty during pragmatic processing than native speakers (e.g., Antoniou, 2019), as enhanced LPC has been linked to pragmatic analysis (Coulson & Van Petten, 2002; Regel & Gunter, 2017; Regel et al., 2011; Weiland et al., 2014; Zhang et al., 2021; Zhang, Pan et al., 2023). Given the decreased cognitive resources available for processing intended meanings of indirect replies (Morishima, 2013), it is also reasonable that non-native speakers are slower in detecting differences between indirect and direct replies. Despite significant differences in time course, neurophysiological findings, together with our behavioral results, support that non-native Mandarin speakers are capable of deriving meaning from indirect replies, albeit at a later stage when compared with native speakers.

Interestingly, we observed that the LPC effect for non-native Mandarin speakers was mainly distributed in the left hemisphere. LPC effects observed in the central posterior area of the left hemisphere were previously reported (Peters & Daum, 2009; Wilding & Rugg, 1996; Zhang et al., 2020), which has been related to the encoding strength of memory (Paller & Wagner, 2002), retrieval of general (Peters & Daum, 2009), and episodic memory (Rugg & Curran, 2007), and controlled semantic retrieval (Martin et al., 2009). In the present study, non-native speakers may be applying more effort in general memory retrieval during online indirect reply processing. When target stimuli were presented, the information preceding it disappeared, requiring readers to retain the question asked by Speaker A in their WM and to retrieve world knowledge

https://doi.org/10.1017/S1366728924000695 Published online by Cambridge University Press

related to the conversation from long-term memory. These processes may be more challenging for non-native speakers but are still necessary for successful comprehension of indirect replies in a second language.

Additionally, for non-native speakers, we found an enhanced delayed positivity for indirect replies in comparison to direct replies. The enhanced delayed positivity has also been previously reported (Jiang & Pell, 2016a, 2016b; Jiang & Zhou, 2015; Zhang et al., 2021). This electrophysiological result is thought to reflect attempts to infer other's intentions and goals, as well as late pragmatic inference for deriving intended meanings. The higher delayed positivity elicited by indirect replies observed in the present study may be associated with pragmatic inference for deriving speakers' meanings and communication intention.

The absence of an N400 effect for indirect replies is consistent with other ERP studies on sentence and discourse comprehension (Foucart, Moreno et al., 2015; Martin et al., 2013; Newman et al., 2012; Pérez et al., 2019). Given differences in cognitive resources available and prediction ability, non-native speakers might adopt a less demanding reactive processing strategy for conversation comprehension, unlike native speakers' proactively processing strategy (Grüter & Rohde, 2013; Grüter et al., 2014; Pérez et al., 2019). Consequently, the reduced processing efficiency of non-native speakers may present only in later-stage components. The absence of an N400 effect may also reflect heightened susceptibility to retrieval interference in non-native speakers (Cunnings, 2017a, 2017b). During the processing of indirect replies, both contextual cues and world knowledge stored in memory are important for identifying the most relevant interpretation. As non-native speakers are prone to retrieval interference, it is possible they could not efficiently use various information to access maximum relevant interpretations for indirect replies. This bears a resemblance to the processing of indirect replies in readers with low WM span (Zhang et al., 2021). When comprehending indirect replies, low-span readers may not benefit from default knowledge, resulting in a delayed understanding of what a speaker is trying to say. In the present study, during indirect reply processing, non-native speakers may decode speakers' meanings for coherent comprehension without initial analysis at an early stage.

4.3. Differences between native and non-native Mandarin speakers

Non-native Mandarin speakers performed worse than native speakers during online indirect reply processing, reflected by a significant interaction between Reply and Group on comprehension accuracy. These findings contrast with those in previous studies using offline multiple-choice tests which reported native-like processing of English direct replies in highly proficient non-native speakers who had relocated to the USA for an extended period of time (Bouton, 1992, 1994). Our findings also differ from those reported in more recent work using a self-paced reading paradigm (Zhang et al., under review), in which no significant interaction between Reply and Group on comprehension accuracy was identified. Under these conditions, which lack the time constraints present during online reading tasks like the one used in the present study, non-native speakers perform similarly to native speakers when reading indirect replies. These conflicting patterns of results support that differences between native and non-native Mandarin speakers may only manifest under certain conditions.

