



Event-related potential (ERP) evidence for the differential cognitive processing of semantic jokes and pun jokes

Yi-Tzu Chang, Li-Chuan Ku, Ching-Lin Wu & Hsueh-Chih Chen

To cite this article: Yi-Tzu Chang, Li-Chuan Ku, Ching-Lin Wu & Hsueh-Chih Chen (2019) Event-related potential (ERP) evidence for the differential cognitive processing of semantic jokes and pun jokes, *Journal of Cognitive Psychology*, 31:2, 131-144, DOI: [10.1080/20445911.2019.1583241](https://doi.org/10.1080/20445911.2019.1583241)

To link to this article: <https://doi.org/10.1080/20445911.2019.1583241>



Published online: 27 Feb 2019.



Submit your article to this journal [↗](#)






Article views: 43



View Crossmark data [↗](#)



Event-related potential (ERP) evidence for the differential cognitive processing of semantic jokes and pun jokes

Yi-Tzu Chang ^a, Li-Chuan Ku^b, Ching-Lin Wu ^c and Hsueh-Chih Chen ^{a,d,e}

^aDepartment of Educational Psychology and Counseling, National Taiwan Normal University, Taipei, Taiwan; ^bDepartment of Psychology, University of Arizona, Tucson, USA; ^cProgramme of Learning Sciences, National Taiwan Normal University, Taipei, Taiwan; ^dInstitute for Research Excellence in Learning Sciences, National Taiwan Normal University, Taipei, Taiwan; ^eChinese Language and Technology Centre, National Taiwan Normal University, Taipei, Taiwan

ABSTRACT

Brain-imaging studies report that separate neural correlates are associated with processing of different types of humorous materials. However, such evidence lacks temporal information. In this study, we examined the temporal dynamics of humour comprehension between two types of jokes: semantic (SEMs) and pun (PUNs) jokes, using electroencephalographic (EEG) techniques. Thirty SEMs and 30 PUNs were presented to 16 healthy subjects, and their EEG data were concurrently recorded. PUNs consequently showed a larger N400 amplitude than did SEMs, without a specified scalp site, which implies that PUNs induce greater surprise and semantic violation. Meanwhile, SEMs induced a larger P600-like amplitude at the posterior site, which implies that, in order to understand SEMs, higher working memory loads are needed to form novel associations and successfully frame-shift. A possible explanation is the differing logical mechanisms used to understand SEMs and PUNs: the former builds on semantic relationships, the latter on phonological causality.

ARTICLE HISTORY

Received 24 November 2017
Accepted 8 February 2019

KEYWORDS

Humour; semantic jokes; puns; N400; P600

Introduction

Humour not only functions as a useful communication tool in social contexts, but also represents a highly adaptive coping strategy that allows individuals to address difficulties they experience in their daily lives. Humour, for humans, is considered to be accessible and implicitly recognisable; however, it is very difficult to explicitly describe how humour leads to laughter-related phenomena. Nevertheless, existing theories can be used as bases from which to investigate the mental operations that underlie the humour process in more detail.

The humour process features the dual involvement of both cognitive (humour comprehension) and emotional components (humour appreciation). The present study investigates and discusses the cognitive component, which involves comprehension of the humour process. In this field, two fundamental theories relating to humour comprehension are frequently cited: Suls' incongruity-resolution theory (1972) and Wyer and Collin's comprehension-elaboration theory of humour (1992). The

former theory posits that the listener has a prior belief, or a predicted likely outcome, that are supposed to be funny in joke setups. When a punchline mismatches with the listener's prediction (an incongruity), the listener experiences surprise and uses cognitive rules to search for methods by which the punchline can conform more closely to the joke setup. Consequently, jokes are only perceived as funny when incongruities are resolved (or do not exist at all). Based on such cognitive logic, the incongruity-resolution theory features a two-stage model (an incongruity-detection stage followed by an incongruity-resolution stage) and views humour comprehension as a type of cognitive problem-solving task (Suls, 1972).

Differing from the incongruity-resolution theory, which emphasises the cognitive aspects of humour comprehension, the comprehension-elaboration theory considers the broader social contexts in which humour occurs. Adopting Apter's concept of "cognitive synergy" (1982), which refers to a cognitive process in which two contradictory images or

conceptions of the same object are concurrently held in one's mind, Wyer and Collins (1992) developed the comprehension-elaboration theory. In this theory, comprehension refers to the initial encoding of a stimulus event, based on its features, in terms of previously formed concepts or schema (which are stored in the memory system), as well as the creation of inferences based on unstated features, which are necessary to understand the event. Meanwhile, elaboration refers to conscious inferences of features that are not explicit during the former comprehension stage, or other thoughts that might be stimulated by the former comprehension stage. Examining how people comprehend humour within social contexts, Wyer and Collins suggested that humour arises from the concurrent activation of two different meanings of a word or phrase that are brought together simultaneously. They posited that in various humorous conditions, humour is elicited only if the second schema is activated, and this leads to interpretations that conform more to humorous contexts than those for the first schema. A classic example of such a schema-based model is the pun (e.g. two cannibals are eating a clown. One says to the other, "Does this taste funny to you?" – cited from Martin, 2007). Thus, humour comprehension seems to be a cognitively sophisticated ability that underlies comprehension-elaboration theory.

According to the incongruity-resolution theory, registration of surprise acts as a facilitator during the first step, and prompts the subsequent process of humour comprehension. However, the two theories propose that different logical mechanisms (with cognitive rules used to resolve the incongruity of a joke or cartoon) are employed to re-establish coherence in the context of humour processing: 1) by removing incongruities or 2) by activating a second schema that allows two incompatible interpretations to exist simultaneously in one's mind. It is notable, however, that both the incongruity-resolution theory and the comprehension-elaboration theory agree that humour comprehension is sequentially processed over multiple cognitive steps.

Recently, researchers have attempted to use neurophysiological methods to dissect and examine the neural mechanisms of humour processing. Harnessing the advantages of the high time resolution afforded by event-related potentials (ERPs), some studies have reported that representative components correspond to each of the multiple stages of humour processing (Coulson & Kutas, 2001; Coulson & Lovett, 2004; Coulson & Severens, 2007;

Du et al., 2013; Feng, Chan, & Chen, 2014; Ku, Feng, Chan, Wu, & Chen, 2017; Mayerhofer & Schacht, 2015; Shibata et al., 2017; Tu et al., 2014). Meanwhile, functional magnetic resonance imaging (fMRI) data has provided some understanding of the segregated neural substrates involved in humour comprehension and humour appreciation (Azim, Mobbs, Jo, Menon, & Reiss, 2005; Bartolo, Benuzzi, Nocetti, Baraldi, & Nichelli, 2006; Chan, 2016; Chan, Chou, Chen, & Liang, 2012, 2013; Moran, Wig, Adams, Janata, & Kelley, 2004; Shibata, Terasawa, & Umeda, 2014; Watson, Matthews, & Allman, 2007). Regarding understanding how the human brain comprehends different types of humorous materials (e.g. cartoon-implied incongruity-resolution humour, nonsense humour, and puns), a number of imaging studies have hypothesised the differential involvement of cognitive functions through topographical divergence (Chan, 2016; Goel & Dolan, 2001; Samson, Hempelmann, Huber, & Zysset, 2009; Samson, Zysset, & Huber, 2008). However, these studies have disregarded the core essence of humour comprehension: an ongoing process that is built on a time series. Hence, the possibility of the involvement of different cognitive functions through topographical divergence is not likely to correspond to the so-called multi-stage humour processing that is suggested by most cognitive humour theories (Coulson & Kutas, 2001; Suls, 1972; Wyer & Collins, 1992). Thus, it is unlikely that such a research approach can identify the potential time-point at which the cognitive processing of humour comprehension initially diverges between different types of humorous materials.

