


## Examining the Online Processing of Satirical Newspaper Headlines

Stephen Skalicky & Scott A. Crossley



To cite this article: Stephen Skalicky & Scott A. Crossley (2017): Examining the Online Processing of Satirical Newspaper Headlines, Discourse Processes

To link to this article: <http://dx.doi.org/10.1080/0163853X.2017.1368332>

 View supplementary material 

 Published online: 21 Sep 2017.

 Submit your article to this journal 

 View related articles 

 View Crossmark data 



# Examining the Online Processing of Satirical Newspaper Headlines

Stephen Skalicky and Scott A. Crossley

Department of Applied Linguistics & ESL Georgia State University

## ABSTRACT

Previous investigations of satire posit that satire comprehension is influenced by prior knowledge, satirical strategies, and other demographic features, such as age. However, these claims have not yet been tested using online processing techniques. In this study we investigate satire processing using newspaper headlines from the satirical American newspaper *The Onion* and nonsatirical headlines from *The New York Times*. Reading times and humor ratings were gathered from 76 participants using behavioral response data. Results demonstrate that a higher age results in significantly higher reading times for satirical compared with nonsatirical headlines. In addition, levels of prior knowledge significantly interacted with different satirical strategies, influencing both reading times and perceptions of humor. Measures of headline familiarity and sincerity also interacted with perceptions of humor and satirical strategy. These results provide additional evidence for the importance that age, prior knowledge, and satirical strategies have on satire comprehension.

## Introduction

Satirical language is an important component of human communication (Simpson, 2003). Although mostly known as a literary style for social or political criticism (Nilsen & Nilsen, 2008; Simpson, 2003), satire is commonly used to critique a variety of targets, as evidenced by modern satirical newspapers and television programming. Satire typically results in two discourse functions: the mockery of a satirical target and a humorous response from readers who agree with the satirical message (Simpson, 2003). Comprehension of satirical language is believed to be determined partly through ironic mechanisms (i.e., an opposition between what is said and what is meant), but also relies heavily on other nonlinguistic information, such as background and genre level knowledge, political beliefs, and awareness of authorial intent (Boukes, Boomgaarden, Moorman, & de Vreese, 2015; Lee & Kwak, 2014; Pfaff & Gibbs, 1997; Simpson, 2003). Moreover, different types of satire may rely on specific linguistic features, such as exaggeration, metaphor, or levels of explicitness (LaMarre, Landreville, Young, & Gilkerson, 2014; Simpson, 2003). However, to date, no research has investigated the role these features play using online behavioral methods.

The purpose of this study is to investigate satirical text and how it interacts with participant variables such as prior knowledge, age, satirical strategies, and perceptions of humor. We do so by comparing reading times for satirical news headlines taken from the satirical American newspaper *The Onion* and nonsatirical headlines taken from the newspaper *The New York Times*.

## Defining satire

Although there is no concrete definition of satirical language, the definition of satire we follow in this study is one that describes satire as a discursive practice that uses both irony and subtlety to

**CONTACT** Stephen Skalicky ✉ [sskalicky1@gsu.edu](mailto:sskalicky1@gsu.edu) 📍 Department of Applied Linguistics & ESL, Georgia State University, 15th Floor, 25 Park Place, Atlanta, GA 30303.

📎 Supplemental data for this article can be access on the [publisher's website](#).

© 2017 Taylor & Francis Group, LLC

critique a specific (and commonly political) target (Johnson, Del Rio, & Kemmitt, 2010; Pfaff & Gibbs, 1997; Simpson, 2003). Satire is ubiquitous in text and television, with examples ranging from Jonathon Swift's 1729 *A Modest Proposal* to contemporary television shows such as *The Daily Show* and *The Colbert Report*. In terms of written satire, one of the better known outlets of satirical content in the United States is *The Onion*.

### **The Onion and satirical news headlines**

*The Onion* is a satirical newspaper that regularly publishes satirical news stories that poke fun at a variety of topics, ranging from the mundane pleasantries of everyday life to pop culture and politics. Masquerading primarily as an online newspaper, *The Onion* follows the genre conventions of news articles and also provides video broadcasts using the accompaniments of serious cable and local news shows, replete with news anchors, reporters, and weather experts. This incongruity between the supposed genre of providing news and the real goal of mocking satirical targets is what defines *The Onion* as a primarily satirical outlet. One of the ways *The Onion* disseminates its stories is through news headlines, published on its main website and also broadcast through social media. For example, *The Onion* published the following three headlines in April of 2014:

- (a) "Area Man Mystified by Layout of Adjacent Town's Kroger"
- (b) "Couple Keeps It Interesting by Bickering in Different Positions, Rooms"
- (c) "Devastated Family Struggling to Cope After Losing Everything on DVR"

Headline (a) uses the news trope of *area man* to set up a story that at first appears to pertain to local news but instead pokes fun at human behavior, a rather mundane finding and not one that would rate highly as news. Headline (b) plays with the cliché of couples engaging in sexual activities in a variety of places but fills the cliché with a different activity: bickering. Headline (c) exaggerates the impact of a family's personal television recordings becoming lost, perhaps as a way to poke fun at television's role in American lives. All these headlines are represented as "news" but are not actually newsworthy, hence the satire.