Results from our ERP analysis during indirect reply processing were similar to those reported in previous studies (Foucart et al., 2015). Specifically, while both N400 and LPC effects were identified in native Mandarin speakers, only LPC effects were observed in non-native speakers. Taken together, these findings support that both native and non-native speakers are capable of online processing of indirect replies, but that the time course of this processing is delayed in non-native speakers.

Overall, the results of the present study support that high-proficiency non-native speakers behave like native speakers with low WM capacity or poor comprehension abilities (Boudewyn et al., 2013; Foucart et al., 2016; Long & Chong, 2001; Zhang et al., 2021). These findings are in alignment with the limited resources hypothesis. Pragmatic interpretation difficulties for L2 learners could be attributed to insufficient cognitive resources because of the more effortful nature of L2 language processing (e.g., Antoniou, 2019). Decoding literal meanings of indirect replies is nearly automatic and requires very few cognitive resources for native speakers. In contrast, non-native speakers require more cognitive resources while using their L2, leaving fewer resources for deriving intended meanings from indirect replies. The inefficient allocation of cognitive resources in non-native speakers may also be behind other patterns of results observed during L2 processing including reduced speed (Dekydtspotter et al., 2006; Hopp, 2010), decreased ability to generate predictions (Grüter & Rohde, 2013; Grüter et al., 2014; Pérez et al., 2019), and susceptibility to retrieval interference (Cunnings, 2017a, 2017b).

Relevance Theory (Sperber & Wilson, 1986) considers relevance-seeking as the cognitive mechanism behind comprehension of indirect replies. When a reply is presented, readers automatically seek relevance for the information by maximizing the use of available contextual cues, with the number of cues determining the strength of implicature. As Taguchi stated, relevance-seeking mechanisms can transfer from L1 to L2 (e.g., Taguchi & Yamaguchi, 2019), thus non-native speakers can comprehend conversation implicature. However, for indirect reply comprehension, the number of relevant contextual cues available for native and non-native Mandarin speakers may differ. It is possible that relevant contextual cues are more quickly identified by native speakers, leading to processing differences for indirect replies at an earlier stage, as evidenced in our observed native speaker-specific N400 effect.

In the field of human communication, the principle of relevance maximization, as proposed by Relevance Theory (Sperber & Wilson, 1986), has inspired a large number of linguistic and philosophical studies (e.g., Levinson, 2000; Sperber & Wilson, 1986). More recently, this interest has manifested in neuropsychological studies (e.g., Hagoort, 2019; Noveck & Reboul, 2008), as well as computational investigations (e.g., Dale & Reiter, 1995; Goodman & Frank, 2016; Mi et al., 2021). Informed by both Relevance Theory and psychological research methods, the present study examined whether non-native and native Mandarin speakers differed in online indirect reply processing. Our findings indicate that non-native Mandarin speakers can understand indirect replies in the absence of immediate pragmatic analysis, suggesting cross-language relevance maximization, albeit with delayed processing. These findings also align with the limited cognitive resources hypothesis for L2 pragmatic processing (Morishima, 2013). Together, findings from the present work have important implications for our understanding of human communication and L2 pragmatic processing.

Findings from the present study should be considered in light of a few limitations. First, we did not consider the potential influence of language proficiency and cultural differences in non-native Mandarin speakers. Given that language proficiency and cultural background may influence both linguistic (e.g., Momenian et al.,

2024) and non-linguistic cognitive function (e.g., Privitera et al., 2023; Samuel et al., 2018), future studies should account for these differences. Additionally, we did not assess participants' general cognitive ability including WM capacity or cognitive control. This prevented us from accounting for differences in domain-general cognitive function in native and non-native Mandarin speakers that may impact the processing of indirect replies (Zhang et al., 2021; Zhang et al., 2024). The ecological validity of the present study is also limited by the use of a paradigm that does not mimic the conditions of real-world reading. Future studies should consider combining EEG with eye tracking (e.g., Wei et al., 2023) or EEG-derived virtual eye tracking (Sun et al., 2023) to test whether findings reported in the present study emerge under natural reading conditions. Finally, reliance on sensor space EEG measures prevented us from drawing conclusions regarding neural generators underlying our observed neurophysiological findings. Future work utilizing source space EEG measures (e.g., Pascual-Marqui, 2002; Privitera & Tang, 2022) can provide insight into the brain's role in the processing of indirect replies.

