LANGUAGE, COGNITION AND NEUROSCIENCE
2023, VOL. 38, NO. 8, 1068–1081
https://doi.org/10.1080/23273798.2023.2192509

Routledge
Taylor & Francis Group

REGULAR ARTICLE

Frame-shifting instead of incongruity is necessary for pun comprehension: evidence from an ERP study on Chinese homophone puns

Wei Zheng [a] and Xiaolu Wang [b,c,d]

[a]School of Foreign Languages, China Three Gorges University, Yichang, People's Republic of China; [b]School of Foreign Languages, Hangzhou City University, Hangzhou, People's Republic of China; [c]School of International Studies, Zhejiang University, Hangzhou, People's Republic of China; [d]School of Humanities & Communication Arts, University of Western Sydney, Sydney, Australia

ABSTRACT

Recent psycholinguistic research has made significant progress in understanding the meaning-access process during pun comprehension. However, to date, little research has directly investigated how the two retrieved meanings are integrated into the pun context afterwards. In the current ERP study, we examined this process by comparing homophone puns with two control conditions. Different from previous ERP studies on jokes, we did not observe significantly enhanced N400 amplitudes (300-500 ms) in the pun condition, indicating no apparent detection of incongruity. However, we observed a sustained positivity around the left anterior regions (500-900 ms) and enhanced LPC amplitudes around the central-parietal regions (600-900 ms). These two components could index the sudden access to the second meaning and the additional integration operations to establish a new cognitive model respectively. These findings were compatible with the Space Structure Model, which emphasises the frame-shifting process as a crucial element for understanding verbal humour.

ARTICLE HISTORY
Received 3 August 2022
Accepted 10 March 2023

KEYWORDS
Frame-shifting; space structure model; incongruity; homophone puns; ERP study

Introduction

Punning is a rhetorical device for humourous effect by simultaneously generating two possible interpretations with words of the same sound (homophone puns) or orthographic form (homograph puns). Due to such a function, puns are used as an important means of verbal humour and are widely found in both literary works and advertisement slogans. For example, Mercutio, a joyful and humourous character in Shakespeare's *Romeo and Juliet*, has left a deep impression on many readers with his final words, "Ask for me tomorrow, and you shall find me a grave man".

Although it is generally agreed that what makes a pun unique is the juxtaposition of two distinct meanings, psycholinguists are more curious about how the two distinct meanings are accessed temporally. In recent years, there is increasing evidence showing that the access to the two meanings of a pun is not simultaneous but modulated by their relative salience[1] in the pun receiver's mind (Coulson & Severens, 2007; Koleva et al., 2019; Zheng et al., 2020). For example, Zheng et al. (2020) investigated whether the two meanings of a pun word (the homograph/homophone that triggers the punning effect) are accessed simultaneously or sequentially, using the visual world paradigm with printed words (Shen et al., 2016). These authors demonstrated that after hearing a pun word, the participants fixated more on the word related to its salient meaning than its less salient meaning in the initial 200 ms, and after that, the words related to both meanings received equal fixations. They concluded that the salient meaning of a pun word was invariantly accessed faster than the less salient ones. A recent study on reading homophone puns also supported this salience effect, showing that meanings saliently related to a homophonic sound were still recoverable even when they were not presented visually (Zheng & Wang, 2022).

Despite the progress in understanding how different meanings are accessed during pun comprehension, little psycholinguistic research has been conducted on how these meanings are integrated into the pun context afterwards. This process is equally important, if not more, in that only when the two retrieved meanings are successfully integrated into the pun context can it bring out the pun humour. After all, the humourous

CONTACT Xiaolu Wang lucywxl@zju.edu.cn
Supplemental data for this article can be accessed online at https://doi.org/10.1080/23273798.2023.2192509.

effect is one of the most important reasons that puns are widely employed in various discourses.

Since punning is an important type of verbal humour, one is likely to expect that humour theories may shed some light on its comprehension process. Suls (1972), for example, proposed a very influential two-stage model for humour appreciation based on the notion of "incongruity". Incongruity refers to such situations when the arrangement of the constituent elements of an event violates the normal or expected pattern (McGhee, 1972). According to this model, humour appreciation involves two essential cognitive processes: detection of the incongruity and its successful resolution. Suls used the following joke to illustrate his point: *O'Riley was on trial for armed robbery. The jury came out and announced, "Not guilty". "Wonderful!" said O'Riley, "Does that mean I can keep the money?"* According to the incongruity-resolution theory, the reader is likely to believe that O'Riley is innocent and does not possess the money after reading the setup sentences. However, this expectation turns out to be violated when the punch line (the last sentence) appears, hence the incongruity. To get the joke, the reader has to revise his/her early interpretation to reach the other possible but less-likely interpretation: the court has made a mistake. The notion of "incongruity" is inherited by many following theories, such as the Semantic-Script Theory of Humour (Attardo & Raskin, 1991) and its expanded version, the General Theory of Verbal Humour (Attardo, 2001).

The incongruity-based accounts have received some support from the so-called N400 joke effect from ERP (event-related potential) studies. The N400, usually with a central-parietal distribution, is a negative shift in the ERP waveform with a peak latency of about 400 ms after the critical stimulus (Kappenman & Luck, 2012). It is generally believed to be an inverse function of the reader's expectancy for a particular word, which is usually measured by its cloze probability[2] (Kutas & Hillyard, 1984; See Dien et al., 2004 for a different view). In many ERP studies on jokes, researchers have reported enhanced N400 amplitudes in jokes compared to the non-funny controls, which was taken as evidence for incongruity detection (Coulson & Kutas, 2001; Feng et al., 2014; Ku et al., 2017). For example, Coulson and Kutas (2001) examined the neural mechanisms for joke comprehension under the incongruity-resolution framework. In the study, they compared the ERPs elicited by one-line jokes with their unfunny controls, which were identical until the sentence-final words. In addition, the experimental sentences were further divided into a high-constraint group and a low-constraint group based on the sentence constraint measured by the cloze probability of the final word. Take one of their high-constraint sentence pairs, for example, *She read so much about the bad effects of smoking that she decided she'd have to give up reading/habit*. The ERPs elicited by a joke ending (e.g. reading) were compared with those elicited by an unfunny ending (e.g. habit), matched in word length and frequency. According to their results, the high-constraint jokes elicited larger N400 amplitudes with a central-parietal focus, supporting the prediction of the incongruity-resolution model. However, it is worth noting that they failed to find such an N400 joke effect in their low-constraint jokes, casting some doubt on the necessity of incongruity for humourous effect (See also Coulson & Lovett, 2004).

More importantly, the problem with the incongruity account for pun humour is that there seems to be no apparent incongruity in puns since both meanings of the pun word need to be compatible with the pun context. Take the following homophone pun, for example, *A bicycle cannot stand on its own because it is two-tired*. The pun word (*two-tired*) is consistent with the overall sentential context about balance, and there seems no need for readers to search for another interpretation. Rather, the meaning of its more salient counterpart (*too tired*) is largely activated through the shared phonology, which needs no "incongruity" to trigger. This line of thought is supported by Dholakia et al.'s (2016) ERP study on homograph puns. Dholakia and collaborators employed four types of sentence context to instantiate the different meanings of an ambiguous word (e.g. crown): (1) the pun context (e.g. *The prince with a bad tooth got a crown*), where both the dominant meaning (*royal ornament*) and the subordinate meaning (*dental device*) were equally possible; (2) the dominant-biasing context (e.g. *The prince with a bad leg got a crown*), where the more-frequently instantiated meaning was supported; (3) the subordinate-biasing context (e.g. *The adult with a bad tooth got a crown*), where the less frequent meaning was preferred; and (4) the neutral context (e.g. *The adult with a bad leg got a crown*), where no context words were semantically related to the ambiguous word. The results showed that the homographs elicited the largest N400 amplitudes (least primed) in the neutral condition, followed by the subordinate, pun, and dominant conditions. Although the dominant condition elicited the smallest N400 component, there was still a significant facilitative effect of the pun context compared to the neutral context where no prime word was present; or in these authors' words, two primes are better than none. As a result, their findings provided further evidence that incongruity could be absent in verbal humour.