### **Satirical strategies**

Simpson (2003) detailed four different satirical strategies that authors may use: attenuation, saturation, negation, and metaphor. Briefly, saturation exaggerates stereotypical qualities or perceptions of a particular person or thing. For example, the headline "Biden Loses Control of Butterfly Knife during Commencement Speech" exaggerates the public portrayal of the former U.S. Vice President's reported gaffes and generally casual demeanor. Attenuation works in the opposite manner by deflating typical elements or highlighting less salient elements of an entity or topic to implicitly draw attention towards something that is left unsaid. For example, the headline "High School Student, Teacher Applying for Same Summer Waitressing Job" is an implicit critique of the relatively low pay that high school teachers in the United States receive. Negation is a strategy that focuses on an event or situation that did not or has not happened and will use words expressing negation (e.g., *not*, *never*) to make this clear. For example, the headline "Obama Not Ruling Out U.S. Military Action in Congress" prompts the reader to consider the alternative: that military action is a feasible response to the gridlock of the American legislative system. Finally, metaphorical satire is the use of metaphor (i.e., construing an entity in a different content domain) to construct incongruous news stories. The headline "Budget Woes Force Heaven to Reduce Eternal Life to 500 Billion Years" construes heaven metaphorically as a corporation, placing heaven outside of its typically religious and spiritual content domain.

## Satire processing

No theoretical models currently exist that make predictions regarding satire processing. However, several different phenomena have been shown to influence satire processing, including background information, age, political beliefs and ideologies, attitudes, and ability to interpret authorial intent. For instance, Boukes et al. (2015) found that providing participants with background knowledge about a topic increased participants' appreciation of satirical political messages (measured through humor ratings) when those messages aligned with their political beliefs. Boukes et al. (2015) also found that younger participants (<35 years old) reported higher levels of concentration and lower levels of distraction while viewing satire than participants over 35 years of age. Becker (2014) and LaMarre, Landreville, and Beam (2009) found similar results, with both studies reporting that comprehension and appreciation of political satire was greater when participants agreed with the political message.

Different types of satire (i.e., satirical strategies) may also influence comprehension. LaMarre et al. (2014) asked participants to view two different types of political satire (Horatian or Juvenalian) and found that political satire with an overt and obvious humorous tone (Horatian) resulted in comprehension of just the humorous message, whereas satire with a less overt humorous tone (Juvenalian) resulted in processing of both a political and humorous message. These findings suggest that different satirical strategies can result in differing levels of cognitive load, which in turn may affect the time course of satire processing. Finally, Pfaff and Gibbs (1997) studied participants' ability to comprehend satirical intent on the part of the author. Their results demonstrated that participants were able to construct an impression of the author's satirical intent while reading the stories, and this impression was influenced by the level of exaggeration, wordiness, and information participants knew about the author. Furthermore, readers who were at first unaware of the author's satirical intent were able to reinterpret the satirical meaning of a text after being informed of the author's satirical intent. Together, these studies all suggest that a variety of variables, most notably levels of prior information, age, and participant attitudes, may influence the comprehension of satire. However, none of the previous studies used online processing methods to investigate the time course of satire processing. In general, features that increase the difficulty of comprehending satire are likely to also increase satire processing time.

Theories of general text processing may also provide predictions regarding satire processing. Situation and mental representation models of discourse processing demonstrate that readers construct mental representations of a text while reading through a combination of background knowledge and semantic inferences (van den Broek & Helder, 2017; Zwaan & Radvansky, 1998). Readers first rely on passive processes of comprehension (i.e., bottom up, automatic, and noneffortful comprehension) to develop a situation model for a particular text. When a text does not meet a reader's standards of coherence (i.e., does not match background knowledge or expectations), readers will turn to more purposeful, top-down reader-initiated processes such as rereading or deliberate consideration of a text's implied meaning (van den Broek & Helder, 2017). In terms of satire processing, if a reader never engages in reader-initiated processes while reading satire, he or she may only glean the surface meaning of the text through passive processes and miss the satirical meaning. Specifically, satirical texts likely possess two thresholds for a reader's standards of coherence: one related to a literal, surface level meaning and one related to a satirical meaning. Readers who do not recognize the satirical intent of a text (due to lack of background knowledge or other reasons) may nevertheless still meet their standards of coherence based on the literal meaning of a satirical text and thus never engage in the purposeful reader-initiated strategies required to understand the satirical message.

In sum, a variety of factors may influence how satire is processed and mentally represented. Specifically, higher levels of prior knowledge should facilitate satire processing, especially if that knowledge is directly related to the content of the text (LaMarre et al., 2009; Pfaff & Gibbs, 1997; Simpson, 2003). Additionally, because younger readers may be more engaged with satire (Boukes et al., 2015), they may have more readily developed passive and reader-initiated processes capable of

facilitating satire processing. Finally, because the different satirical strategies of attenuation, saturation, metaphor, and negation all range from more-to-less explicit in the satirical targets they name, strategies such as attenuation may be processed slower than more obvious strategies such as saturation, as they would require more effortful, reader-initiated processes of reading comprehension.