5. Conclusions

To summarize, the present study examined how native Mandarin speakers differed from non-native speakers in the online processing of indirect replies using both behavioral and neurophysiological measures. Our findings support that, like native speakers, highly proficient non-native speakers can comprehend indirect replies online, though the time course of processing is delayed. These findings contribute to our understanding of how native and non-native speakers differ in online language processing and cognitive resource limits associated with L2 use, and provide evidence in support of cross-language relevance maximization.

Supplementary material. To view supplementary material for this article, please visit http://doi.org/10.1017/S1366728924000695.

Data availability statement. The datasets generated for this study are available on request to the corresponding author.

Acknowledgments. This work was supported by the Science Foundation of Beijing Language and Culture University (supported by "the Fundamental Research Funds for the Central Universities") (Approval number 24QN47), the Fundamental Research Funds for the Central Universities, and the Research Funds of Beijing Language and Culture University (24YCX045), the Discipline Team Support Program of Beijing Language and Culture University (grant number 2023YGF07), the Research Project "the Construction of the Advanced Disciplines in Universities in Beijing," and the National Natural Science Foundation of China (31800914).

Authorship contribution. Xiuping Zhang: the conception and design of the work, the interpretation of data for the work, Funding Acquisition, drafting and revising the work, and final approval of the version to be published. Xiaoxi Pan: the acquisition and analysis of data for the work, drafting and revising the work, and final approval of the version to be published. Yizhu Wang: drawing, the analysis of data for the work, revising the work, and final approval of the version to be published. Maoyao Xu: the analysis of data for the work, revising the work, and final approval of the version to be published. Adam John Privitera: revising the work, final approval of the version to be published.

Competing interests. The author(s) declare none.

Ethical standard. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. Approval for this study was granted by the ethics committee of the Beijing Language and Culture University.

References

Amoruso, L., **Gelormini, C.**, **Aboitiz, F.**, **Alvarez González, M.**, **Manes, F.**, **Cardona, J.**, & **Ibanez, A.** (2013). N400 ERPs for actions: Building meaning in context. *Frontiers in Human Neuroscience*, **7**, 57. https://doi.org/10.3389/fnhum.2013.00057

Antoniou, K. (2019). Multilingual pragmatics: Implicature comprehension in adult L2 learners and multilingual children. In Taguchi N (ed), *The Routledge handbook of second language acquisition* (pp. 495–510). New York: Routledge

Arzouan, Y., **Goldstein, A.**, & **Faust, M.** (2007). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Research*, **1160**, 69–81. https://doi.org/10.1016/j.brainres.2007.05.034

Bambini, V., **Canal, P.**, **Resta, D.**, & **Grimaldi, M.** (2019). Time course and neurophysiological underpinnings of metaphor in literary context. *Discourse Processes*, **56**(1), 77–97. doi:10.1080/0163853X.2017.1401876

Bašnáková, J., **Weber, K.**, **Petersson, K. M.**, **van Berkum, J.**, & **Hagoort, P.** (2014). Beyond the language given: The neural correlates of inferring speaker meaning. *Cerebral Cortex*, **24**(10), 2572–2578. https://doi.org/10.1093/cercor/bht112

Boudewyn, M. A., **Long, D. L.**, & **Swaab, T. Y.** (2013). Effects of working memory span on processing of lexical associations and congruence in spoken discourse. *Frontiers in Psychology*, **4**, 60. https://doi.org/10.3389/fpsyg.2013.00060.

Bouton, L. F. (1992). The interpretation of implicature in English by NNS: Does it come automatically--without being explicitly taught? *Pragmatics and Language Learning*, **3**, 53–65.