Unlike the incongruity-based account, the Space Structuring Model provides a new perspective on the cognitive processes of humour comprehension (Coulson et al., 2006). This general theory of language comprehension is developed under the framework of cognitive theories, such as mental space theory, conceptual blending, and frame semantics. It claims that background and contextual information do not merely help to disambiguate the meaning of a lexical item but rather contribute to the meaning-construction process at a discourse level. According to this model, humour comprehension involves a cognitive process called "frame-shifting", in which the currently established discourse representation is reorganised into a new frame retrieved from long-term memory (Coulson, 2015).

The psychological reality of the frame-shifting process has received support from studies using different experimental techniques, such as self-paced reading (Coulson & Kutas, 1998), ERP (Coulson & Kutas, 2001) and eye-tracking (Coulson et al., 2006). More relevant to the current study, Coulson and Kutas (2001) found among the good joke comprehenders that jokes elicited a sustained negativity over left anterior lateral sites 500–900 ms after the onset of the punch word, which they assumed to be an index reflecting the frame-shifting process needed to re-establish coherence. In a later study exploring the hemispheric asymmetries in joke comprehension, Coulson and Lovett (2004) also observed such a sustained frontal negativity within the same time window, which was more prominent in the right-handed participants than the left-handers. According to their explanation, this sustained frontal negativity might reflect the updating of mental models in working memory and/or suppression of joke-irrelevant information.

Besides the two above-mentioned ERP components, the LPC (late positive complex) is the other ERP component of particular interest to the current investigation. The LPC is a positive-going wave with a central posterior topography that reaches the maximum value between 600 and 800 ms post-stimulus onset. It is usually taken as an index for extra integration efforts to integrate additional information from semantic memory (Coulson & Van Petten, 2002; Rataj et al., 2018). The LPC and P600 are sometimes used interchangeably in the literature. However, increasing evidence is showing that the former is more semantic in nature while the latter reflects extra syntactic effort (Bornkessel-Schlesewsky & Schlesewsky, 2008; van Herten et al., 2005). To avoid confusion, hereafter, we use LPC as a cover term for a family of semantic-related positive deflections after the N400 time window, and P600 as the classic syntactically-elicited positivity (Osterhout & Holcomb, 1992). Previous ERP studies on verbal jokes have reported that jokes elicited enhanced LPC amplitudes relative to their non-joke controls, indicating that LPC amplitudes could be related to the resolution of incongruity or processes related to the "frame-shifting" in the integration stage (Coulson, 2001; Feng et al., 2014; Ku et al., 2017).

Inspired by the SSM theory and the salience effect reported in previous studies (e.g. Zheng et al., 2020), we postulate that successful pun comprehension involves sequential access to the two relevant meanings of the pun word, the second of which (usually the more salient one) triggers the establishment of a new cognitive model in the working memory (i.e. frame-shifting) in parallel with the first one. Two points need to be further illustrated here. Firstly, the current study adopted the stance proposed by the graded salience hypothesis that the meaning-access process of a lexical item is sequential, which is modulated by the saliency of its different meanings (Giora, 1997). Accordingly, we believe that both meanings of the pun word are sequentially accessed and then automatically integrated into the pun context, generating two plausible propositions rather than a more conscious process triggered by incongruity (Zheng et al., 2020). In addition, we believe that both relevant meanings of the pun word should be retained and integrated into the sentence context, and these extra cognitive operations should be captured by the ERP technique.

To test our assumptions, in the current study, we directly recorded the ERPs elicited by the sentence-final homophones in three conditions: homophone-salient sentences, homophone-pun sentences, and homophone-error sentences (Refer to Table 1 in the Methods section for the exemplar sentences). Among these three conditions, the homophone-salient condition involves neither incongruity nor frame-shifting; the homophone-pun condition is supposed to trigger a frame-shifting process but incongruity; the homophone-error condition is expected to elicit incongruity responses. As a result, we expected that the homophone-error condition would elicit larger N400 amplitudes than the other two conditions where no incongruity was involved. In addition, we predicted that the frame-shifting process involved in the pun condition would be reflected in the emergence of a sustained frontal negativity and that the additional integration effort would elicit enhanced LPC amplitudes.

To date, the investigation of Dholakia et al. (2016) is the only study that has directly measured the ERPs evoked by a pun word. However, as their primary research interest is the so-called subordinate effect (Kambe et al., 2001; Sheridan et al., 2009), they only analysed the N400 component to examine the meaning-

Table 1. Exemplar sentences for the three sentence conditions.

Condition	Example sentence	Context noun	Homophone	Question
Homophone-salient	陈氏男科医院，我们解决您的难题。(Chen's andrology hospital, we solve your difficulties.)	男科 (andrology)	难题 (difficulty)	这家医院专门治疗儿童疾病？(This hospital specialises in children's diseases?)
Homophone-pun	陈氏男科医院，我们解决您的男题。(Chen's andrology hospital, we solve your male problems.)	男科 (andrology)	男题 (male problem)	
Homophone-error	陈氏牙科医院，我们解决您的男题。(Chen's dentistry hospital, we solve your male problems.)	牙科 (dentistry)	男题 (male problem)	

Note: English translations were given in parentheses.

access process without any analysis of later components that may shed more light on the following integration process. Therefore, the current work would be the first ERP report focusing on the integration process of pun comprehension. In addition, findings from the present study could provide new evidence to verify humour theories, especially those built upon findings from joke studies. Although jokes and puns are two different kinds of verbal humour, both of their comprehension involve a frame-shifting process. Namely, the receiver of a joke or a pun needs to mentally switch from one semantic frame to the other. As a result, humour theories built upon jokes can shed light on the processing of puns, especially given the little ERP literature on pun comprehension. On the other hand, a critical difference between jokes and puns should be noted as well: when understanding a joke, the original meaning/interpretation needs to be suppressed while no such operations are needed in the case of puns because both relevant meanings need to be retained to get the pun. From this perspective, it can be beneficial to compare the ERP components elicited in these two types of humour, which may shed new light on the current humour theories.

Methods

Participants

A group of 34 native Chinese speakers (20 males and 14 females, mean age = 23.2, *SD* = 2.0) participated in the ERP experiment. All the participants were students studying at a key university in China and were recruited through the campus forum. On the recruiting post, the potential participants were informed about the typical ERP experiment procedures and requirements for a suitable participant. None of the participants reported any history of neurological or psychiatric disorders. They were all right-handed with normal or corrected-to-normal vision. After the experiment, they were paid 70 *yuan* for their participation. The experiment received approval from the Research Ethics Board of the university. Data from three participants were removed because of excessive artefacts or slow drifts; hence, 31 data sets were kept for further analysis (20 males and 11 females, mean age = 23.3, *SD* = 2.1).

Materials

A total of 72 homophone puns (e.g. 陈氏男科医院，我们解决您的男题。*Chen's andrology hospital, we solve your male problems*) were adopted from Zheng and Wang's (2022) eye-tracking study with slight modifications to suit the current research paradigm: the critical homophones were placed at the sentence-final position. The homophone puns were separated with a comma into two sections: the setup section and the homophone section. The setup section introduces the overall background of the whole sentence, containing a critical noun (e.g. 男科, *andrology*) that supports the meaning of the less salient homophone (e.g. 男题, *male problem*); the homophone section (similar to a punch line) generates the punning effect, containing a less salient homophone whose unpresented salient homophone mate (e.g. 难题, *difficulty*) is also semantically congruent with the overall sentence context. To decide which of the two involved homophones is more saliently related to the homophonic sound (e.g. /nan$_2$ ti$_2$/), a group of 30 college students who did not attend the ERP experiment was recruited to write down the first word that they could think of according to the Chinese *pinyin*. The homophone written down by more than 90% of the participants was regarded as salient and the others as less salient.

The other two conditions were created by changing only one Chinese character or word from the homophone-pun condition to minimise confounding effects that could arise from lexical differences. The homophone-salient condition was created by replacing the less salient homophone with its salient homophone mate, and the homophone-error condition was created by changing the critical noun (e.g. 男科, *andrology*) into an unrelated noun (牙科, *dentistry*). See Table 1 for an exemplar sentence for each condition.

Table 2. Properties of the critical context noun used in the three conditions.