### Current study

In this study we investigate satire processing using newspaper headlines taken from *The Onion* and nonsatirical headlines from *The New York Times*. The following research questions guide our study:

- (1) Do reading times differ between satirical and nonsatirical newspaper headlines or among different satire strategies?
- (2) Do individual differences, such as prior knowledge, age, and perceptions of humor, influence satire processing?
- (3) Do perceptions of humor differ among different satirical strategies?

## Methods

### Participants

Seventy-eight undergraduate students were recruited from applied linguistics courses at a large southeastern public university in the United States. Of these 78 participants, 55 were women and 23 men. Eleven participants reported left-hand dominance. Average participant age was 24 years old, with lowest age 18 and highest age 66. Sixty-six participants spoke English as their first language, and 55 participants reported they spoke more than one language. Ten participants were in their first year of college, 18 in their second year, 22 in their third year, and the remaining 28 were in their fifth year or greater. All students were provided with extra credit in their applied linguistics courses in exchange for their participation in the study.

### Materials and study design

We collected news headlines from the Twitter feeds of *The Onion* and *The New York Times*. We began by choosing the first 100 headlines found on *The Onion*'s twitter feed during early 2015. We then isolated any headlines that belonged to one of the four satirical strategies of attenuation, saturation, negation, and metaphor using the following definitions and examples, shown in [Table 1](#).

Two raters were trained using these categories. The raters first worked individually to code the headlines, and any disagreements between category membership resulted in the headline being replaced. This procedure was followed until both raters expressed 100% agreement for 10 headlines in each strategy condition, resulting in a total of 40 satirical headlines. To keep the content of the

**Table 1.** Satirical Strategy Coding Descriptions with Representative Samples.

Strategy	Definition	Example
Attenuation	Headlines that are purposefully ambiguous	<i>'State of the Union Strong,' Says Man Burdened with Protecting Nation from Truth</i>
Metaphor	Headlines that invite new perspectives though cross-domain mappings	<i>Drone Places Fresh Kill on Steps of White House</i>
Negation	Headlines that imply or highlight negated events	<i>Chinese Factory Workers Terrified Machines Never to Replace Them</i>
Saturation	Headlines that exaggerate stereotypes	<i>North Korea Successfully Harvests Wheat in Show of Growing Strength</i>

satirical and nonsatirical headlines as contextually similar as possible, we then chose the first 100 headlines from *The New York Times*' Twitter feed and selected 40 headlines from this list that matched, as much as possible, in content with the headlines from *The Onion*.

Because a variety of linguistic features may influence the reading times of sentences, we measured the headlines for lexical, syntactic, and affective features using automatic text analysis tools. We used TAALES (Kyle & Crossley, 2015) to examine the headlines for measures of lexical frequency, meaningfulness, concreteness, imagability, and age of acquisition, TAASC (Kyle, 2016) to measure the syntactic complexity of the headlines, and SEANCE (Crossley, Kyle, & McNamara, 2016) to conduct sentiment analysis of the headlines. The headlines initially demonstrated significant differences in total number of content words and content word frequency. Where possible, headlines from *The New York Times* were replaced to reduce linguistic differences among the headlines types. However, it was still necessary to manipulate some of the satirical headlines to include more or less frequent content words and total numbers of content words. The final list of manipulated headlines demonstrated no significant differences for any of the selected linguistic features and are presented in Supplemental Appendix A.

After controlling the headlines for linguistic features, we used Amazon Mechanical Turk (a website where online workers complete tasks for small fees) to collect human ratings of the headlines for five additional features: understandability of the content, familiarity of the content, positivity of the headlines, sincerity of the headlines, and humor of the headlines. We included these measures to capture additional affective and cognitive features to complement the results from the linguistic tools above and to further control for additional influences on the headline processing. Workers were randomly presented with 10 headlines (5 satirical and 5 nonsatirical) and answered *yes* or *no* for the following questions for each of the headlines: (1) I am familiar with the content, (2) I understand this headline, (3) The headline is positive, (4) The headline is sincere, and (5) The headline is funny. Each headline was rated by at least 30 different workers, and workers were paid \$.10 to rate 10 headlines one time (workers could not complete the job more than once). Percentage scores for each feature per headline were computed by dividing the total number of *yes* answers by the amount of raters per headline. Nonparametric Welch test ANOVAs indicated that significant differences existed among satirical and nonsatirical headlines for ratings of understandability ( $F[4, 20.41] = 20.36, p < .001$ ), familiarity ( $F[4, 26.43] = 11.30, p < .001$ ), sincerity ( $F[4, 22.19] = 92.13, p < .001$ ), and humor ( $F[4, 19.83] = 92.57, p < .001$ ). Post-hoc comparisons revealed that all satirical headlines differed significantly from the nonsatirical headlines, in that nonsatirical headlines were significantly less funny and significantly more understood, familiar, and sincere than the satirical headlines. None of the satirical headline strategies differed significantly from one another for any of these features. Table 2 displays descriptive statistics for the linguistic features and crowd-sourced human ratings.