Bouton, L. F. (1994). Can NNS skill in interpreting implicature in American English be improved through explicit instruction?-A pilot study. *Pragmatics and Language Learning*, **5**, 89–109.

Boux, I., **Tomasello, R.**, **Grisoni, L.**, & **Pulvermüller, F.** (2021). Brain signatures predict communicative function of speech production in interaction. *Cortex*, **135**, 127–145. https://doi.org/10.1016/j.cortex.2020.11.008

Brouwer, H., **Fitz, H.**, & **Hoeks, J. J. B. r**. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, **1446**, 127–143. https://doi.org/10.1016/j.brainres.2012.01.055

Clahsen, H., & **Felser, C.** (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, **27**, 107–126. https://doi.org/10.1017/S0142716406060206

Clahsen, H., & **Felser, C.** (2006b). How native-like is non-native language processing? *Trends in Cognitive Sciences*, **10**(12), 564–570. https://doi.org/10.1016/j.tics.2006.10.002.

Clahsen, H., & **Felser, C.** (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, **40**(3), 693–706. https://doi.org/10.1017/S0272263117000250

Coulson, S. (2004). Electrophysiology and pragmatic language comprehension. In I. A. Noveck & D. Sperber (Eds.), *Experimental pragmatics* (pp. 187–206). London: Palgrave Macmillan UK.

Coulson, S., & **Kutas, M.** (2001). Getting it: Human event-related brain response to jokes in good and poor comprehenders. *Neuroscience Letters*, **316**(2), 71–74. https://doi.org/10.1016/S0304-3940(01)02387-4

Coulson, S., & **Lovett, C.** (2010). Comprehension of nonconventional indirect requests: An event-related brain potential study. *Italian Journal of Linguistics*, **22**(1), 107–124.

Coulson, S., & **Van Petten, C.** (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition*, **30**(6), 958–968. https://doi.org/10.3758/bf03195780

Cunnings, I. (2017a). Interference in native and non-native sentence processing. *Bilingualism: Language and Cognition*, **20**(4), 712–721. https://doi.org/10.1017/S1366728916001243

Cunnings, I. (2017b). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, **20**(4), 659–678. https://doi.org/10.1017/S1366728916000675

Dale, R., & **Reiter, E.** (1995). Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science*, **19**(2), 233–263.

Dekydtspotter, L., **Schwartz, B. D.**, & **Sprouse, R. A.** (2006). The comparative fallacy in L2 processing research. In M. Grantham O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th generative approaches to second language acquisition conference (GASLA 2006)* (pp. 33–40). Somerville, MA: Cascadilla Proceedings Project.

Delorme, A., & **Makeig, S.** (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, **134**(1), 9–21.

Feng, W., **Wu, Y.**, **Jan, C.**, **Yu, H.**, **Jiang, X.**, & **Zhou, X.** (2017). Effects of contextual relevance on pragmatic inference during conversation: An fMRI study. *Brain and Language*, **171**, 52–61. https://doi.org/10.1016/j.bandl.2017.04.005

Feng, W., **Yu, H.**, & **Zhou, X.** (2021). Understanding particularized and generalized conversational implicatures: Is theory-of-mind necessary? *Brain and Language*, **212**, 104878. https://doi.org/10.1016/j.bandl.2020.104878

Filik, R., **Leuthold, H.**, **Wallington, K.**, & **Page, J.** (2014). Testing theories of irony processing using eye-tracking and ERPs. *Journal of Experimental Psychology Learning Memory & Cognition*, **40**(3), 811.

Foucart, A., **Garcia, X.**, **Ayguasanosa, M.**, **Thierry, G.**, **Martin, C.**, & **Costa, A.** (2015). Does the speaker matter? Online processing of semantic and pragmatic information in L2 speech comprehension. *Neuropsychologia*, **75**, 291–303. https://doi.org/10.1016/j.neuropsychologia.2015.06.027

Foucart, A., **Martin, C. D.**, **Moreno, E. M.**, & **Costa, A.** (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, Cognition*, **40**(5), 1461.