Considering the above, examining brain activity in a time-series may allow us to explore the points at which neural components of different types of jokes are shared, as well as when they begin to differ, during the multiple stages of humour processing. Therefore, to address our research question, ERPs were considered to serve as an appropriate approach for identifying the time-period at which cognitive processing begins to segregate, and to compare the subsequent neural activities and stages involved in comprehending different types of jokes.

The temporal dynamics of cognitive processing for humour comprehension

No previous study has examined and compared the real-time dynamics underlying different types of

jokes. In the following section, we present an overview of the results of previous studies of humour processing that have been based on consideration of ERPs.

N400: registration of surprise

Coulson and Kutas (2001) were the first to measure ERPs elicited in response to verbal jokes with the aim of examining the cognitive processing that occurs during humour comprehension, from registration of surprise to reestablishment of coherence. These authors found that, compared to non-jokes, jokes elicit a prominent negative-going waveform in a manner characteristic of the N400 component (a negative-going deflection observed at approximately 300–700 ms), followed by enhanced positivity (500–900 ms). The N400 was interpreted to reflect the registration of a surprise component resulting from the jokes' violation of frame-level expectations. The authors suggested that the observed N400, and the subsequent positivity, were mediated by sentence constraints (e.g. high-constraint jokes elicit larger N400 amplitudes), and the working memory operations involved in different levels of joke comprehension, respectively.

In agreement with Coulson and Kutas' findings, some recent studies of the temporal dynamics of humour processing have re-verified the prominence of the N400, characterising it as showing larger magnitudes for sentences that end in a funny manner than for those with non-funny and unrelated endings (Du et al., 2013; Feng et al., 2014; Mayerhofer & Schacht, 2015; Tu et al., 2014). Consistent observation of the N400 across humour-processing studies has established that this parameter is representative of the incongruity-detection stage suggested by Suls (Chan et al., 2013; Du et al., 2013; Feng et al., 2014; Ku et al., 2017; Samson et al., 2008; Suls, 1972). Further, Ku et al. (2017) categorised subjects into subgroups based on their subjective ratings of their surprise in response to joke materials. Consequently, the highly surprised group was determined to elicit a larger N400 than did those who were less surprised. Based on this, the authors suggested that the larger the perceived semantic incongruence between the setup sentences and punchlines, the greater the semantic-expectation violation. Generally, the N400 has

been suggested to act as a neural index of lexical or semantic integration difficulties across linguistic and non-linguistic (e.g. actions and gestures) domains; the more difficult the lexical or semantic integration task, the larger the N400 amplitude (Coulson & Williams, 2005; Kutas & Federmeier, 2011; Kutas & Hillyard, 1980; Mayerhofer & Schacht, 2015). Further, Coulson and Kutas (2001) suggested that the amplitude of the N400 is mediated by sentence constraints, and that the level of difficulty integrating current lexical-semantic information into a previously established frame influences the amplitudes of the N400 in the context of humour processing (Nieuwland & Van Berkum, 2005; Van den Brink et al., 2012).

However, some studies have reported a different N400 performance during humour processing (Marinkovic et al., 2011; Mayerhofer & Schacht, 2015; Shibata et al., 2017). In contrast to the findings of Coulson and Kutas and the other studies mentioned above, Marinkovic et al. (2011) used magnetoencephalography (MEG) measurements and found that punchlines with funny endings generate lower N400 amplitudes than do sentence endings that are simply incongruous. These authors suggested that this reduced N400 may result from "surface congruity", which means that the lexical-semantic association is easier to identify as a result of the strong priming effect exerted on funny punchlines by the preceding setup questions. Supporting Marinkovic et al.'s (2011) findings, Shibata et al. (2017) also reported the lack of an effect on the N400 when reading humorous scenarios. These researchers' experimental approach involved jokes such as the following:

A woman who'd finally landed her first boyfriend proudly boasted about her achievement to a close friend. "I'm getting begged practically every day to marry him!" "That's amazing! Are you serious? Wait – Who's doing the asking?"- "My parents." Shibata et al. (2017) consequently concluded that the reason for the lack of N400 in this regard was due to the absence of semantic violation in their experimental paradigm.

Further supporting this conclusion, Mayerhofer and Schacht (2015) used garden-path jokes¹ with three different endings (a coherent ending, a joke ending, and a discourse-incoherent ending) to test

¹Garden-path jokes: A garden-path (GP) joke is a joke for which the first dominant interpretation is that the joke is incoherent, but this is subsequently substituted by a hidden joke interpretation. Two important factors for the processing of GP jokes are saliency of the initial interpretation and the accessibility of the hidden interpretation. Both factors are assumed to be affected by contextual embedding (Mayerhofer, Maier, & Schacht, 2016).

whether the N400 could be reliably found across conditions. Contrary to the authors' expectations, however, the N400 did not occur when ERP comparisons were conducted between the joke and coherent conditions (the incoherent condition was excluded in Experiment 3). The authors, therefore, concluded that the N400 is strongly affected by experimental contexts.

In summary, the majority of studies agree that the N400 has a representative role as the neural index representing the incongruity-detection stage, but some studies have reported an absence of the N400 during joke comprehension. The activation of the N400 in humorous discourses is functionally correlated to the registration of surprise in response to expectation violation. Further, the magnitudes of the N400 are mediated by the different levels of semantic incongruence, or difficulties in integrating semantic information between the preceding contexts and the punchlines.

P600: reestablishment of coherence

After observing the prominence of the N400, Coulson and Kutas (2001) interpreted this protracted positive deflection as 1) a frame-shifting process that activates a new frame from long-term memory and reinterprets information that is already active in the working memory and/or 2) a suppression of joke-irrelevant information in the context of humour comprehension with the aim of re-establishing semantic coherence (Coulson, 2000; Coulson & Lovett, 2004; Kutas & King, 1996). More recent studies have provided consistent findings of this "P600-like" deflection, which is characterised by larger magnitudes for joke than for non-joke conditions. These studies have consistently agreed that this "P600-pattern" in joke processing represents the mental process used to search for attainable cues with which to form possible novel associations, as well as the process of integrating semantic information to build global coherence in response to the violation of the expectation created from the setup sentences (Du et al., 2013; Feng et al., 2014; Ku et al., 2017; Marinkovic et al., 2011; Mayerhofer & Schacht, 2015; Shibata et al., 2017; Tu et al., 2014). Ku et al. (2017) further hypothesised that a larger P600 amplitude indicates a more successful resolution of the detected incongruities, which would distinguish individuals with good comprehensive ability from those with poor comprehensive ability. Taken together, this "P600-like" pattern, which is referred to by different names across

related studies ("P600": Feng et al., 2014; Ku et al., 2017; Shibata et al., 2017; "P600–800": Du et al., 2013; "P800–1000" and "P1000–1600" by Tu et al., 2014; "P600m": Marinkovic et al., 2011), has been suggested to be an indicator of the incongruity-resolution stage of humour comprehension (referenced in Suls' incongruity-resolution theory and Wyer and Collins' comprehension-elaboration theory). In addition, this P600-like pattern in humour processing represents a process that is distinct from the original P600, indicating syntactic violation, as proposed by Osterhout and Holcomb (1992). Moreover, Mayerhofer and Schacht (2015) suggested that in humour processing, rather than the syntactic repair process elicited by syntactic anomalies, the P600-like pattern resembles a semantic repair process that leads to incongruity resolution.