After the list of headline stimuli was finalized, we developed comprehension questions for each headline using the format "The headline was about \_\_\_\_." We chose comprehension topics that related to the surface level of the sentence only and did not relate to any of the implied satirical meanings for any of the satirical headlines. Thus, the comprehension questions served to measure participant attention rather than ability to interpret satirical meaning. Comprehension questions were balanced so that correct answers for half of the headlines in each of the five conditions were true and the other half were false.

A prior knowledge test was included to control for the general role that background knowledge plays in text processing and comprehension (McNamara & Kintsch, 1996; McNamara, Kintsch, Songer, & Kintsch, 1996) and to test if prior knowledge influences satire processing specifically (Boukes et al., 2015; Simpson, 2003). The prior knowledge test consisted of 30 multiple-choice questions (with four answer options) about history, science, and literature (10 questions each). The questions for this test were originally designed to prepare high school students in New York for state mandated examinations.

**Table 2.** Linguistic Features of Headlines by Condition.

Linguistic Feature	Attenuation		Metaphor		Negation		Saturation		Nonsatirical	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Mean age of acquisition score, content words	6.56	1.01	6.53	1.36	7.09	0.68	6.48	0.37	7.12	0.79
Mean unigram concreteness score, content words	449.25	46.51	464.96	72.94	415.52	48.47	393.84	140.12	464.96	72.94
Mean unigram familiarity score, content words	568.35	20.91	565.66	35.05	562.34	23.72	572.60	20.56	556.77	20.87
Total number of words in headline	10.20	1.93	10.20	2.14	11.10	2.28	11.40	1.95	10.52	1.39
Composite VADER score	-0.01	0.04	-0.04	0.38	-0.12	0.36	0.04	0.41	-0.14	0.37
Mean unigram imagability score, content words	475.79	37.34	480.26	51.09	445.74	34.44	456.03	75.25	459.46	48.63
Mean length of clause	10.50	1.95	8.95	2.03	9.85	2.51	11.40	1.95	10.25	2.08
Mean frequency score, content words (COCA)	212.21	127.16	219.59	104.95	287.44	138.93	185.58	74.76	296.94	295.10
Number of content words	8.20	1.61	8.10	1.66	9.00	1.49	8.60	1.17	8.27	0.96
Percentage of raters rating headline as positive	26.62	28.82	22.65	22.70	34.78	23.19	21.76	15.90	33.62	35.06
Percentage of raters rating headline as familiar <sup>a</sup>	18.86	13.60	17.42	6.41	21.50	12.48	19.64	8.02	42.50	19.93
Percentage of raters rating headline as understandable <sup>a</sup>	70.98	9.05	66.88	11.61	71.78	18.59	76.55	14.13	91.51	7.03
Percentage of raters ratings headline as funny <sup>a</sup>	41.65	22.41	57.98	18.79	59.55	9.24	51.57	27.93	3.58	7.51
Percentage of raters rating headline as sincere <sup>a</sup>	29.14	21.94	21.41	9.45	29.23	8.85	32.98	25.50	76.85	10.61

<sup>a</sup>Significant differences among headlines types exist for these measures based on Welch ANOVAs. All four satire conditions differed significantly from the nonsatirical condition based on Tukey post-hoc analyses (all  $p < .001$ ). All satirical headlines were rated as less familiar, less understandable, less sincere, and more funny compared with nonsatirical headlines. There were no significant differences between these measures among the separate satire conditions. No significant differences existed for any of the other linguistic features.

### Apparatus and procedure

All participants completed the experiment in a single session seated in front of a desktop computer in a soundproof room. Participants first completed the prior knowledge test using the online survey platform Qualtrics (Qualtrics, Provo, UT). After completing the prior knowledge test, participants entered their demographic information (e.g., handedness, gender, age, language knowledge) using E-Prime (Psychology Software Incorporated, Sharpsburg, PA) before starting the experiment. The experiment stimuli were also presented to participants using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). Participants were first shown a screen displaying written instructions. The instructions informed participants they would complete two phases of the experiment. In the first phase of the experiment participants were to read a series of news headlines and then press the space bar on the computer keyboard once they were done reading each headline. They would then answer a true or false comprehension question about the headline they had just read. In the second phase participants were to rate some of the same headlines for humor, using a scale of 1 to 4. The researchers read these instructions to the participants and answered any questions before initiating the experiment.

In the first phase of the experiment, participants initiated each trial by pressing the spacebar on the keyboard in front of them. Each trial began with a fixation element comprised of a row of asterisks (“\*”) displayed on the center of the screen for 250 ms. After the fixation period, a randomly chosen headline appeared on the center of the screen. Participants read the headline and pressed the spacebar when finished. After pressing the spacebar, a comprehension statement about the headline was displayed on the screen (e.g., *The headline was about the United States*). Participants pressed “1” if the statement was true and “2” if the statement was false. After completing a practice session of five trials, participants were presented with the 80 stimuli for the first phase of the experiment.