Foucart, A., **Moreno, E.**, **Martin, C. D.**, & **Costa, A.** (2015). Integration of moral values during L2 sentence processing. *Acta Psychologica*, **162**, 1–12. doi:https://doi.org/10.1016/j.actpsy.2015.09.009

Foucart, A., **Romero-Rivas, C.**, **Gort, B. L.**, & **Costa, A.** (2016). Discourse comprehension in L2: Making sense of what is not explicitly said. *Brain and Language*, **163**, 32–41. doi:10.1016/j.bandl.2016.09.001

Goldstein, A., **Arzouan, Y.**, & **Faust, M.** (2012). Killing a novel metaphor and reviving a dead one: ERP correlates of metaphor conventionalization. *Brain and Language*, **123**(2), 137–142. https://doi.org/10.1016/j.bandl.2012.09.008

Goodman, N. D., & **Frank, M. C.** (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, **20**(11), 818–829. https://doi.org/10.1016/j.tics.2016.08.005

Grice, H.P., 1975. Logic and conversation. In: P. Cole, J. Morgan (Eds.), *Syntax and semantics: Speech acts* (pp. 41–58). New York: Academic Press.

Grisoni, L., **Mohr, B.**, & **Pulvermüller, F.** (2019). Prediction mechanisms in motor and auditory areas and their role in sound perception and language understanding. *NeuroImage*, **199**, 206–216. https://doi.org/10.1016/j.neuroimage.2019.05.071

Grüter, T., & **Rohde, H.** (2013). *L2 processing is affected by RAGE: Evidence from reference resolution.* Paper presented at the proceedings of talk presented at the 12th conference on Florida, FL.

Grüter, T., **Rohde, H.**, & **Schafer, A.** (2014). *The role of discourse-level expectations in non-native speakers' referential choices.* In Paper presented at the proceedings of the annual Boston university conference on language development.

Guo, X., **Li, X.**, & **Yang, Y.** (2023). How the brain processes emotional meaning of indirect reply: Evidence from EEG. *Language, Cognition and Neuroscience*, **38**(9), 1302–1317. https://doi.org/10.1080/23273798.2023.2232055

Hagoort, P. (2019). The neurobiology of language beyond single-word processing. *Science*, **366**(6461), 55–58. https://doi.org/10.1126/science.aax0289

Harrington, M., & **Sawyer, M.** (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, **14**(1), 25–38. https://doi.org/10.1017/S0272263100010457

Holtgraves, T. (1999). Comprehending indirect replies: When and how are their conveyed meanings activated? *Journal of Memory and Language*, **41**(4), 519–540. https://doi.org/10.1006/jmla.1999.2657

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, **120**(4), 901–931. http://doi.org/10.1016/j.lingua.2009.06.004

Jang, G., **Yoon, S. A.**, **Lee, S. E.**, **Park, H.**, **Kim, J.**, **Ko, J. H.**, & **Park, H. J.** (2013). Everyday conversation requires cognitive inference: Neural bases of comprehending implicated meanings in conversations. *NeuroImage*, **81**, 61–72. https://doi.org/10.1016/j.neuroimage.2013.05.027

Jiang, X., & **Liu, Y.** (2004). A study of character writing errors by foreign learner using alphabetic writing (in Chinese). *Chinese Teaching in the World*, (01), 60–70.

Jiang, X., & **Pell, M. D.** (2016a). Neural responses towards a speaker's feeling of (un)knowing. *Neuropsychologia*, **81**, 79–93. https://doi.org/10.1016/j.neuropsychologia.2015.12.008

Jiang, X., & **Pell, M. D.** (2016b). The feeling of another's knowing: How "mixed messages" in speech are reconciled. *Journal of Experimental Psychology: Human Perception and Performance*, **42**(9), 1412–1428. https://doi.org/10.1037/xhp0000240

Jiang, X., & **Zhou, X.** (2015). Who is respectful? Effects of social context and individual empathic ability on ambiguity resolution during utterance comprehension. *Frontiers in Psychology*, **6**, 1588, 1–16. https://doi.org/10.3389/fpsyg.2015.01588