Sentence condition	Word frequency	Stroke number	Semantic relatedness
Homophone-salient	20.54	16.93	2.35
Homophone-pun	20.54	16.93	3.69
Homophone-error	22.03	17.18	1.95

Note: "Semantic Relatedness" refers to the extent to which the critical context noun (e.g. 牙科, *dentistry*) is semantically related to the visually-presented homophone (e.g. 男题, *male problem*). Word frequency is measured in occurrences per million.

Word frequency and stroke number of the critical context noun were closely matched between the homophone-error and homophone-pun conditions ($ps > .1$).[3] In addition, rating scores of their semantic relatedness with the homophones were acquired through an online questionnaire from another group of 30 college students (19 males and 11 females, mean age = 21.6, $SD = 1.9$) using a 5-point Likert scale (1 for highly unrelated and 5 for highly related). According to the results, the semantic relatedness between the critical context nouns and the visually-presented homophones was rated significantly higher in the pun condition than those in the homophone-salient and homophone-error conditions ($ps < .001$; Table 2). To test whether the homophone-pun sentences could be understood equally well as the homophone-salient sentences, readability ratings of the three sentence conditions were also obtained through an online questionnaire from a third group of 30 college students (17 males and 13 females, mean age = 22.5, $SD = 2.1$). The rating results showed that the homophone-salient sentences were rated equally understandable ($M = 3.87$, $SD = .31$) as the pun sentences ($M = 3.96$, $SD = .37$, $p > .10$). And as expected, the homophone-error sentences were rated significantly more difficult to understand ($M = 3.46$, $SD = .43$, $p < .001$) than the other two conditions.

Three counterbalanced lists were created based on the 72 sentence-triads using a Latin square design. In addition to the experimental trials, 88 filler sentences were added to each list to lower the probability of homophone puns. They were of various types chosen from newspaper headlines or the internet. (e.g. 陪娃儿写作业固然重要，但也不能做狮吼妈妈。*Accompanying children to do homework is important, but you should not become a roaring mom*). As a result, each participant read 160 sentences in the experiment (24 homophone-salient sentences, 24 homophone-pun sentences, 24 homophone-error sentences, and 88 filler sentences).

Procedures

After the participants arrived at the lab, they were required to wash and blow-dry their hair to improve the electric conduction between the electrodes and the scalp. The participants were seated comfortably in a dimly lit and sound-attenuated booth facing a computer screen around 85 cm away, resting their left hand on the table and right hand on a computer mouse.

In each trial, a fixation sign "+" would appear for 500 ms at the centre of the computer screen to indicate the beginning of the trial. After a 500-ms blank screen, the first half of the experimental sentence was presented to the participants for 1,800–3,000 ms based on its length. The second part of the sentence was then presented on a word-by-word basis using the RSVP (Rapid Sequential Visual Presentation) technique to better time-lock the critical final word. Each word segment lasted for 600 ms with a blank screen of 500 ms between them, except for the last one, which lasted for 1,200 ms (see Figure 1 for the illustration of an experiment trial). After the presentation of the whole sentence, a screen of the 5-point Likert scale would appear, on which the participants rated the funniness of the sentence. Once they mouse-clicked on any of the five scores, a yes/no comprehension question concerning the overall sentence meaning would appear on one-quarter of the trials to encourage a better comprehension of the texts (See also Table 1 for an example question). The participants were required to respond by clicking the "yes" or "no" button beneath the question.

The participants were first given six practice trials to familiarise themselves with the experimental procedures. Then the same set of practice trials was run again, during which the participants were informed about how refrainment from blinking and bodily moments could help the experimenter collect better EEG data. The experimental trials were randomly separated into four 40-trial blocks, and the participants took a short break after each block. For each participant, the experiment tasks lasted for approximately 40 minutes. E-Prime 3.0 was used to present the experimental materials, and an in-line script was added to the experiment script so that the experiment could be conveniently interrupted when necessary.

Electrophysiological recording and data processing

Scalp voltages were continuously recorded using the actiCHamp Plus amplifier (Brain Product, Inc.) with an elastic 64-channel cap. The electrode sites followed the international 10–20 system. The impedance in all electrodes was kept below 5 kΩ, and the amplified analogue voltages were sampled at 1000 Hz without any online filters. EEG was measured online with reference to the FCz electrode and with a ground electrode on the medial frontal site. Vertical EOGs were recorded with

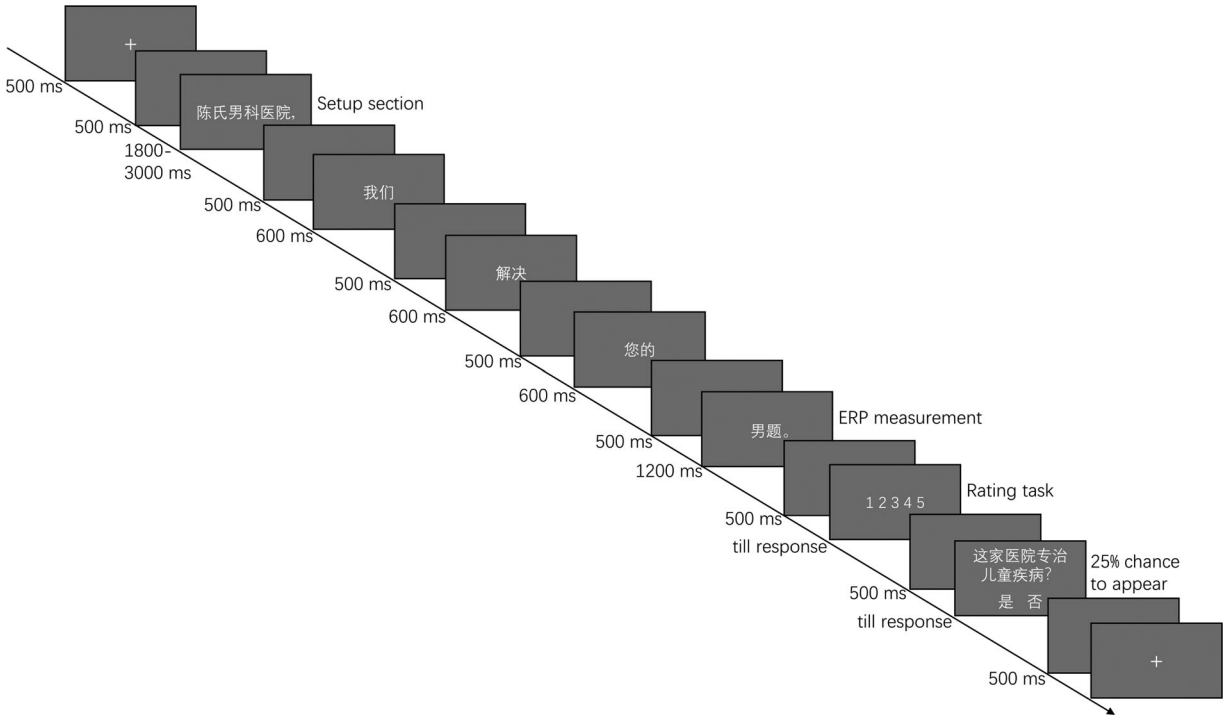

Figure 1. An illustration of the experimental paradigm.

an electrode (AF1) placed below the right eye and the horizontal EOGs (AF2) at the left canthus. The EEG data were then analysed offline with a re-reference to an average of the left and right mastoids (TP9 and TP10). As a result, data from the 60 channels other than AF1, AF2, TP9, and TP10 were used for further ERP analysis.

The EEG signal was analysed with the EEGLAB software (version 13.0.0; Delorme & Makeig, 2004) in the following steps. The EEG signal was first digitally filtered using a bandpass of 0.1–30 Hz (24 dB/oct). The epochs for ERP analyses were then extracted for the last word of each sentence, with a 200 ms pre-stimulus baseline and 1,200 ms total duration. Then, Independent Component Analysis (ICA) was performed on the segmented data. Components indicating artefacts of eye movement were visually identified and removed using the "ICLabel" plugin. In addition, artefacts with amplitudes greater than ±75 mV were removed from further analyses. The ERP waveforms for each condition were separately computed within each participant and then grand-averaged across participants.