The second phase of the experiment began as soon as participants completed the first phase. Before starting, participants were again provided with written instructions explaining they would now see some of the same headlines again and rate how funny they perceived the headlines to be. After completing a practice session of three trials, participants initiated the experiment by pressing the spacebar. As in the first phase of the experiment, a 250-ms fixation element of asterisks was displayed on the screen before being replaced with one of the headlines. On the screen and underneath the headline, a four-point scale displayed the four humor rating options: not funny, somewhat funny, funny, or very funny. Participants made their choices using the 1 to 4 number keys on their keyboards, with 1 being “not funny” and 4 being “very funny.” Only the 40 satirical headlines were repeated during this phase of the experiment and were presented in random order.

### Statistical analysis

We analyzed the data using linear mixed effect (LME) models due to their statistical advantages over traditional repeated-measures ANOVAs and regression models. LMEs can include both fixed and random effects, allowing for individual variation among participants and items to be retained in the models and therefore providing a more accurate measure of the influence of independent variables on a dependent variable. We used R (R Core Team, 2016) and the *lme4* (Bates, Mächler, Bolker, & Walker, 2015) and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2016) packages to build and interpret our models. To derive measures of effect sizes, we used the *MuMIn* package (Nakagawa & Schielzeth, 2013), which provides two  $R^2$  measures: a marginal  $R^2$  measuring the variance explained by the fixed effects and a conditional  $R^2$  measuring the variance explained by the fixed and random effects.

To systematically test our research questions, we opted for maximal models in which we entered all random effects and slopes justified by the experiment design (Barr, Levy, Scheepers, & Tily, 2013) as well as a model testing only the effects of interest without any control variables included (e.g., satirical vs. nonsatirical headline reading times). All categorical variables were entered using treatment coding. We constructed three different LME models: one comparing the reading times for satirical and nonsatirical headlines, one analyzing only the four different satirical headline strategies, and one analyzing the humor ratings assigned to the satirical headlines by participants. In the reading time model for satirical and nonsatirical headlines, we entered reading times (in milliseconds) as the dependent variable and the following fixed effects related to our first research question: participants' age, prior knowledge scores, and the headline type (satirical or non-satirical). To answer the second research question, we kept the fixed effects from the first model (except for headline type of satire or nonsatire) and added satirical strategy (attenuation, metaphor, negation, or saturation) and participant humor ratings (scale of 1 to 4). To control for features that may influence general text processing, we also included fixed factors of sex (male or female); handedness (left or right); total number of languages known; whether English was a participant's first language (yes or no); college year and grade point average; the crowd-sourced ratings of headline familiarity, positivity, sincerity, and understandability; trial order of the headlines; and the total number of content words<sup>1</sup> to both models. We also included interactions between headline type or satirical strategy and age, total number of languages, prior knowledge scores, first language, humor ratings, and the crowd-sourced ratings of familiarity, positivity, sincerity, and understandability, only retaining significant interactions in the final models to properly interpret the coefficients for other main effects in the models. Subjects and headlines were entered as random effects, with a random slope of headline type (first model) or satirical strategy (second model) added to subjects.

---

<sup>1</sup>Despite showing no significant differences among headline conditions using Welch ANOVAs, exploratory modeling demonstrated that number of content words was still a significant predictor of reading times and was thus retained in the final LME models. None of the other linguistic features reported in Table 2 were significant according to exploratory LME models and were thus not included.



For the third model examining the humor rating data, we entered humor ratings as the dependent variable (a scale of 1 to 4) and the following variables as fixed effects: age, prior knowledge scores, and satirical headline strategy (attenuation, metaphor, negation, or saturation). These variables are also related to our second and third research questions. Like the first two models we also included control variables of sex (male or female), whether English was a participant's first language (yes or no), total languages, college year and grade point average, and the crowd-sourced ratings of familiarity, positivity, sincerity, and understandability. We also tested interactions between satirical headline strategy and age, prior knowledge, first language, total number of languages, and the crowd-sourced ratings of familiarity, positivity, sincerity, and understandability, only retaining significant interactions in the final model. Subjects and headlines were entered as random effects, with satirical headline entered as a random slope for subjects.