As the humour process at the cognitive level comes to an end, subjects are closer to the emotional and enjoyable aspects of the joke. However, researchers have experienced difficulty pinpointing the time point at which humour comprehension and humour appreciation segregates. Indeed, Marinkovic et al. (2011), using a MEG approach, reported a slow-rising and ramp-like waveform during the late stage of joke appreciation, and claimed that because humour comprehension and humour appreciation are integrated at different latencies for each subject, the temporal overlap of such activities leads to difficulty time-locking the moment at which people "get" a joke. Such difficulties were reconfirmed by the wide range of time ranges (from 800 to 2000 ms) that have been reported in previous studies that have postulated that the third stage comprises humour elaboration (Du et al., 2013; Feng et al., 2014; Ku et al., 2017; Mayerhofer & Schacht, 2015; Tu et al., 2014); humour elaboration, due to its varying temporal characteristics, therefore, only plays a minor role in this study.

Humour comprehension for puns

Long and Graesser (1988) categorised types of jokes based on the intention or use of the humour involved, and defined puns as the "humorous use of a word that evokes a second meaning, usually based on a homophone (in other words, a word with a different meaning that sounds the same)". It is worth noting that, of the previous research studies involving ERPs, only Coulson and Severens (2007) clearly stated that they used puns as their main stimuli (including homographic puns, puns

that exploit the multiple meanings of the same word form, and ideophonic puns), instead of the more commonly used category of semantic jokes (e.g. Chan et al., 2013; Coulson & Kutas, 2001; Du et al., 2013; Feng et al., 2014; Mayerhofer & Schacht, 2015; Shibata et al., 2017). Based on their findings, Coulson and Severens argued that the word play involved in pun jokes is markedly different from that present in most semantic jokes. Crucially, puns provide contextual support for two opposing meanings of an ambiguity, which allow one's thoughts to playfully oscillate back and forth between two incompatible interpretations of the concept. In this regard, the logical mechanism employed in the incongruity resolution necessary for understanding puns relates strongly to Wyer and Collins' comprehension-resolution theory.

Some fMRI studies have suggested that there are differing neural correlates involved in humour comprehension, especially in regard to differentiating puns from other types of jokes. For instance, Goel and Dolan (2001) were the first to report, using an fMRI, the existence of differing modality-specific pathways between semantic jokes and puns. In their study, they found that semantic jokes induced greater blood-oxygen-level dependent responses in the right posterior middle temporal gyrus (BA 21/37) and the left posterior inferior temporal gyrus (BA 20/37) than did puns. In contrast, puns showed stronger activation of the left insula/precentral gyrus. The authors consequently inferred that activation of the posterior temporal lobe is a specific requirement for the integrative processes necessary for understanding semantic jokes. On the other hand, activation of the left inferior prefrontal cortex and insula were considered to be the requirements for the processing of speech sounds necessary to comprehend a phonological punchline. Based on these findings, Goel and Dolan (2001) suggested the existence of differing modality-specific pathways for the cognitive processing of the humour concerned in semantic jokes and puns, respectively.

Samson et al. (2008) subsequently simulated Goel and Dolan's (2001) study, but replaced verbal jokes with three types of humorous cartoons (visual puns, semantic cartoons, and theory of mind cartoons) which, they hypothesised, would be resolved by different logical mechanisms. They consequently observed, across the different types of joke cartoons, involvement of the left-sided network for pure incongruity resolution (including the temporo-

parietal junction, inferior frontal gyrus and ventromedial prefrontal cortex); however, they also noted that puns stimulated particularly strong activation in the extrastriate cortex. The authors inferred, as a result of the lower requirements regarding building a situation model, that puns are processed differently than semantic and theory of mind cartoons. Instead, they suggested that puns are realised by placing a greater emphasis on recognising the compatibility of the presented elements. In other words, the necessity of establishing semantic relationships in order to understand semantic cartoons, or to use mental abilities to "get" theory of mind cartoons, plays a subordinate role to understanding puns. The authors concluded that the different logical mechanisms used to comprehend humorous cartoons depend upon separate neural networks. More importantly, puns are clearly processed in a different manner than are other types of jokes, with less sophisticated and less profound characteristics than semantic jokes. The authors suggested that understanding puns is more technical than understanding semantic jokes (Samson et al., 2008).

The brief summary presented above, derived from previous studies, has yielded two important viewpoints:

- (1) The N400 and P600-like processes are neural indices that are related to the temporal dynamics of brain activity for humour comprehension during the incongruity-detection stage of surprise registration and the incongruity-resolution stage of reestablishment of coherence, respectively;
- (2) Both functional neuro-imaging studies and ERP studies highlight that puns involve markedly different cognitive processing than do other types of jokes.

In this study, which represents a preliminary study for investigating different types of jokes in real-time dynamics, we chose the relatively stable neural indexes of the N400 and P600-like processes as bases for our hypotheses. Moreover, as a result of the fact that previous studies have shown that puns require a particularly different form of processing, we chose both semantic jokes (SEMs) and pun jokes (PUNs) as the two main conditions for this present study. Collectively, the main goal of the present study focusses primarily upon comparing humour comprehension for SEMs and PUNs at the incongruity-detection stage and the incongruity-resolution stage. The findings from the present study

should contribute to providing a greater understanding of how different humorous materials are processed in a step-wise fashion.

Study hypotheses

Considering the particularity of processing pun jokes, the first hypothesis of this study is that PUNs evoke a larger N400 than do SEMs. Although the phonological features of pun jokes create a strong priming effect between the setup and the punchline, it is much more difficult to integrate, between the preceding context and the punchline, semantics into puns, than it is for semantic jokes. We hypothesised that such semantic integration difficulties lead to a greater sense of surprise and a greater extent of perceived incongruities for PUNs than for SEMs, thereby inducing larger N400 amplitudes for PUNs.

Furthermore, because of the different logical mechanisms required to understand PUNs and SEMs, the second hypothesis of this study is that PUNs evoke smaller positive deflections subsequent to the N400, in a manner potentially analogous to the P600, than do SEMs. As suggested by previous studies (Coulson & Severens, 2007; Goel & Dolan, 2001; Samson et al., 2008), the key to understanding puns relies on the compatibility of the phonological elements, namely features based on phonological causalities. Furthermore, Coulson and Severens (2007) noted that puns are based on simply promoting the two (or more) meanings of an ambiguous word or phrase. Thus, this process places less requirements on the working memory in regard to resolving ambiguities and establishing semantic coherence through successful frame-shifting, which would consequently result in a reduced P600-like waveform for PUNs.