Statistical analyses

Since a between-item design has been used between the homophone-salient and homophone-pun conditions, single-trail analyses on the EEG data were conducted in the R environment (Version 4.2.2; R Development Core Team, 2022) using linear mixed effect models (LMEs) with the lme4 package (Bates et al., 2015). Compared with traditional ANOVA analyses, this approach is more advantageous because it can simultaneously model the fixed effects and random variations associated with participants and stimuli, making estimates of the fixed effects more accurate (Brown, 2021). In addition, LMEs can fit an individual regression model for each participant and assign different weights to participants according to their contribution to the mean. This partial pooling algorithm in LMEs allows a researcher to include participants who have only a few usable trials, hence more statistical power (Heise et al., 2022).

In the analyses, sentence *Condition* (homophone-salient, homophone-pun and homophone-error) and scalp distribution *ROI* were defined as the fixed effects, and *Participant* and *Item* as the random factors. Following the recommendation of Barr et al. (2013) and the current experimental design, we started the model with a random intercept by participant and item, and a random slope by item. The models were built using the following formula: DV ~ Condition*ROI + (1 + Condition | Participant) + (1 | Item). The significance of predictors and their interactions were computed using the *Mixed* function from the *afex* package (Singmann et al., 2023) and post hoc pairwise comparisons were performed with Bonferroni correction using the *emmeans* function from the *emmeans* package (Lenth, 2022). Following conventions, *t* values greater than 2 were treated as significant. This is because the *t*-statistic

generally corresponds to the *z*-statistic, considering the small number of fixed and random effects and the large number of observations estimated.

Results

Behavioural results

The mean accuracy of the comprehension questions was 90.1% ($SD = 5.2$), suggesting that the participants had generally understood the contents of the sentences. The mean humour-rating score for all the experimental sentences was 2.97 ($SD = 1.00$), indicating that the experimental materials do not skew toward the humourous sentences.

Repeated-measure ANOVA reveals significant differences between the three sentence conditions [$F(2, 60) = 70.74$, $p < .001$, $\eta_p^2 = .70$]. Pairwise comparison with Bonferroni correction revealed that the participants rated the homophone puns as significantly funnier ($M = 3.29$, $SD = .51$) than both the homophone-salient ($M = 2.53$, $SD = .39$, $p < .001$, Cohen's $d = 1.67$) and the homophone-error conditions ($M = 3.04$, $SD = .49$, $p = .001$, Cohen's $d = .50$).

ERP results

Based on the previous study by Coulson and Severens (2007), two time windows time-locked to the onset of sentence-final words were selected for the examination of the N400 effect (300–500 ms) and the LPC effect (600–900 ms). To examine the distribution of these two effects, 7 ROI were set up (left anterior: AF3, F1, F3, F5, FC1, FC3; right anterior: AF4, F2, F4, F6, FC2, FC4; left central: C1, C3, C5, CP1, CP3, CP5; right central: C2, C4, C6, CP2, CP4, CP6; left posterior: P3, P5, P7, PO3, PO7; right posterior: P4, P6, P8, PO4, PO8; midline: Fz, FCz, Cz, CPz, POz). For the analysis of the sustained negativity (500–900 ms), two left lateral channels (F7, FT7) were chosen to compare with the previous studies (Coulson & Kutas, 2001; Coulson & Lovett, 2004).

The mean amplitudes of each ROI during the three time windows were computed and used as the dependent variables. The mean amplitudes, grand-averaged ERP waveforms and topographic maps for the three sentence conditions (homophone-salient, homophone-pun and homophone-error) are illustrated in Figure 2, Figure 3, and Figure 4 respectively.

The N400 component (300–500 ms)

Likelihood-ratio tests indicated that the main effect of Condition ($\chi^2(2) = 10.74$, $p = .005$) and ROI ($\chi^2(6) = 992.38$, $p < .001$) were significant. More importantly, the interaction between Condition and ROI ($\chi^2(12) = 22.38$, $p = .033$) also provided a better fit for the data than a model without it.

The Bonferroni-corrected pairwise comparisons revealed that the homophone-error condition elicited more negative-going N400 amplitudes than the homophone-pun condition in the ROIs of left central ($z = -3.04$, $p = .007$), right central ($z = -2.97$, $p = .009$), and the midline ($z = -2.51$, $p = .037$), in line with the typical central distribution of the N400. Meanwhile, there was no significant difference between the homophone-pun sentences and the homophone-salient sentences in all the ROIs (See Table 3 for more details).

The sustained negativity/positivity (500–900 ms)

A likelihood-ratio test showed that there was a significant main effect of Condition ($\chi^2(2) = 14.75$, $p < .001$). However, Bonferroni-corrected pairwise comparisons revealed an unexpected pattern that the homophone-salient condition elicited more negative amplitudes than the homophone-pun condition ($z = -3.20$, $p = .004$), while there was no significant difference in amplitudes between the homophone-pun and homophone-error condition ($z = .23$, $p = 1.00$). To put it differently, both the homophone-pun and homophone-error conditions elicited a sustained positivity in the left anterior lateral sites.

To better understand the possible relationship of this sustained positivity with frame-shifting, we built another linear mixed-effect model. For simplicity, we collapsed the 5-point funniness rating scores into three groups: Low (1, 2), Middle (3) and High (4, 5). Then the rating scores were used as the predictor (*Fun*) to model the amplitudes of the sustained positivity in this time window, using the "Low" group as the baseline. According to the likelihood-ratio test, there was a significant main effect of Fun ($\chi^2(2) = 6.41$, $p = .041$). Compared with the baseline ($\beta = -1.00$, $SE = 0.47$, $t = -2.14$, $p = .039$), the "High" group elicited greater positive amplitudes ($\beta = 1.10$, $SE = .45$, $t = 2.43$, $p = .022$), followed by the "Middle" group ($\beta = .74$, $SE = .36$, $t = 2.07$, $p = .047$). In another word, the sustained positivity was positively correlated with the funniness rating scores.

The LPC component (600–900 ms)

Likelihood-ratio tests indicated that the main effect of Condition ($\chi^2(2) = 19.89$, $p < .001$) and ROI ($\chi^2(6) = 1532.43$, $p < .001$) were significant. In addition, the interaction between Condition and ROI ($\chi^2(12) = 53.40$, $p < .001$) also provided a better fit for the data than a model without it.

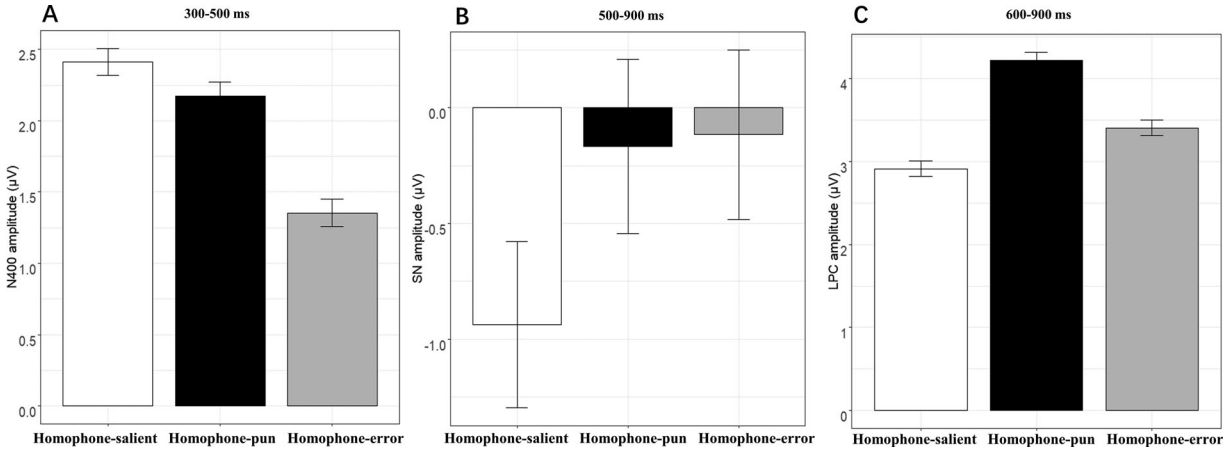

Figure 2. Mean amplitudes of the N400 (Panel A), sustained positivity (Panel B) and LPC (Panel C) for the three sentence conditions. The N400 and LPC are computed by averaging values in the central, parietal and midline ROIs. The sustained positivity is computed by averaging values at the F7 and FT7 electrodes. Error bars indicate the standard errors of the mean.