## Results

We performed initial trimming of the reading time data based on extremely quick responses to the headlines and also the accuracy of responses to the headline comprehension questions. We first calculated total accuracy for all participants and removed two participants whose overall accuracy was below 80%. We then removed any trials where participants responded to the headline in 1 second or less, removing 16 trials (.002% of the data). Next, we removed all inaccurate trials, deleting 467 more trials (.08% of the data). After this initial trimming, visual inspection indicated the reading time data remained positively skewed. To account for these outliers, we opted to winsorize the data, wherein we replaced outliers with the next highest nonoutlier score (Field, 2009). We did so because it is difficult to know which of the positive outliers are due to participant inattention or legitimate reading of the sentences (Ratcliff, 1993) and also to retain the original reading time scale for more accurate interpretation of results. To winsorize, we calculated means and standard deviations for each participant and classified all outliers as reading times that were greater than 2.5 standard deviations from each participants' individual mean (i.e., all outliers were unique to each participant). We then replaced all participants' outliers with their next highest nonoutlier reading time, which included a total of 487 values (.09% of the data). In addition, before constructing our models, we centered all numerical predictor variables by subtracting the grand mean from each variable so that each variable's mean value was zero. Finally, we checked the predictor variables for multicollinearity using variance inflation values and correlations and found that the crowd-sourced ratings of sincerity and humor were strongly multicollinear ( $r = -.87$ , variance inflation value = 8.91). We thus removed the crowd-sourced ratings of humor because they represented a more subjective measure and also because we obtained ratings of humor from the participants themselves. After removing the crowd-sourced humor ratings variance inflation values for all remaining predictors were  $< 3$ .

### **Headline reading times: Satirical and nonsatirical headlines**

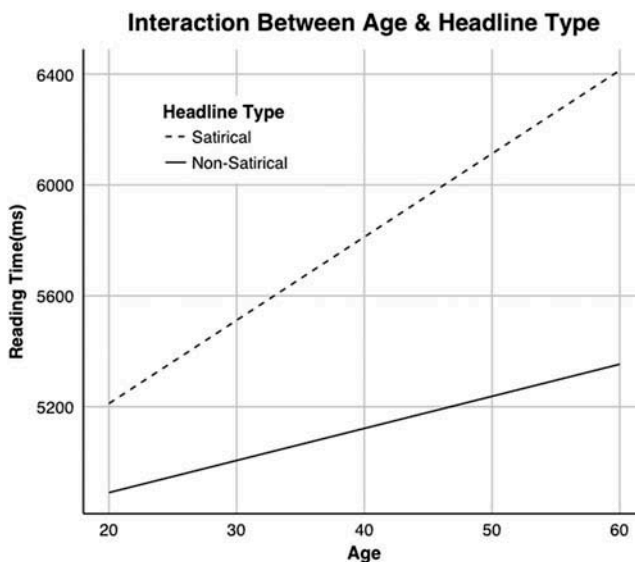
The mean reading time for the satirical headlines was 5329 ms ( $SD = 2301$ ) and for the nonsatirical headlines 4923 ms ( $SD = 2219$ ). In an LME model containing headline type as the only fixed effect, satirical condition significantly increased reading times by 427 ms ( $SE = 190.02$ ,  $df = 82.41$ ,  $t = 2.25$ ,  $p = .027$ ), suggesting that satire does take longer to process than nonsatire (RQ1). However, in a fully defined LME model containing all predictors listed above, this effect was no longer significant when taking into account hypothesis- and control-related fixed effects. In this model there was a significant, ordinal interaction between age and headline type such that a higher age significantly increased reading times for satirical headlines when compared with nonsatirical headlines times. This indicates that although the mean reading time for satirical headlines was higher than the nonsatirical headlines, the relative age of the participants played a significantly stronger role during the processing of satirical compared with nonsatirical headlines (RQ1 and RQ2). In other words,

higher age was associated with more difficulty in processing satirical headlines when compared with nonsatirical headlines, as evidenced by significantly longer reading times for satirical headlines at a higher age when compared to nonsatirical headlines at a higher age. This interaction is plotted in Figure 1.

No other variables related to the research questions demonstrated significant main effects or interactions. Several control variables significantly affected reading times but did not significantly interact with headline type, suggesting they accounted for features related to general text processing. Specifically, a higher number of total known languages, higher understandability and positivity ratings, and higher trial order all significantly decreased reading times, whereas not having English as a first language and a higher number of content words served to significantly increase headline reading times. The marginal  $R^2$  for this model was .15 and the conditional  $R^2$  was .43. Table 3 displays the coefficients, standard error, degrees of freedom, and  $t$  and  $p$  values for the predictors in this model.

### Headline reading times: Satirical strategies

The mean reading times for the different satirical strategies were as follows: attenuation = 5239 ms ( $SD = 2259$ ), metaphor = 5337 ms ( $SD = 2321$ ), negation = 5469 ms ( $SD = 2334$ ), and saturation = 5270 ms ( $SD = 2289$ ). In an LME model containing satirical strategy as the only fixed effect, there were no significant differences in reading times among all four satire strategies,<sup>2</sup> suggesting the different satirical strategies did not significantly increase or decrease headline reading times (RQ1). In a fully defined model containing all predictors of interest, several subject- and item-level variables significantly affected reading times. Age was a significant positive predictor of headline reading times in that an increase in age resulted in an increase in reading times, but there was no significant



**Figure 1.** Interaction plot displaying the effect of age on headline reading times when all other predictor variables were controlled for in an LME model. The distance between the two lines at a higher age demonstrates a significant difference in that a higher age results in significantly higher reading times for satirical compared with nonsatirical headlines. In the model, age was centered by subtracting the grand mean from each value so that the mean of age was equal to zero.