Method

Participants

Sixteen undergraduate students (eight males) aged 20–25 years (mean age \pm standard deviation: 21.7 \pm 1.62 years) were recruited for this study. Written informed consent was obtained from all subjects prior to the formal experimental session. All participants were right-handed, neurologically healthy, and native Chinese speakers, and all had either normal or corrected-to-normal vision. This study was approved by the Institutional Review Board of National Taiwan University (201208HS012, Taipei, Taiwan).

Materials

Thirty semantic jokes and 30 pun jokes (referred to as “SEMs” and “PUNs” hereafter) were used in this study. In addition, 30 non-jokes served as a baseline control condition (referred to as “Controls” hereafter). Thus, a total of 90 trials were presented in a random order to each subject. The SEMs and Controls materials used in the current study were selected from a pool of Chinese riddles collated by Feng et al. (2014). All trials comprised a setup question and a punchline; the word length of the setup question was controlled to between 11 and 14 characters, while the punchlines were compound words comprising two to four characters. The major characteristic separating the PUNs from SEMs in this study was that the punchlines for the PUNs were phonologically correlated to the preceding setups, while those for the SEMs were not.

When reading pun jokes, the reader firstly detects the semantic incongruence, and then the pronunciation-related correspondence between the setup question and the punchline. To “get” pun jokes, the reader must build a model, predominantly through phonological causality, that allows two incongruent concepts to exist simultaneously in his/her mind (Marinkovic et al., 2011). An example of joke used in the PUNs condition is the following: “What would a kylin become if he flew to the North Pole? ... Ice cream”. A kylin (pronounced /qi2 lin2/) is a mythical Chinese creature, and if this creature were to fly to the cold environment of the North Pole, it would then become an “iced kylin”. The semantic meaning of “iced kylin” (pronounced /bing1 ci2 lin2/), is very different from the meaning of “ice cream”, but has the same pronunciation in Chinese. In this example, incongruities are resolved based on phonological features, rather than semantic relationships.

In the SEMs condition, to get the jokes it was necessary to merely resolve semantic incongruences between the setup question and the punchline. Confounding factors, such as the phonological causality (or homophones) that features in the PUNs condition in this study, or any content involving sexual or violent issues that might cause sexual preferences in regard to jokes, were avoided. An example from the SEMs condition is as follows: “Where’s the place with the cheapest rent? ... The prison”.

The Controls condition featured sets of general knowledge questions and answers, without induced incongruity detection or a subsequent emotional

response (for instance, “what is the most intelligent mammal in the ocean? ... The dolphin”). In other words, these were neither funny nor surprising.

Pre-experimental ratings were conducted of the semantic jokes, pun jokes, and non-jokes used in this study. Here, 18 participants (nine males) aged between 20 and 27 years (mean age: 22.67 ± 2.4 years) evaluated the jokes based on three domains: (a) surprise: whether the punchline surprised them; (b) comprehensibility: how well they understood the punchline’s meaning in relation to the corresponding setup; and (c) funniness: how funny they considered each trial to be. Subjects were asked to provide their evaluations for each trial, in order of surprise, comprehensibility, and funniness, using a nine-point Likert scale, ranging from “1” (“totally not surprised”, “totally incomprehensible”, or “not funny at all”), through “5” (neutral response for each domain), to “9” (“totally surprised”, “totally comprehensible”, or “very funny”). One-way analyses of variance (ANOVAs) were conducted for the ratings for all three experimental domains, and this showed significant effects of condition [surprise: $F(2, 87) = 127.718, p < 0.001$; comprehensibility: $F(2, 87) = 9.182, p < 0.001$; funniness: $F(2, 87) = 256.388, p < 0.001$]. For the surprise domain, post-hoc Bonferroni comparisons demonstrated that both the SEMs and PUNs showed significantly larger surprise ratings than did the Controls ($p < 0.01$ for both SEMs and PUNs), but equivalent surprise ratings to each other ($p = 1.00$) (surprise scores: 6.27 ± 0.75 for SEMs, 6.13 ± 0.68 for PUNs, and 3.63 ± 0.76 for Controls). In terms of comprehensibility, post-hoc Bonferroni comparisons demonstrated that both the SEMs and Controls showed significantly larger comprehensibility ratings than did the PUNs ($p < 0.01$ in each case), but equivalent comprehensibility ratings to each other (comprehensibility scores: 8.05 ± 0.41 for SEMs, 7.30 ± 1.06 for PUNs, 8.04 ± 0.71 for Controls). In terms of funniness, post-hoc Bonferroni comparisons demonstrated that SEMs showed the highest funniness scores, followed by PUNs, with the Controls scored the lowest (funniness scores: 6.76 ± 0.36 for SEMs, 5.91 ± 0.77 for PUNs, 3.65 ± 0.40 for Controls; $p < 0.01$ in each case). The results of the pre-experiment showed that the SEMs and PUNs were associated with more surprising and funnier elements than were the Controls, thus verifying the humour and emotionally arousing and pleasurable sensations induced by the materials in the SEMs and PUNs conditions. Additionally, PUNs

were determined to be significantly less comprehensible than were the other two conditions.

Procedure

The experimental procedure was conducted in a dimly lit and sound-proof room, and featured the use of electroencephalogram (EEG) recording. Participants were seated in front of a computer monitor, at a distance of approximately 110 cm. The task required the participants to silently read the materials in a set order: first the setup questions and then the following punchlines. Then, they were required to provide ratings, using a four-point Likert scale, of their subjective judgements of the level of surprise, comprehensibility, and funniness of each trial, respectively (“1” for “totally not surprised”, or “totally incomprehensible”, or “not funny at all”; “2” for “not surprised”, or “incomprehensible”, or “not funny”; “3” for “surprised”, or “comprehensible”, or “funny”; and “4” for “totally surprised”, or “totally comprehensible”, or “very funny”). Before formal testing, each participant performed 10 practice trials. The presentation order, and the duration of each trial, was as follows: (1) a fixation cross was presented at the centre of the screen for 1000 ms (to act as a cue for the beginning of the new trial); (2) a setup question was presented for 4000 ms; (3) another fixation cross was presented for 4000 ms (to minimise eye movements between the setup question and the punchline), (4) the punchline was presented for 2000 ms; and (5) self-paced subjective ratings were presented in the following order: surprise, comprehensibility, and funniness.

Data recording and analysis

EEG data were recorded using 32 channel-sintered Ag/AgCl electrodes arranged in a standard 10–20 system (QuickCap, Neuromedical Supplies, Sterling, Texas, USA). The signals were amplified using SynAmps2 (Neuroscan Inc., El Paso, Texas, USA) and were recorded continuously with a 500 Hz sampling rate and 0.05–70 Hz filtering. References were set to the bilateral mastoids. Meanwhile, eye movements and blinks were monitored via vertical electrooculogram and horizontal electrooculogram electrodes placed on the supraorbital and infraorbital ridges of the left eye, and lateral to the outer canthi of the right and the left eyes, respectively. Impedances were maintained below 5 k Ω .