The pairwise comparisons with Bonferroni correction revealed that the homophone-pun condition elicited significantly larger LPC amplitudes than both the homophone-salient and homophone-error conditions in the central, posterior and midline regions[4] ($ps < .05$). Numerically, the LPC component was larger over the right hemisphere ($M = 5.00$, 4.74 and 4.16 on the right central, right posterior and midline ROIs respectively)

We also ran an analysis on the relationship between the LPC and funniness rating scores using data from

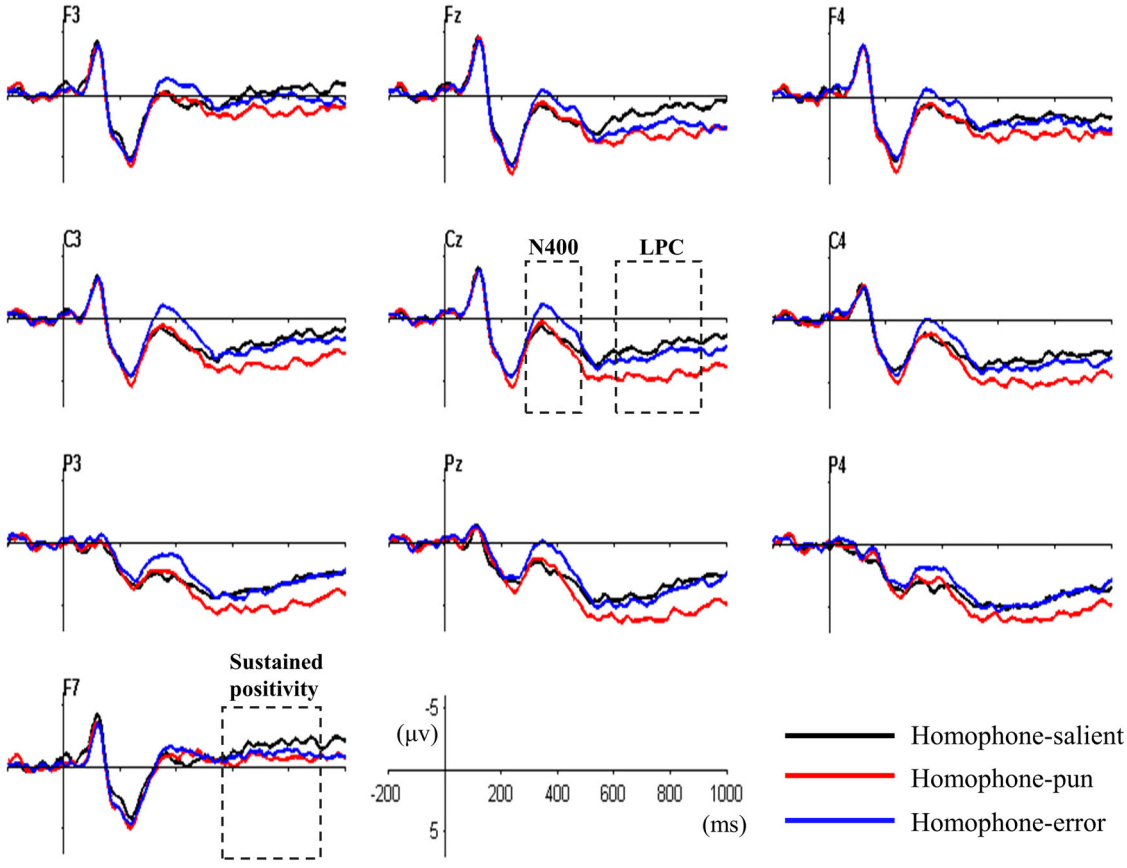

Figure 3. Grand average ERP response to the sentence-final homophones at ten representative electrodes for the three sentence conditions. Nine electrodes were chosen to reflect laterality (left: F3, C3, P3; middle: Fz, Cz, Pz; and right: F4, C4, P4) and brain region (anterior: F3, Fz, F4, central: C3, Cz, C4; and posterior: P3, Pz, P4). One additional electrode (F7) was chosen to capture the sustained negativity reported by Coulson and Kutas (2001).

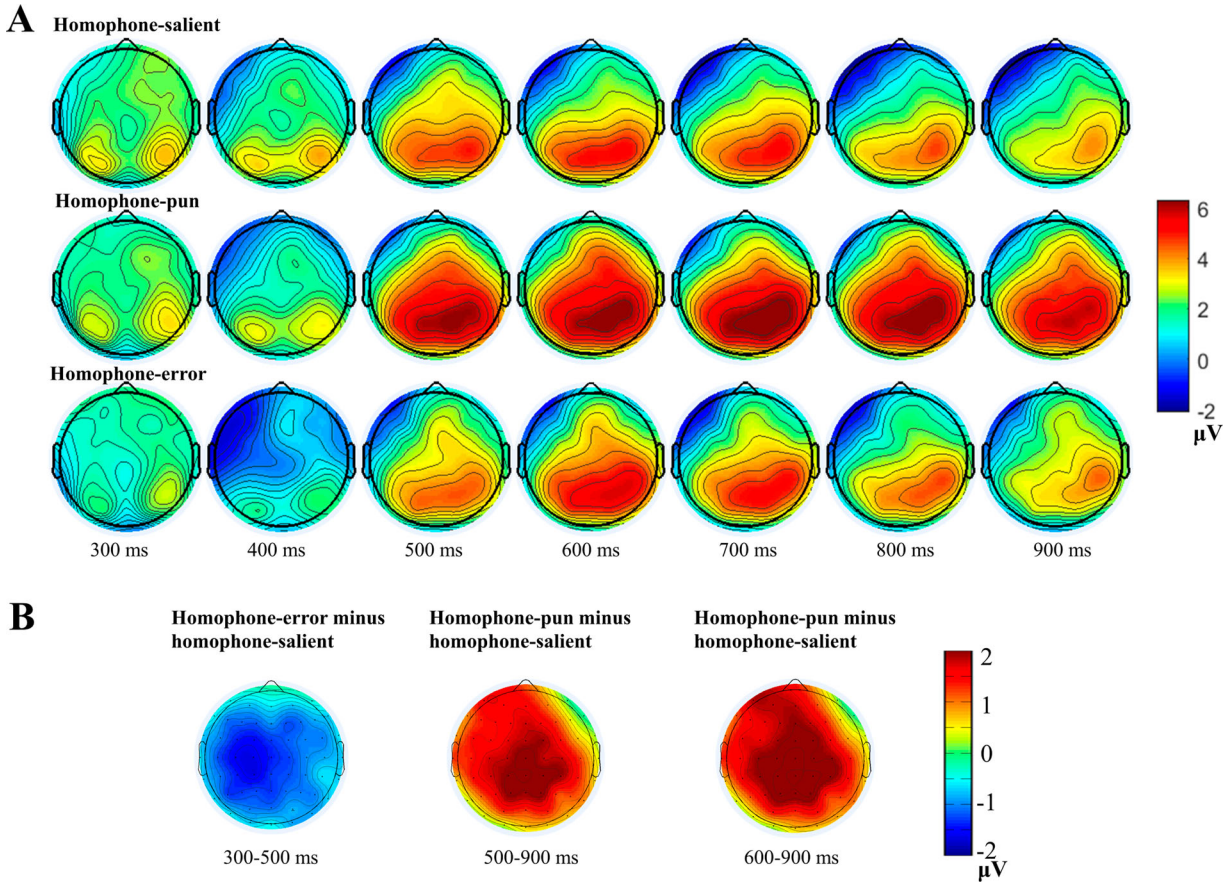

Figure 4. Topographic plots for the three sentence conditions from 300 to 900 ms after stimulus onset (Panel A). Topographic plots for the ERP difference between different sentence conditions during the N400, sustained positivity and LPC time windows (Panel B).

three representative ROIs (right central, right posterior and midline), in consideration of the right-lateralised scalp distribution of the LPC. A likelihood-ratio test suggested no significant main effect of Fun ($\chi^2(2) = 4.49$, $p = .106$). However, there was still a trend showing that the LPC amplitudes were also positively correlated with the funniness rating scores. More specifically, compared with the baseline ($\beta = 3.55$, $SE = .388$, $t = 9.44$, $p < .001$), both the "High" group ($\beta = .49$, $SE = .26$, $t = 1.92$, $p = .064$) and the "Middle" group ($\beta = .52$, $SE = .29$, $t = 1.77$, $p = .087$) elicited larger LPC amplitudes.