Lai, V. T., & **Curran, T.** (2013). ERP evidence for conceptual mappings and comparison processes during the comprehension of conventional and novel metaphors. *Brain and Language*, **127**(3), 484–496. https://doi.org/10.1016/j.bandl.2013.09.010

Lai, V. T., **Curran, T.**, & **Menn, L.** (2009). Comprehending conventional and novel metaphors: An ERP study. *Brain Research*, **1284**, 145–155. https://doi.org/10.1016/j.brainres.2009.05.088

Leuthold, H., **Filik, R.**, **Murphy, K.**, & **Mackenzie, I. G.** (2012). The on-line processing of socio-emotional information in prototypical scenarios: Inferences from brain potentials. *Social Cognitive and Affective Neuroscience*, **7**(4), 457–466. https://doi.org/10.1093/scan/nsr029

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*: MIT Press.

Li, X., **Zhang, Y.**, **Xia, J.**, & **Swaab, T. Y.** (2017). Internal mechanisms underlying anticipatory language processing: Evidence from event-related-potentials and neural oscillations. *Neuropsychologia*, **102**, 70–81. https://doi.org/10.1016/j.neuropsychologia.2017.05.017

Long, D. L., & **Chong, J. L.** (2001). Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, **27**(6), 1424–1429. https://doi.org/10.1037/0278-7393.27.6.1424

Luck, S. J., & **Gaspelin, N.** (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, **54**(1), 146–157. https://doi.org/10.1111/psyp.12639

Maris, E., & **Oostenveld, R.** (2007). Nonparametric statistical testing of EEG- and MEGdata. *Journal of Neuroscience Methods*, **164**(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Martin, C. D., **Dering, B.**, **Thomas, E. M., and Thierry, G.** (2009). Brain potentials reveal semantic priming in both the "active" and the "non-attended" language of early bilinguals. *NeuroImage*, **47**, 326–333. http://doi.org/10.1016/j.neuroimage.2009.04.025

Martin, C. D., **Thierry, G.**, **Kuipers, J.-R.**, **Boutonnet, B.**, **Foucart, A.**, & **Costa, A.** (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, **69**(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

Mi, Q., **Wang, C.**, **Camerer, C. F.**, & **Zhu, L.** (2021). Reading between the lines: Listener's vmPFC simulates speaker cooperative choices in communication games. *Science Advances*, **7**(10), eabe6276. http://doi.org/10.1126/sciadv.abe6276

Momenian, M., **Privitera, A. J.**, & **Weekes, B.** (2024). Nouns are not always processed faster than verbs in bilingual speakers: Effects of language distance. *Bilingualism: Language and Cognition*, **27**(1), 75–83.

Morishima, Y. (2013). Allocation of limited cognitive resources during text comprehension in a second language. *Discourse Processes*, **50**(8), 577–597. https://doi.org/10.1080/0163853X.2013.846964

Newman, A. J., **Tremblay, A.**, **Nichols, E. S.**, **Neville, H. J.**, & **Ullman, M. T.** (2012). The influence of language proficiency on lexical semantic processing in native and late learners of English. *Journal of Cognitive Neuroscience*, **24**(5), 1205–1223. https://doi.org/10.1162/jocn_a_00143

Noveck, I. A., & **Reboul, A.** (2008). Experimental pragmatics: A Gricean turn in the study of language. *Trends in Cognitive Sciences*, **12**(11), 425–431. http://doi.org/10.1016/j.tics.2008.07.009

Otten, M., & **Van Berkum, J. J. A.** (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes*, **45**(6), 464–496. https://doi.org/10.1080/01638530802356463

https://doi.org/10.1017/S1366728924000695 Published online by Cambridge University Press

Paller, K. A., and Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Science*, **6**, 93–102. http://doi.org/10.1016/s1364-6613(00)01845-3

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *Methods and Findings in Experimental and Clinical Pharmacology*, **24**(Suppl D), 5–12.

Pérez, A., **Hansen, L.**, & **Bajo, T.** (2019). The nature of first and second language processing: The role of cognitive control and L2 proficiency during text-level comprehension. *Bilingualism: Language and Cognition*, **22**(5), 930–948. https://doi.org/10.1017/S1366728918000846

Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.