For off-line analysis, the EEG data were analysed using Scan 4.5 software (Compumedics Ltd., Australia). The ERP epochs were time-locked to the onset of the punchline, including 100 ms pre-stimulus (which served as the baseline) and 2000 ms after the target onset. For each epoch, baseline correction for the data was performed 100 ms prior to the stimulus, and filtered with a band-pass of 0.1–30 Hz and a notch filter of 60 Hz. Trials contaminated by eye blinks, or any other artefacts, with voltage variations larger than $\pm 100 \mu\text{V}$, were excluded from averaging. Moreover, trials for which participants provided a rating of “not funny at all” (value “1” from the four-point Likert scale) were also excluded from further analysis. Finally, at least 20 trials were retained for each condition (a rejection rate of 33%). Average waves were separated by conditions.

Regarding the grand average ERP waveforms, we analysed the mean amplitudes of the N400 (280–500 ms) and the positive deflection in a manner potentially analogous to the P600 (500–700 ms). In addition, our topography map of grand average waveforms demonstrated that the most prominent effect was found at the midline electrodes (Figure 1); therefore, a cluster of five electrodes at the midline sites (FZ, FCZ, CZ, CPZ, and PZ) were selected as the electrodes of interest through which to compare mean N400 and P600-like amplitudes between SEMs and PUNs. No other electrodes were used for further analysis. A two-way repeated measures ANOVA using the factors of “Condition” (SEMs, PUNs, Controls) and “Electrodes” (FZ, FCZ, CZ, CPZ, PZ) was performed to compare mean

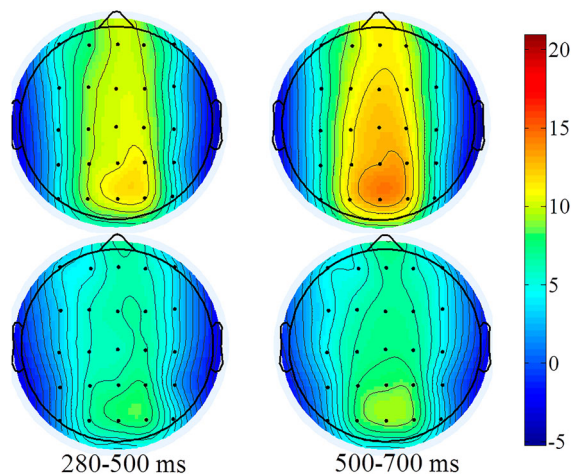


Figure 1. Topography map of the grand average waveforms for SEMs (upper column) and PUNs (lower column) between 280–500 ms (N400) and 500–700 ms (P600-like waveform).

N400 and P600-like amplitudes between conditions. Meanwhile, Bonferroni correction was used for the post-hoc comparisons. A statistical difference of $p < 0.05$ was considered to represent significance. Greenhouse-Geisser correction was applied when sphericity was violated. Behavioural rating scores were compared using a one-way ANOVA. Additionally, correlations between the subjective ratings of surprise, comprehensibility, and funniness were also tested as possible indices in order to understand the logical mechanisms between PUNs and SEMs.

Results

Subjective ratings

The mean rating scores of surprise, comprehensibility, and funniness for PUNs, SEMs, and Controls are shown in Table 1. Statistical analysis indicated that both SEMs and PUNs showed significantly higher surprise and funniness ratings than did Controls ($p < 0.01$ for both SEM and PUN), but equivalent surprise and funniness ratings to each other ($p = 1.00$). In terms of comprehensibility rating, post-hoc Bonferroni comparisons demonstrated that both the SEMs and Controls showed significantly larger comprehensibility ratings than did the PUNs ($p < 0.01$ in each case), but equivalent comprehensibility ratings to each other ($p = 1.00$). Except for the funniness rating (which showed the highest scores for SEMs), the behavioural patterns of the subjective ratings of surprise and comprehensibility in the formal test were the same as the results for the pre-experimental ratings.

Correlation testing showed neither significant nor marginal correlations between any pair of the surprise, comprehensibility, or funniness scores for PUNs (all $p > 0.1$). On the other hand, for SEMs, there were marginally significant correlations between the subjective ratings for comprehensibility and funniness ($r = 0.49$, $p = 0.054$). The Control condition showed a positive correlation between the surprise and funniness ratings ($r = 0.69$, $p < 0.01$).

Table 1. Subjective rating scores of SEMs, PUNs and controls.

Rating scores	SEMs		PUNs		Controls	
	Mean	SD	Mean	SD	Mean	SD
Surprise	2.65	0.31	2.80	0.42	1.88	0.27
Comprehensibility	3.29	0.30	3.08	0.32	3.40	0.34
Funniness	2.74	0.44	2.75	0.41	1.95	0.37

ERP results

Figure 2 depicts the grand average waveforms for the SEM, PUN, and Control conditions at the midline positions. During the time window from 280 to 500 ms, no interaction between Condition and Electrode was observed for the N400 component [$F(8,120) = 0.446$, $p = 0.751$], although

analysis revealed a significant main effect of Condition [$F(2,30) = 12.954$, $p < 0.001$, $\eta^2 = 0.46$], but not of Electrode [$F(4,60) = 3.514$, $p = 0.068$]. However, post-hoc comparison illustrated that PUNs ($6.55 \pm 1.12 \mu\text{V}$) showed larger N400 amplitudes than did both SEMs ($10.15 \pm 1.14 \mu\text{V}$) and Controls ($9.80 \pm 0.91 \mu\text{V}$) ($p < 0.01$ in each case). Meanwhile, SEMs