Discussion

In the current study, we investigated the temporal dynamics of pun comprehension by directly recording the ERP amplitudes elicited by Chinese homophonic words. Unlike previous studies on pun comprehension, the current study aimed at exploring how the two meanings related to a pun word were integrated into the pun context. The behaviour results showed that the participants rated the homophone puns as more humourous than the other two conditions, consistent with the idea that puns are an important type of verbal humour. The ERP findings revealed that only the less salient homophone in the homophone-error condition but not the homophone-pun condition evoked significantly larger N400 amplitude than its salient homophone mate in the homophone-salient condition. Unexpectedly, we failed to observe a sustained negativity (500–900 ms after stimulus onset), which was assumed to reflect the framing-shifting process in understanding jokes. Instead, we found in this time window a sustained positivity in both the homophone-pun and homophone-error conditions. In addition, a larger LPC component was found in the homophone-pun condition with a central-parietal distribution, probably suggesting more integration operations were involved. The current results were consistent with the Space Structure Model, which suggests "frame-shifting" as a crucial component in understanding humorous materials.

Incongruity and pun humour

Incongruity is a crucial component in many modern humour theories (Attardo & Raskin, 1991; Suls, 1972).

Table 3. Pairwise comparisons between the homophone-pun condition and the other two conditions for each ROI in three time-windows.

			Homophone-salient minus Homophone-pun					Homophone-error minus Homophone-pun				
ERP component	Time (ms)	ROI	Mean diff.	*SE*	*Cohen's d*	*z*	*p*	Mean diff.	*SE*	*Cohen's d*	*z*	*p*
N400	300–500	left anterior	0.57	0.29	0.09	1.95	0.155	−0.48	0.30	−0.08	−1.59	0.333
		right anterior	0.33	0.29	0.05	1.14	0.760	−0.59	0.30	−0.09	−1.95	0.155
		left central	0.49	0.29	0.08	1.67	0.287	−0.91	0.30	−0.15	−3.04	0.007**
		right central	0.11	0.29	0.02	0.36	1.000	−0.90	0.30	−0.15	−2.97	0.009**
		left posterior	0.55	0.30	0.09	1.86	0.190	−0.32	0.31	−0.05	−1.03	0.907
		right posterior	0.26	0.30	0.04	0.88	1.000	−0.56	0.31	−0.09	−1.83	0.202
		midline	0.46	0.30	0.07	1.53	0.381	−0.77	0.31	−0.12	−2.51	0.037*
Sustained positivity	500–900	F7 & FT7	−0.82	0.26	−0.13	−3.20	0.004**	0.06	0.26	0.01	0.23	1.000
LPC	600–900	Left anterior	−1.42	0.30	−0.22	−4.74	0.000**	−0.35	0.26	−0.05	−1.35	0.530
		Right anterior	−1.34	0.30	−0.20	−4.45	0.000**	−0.56	0.26	−0.08	−2.13	0.100
		Left central	−1.59	0.30	−0.24	−5.29	0.000**	−0.81	0.26	−0.12	−3.11	0.006**
		Right central	−1.64	0.30	−0.25	−5.48	0.000**	−0.89	0.26	−0.14	−3.39	0.002**
		Left posterior	−0.84	0.31	−0.13	−2.72	0.019*	−0.70	0.27	−0.11	−2.61	0.027*
		Right posterior	−0.92	0.31	−0.14	−3.00	0.008**	−1.04	0.27	−0.16	−3.86	0.000**
		Midline	−1.58	0.31	−0.24	−5.13	0.000**	−0.75	0.27	−0.11	−2.78	0.016*

Note: The pairwise comparisons were adjusted with Bonferroni correction.
*$p < .05$; **$p < .01$.

Previous ERP studies on jokes have reported enhanced N400 amplitudes in jokes relative to non-funny controls, which are usually taken as evidence for incongruity detection (Coulson & Kutas, 2001; Coulson & Williams, 2005; Feng et al., 2014). However, Coulson and Lovett (2004) failed to find such an effect using one-line jokes of low constraint. In addition, Dholakia et al. (2016) also did not find such an "N400 joke effect" when comparing the same sentence-final homographs embedded in pun contexts and neutral contexts. On the contrary, they observed that the homographs in puns elicited even smaller N400 amplitudes.

In line with Dholakia et al. (2016), we also observed a similar facilitative effect in homophone puns. To be more specific, we found in the central regions that the less salient homophones in the homophone-pun sentences elicited significantly smaller N400 amplitudes than they were in the homophone-error sentences (the "incongruity" condition). In the current experiment, many of the less salient homophones were temporarily-composed compound words (e.g. 男题, *male problem*). Therefore, the increased N400 amplitudes probably indicated that the participants had registered these novel stimuli as "incongruity" due to the lack of contextual support. While the much smaller N400 component in the homophone-pun condition indicated that the participants were not so surprised by the same homophones. Moreover, it was interesting to find that the less salient homophones in the homophone-pun condition did not elicit larger N400 amplitudes than their salient homophone mates in the homophone-salient condition. Many previous ERP studies have demonstrated that less salient or low-frequency words elicited a larger N400 component than more salient or frequent words (Besson et al., 1992; Van Petten & Kutas, 1990). In the current study, we believe that the facilitative effect of the pun context may have boosted the processing of the less salient homophones, making them as easy to be processed as their salient homophone mates.

In light of the absence of the N400 effect in the homophone-pun condition, the current results suggested no evident incongruity detection during pun comprehension. As a result, this study supported the idea that incongruity may not be an antecedent for humourous effects (Veale, 2004). In a recent meta-analysis of humour research, Warren et al. (2021) also noted that the term "incongruity" was used among scholars with different connotations. Many theories use this term as a synonym for "surprise", referring to clashes with expectations or previously established schemas (Suls, 1972). Instead, these authors argued that "simultaneity" (entertaining contrasting perceptions or interpretations at the same time) rather than "incongruity" is an antecedent for humour. As a result, the "incongruity" in incongruity-based accounts should not be taken as a synonym for "surprise" to account for the current findings from homophone puns.

Instead of "incongruity", we believe that the cognitive advantage of the salient meaning plays a crucial role in triggering a pun interpretation. The salience advantage is highlighted in the Graded salience hypothesis (Giora, 1997, 2003). According to this hypothesis, the salient meaning of a word or phrase can become so entrenched in our mind (due to its high frequency, familiarity, or conventionality) that it will always be activated automatically, regardless of its contextual fit. This prediction has been supported by most recent psycholinguistic studies on puns (e.g. Coulson & Severens, 2007; Koleva et al., 2019), as well as the humour rating scores in this

study. In the current experiment, the homophone-salient and homophone-pun conditions differed only in the visually-presented homophones: salient homophones were presented in the homophone-salient condition, and the less salient ones in the homophone-pun condition. Therefore, the much higher humour rating scores towards the homophone puns indicated that meanings of the salient homophones were more recoverable through shared phonology even if they were not presented visually, leading to the greater humour response to the homophone-pun condition than the homophone-salient condition (See also Zheng & Wang, 2022).

In this line of thought, what is unique about a pun context is that it creates a delicate balance between the cognitive advantage of the salient meaning and the contextual support for the less salient meaning through visual, aural, or semantic cues. This special contextual configuration helps to generate and retain two plausible interpretations, in contrast with jokes where only one interpretation is eventually adopted.

