

Emotion Regulation through Irony: Evidence from Behavior and ERPs

Valeria A. Pfeifer & Vicky Tzuyin Lai (University of Arizona)

vpfeifer@email.arizona.edu

Verbal irony is when a speaker uses words that mean the opposite of what the speaker intends to say. What does this convey about the speaker's mental state? While Roberts & Kreuz (1994) claim that irony is primarily used to convey negative emotions, Filik et al. (2017) suggested that irony is perceived as less negative, meaning less hurtful and more amusing than literal statements. The current study used behavioral ratings and ERPs to investigate the role of irony in perceived emotional states. We hypothesized that a speaker using irony is perceived as being in a less negative mental state, and that this depends on how emotional the context is. We predicted that the effect of irony is enhanced only when a context is not strongly negative.

Stimuli consisted of 121 short stories in a 2 (emotion; high, low) x 2 (literality; literal, irony) design. In each story, a social context is described, where something more (high emotion) or less (low emotion) negative happens to one person. In the target sentence, the person makes an either literal or ironic comment. All stories were normed for valence, arousal and literality in a previous norming task. All critical words were used for both, literal and ironic conditions across contexts to control for psycholinguistic factors (see Tab.1).

Experiment 1 was a web-based rating study. 83 participants were asked to judge the emotional state of the speaker for valence and arousal on 5-point Likert-scales, respectively. A RM-MANOVA with valence and arousal as DVs revealed significant main effects of emotionality ($F(2,119)=48.5$, $p<.001$) and literality ($F=35.9$, $p<.001$), but no interaction ($p=.74$). Given the significance of the overall test, univariate main effects were examined with RM-ANOVAs. Significant main effects were obtained for literality in valence ($F=7.34$ (476), $p=.007$) and arousal ratings ($F=32.59$, $p<.001$), and a main effect of emotionality for arousal ($F=9.47$, $p=.002$).

In Experiment 2, 19 right-handed, native-English participants with a mean age of 18.8 read stimuli while their EEG was recorded. The story context was presented as a whole and read in a self-paced manner. The target utterance was presented word-by-word, with each word presented for a length-dependent duration (average: 386ms) followed by a 500ms blank. Comprehension questions after 20% of trials verified participation. ERPs were time-locked to the critical word, which was either literal or ironic. A RM-ANOVA revealed that ironic statements elicited a larger negativity from 450-650ms at the posterior sites than literal statements ($F(1,18)=4.76$, $p=.0426$), identified as a late N400. An interaction of literality and emotionality over the posterior sites in an P200 time-window from 200 to 300ms reached marginal significance ($F=3.98$, $p=.0615$). Given this result, we did pairwise comparisons and found ironic statements eliciting a marginally larger P200 than literal statements over central sites in the low emotion condition ($F=3.21$, $p=.09$).

Results from the behavioral rating task confirm our hypothesis that a speaker using irony is perceived as being in a less negative mental state, consistent with Filik et al. (2017). Similarly, regardless of literality of the utterance, a speaker in a more negative situation is perceived as being in a more negative state of mind compared to a speaker in a less negative situation. Results from ERPs suggest that in mildly negative situations only, ironic meaning receives more attentional resources early on, as reflected in a (marginally) larger P200. This shows that irony is perceived differently in mildly negative situations compared to strongly negative situations and could imply that the emotionality of the situation modulates when in the processing stream irony is detected. In addition, we found a larger N400 effect for ironic relative to literal statements, inconsistent with studies that found a P2-P6 complex for irony (e.g. Regel et al., 2010). This may be due to the overall negative valence of our stimuli, which may have created a global positivity that pulled waveforms more towards positive polarity. Our results are consistent with Baptista et al. (2018), who identified N400 as a marker of irony through its disruption via tDCS. In conclusion, speakers using irony are perceived as being in a less negative mental state compared to the same speaker using literal language. ERPs suggest that how emotional a context is can modulate irony perception, with low emotional situations facilitating irony detection.