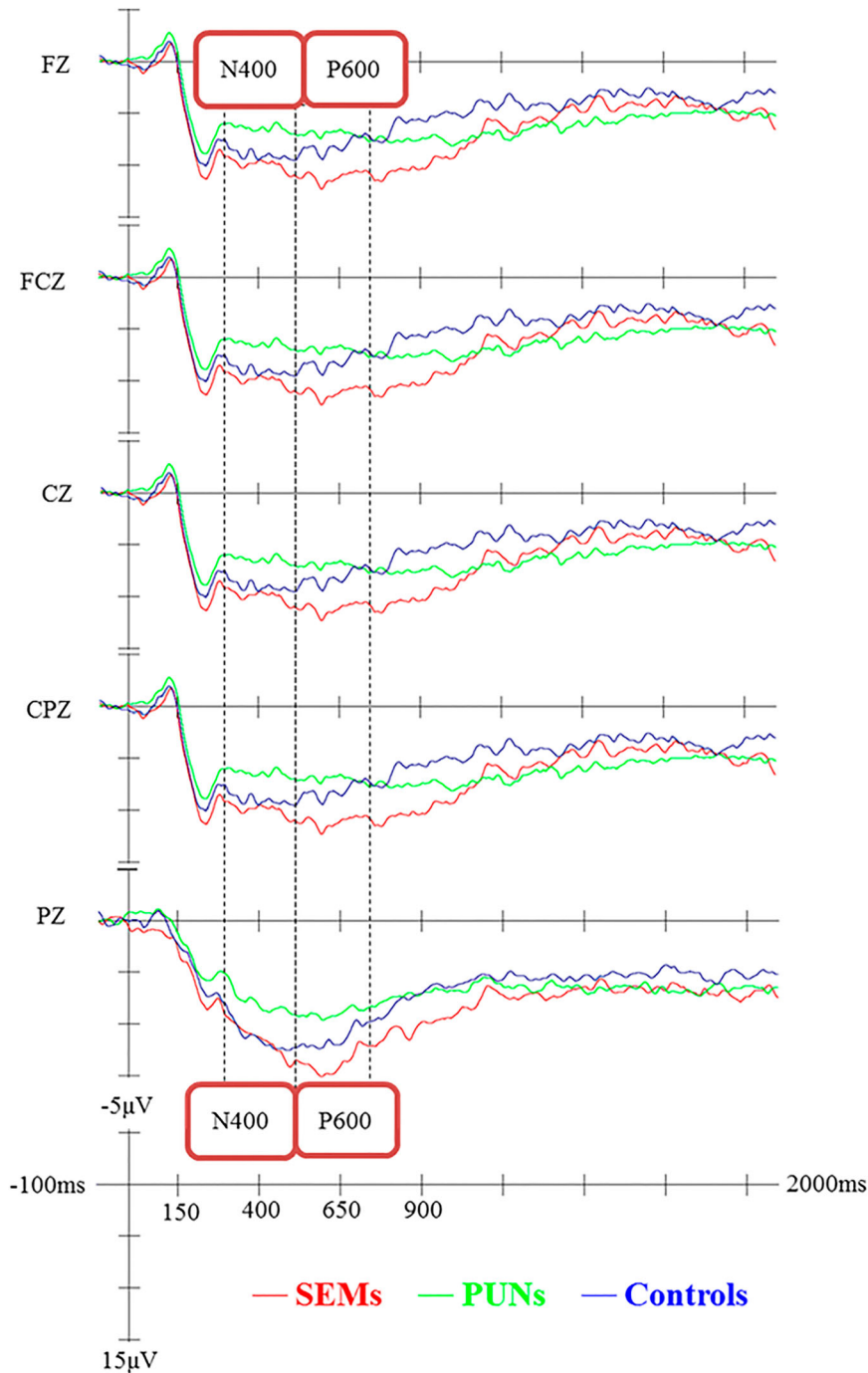


Figure 2. (Colour online) Grand average waveforms for SEMs (red), PUNs (green), and Controls conditions (blue) at the midline positions. Voltage scales in μV . Time scales in milliseconds. PUNs demonstrated a larger N400 without specified scalp site than SEMs. SEMs demonstrated a larger P600-like amplitude at the parietal site than PUNs.

showed equivalent N400 amplitudes to Controls ($p > 0.05$).

During the time window from 500 to 700 ms, no interaction between Condition and Electrode was reported for the P600-like component [$F(8,120) = 0.770$, $p = 0.549$]. A significant main effect was reported for Condition [$F(2,30) = 10.433$, $p < 0.001$, $\eta^2 = 0.41$] and for Electrode [$F(4,60) = 10.026$, $p < 0.001$, $\eta^2 = 0.40$]. Post-hoc comparison showed that SEMs ($12.26 \pm 1.34 \mu\text{V}$) elicited significantly larger P600-like amplitudes than did PUNs ($7.73 \pm 1.23 \mu\text{V}$, $p < 0.001$), but equivalent P600-like amplitudes to Controls ($10.05 \pm 0.90 \mu\text{V}$, $p > 0.05$). PUNs showed equivalent P600-like amplitudes to Controls ($p > 0.05$). Furthermore, the PZ electrode ($11.44 \pm 0.88 \mu\text{V}$) showed significantly larger P600-like amplitudes than did the FZ ($8.77 \pm 1.21 \mu\text{V}$), the FCZ ($9.22 \pm 1.12 \mu\text{V}$), and the CZ ($9.90 \pm 1.07 \mu\text{V}$) electrodes ($p < 0.05$ in each case).

Discussion

This study revealed the differential temporal dynamics induced by different types of jokes at both the incongruity-detection stage and incongruity-resolution stage of humour comprehension. PUNs evoked a larger N400 amplitude without specified scalp sites, and a smaller P600-like amplitude at the parietal site. Both of these findings accord with our hypotheses. In addition, the results of the subjective ratings provided by the participants also revealed important findings. In both the pre-experimental ratings and formal testing, PUNs were reported to have significantly less comprehensibility than had SEMs. This result conforms with the findings of Samson et al. (2008), and confirms that there are distinct differences in the structural characteristics of the resolvability processes required to understand PUNs and SEMs (Ruch & Hehl, 2007). Moreover, results from the correlation tests may underline this point, also serving as potential candidates for illustrating that the logical mechanism necessary for understanding PUNs differs from that required for understanding SEMs (as manifested by the marginal significance between the scores for comprehensibility and funniness in SEMs, a pattern that was absent for PUNs).

While both semantic jokes and puns are associated with perceived incongruity at the semantic level and the necessity of resolving incongruity, the greater extent of semantic violation between the preceding setups and the punchlines of puns

leads to greater difficulty in integrating the semantics involved. In accordance with previous studies' interpretations of the N400's indication of humour comprehension (discussed in the introduction), we suggest that the greater degree of perceived incongruities and expectation violation associated with PUNs results in greater difficulty integrating semantics at the lexical level, which results in enhanced N400 amplitudes. Interestingly, subjects may feel that puns trick or fool them "twice", not only by the violation of semantics presented through the immediate "surface congruity", but also by the phonological resemblance. This makes incongruity detection more arousing and surprising for PUNs. Additionally, our results showed that there are implicit differences between SEMs and PUNs regarding the cognitive processing of the N400, but also that there are comparable levels of explicit surprise between the two joke types at the incongruity-detection stage. The necessity of disclosing these subtle but important differences through neurophysiological measurement was, therefore, re-verified in the present study, as it afforded improved understanding of the differential cognitive processing required for the different types of jokes.

The finding of higher P600-like amplitudes for SEMs than for PUNs is in accordance with the proposal that SEMs require subjects to use more working memory to retrieve word meanings, form novel associations, and build global semantic coherence (Coulson & Kutas, 2001; Coulson & Lovett, 2004; Coulson & Severens, 2007; Kutas & King, 1996). As mentioned above, the premise of understanding semantic jokes relies on deep processing of semantic relationships, and this places greater inferential demands on working memory operations (Hempelmann & Samson, 2007). In contrast, the key to understanding a pun lies in the simultaneous perception and acceptance that there are two ambiguous meanings for the same word (Coulson & Severens, 2007). In this case, phonological causality may be sufficient to resolve puns (Samson et al., 2008). Meanwhile, the lower comprehensibility ratings provided for PUNs than SEMs in this study indicate the differing humorous nature of the logical mechanisms used to understand puns and semantic jokes. Taken together, the reduced P600-like amplitudes for PUNs observed in the current study may serve as a potential index for illustrating the strategic trade-off between SEMs and PUNs that occurs in a time-sensitive manner during cognitive processing at the incongruity-resolution stage.

Hempelmann (2004) noted that the commonality of the incongruity resolution between puns and semantic jokes is semantic, but that puns manipulate the incongruity resolution of both phonological features and semantic meanings, while semantic jokes only manipulate the incongruity of semantic meanings. Building on this, Samson et al. (2008) claimed that contrasting visual semantic cartoons with visual puns should reveal the differing incongruity-resolution stages involved for both (pure semantic or visual/semantic). However, the temporal dynamics obtained in this study revealed that the cognitive processing of PUNs and SEMs diverges at the incongruity-detection stage, which precedes the incongruity-resolution stage. Thus, taking both the behavioural and neural results obtained in this study together, our research provides insightful perspectives regarding the inherent differences between the cognitive processing performed for semantic jokes and puns at both the incongruity-detection and incongruity-resolution stages during humour comprehension.