Sustained positivity, LPC and frame-shifting

The absence of sustained frontal negativities in the current study seemed surprising at the first glance, but it could reflect the distinct difference between the comprehension of jokes and puns. To understand the following joke, for example, *My grandfather has the heart of a lion and a lifetime ban at the zoo,* the reader not only has to shift from the "bravery" frame to the "animal-killing" frame but also has to suppress the metaphorical meaning of "heart of a lion". In contrast, the reader only needs to shift from the "balance" frame to the "physical energy" frame to get the pun mentioned at the beginning of this paper (*A bicycle cannot stand on its own because it is two-tired*). In their studies on joke comprehension, Coulson and colleagues reported a sustained negativity over the left lateral frontal sites, which they assumed to reflect the frame-shifting process (Coulson & Kutas, 2001; Coulson & Lovett, 2004). However, the wave shape and topography of this sustained negativity were similar to the slow-rising positive drift around the same left anterior sites, associated with the construction of a mental model in the working memory (Kutas & King, 1996). According to Coulson and Lovett (2004), this sustained effect could reflect not only the manipulation of information in working memory but also the suppression of earlier expectations derived from the joke context, hence the negative modulation of the positivity. As a result, the sustained positivities we observed in the pun condition relative to the homophone-salient condition could reflect the frame-shifting process without the suppression manipulation in the working memory. This speculation received some support from the significant positive correlation between the funniness rating scores and the amplitudes of this sustained positivity.

Meanwhile, the sustained positivity in the homophone-error condition implies that the participants could have undergone a similar frame-shifting process as in the homophone-pun condition. In the homophone-error condition, most of the less salient homophones were temporary words composed of two Chinese characters, which were difficult to understand without sufficient support from the sentence context. Given that the participants were required to finish a comprehension task, it was quite likely that they had made some sense of the homophone-error sentences with the meaning of the salient homophones through their shared homophonic sound. Evidence for this speculation can also be found in the readability rating scores. Although the readability ratings towards the homophone-error condition ($M = 3.46$) were significantly lower than those of the homophone-salient condition and the homophone-pun conditions, it still indicated that the homophone-error sentences were understandable to some extent. Indeed, during a conversation after the experiment, one participant mentioned that there may be some misspellings in the experimental materials; however, he could still figure out the meanings of those sentences. Of course, there are still other possibilities for the sustained positivity in the homophone-error condition. As one of the reviewers suggested that this frontal positivity could simply reflect the participants' effort to access the meaning of the salient homophones without the frame-shifting process. Therefore, future studies are still needed to further explore the nature of this ERP component.

The increased LPC amplitudes in response to the pun sentences relative to the other two conditions were expected, which could reflect the extra cognitive effort to integrate the second meaning into the sentence context. Coulson and Van Petten (2002) also reported increased LPC amplitudes on words when used metaphorically rather than literally (e.g. *Their last shot in that game was only a prayer*; *Before dinner, there was almost always a prayer*). They believed that the comprehension of a metaphor involved the retrieval of some new aspects of the conceptual structure different from its corresponding literal sentence. As a result, the increased LPC amplitudes were assumed to index the recovery and integration of additional information from semantic memory. In line with these authors, we interpret the current findings on the increased LPC amplitudes in the homophone-pun condition as reflecting the extra integration effort that the participants had

to make for integrating the second meaning into the context (See also Brouwer et al., 2017).

Enhanced LPC amplitudes were also found in jokes relative to their unfunny controls in some previous ERP studies, which were assumed to reflect the resolution of incongruity (Coulson & Kutas, 2001; Feng et al., 2014; Marinkovic et al., 2011). For example, Ku et al. (2017) found that jokes elicited a larger LPC among the good comprehenders, who gave higher scores in the comprehensibility rating task. They speculated that it could reflect the good comprehenders' success in re-establishing coherence to resolve the incongruity at punch lines. In the current study, we found a trend of positive correlation between the funniness rating scores and the LPC amplitudes. Namely, the funnier the participants rated the sentence, the larger the LPC component would be. However, we assume the enhanced positive amplitudes reflect the integration of the second meaning into the pun context, i.e. the establishment of an additional frame needed to get the pun, rather than the resolution of incongruity. And the extra cognitive effort to find relevance between the second meaning and sentence context could be the cause leading to the greater amusement in the pun condition (Sperber & Wilson, 1986).

To put it together, the sustained frontal positivity and the enhanced posterior LPC amplitudes could index two crucial processes in understanding a pun. More specifically, the sustained positivity could reflect the retrieval of the second meaning from a different frame (frame-shifting), without suppression of the first meaning. According to its waveform and distribution, this sustained positivity is similar to the slow positive drift associated with working memory operations (Kutas & King, 1996). Meanwhile, the enhanced LPC amplitudes could reflect the increased cognitive load to integrate the additional meaning.

Two limitations should be noted about the current study. Firstly, we only examined ERP components within a comparatively short time window (300–900 ms). Feng et al. (2014) proposed a three-stage model for humour processing, including incongruity detection (indexed by the N400, 350–500 ms), incongruity resolution (indexed by the LPC, 500–700 ms), and humour appreciation (reflected by the LPC, 800–1500 ms). In fact, this cognition-affection distinction is reasonable. There has already been neurological imaging evidence indicating that the regions related to language processing, such as the bilateral inferior frontal gyri, are more activated at an early stage when reading jokes; while the regions connected with emotions (e.g. the bilateral amygdalae) are more involved at a later stage (Chan et al., 2013). In this sense, the current investigation has mainly examined the cognitive processes during pun comprehension without much investigation of its affective aspects. As a result, future studies could further explore the affective processes of pun comprehension. Secondly, the funniness rating task could, to some extent, have lowered the participants' sensitivity to incongruity. In the current study, the participants were required to rate the funniness of each sentence so that the possible relationships between pun appreciation and ERP components could be assessed. However, this requirement could have promoted a humourous mindset in some participants and lowered their sensitivity to incongruity. Hence, the funniness rating task could have reduced the overall N400 amplitudes. As a result, future studies can use no explicit behaviour task or just a comprehension task to compare with the current findings.

Conclusion

In summary, the absence of the N400 effect in the homophone-pun condition indicates that "incongruity" may not be a prerequisite for the generation of pun humour or at least this term should be distinguished from "surprise". In contrast, the frame-shifting process proposed by the Space Structure model is more consistent with the current findings. The current results suggest that the special configuration of the pun context enables the reader to switch from one meaning to the other (usually from the less salient meaning to the salient one), which could be reflected by the sustained positivity in the left anterior regions. While operations to integrate the additional meaning could be reflected by the enhanced LPC amplitudes around the central-parietal regions. It should be noted that the current study is an initial step in the exploration of the frame-shifting process in pun comprehension, which still awaits cross-validation from future studies, especially those on homograph puns.

Notes

1. The salience of a meaning is positively correlated with its conventionality, familiarity, frequency or giveness status in a certain context (Giora, 1997).
2. The cloze probability is usually operationalised as the percentage of participants who have written down a particular word in a sentence context. For example, if 80 out of 100 participants write down the word "butter" in the sentence frame "He spread the warm bread with ___", then the cloze probability of "butter" is 80%.
3. Early eye-tracking measures (i.e. first fixation and gaze duration) have also indicated that the critical context

noun were well-matched among the three conditions at the lexical level (Zheng & Wang, 2022).

4. The differences in the anterior regions were also significant; however, the time-window was largely overlapped with the sustained positivity window, so the differences will not be discussed further.

Acknowledgements

The study was completed in partial fulfilment of the requirements for the first author's doctoral dissertation at Zhejiang University. We would like to thank Professor Phaedra Royle for her suggestions on the design of the ERP experiment and Professor Nai Ding for his help in the data analysis. We would also like to thank the two anonymous reviewers for their insightful comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study is supported by the Major Project of National Social Science Foundation of China (14ZDB155).

ORCID

Wei Zheng http://orcid.org/0000-0002-4133-0714
Xiaolu Wang http://orcid.org/0000-0001-6187-3516

References

Attardo, S. (2001). *Humorous texts: A semantic and pragmatic analysis*. Mouton de Gruyter.