Peters, J., and Daum, I. (2009). Frontal but not parietal positivity during source recollection is sensitive to episodic content. *Neuroscience Letters*, **454**, 182–186. http://doi.org/10.1016/j.neulet.2009.03.019

Privitera, A. J., **Momenian, M.**, & **Weekes, B. S.** (2023). Graded bilingual effects on attentional network function in Chinese high school students. *Bilingualism: Language and Cognition*, **26**(3), 527–537. https://doi.org/10.1017/S1366728922000803

Privitera, A. J., & **Tang, A. C.** (2022). Functional significance of individual differences in P3 network spatial configuration. *Journal of Psychophysiology*, **36**(4), 215–230. https://doi.org/10.1027/0269-8803/a000295

Pynte, J., **Besson, M.**, **Robichon, F.-H.**, & **Poli, J.** (1996). The time course of metaphor comprehension: An event-related potential study. *Brain and Language*, **55**(3), 293–316. https://doi.org/10.1006/brln.1996.0107

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rai, M. K., **Loschky, L. C.**, **Harris, R. J.**, **Peck, N. R.**, & **Cook, L. G.** (2011). The effects of stress and working memory capacity on foreign language comprehenders' inferential processing during comprehension. *Language Learning*, **61**, 187–218. https://doi.org/10.1111/j.1467-9922.2010.00592.x

Regel, S., & **Gunter, T. C.** (2017). Don't get me wrong: ERP evidence from cueing communicative intentions. *Frontiers in Psychology*, **8**, 1465. https://doi.org/10.3389/fpsyg.2017.01465

Regel, S., **Coulson, S.**, & **Gunter, T. C.** (2010). The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. *Brain Research*, **1311**, 121–135.

Regel, S., **Gunter, T. C.**, & **Friederici, A. D.** (2011). Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience*, **23**(2), 277–293. https://doi.org/10.1162/jocn.2010.21411

Regel, S., **Meyer, L.**, & **Gunter, T. C.** (2014). Distinguishing neurocognitive processes reflected by P600 effects: Evidence from ERPs and neural oscillations. *Plos One*, **9**(5), e96840. https://doi.org/10.1371/journal.pone.0096840

Rugg, M. D., and Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Science*, **11**, 251–257. https://doi.org/10.1016/j.tics.2007.04.004

Samuel, S., **Roehr-Brackin, K.**, **Pak, H.**, & **Kim, H.** (2018). Cultural effects rather than a bilingual advantage in cognition: A review and an empirical study. *Cognitive Science*, **42**(7), 2313–2341.

Shibata, M., **Abe, J. I.**, **Itoh, H.**, **Shimada, K.**, & **Umeda, S.** (2011). Neural processing associated with comprehension of an indirect reply during a scenario reading task. *Neuropsychologia*, **49**(13), 3542–3550. https://doi.org/10.1016/j.neuropsychologia.2011.09.006

Sperber, D., & **Wilson, D.** (1986). *Relevance: Communication and cognition* (Vol. **1**). Cambridge, MA: Harvard University Press.

Spotorno, N., **Cheylus, A.**, **Van Der Henst, J. B.**, & **Noveck, I. A.** (2013). What's behind a P600? Integration operations during irony processing. *Plos One*, **8**(6). https://doi.org/10.1371/journal.pone.0066839

Sun, R., **Cheng, A. S.**, **Chan, C.**, **Hsiao, J.**, **Privitera, A. J.**, **Gao, J.**, … & **Tang, A. C.** (2023). Tracking gaze position from EEG: Exploring the possibility of an EEG-based virtual eye-tracker. *Brain and Behavior*, **13**(10), e3205. https://doi.org/10.1002/brb3.3205

Taguchi, N. (2005). Comprehending implied meaning in English as a foreign language. *The Modern Language Journal*, **89**(4), 543–562. https://doi.org/10.1111/j.1540-4781.2005.00329.x

Taguchi, N. (2008a). Cognition, language contact, and the development of pragmatic comprehension in a study-abroad context. *Language Learning*, **58**(1), 33–71. https://doi.org/10.1111/j.1467-9922.2007.00434.x