The differential logical mechanisms used for understanding different types of jokes, such as for understanding SEMs and PUNs, may also be associated with the differing structural characteristics of the resolvability for each type of joke. Using factorial analysis, Ruch and colleagues (Ruch, 1992; Ruch & Hehl, 2007) segregated incongruity-resolution stimuli from nonsense stimuli, predominantly based upon their resolvability of incongruities. In most cases, the incongruities in the incongruity-resolution stimuli were largely resolvable, whereas in the nonsense stimuli the incongruities were only partially resolvable, or could even create more confusion. Based on this, Samson et al. (2009) then used fMRI to explore the neural substrates between incongruity-resolution humour and nonsense humour. They consequently found that incongruity-resolution cartoons induce activation in multiple regions, such as the bilateral temporo-parietal junction (allowing more information to be integrated), the bilateral superior frontal gyrus (allowing information or scripts to be manipulated), and the right anterior medial prefrontal cortex (amPFC; for closer reference to reality). However, nonsense cartoons showed no specific activation of these areas. The authors, thus, inferred that, in most cases, nonsense cartoons require less self-reference and leave more residual incongruity than do incongruity-resolution cartoons. In other words, the absurdity, or the exaggerated violation of reality, in nonsense jokes,

leads to a mere search for a means of resolving the incongruity, thus contributing to the differing patterns of activation observed between incongruity-resolution jokes and nonsense jokes (Samson et al., 2009).

It is worth noting that SEMs and PUNs might be classified based on the structural characteristics of the resolvability, as suggested by Ruch and colleagues (Ruch, 1992; Ruch & Hehl, 2007). From our viewpoint, the structural characteristics of SEMs are more closely inclined to incongruity-resolution jokes because of their higher comprehensibility scores (the higher comprehensibility ratings parallel their higher resolvability). Meanwhile, the marginal significance we observed between comprehensibility and funniness scores may imply a potentially traceable path, and that SEMs are resolved, at least to a certain degree, through applying some cognitive rules. In this regard, humour comprehension of SEMs, as well as of incongruity-resolution jokes, is closely related to the incongruity-resolution theory proposed by Suls (1972). In contrast, based on their lower resolvability and the more abstract cause-and-effect between their comprehensibility and subjective funniness ratings, we suggest that PUNs are comparatively more inclined towards nonsense jokes. As mentioned above, nonsense jokes and PUNs have the commonality of a lower necessity of establishing semantic relationships with the preceding contexts. Instead, allowing two scripts to be accepted simultaneously and to be compatible in one's mind, rather than requiring removal or resolution of incongruities, pertains to the core of understanding PUNs and nonsense jokes (Apter, 1982; Hempelmann & Samson, 2007; Wyer & Collins, 1992). Thus, humour comprehension of PUNs, as well as of nonsense jokes, is more closely related to the comprehension-elaboration theory proposed by Wyer and Collins (1992).

The strategy of merely searching for a means of resolving the incongruities used in nonsense jokes may serve as another contributor to the reduced P600-like waveform observed for PUNs. That is, the mere search for a means of resolving incongruities may influence the cognitive loads the listeners decide to devote to incongruity resolution, especially as considering the semantic relationship is not a prerequisite for understanding a pun. Thus, this lower implementation of working memory operations could lead to the reduced P600-like waveform observed for PUNs. In summary, to understand jokes such as nonsense jokes and PUNs, subjects should

activate the strategy of removing strict semantic boundaries, as is required to understand the example of “iced kylin” and ice-cream provided in the Materials section, which involves both nonsense and homophone elements.

There are several limitations to this study. The first and most significant limitation is the small number of trials involved in the experiment. As a result of the difficulty designing and creating pun jokes, the present study included only 30 trials for each condition. Further, after data cleaning, the number of valid trials was even lower (approximately 20–30 trials). Increasing the number of trials should provide more valid and robust ERP effects. Secondly, although the word length of the setup questions and punchlines were controlled to between 11 and –14 and two to four characters, respectively, stricter control of the lexical-semantic properties, such as word frequency and the number of strokes, should be taken into consideration. Third, our study lacked behavioural measurements pinpointing the exact time-point when subjects “got” the jokes. However, to date, this remains a very difficult problem, and has not yet been resolved by any previous ERP study on humour processing. Future studies should endeavour to solve this, perhaps by adding behavioural records such as reaction times. Fourth, a series of studies conducted by Coulson and colleagues has addressed important issues regarding the associations between the hemispheric differences of language lateralisation, semantic activations, and handedness in regard to joke comprehension (Coulson & Lovett, 2004; Coulson & Severens, 2007; Coulson & Williams, 2005; Coulson & Wu, 2005); these factors were not addressed in the current study, despite the fact that these factors represent a valuable concept with respect to understanding the cognitive processing of humour comprehension. Future studies should take these factors into consideration.

In conclusion, the present study provides empirical evidence and integrative explanations, from both behavioural and neural viewpoints, elaborating the differential cognitive processing between semantic jokes and puns in terms of both incongruity-detection and -resolution stages during humour comprehension, manifesting as larger N400 and smaller P600-like amplitudes for PUNs than for SEMs, respectively. Most notably, puns were determined to be processed especially differently, as the logical mechanism for incongruity resolution of puns mainly lies in the phonological causality,

rather than the semantic relationship, which is required for understanding semantic jokes.

Acknowledgements

We would like to express our appreciation to Prof. Linden Ball and Dr. Esther Fujiwara, the editors of the *Journal of Cognitive Psychology*, for their numerous suggestions regarding the constructions and inferences of our research work, and to anonymous reviewers for their valuable comments on earlier versions of this article.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was financially supported by the Institute for Research Excellence in Learning Sciences and the Chinese Language and Technology Centre of National Taiwan Normal University (NTNU) as part of the Featured Areas Research Centre Programme within the framework of the Higher Education Sprout Project of the Ministry of Education (MOE) in Taiwan.

ORCID

Yi-Tzu Chang  <http://orcid.org/0000-0002-1474-2406>
Ching-Lin Wu  <http://orcid.org/0000-0001-8211-0446>
Hsueh-Chih Chen  <http://orcid.org/0000-0003-2043-0190>

References

- Apter, M. J. (1982). *The experience of motivation: The theory of psychological reversals*. London, England: Academic Press.
- Azim, E., Mobbs, D., Jo, B., Menon, V., & Reiss, A. L. (2005). Sex differences in brain activation elicited by humor. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 16496–16501. doi:10.1073/pnas.0408456102
- Bartolo, A., Benuzzi, F., Nocetti, L., Baraldi, P., & Nichelli, P. (2006). Humor comprehension and appreciation: An fMRI study. *Journal of Cognitive Neuroscience*, 18, 1789–1798. doi:10.1162/jocn.2006.18.11.1789
- Chan, Y. C. (2016). Neural correlates of sex/gender differences in humor processing for different joke types. *Frontiers in Psychology*, 7. doi:10.3389/fpsyg.2016.00536
- Chan, Y. C., Chou, T. L., Chen, H. C., & Liang, K. C. (2012). Segregating the comprehension and elaboration processing of verbal jokes: An fMRI study. *NeuroImage*, 61, 899–906. doi:10.1016/j.neuroimage.2012.03.052
- 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 humor processing: An fMRI study of the neural substrates of incongruity detection