Attardo, S., & Raskin, V. (1991). Script theory revis(it)ed: Joke similarity and joke representation model. *Humour: International Journal of Humour Research*, *4*(3–4), 293–347. https://doi.org/10.1515/humr.1991.4.3-4.293

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. https://doi.org/10.1016/j.jml.2012.11.001

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

Besson, M., Kutas, M., & Vanpetten, C. (1992). An event-related potential (ERP) analysis of semantic congruity and repetition effects in sentences. *Journal of Cognitive Neuroscience*, *4*(2), 132–149. https://doi.org/10.1162/jocn.1992.4.2.132

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews*, *59*(1), 55–73. https://doi.org/10.1016/j.brainresrev.2008.05.003

Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. J. (2017). A neurocomputational model of the N400 and the P600 in language processing. *Cognitive Science*, *41*, 1318–1352. https://doi.org/10.1111/cogs.12461

Brown, V. A. (2021). An introduction to linear mixed-effects modelling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1–19. https://doi.org/10.1177/2515245920960351

Chan, Y. C., Chou, T. L., Chen, H. C., Yeh, Y. C., Lavallee, J. P., Liang, K. C., & Chang, K. E. (2013). Towards a neural circuit model of verbal humour processing: An fMRI study of the neural substrates of incongruity detection and resolution. *Neuroimage*, *66*, 169–176. https://doi.org/10.1016/j.neuroimage.2012.10.019

Coulson, S. (2001). *Semantic leaps: Frame-shifting and conceptual blending in meaning construction*. Cambridge University Press.

Coulson, S. (2015). Frame-shifting and frame semantics: Joke comprehension on the space structuring model. In G. Brône, K. Feyaerts, & T. Veale (Eds.), *Cognitive linguistics and humour research* (pp. 167–190). Mouton de Gruyter.

Coulson, S., & Kutas, M. (1998). *Frame-shifting and sentential integration*. Cognitive Science Technical Report (Vol. 98).

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. (2004). Handedness, hemispheric asymmetries, and joke comprehension. *Cognitive Brain Research*, *19*(3), 275–288. https://doi.org/10.1016/j.cogbrainres.2003.11.015

Coulson, S., & Severens, E. (2007). Hemispheric asymmetry and pun comprehension: When cowboys have sore calves. *Brain and Language*, *100*(2), 172–187. https://doi.org/10.1016/j.bandl.2005.08.009

Coulson, S., Urbach, T. P., & Kutas, M. (2006). Looking back: Joke comprehension and the space structuring model. *Humor – International Journal of Humor Research*, *19*(3), 229–250. https://doi.org/10.1515/HUMOR.2006.013

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

Coulson, S., & Williams, R. F. (2005). Hemispheric asymmetries and joke comprehension. *Neuropsychologia*, *43*(1), 128–141. https://doi.org/10.1016/j.neuropsychologia.2004.03.015

Delorme, A., & Makeig, S. (2004). EEGLAB: An open-source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dholakia, A., Meade, G., & Coch, D. (2016). The N400 elicited by homonyms in puns: Two primes are not better than one. *Psychophysiology*, *53*(12), 1799–1810. https://doi.org/10.1111/psyp.12762

Dien, J., Spencer, K. M., & Donchin, E. (2004). Parsing the late positive complex: Mental chronometry and the ERP components that inhabit the neighborhood of the P300. *Psychophysiology*, *41*(5), 665–678. https://doi.org/10.1111/psyp.2004.41.issue-5

Feng, Y. J., Chan, Y. C., & Chen, H. C. (2014). Specialization of neural mechanisms underlying the three-stage model in humour processing: An ERP study. *Journal of Neurolinguistics*, *32*, 59–70. https://doi.org/10.1016/j.jneuroling.2014.08.007

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, *8*(3), 183–206. https://doi.org/10.1515/cogl.1997.8.3.183

Giora, R. (2003). *On our mind: Salience, context, and figurative language*. Oxford University Press.

Heise, M. J., Mon, S. K., & Bowman, L. C. (2022). Utility of linear mixed effects models for event-related potential research with infants and children. *Developmental Cognitive Neuroscience*, *54*, Article 101070. https://doi.org/10.1016/j.dcn.2022.101070

Kambe, G., Rayner, K., & Duffy, S. A. (2001). Global context effects on processing lexically ambiguous words: Evidence from eye fixations. *Memory & Cognition*, *29*(2), 363–372. https://doi.org/10.3758/BF03194931

Kappenman, E. S., & Luck, S. J. (2012). ERP components: The ups and downs of brainwave recordings. In S. J. Luck & E. S. Kappenman (Eds.), *The oxford handbook of event-related potential components* (pp. 3–30). Oxford University Press.

Koleva, K., Mon-Williams, M., & Klepousniotou, E. (2019). Right hemisphere involvement for pun processing – Effects of idiom decomposition. *Journal of Neurolinguistics*, *51*, 165–183. https://doi.org/10.1016/j.jneuroling.2019.02.002

Ku, L.-C., Feng, Y.-J., Chan, Y.-C., Wu, C.-L., & Chen, H.-C. (2017). A re-visit of three-stage humour processing with readers' surprise, comprehension, and funniness ratings: An ERP study. *Journal of Neurolinguistics*, *42*, 49–62. https://doi.org/10.1016/j.jneuroling.2016.11.008

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–163. https://doi.org/10.1038/307161a0

Kutas, M., & King, J. W. (1996). The potentials for basic sentence processing: Differentiating integrative processes. In T. Innue & J. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 501–546). MIT Press.

Lenth, R. (2022). *emmeans: Estimated marginal means, aka least-squares means* (Version 1.8.3) [Computer software]. https://CRAN.R-project.org/package=emmeans

Marinkovic, K., Baldwin, S., Courtney, M. G., Witzel, T., Dale, A. M., & Halgren, E. (2011). Right hemisphere has the last laugh: Neural dynamics of joke appreciation. *Cognitive, Affective, & Behavioral Neuroscience*, *11*(1), 113–130. https://doi.org/10.3758/s13415-010-0017-7

McGhee, P. E. (1972). Chapter 3 – On the cognitive origins of incongruity humour: Fantasy assimilation versus reality assimilation. In J. H. Goldstein & P. E. McGhee (Eds.), *The psychology of humour* (pp. 61–80). Academic Press.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z

Rataj, K., Przekoracka-Krawczyk, A., & van der Lubbe, R. H. J. (2018). On understanding creative language: The late positive complex and novel metaphor comprehension. *Brain Research*, *1678*, 231–244. https://doi.org/10.1016/j.brainres.2017.10.030

R Development Core Team. (2022). *R: A language and environment for statistical computing* (Version 4.2.2) [Computer software]. R Foundation for Statistical Computing. http://www.R-project.org/

Shen, W., Qu, Q. Q., & Li, X. S. (2016). Semantic information mediates visual attention during spoken word recognition in Chinese: Evidence from the printed-word version of the visual-world paradigm. *Attention, Perception, & Psychophysics*, *78*(5), 1267–1284. https://doi.org/10.3758/s13414-016-1081-z

Sheridan, H., Reingold, E. M., & Daneman, M. (2009). Using puns to study contextual influences on lexical ambiguity resolution: Evidence from eye movements. *Psychonomic Bulletin & Review*, *16*(5), 875–881. https://doi.org/10.3758/PBR.16.5.875

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. (2023). *afex: Analysis of factorial experiments* (Version 1.2.1) [Computer software]. https://CRAN.R-project.org/package=afex

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

Suls, J. M. (1972). A two-stage model for the appreciation of jokes and cartoons: An information-processing analysis. In J. H. Goldstein & P. E. McGhee (Eds.), *The psychology of humour* (pp. 81–100). Academic Press.

van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, *22*(2), 241–255. https://doi.org/10.1016/j.cogbrainres.2004.09.002

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, *18*(4), 380–393. https://doi.org/10.3758/BF03197127

Veale, T. (2004). Incongruity in humor: Root cause or epiphenomenon? *Humor – International Journal of Humor Research*, *17*(4), 419–428. https://doi.org/10.1515/humr.2004.17.4.419

Warren, C., Barsky, A., & McGraw, A. P. (2021). What makes things funny? An integrative review of the antecedents of laughter and amusement. *Personality and Social Psychology Review*, *25*(1), 41–65. https://doi.org/10.1177/1088868320961909

Zheng, W., & Wang, X. (2022). Contextual support for less salient homophones and pun humor appreciation: Evidence from eye movements in reading Chinese homophone puns. *Frontiers in Psychology*, *13*, Article 875479. https://doi.org/10.3389/fpsyg.2022.875479

Zheng, W., Wang, Y., & Wang, X. (2020). The effect of salience on Chinese pun comprehension: A visual world paradigm study. *Frontiers in Psychology*, *11*, Article 00116. https://doi.org/10.3389/fpsyg.2020.00116