Taguchi, N. (2008b). The role of language environment in the development of pragmatic comprehension: A comparison of gains between EFL and ESL learners. *Studies in Second Language Acquisition*, **30**(4), 423–452. https://doi.org/10.1017/S0272263108080716

Taguchi, N. (2011). The Effect of L2 proficiency and study-abroad experience on pragmatic comprehension. *Language Learning*, **61**(3), 904–939. https://doi.org/10.1111/j.1467-9922.2011.00633.x

Taguchi, N., & **Yamaguchi, S.** (2019). Implicature comprehension in l2 pragmatics research. In *The Routledge handbook of second language acquisition and pragmatics* (1st ed., pp. 31–46). Routledge.

Taguchi, N., **Gomez-Laich, M. P.**, & **Arrufat-Marques, M.-J.** (2016). Comprehension of indirect meaning in Spanish as a foreign language. *Foreign Language Annals*, **49**(4), 677–698. https://doi.org/10.1111/flan.12230

van Ackeren, M. J., **Smaragdi, A.**, & **Rueschemeyer, S.-A.** (2016). Neuronal interactions between mentalizing and action systems during indirect request processing. *Social Cognitive and Affective Neuroscience*, **11**(9), 1402–1410. https://doi.org/10.1093/scan/nsw062.

van Elk, M., **van Schie, H. T.**, & **Bekkering, H.** (2010). The N400-concreteness effect reflects the retrieval of semantic information during the preparation of meaningful actions. *Biological Psychology*, **85**(1), 134–142. https://doi.org/10.1016/j.biopsycho.2010.06.004

Wei, Y., **Tang, Y.**, & **Privitera, A. J.** (2023). Functional priority of syntax over semantics in Chinese 'ba'construction: Evidence from eye-tracking during natural reading. *Language and Cognition*, 1–21. doi:10.1017/langcog.2023.42

Weiland, H., **Bambini, V.**, & **Schumacher, P. B.** (2014). The role of literal meaning in figurative language comprehension: Evidence from masked priming ERP. *Frontiers in Human Neuroscience*, **8**, 583. https://doi.org/10.3389/fnhum.2014.00583

Wilding, E. L., and Rugg, M. D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain*, **119**, 889–905. http://doi.org/10.1093/brain/119.3.889

Zhang, X., **Huang, D.**, **Wang, Y.**, **Yang, Y.**, **Pan, X.**, **Shen, M.** (under review). Proficient non-native Mandarin speakers comprehend informative indirect replies similarly to native speakers during Chinese dialogue comprehension. *Applied Psycholinguistics.*

Zhang, X., **Pan, X.**, **Yang, X.**, & **Yang, Y.** (2023). Conventionality determines the time course of indirect replies comprehension: An ERP study. *Brain and Language*, **239**, 105253. https://doi.org/10.1016/j.bandl.2023.105253

Zhang, X., **Wang, Y.**, **Yang, X.**, & **Yang, Y.** (2024). The spontaneous activities of the multiple demand network are related to individual differences in indirect replies comprehension. *Behavioural Brain Research*, **469**, 115021.

Zhang, X., **Xu, M.**, **Yang, X.**, & **Yang, Y.** (2023). Individual difference in emotion attenuation brought by indirect replies is related to resting-state brain activity. *Brain Sciences*, **13**(7). https://doi.org/10.3390/brainsci13071053

Zhang, X., **Zhang, Y.**, **Zhang, Z.**, **Yang, X.**, & **Yang, Y.** (2021). How working memory capacity modulates the time course of indirect replies comprehension: An event-related potential study. *Language, Cognition and Neuroscience*, **36**(10), 1246–1257. https://doi.org/10.1080/23273798.2021.1924386

Zhang, Y., **Lu, Y.**, **Liang, L.**, & **Chen, B.** (2020). The effect of semantic similarity on learning ambiguous words in a second language: An event-related potential study. *Frontiers in Psychology*, **11**, 1633. http://doi.org/10.3389/fpsyg.2020.01633