<sup>2</sup>Comparisons across all four levels of the satirical strategies were conducted using contrast coding and the *multcomp* package (Hothorn, Bretz, & Westfall, 2008).

**Table 3.** LME Model Comparing Reading Times for Satirical and Nonsatirical Headlines.

Fixed Effect	Coefficient	SE	df	t	p
Intercept	5335.25	397.12	92	13.44	<.001
Subject-level features					
Trial order <sup>a</sup>	-14.89	1.02	6442	-14.64	<.001
Total languages <sup>a</sup>	-376.57	118.62	67	-3.18	.002
English L1: no <sup>a</sup>	866.09	375.53	67	2.31	.024
College GPA	-159.16	175.22	67	-0.91	.367
Prior knowledge score	27.17	35.58	67	0.76	.447
Age	11.58	15.48	67	0.75	.457
College year	89.36	125.79	67	0.71	.479
Handedness: right	40.83	370.62	67	0.11	.912
Sex: male	158.53	287.40	67	0.55	.583
Item-level features					
Number of content words <sup>a</sup>	392.25	57.21	73	6.86	<.001
Familiarity ratings	0.84	4.75	72	0.18	.859
Positivity ratings <sup>a</sup>	-5.81	2.64	73	-2.20	.031
Sincerity ratings	9.86	5.80	74	1.70	.093
Understandability ratings <sup>a</sup>	-21.25	7.26	75	-2.93	.004
Headline type: satirical	391.23	294.93	75	1.33	.188
Significant interactions (nonsatirical baseline)					
Age × satirical headlines	18.47	6.03	73	3.07	.003

Baseline levels for categorical variables: English L1 = yes, handedness = left, sex = female, headline type = nonsatirical. All other predictors were centered by subtracting the grand mean from each variable. Adjusted degrees of freedom (*df*) were calculated using the *lmerTest* package. No significant interactions existed between headline type and total number of languages, prior knowledge scores, and crowd-sourced ratings of familiarity, positivity, sincerity, and understandability.

<sup>a</sup>This variable had a significant influence on reading times.

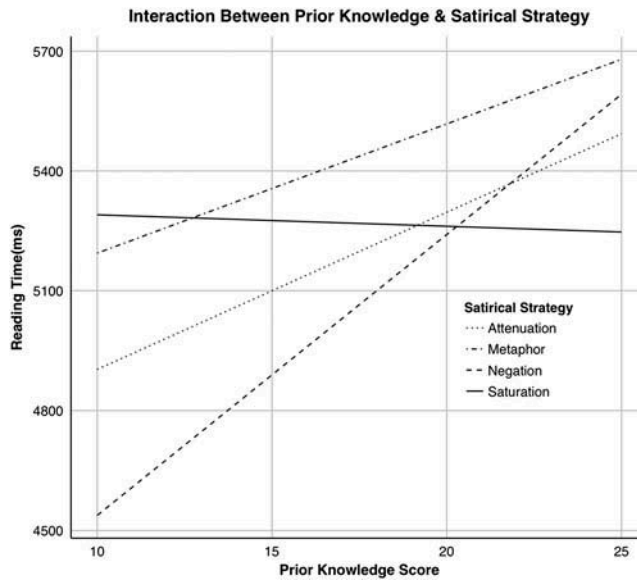
interaction among age and the different satirical strategies. This further suggests that higher age is associated with more difficulty for satire processing (RQ2). In addition to age, there was also a significant interaction between prior knowledge and satirical strategy. Specifically, higher levels of prior knowledge resulted in significantly higher reading times for negation headlines when compared with saturation headlines, suggesting that a difference in the processing of satirical strategies does exist if participants have higher levels of prior knowledge (RQ2), but only for specific satirical strategies. This interaction is plotted in [Figure 2](#).

No other variables related to the research questions demonstrated significant main effects or interactions. As with the first model, several of the same control variables significantly affected reading times in the same manner but did not interact significantly with satirical strategy (total number of languages, English as a first language, understandability ratings, number of content words, and trial order). This marginal  $R^2$  for this model was .18 and the conditional  $R^2$  was .48. [Table 4](#) displays the coefficients, standard error, degrees of freedom, and *t* and *p* values for the predictors entered into the model.

### **Humor ratings**

The mean humor ratings for the different satirical strategies were as follows: attenuation = 1.94 ( $SD = 1.05$ ), metaphor = 1.934 ( $SD = 1.04$ ), negation = 2.06 ( $SD = 1.07$ ), and saturation = 2.07 ( $SD = 1.06$ ). In an LME model containing satirical strategy as the only fixed effect, there were no significant differences in humor ratings among the four satirical strategies, suggesting that perceptions of humor do not significantly differ among satirical strategies (RQ3). In a fully defined model containing all predictors of interest, there was a significant, positive main effect for prior knowledge scores in that higher levels of prior knowledge resulted in higher humor ratings for all satirical strategies. This suggests that higher levels of prior knowledge facilitate perceptions of satirical humor, which may serve as a proxy of satire comprehension (RQ2).