- and resolution. *NeuroImage*, 66, 169–176. doi:10.1016/j.neuroimage.2012.10.019
- Coulson, S., & Kutas, M. (2001). Getting it: Human event-related brain response to jokes in good and poor comprehenders. *Neuroscience Letters*, 316, 71–74. doi:10.1016/S0304-3940(01)02387-4
- Coulson, S., & Lovett, C. (2004). Handedness, hemispheric asymmetries, and joke comprehension. *Cognitive Brain Research*, 19, 275–288. doi: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, 172–187. doi:10.1016/j.bandl.2005.08.009
- Coulson, S., & Williams, R. W. (2005). Hemispheric asymmetries and joke comprehension. *Neuropsychologia*, 43, 128–141. doi:10.1016/j.neuropsychologia.2004.03.015
- Coulson, S., & Wu, Y. C. (2005). Right hemisphere activation of joke-related information: An event-related brain potential study. *Journal of Cognitive Neuroscience*, 17, 494–506. doi:10.1162/0898929053279568
- Coulson, S. (2000). *Semantic leaps: Frame-shifting and conceptual blending in meaning construction*. Cambridge, UK: Cambridge University Press.
- Du, X., Qin, Y., Tu, S., Yin, H., Wang, T., Yu, C., & Qiu, J. (2013). Differentiation of stages in joke comprehension: Evidence from an ERP study. *International Journal of Psychology*, 48, 149–157. doi:10.1080/00207594.2012.66516
- Feng, Y. J., Chan, Y. C., & Chen, H. C. (2014). Specialization of neural mechanisms underlying the three-stage model in humor processing: An ERP study. *Journal of Neurolinguistics*, 32, 59–70. doi:10.1016/j.jneuroling.2014.08.007
- Goel, V., & Dolan, R. J. (2001). The functional anatomy of humor: Segregating cognitive and affective components. *Nature Neuroscience*, 4, 237–238. doi:10.1038/85076
- Hempelmann, C. F. (2004). Script opposition and logical mechanism in punning. *Humor-International Journal of Humor Research*, 17, 381–392. doi:10.1515/humr.2004.17.4.381
- Hempelmann, C. F., & Samson, A. C. (2007). Visual puns and verbal puns: Descriptive or false analogy? In S. Attardo & D. Popa (Eds.), *New approaches to the linguistics of humor* (pp. 180–196). Galati: Dunarea de Jos.
- Ku, L. C., Feng, Y. R., Chan, Y. C., Wu, C. L., & Chen, H. C. (2017). A re-visit of three-stage humor processing with readers' surprise, comprehension, and funniness ratings: An ERP study. *Journal of Neurolinguistics*, 42, 49–62. doi:10.1016/j.jneuroling.2016.11.008
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647. doi:10.1146/annurev.psych.093008.131123
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203–205.
- Kutas, M., & King, J. W. (1996). The potentials for basic sentence processing: Differentiating integrative processes. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 501–546). Cambridge, MA: MIT Press.
- Long, D. L., & Graesser, A. C. (1988). Wit and humor in discourse processing. *Discourse Processes*, 11, 35–60. doi:10.1080/01638538809544690
- 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, 113–130. doi:10.3758/s13415-010-0017-7
- Martin, R. A. (2007). *The psychology of humor: An integrative approach*. Burlington, MA: Elsevier.
- Mayerhofer, B., Maier, K., & Schacht, A. (2016). Priming interpretations: Contextual impact on the processing of garden path jokes. *Discourse Processes*, 53, 675–694. doi:10.1080/0163853X.2015.1092373
- Mayerhofer, B., & Schacht, A. (2015). From incoherence to mirth: Neuro-cognitive processing of garden-path jokes. *Frontiers in Psychology*, 6, 550. doi:10.3389/fpsyg.2015.00550
- Moran, J. M., Wig, G. S., Adams, R. B., Janata, P., & Kelley, W. M. (2004). Neural correlates of humor detection and appreciation. *NeuroImage*, 21, 1055–1060. doi:10.1016/j.neuroimage.2003.10.017
- Nieuwland, M. S., & Van Berkum, J. J. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, 24, 691–701. doi:10.1016/j.cogbrainres.2005.04.003
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785–806. doi:10.1016/0749-596X(92)90039-Z
- Ruch, W. (1992). Assessment of appreciation of humor: Studies with the 3WD humor test. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment* (pp. 27–75). Hillsdale, NJ: Erlbaum.
- Ruch, W., & Hehl, F. J. (2007). A two-mode model of humor appreciation: Its relation to aesthetic appreciation and simplicity-complexity of personality. In W. Ruch (Eds.), *The sense of humor: Explorations of a personality characteristic* (2nd ed., pp. 109–142). Berlin, Germany: Mouton de Gruyter.
- Samson, A. C., Hempelmann, C. F., Huber, O., & Zysset, S. (2009). Neural substrates of incongruity-resolution and nonsense humor. *Neuropsychologia*, 47, 1023–1033. doi:10.1016/j.neuropsychologia.2008.10.028
- Samson, A. C., Zysset, S., & Huber, O. (2008). Cognitive humor processing: Different logical mechanisms in non-verbal cartoons: An fMRI study. *Social Neuroscience*, 3, 125–140. doi:10.1080/17470910701745858
- Shibata, M., Terasawa, Y., Osumi, T., Masui, K., Ito, Y., Sato, A., & Umeda, S. (2017). Time course and localization of brain activity in humor comprehension: An ERP/sLORETA study. *Brain Research*, 1657, 215–222. doi:10.1016/j.brainres.2016.12.010
- Shibata, M., Terasawa, Y., & Umeda, S. (2014). Integration of cognitive and affective networks in humor

- comprehension. *Neuropsychologia*, *65*, 137–145. doi:10.1016/j.neuropsychologia.2014.10.025
- Suls, J. M. (1972). A two-stage model for the appreciation of jokes and cartoons. In P. E. Goldstein & J. H. McGhee (Eds.), *The psychology of humour. Theoretical perspectives and empirical issues* (pp. 81–100). New York, NY: Academic Press.
- Tu, S., Cao, X. J., Yun, X. J., Wang, K. C., Zhao, G., & Qiu, J. (2014). A new association evaluation stage in cartoon apprehension: Evidence from an ERP study. *Journal of Behavioral and Brain Science*, *4*, 75–83. doi:10.4236/jbbs.2014.42010
- Van den Brink, D., Van Berkum, J. J., Bastiaansen, M. C., Tesink, C. M., Kos, M., Buitelaar, J. K., & Hagoort, P. (2012). Empathy matters: ERP evidence for inter-individual differences in social language processing. *Social Cognitive and Affective Neuroscience*, *7*, 173–183. doi:10.1093/scan/nsq094
- Watson, K. K., Matthews, B. J., & Allman, J. M. (2007). Brain activation during sight gags and language-dependent humor. *Cerebral Cortex*, *17*, 314–324. doi:10.1093/cercor/bhj149
- Wyer Jr, R. S., & Collins 2nd, J. E. (1992). A theory of humor elicitation. *Psychological Review*, *99*, 663–688.