Emotionality	Literality
<i>Example 1:</i>	
<u>High</u> : Max is helping Jenny with her computer when he accidentally spills a glass of water over the open computer. Jenny says:	<u>Literal</u> : How clumsy of you!
<u>Low</u> : Max is helping Jenny with her computer when he accidentally spills a glass of water over the closed computer. Jenny says:	<u>Ironic</u> : How considerate of you!
<i>Example 2:</i>	
<u>High</u> : When getting into the car, Ricarda accidentally sits on her boyfriend's new expensive sunglasses. Ricarda's boyfriend says:	<u>Literal</u> : I knew you were not as considerate as you said!
<u>Low</u> : When getting into the car, Ricarda accidentally sits on her boyfriend's new expensive scarf. Ricarda's boyfriend says:	<u>Ironic</u> : I knew you were not as clumsy as you said!

Table 1: Example of stimuli with critical words in bold. While paired with one context, words were used as literal, paired with a second context, words were used as ironic and vice versa.

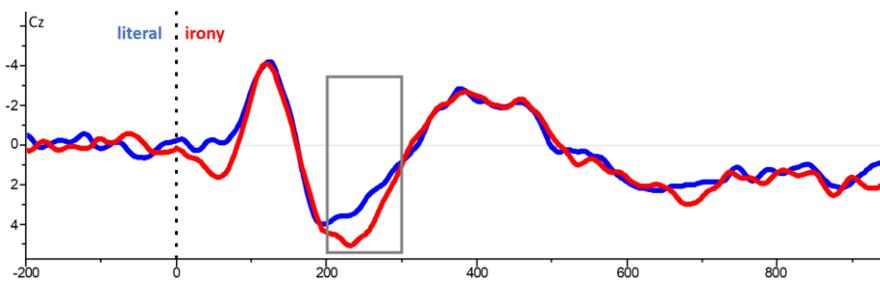
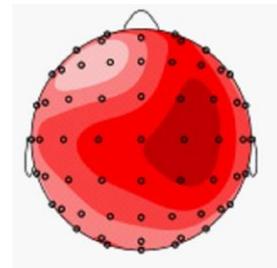


Fig. 1: Grand averaged waveforms for ironic and literal statements in low emotion condition at Cz.



200-300ms: irony minus literal in low emotion (P200)

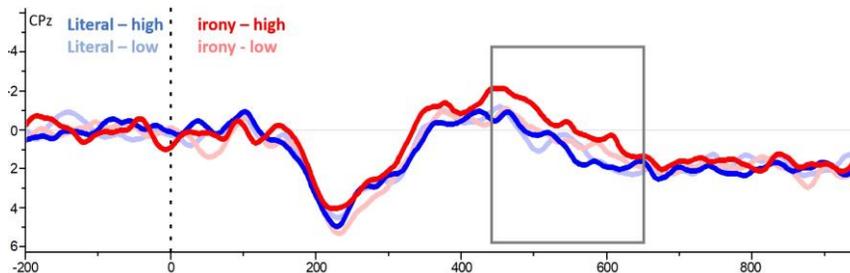
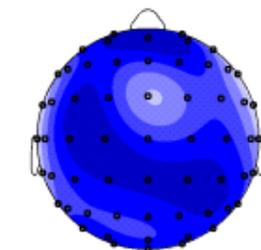


Fig. 2: Grand averaged waveforms at CPz. Ironic statements in red, literal statements in blue.



450-650ms: irony minus literal across conditions (N400)

References:

- Baptista, N. I., Manfredi, M., & Boggio, P. S. (2018). Medial prefrontal cortex stimulation modulates irony processing as indexed by the N400. *Social neuroscience*, 13(4), 495-510.
- Filik, R., Brightman, E., Gathercole, C., & Leuthold, H. (2017). The emotional impact of verbal irony: Eye-tracking evidence for a two-stage process. *Journal of Memory and Language*, 93, 193-202.
- Regel, S., Coulson, S., & Gunter, T. C. (2010). The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. *Brain research*, 1311, 121-135.
- Roberts, R. M., & Kreuz, R. J. (1994). Why do people use figurative language?. *Psychological science*, 5(3), 159-163.