Additionally, there were significant interactions between two control variables and satirical strategy. Specifically, attenuation headlines with higher familiarity ratings were rated as significantly



**Figure 2.** Interaction plot displaying the effect of prior knowledge score on reading times when all other variables were controlled for in an LME model. The distance between negation (dashed line) and saturation headlines (solid line) for reading times became significant as levels of prior knowledge increased. There were no other significant differences among any of the other satirical strategies. In the model, prior knowledge was centered by subtracting the grand mean from each value so that the mean was equal to zero.

**Table 4.** LME Model Comparing Reading Times for Satirical Strategies.

Fixed Effect	Coefficient	SE	df	t	p
Intercept	4940.49	447.07	94.20	11.05	<.001
Subject-level features					
Trial order <sup>a</sup>	-14.94	1.44	2588.80	-10.36	<.001
Total languages <sup>a</sup>	-399.28	125.78	66.60	-3.17	.002
English L1: no <sup>a</sup>	1159.85	398.71	67.10	2.91	.004
Age <sup>a</sup>	38.16	16.39	66.50	2.33	.022
Prior knowledge score	70.20	40.67	68.50	1.73	.088
College GPA	-223.98	185.69	66.50	-1.21	.232
Sex: male	304.58	304.94	66.80	1.00	.321
College year	35.06	133.43	66.70	0.26	.793
Handedness: right	98.93	392.59	66.20	0.25	.801
Item-level features					
Number of content words <sup>a</sup>	414.67	86.45	29.90	4.80	<.001
Familiarity ratings	21.76	12.89	30.30	1.69	.101
Positivity ratings	-4.21	5.80	30.30	-0.72	.473
Sincerity ratings	6.25	8.32	30.40	0.75	.459
Understandability ratings <sup>a</sup>	-18.55	9.27	30.60	-2.00	.054
Participant humor ratings	18.49	124.91	29.80	0.15	.993
Strategy: saturation	-29.86	333.64	30.20	-0.09	.929
Strategy: attenuation	56.98	325.91	30.30	0.18	.862
Strategy: metaphor	207.51	331.02	30.40	0.63	.535
Significant interactions (negation baseline)					
Prior knowledge × saturation	-73.14	25.26	1938.40	-2.90	.003

Baseline levels for categorical variables: English L1 = yes, handedness = left, sex = female, satirical strategy = negation. All other predictors were centered by subtracting the grand mean from each variable. Adjusted degrees of freedom (*df*) were calculated using the *lmerTest* package. No significant interactions existed between satirical strategy type and total number of languages, prior knowledge scores, humor ratings, age, and crowd-sourced ratings of sincerity, familiarity, and understandability.

<sup>a</sup>This variable had a significant influence on reading times.

funnier than metaphor headlines with higher familiarity ratings, and attenuation headlines with higher sincerity ratings were rated as significantly less funny than negation headlines with higher sincerity ratings (these findings are discussed in more detail below). No other main effects or interaction effects were significant. Table 5 displays the coefficients, standard error, degrees of freedom, and  $t$  and  $p$  values for the predictors entered into the model. The marginal  $R^2$  for this model was .08 and the conditional  $R^2$  was .31.

## Discussion

In this study we investigated the online processing of satire by comparing reading times for satirical and nonsatirical newspaper headlines. In addition, we compared reading times and participant humor ratings for four different satirical strategies (attenuation, metaphor, negation, and saturation; Simpson, 2003). Based on previous research demonstrating that satire comprehension relies in part on factors such as prior knowledge and age (Boukes et al., 2015; LaMarre et al., 2014), we also measured participants' age and prior knowledge using a 30-question multiple-choice prior knowledge test, along with several other control variables designed to account for features that affect text processing in general.

The first research question asked whether satirical and nonsatirical headlines were processed differently when measured using behavioral reading response times and also whether there were differences among satire strategies for reading times. Results from our analysis suggested that when taking no other variables into consideration, satirical headlines were read significantly slower than nonsatirical headlines and that there were no significant differences in reading times among the four satirical strategies. These results suggest that satire takes longer to process than nonsatire and that no satirical strategy was more or less difficult to process than another. However, the results from the second research question altered this interpretation.

**Table 5.** LME Predicting Humor Ratings for Satirical Headlines.

Fixed Effect	Coefficient	SE	df	$t$	$p$
Intercept	2.10	0.16	92.88	13.29	<.001
Subject-level features					
Prior knowledge score <sup>a</sup>	0.03	0.01	70.00	2.35	.021
Age	<0.01	0.01	70.00	0.15	.883
Sex: male	0.20	0.11	70.00	1.67	.099
College GPA	-0.11	0.07	70.00	-1.50	.138
College year	-0.05	0.05	70.00	-0.94	.349
English L1: no	0.21	0.15	70.00	1.44	.154
Total languages	0.02	0.05	70.00	0.52	.608
Item-level features					
Sincerity ratings	-0.02	0.01	26.00	-3.15	.004
Familiarity ratings	0.02	0.01	26.00	1.77	.088
Positivity ratings	0.06	0.06	26.00	1.05	.303
Understandability ratings	<0.01	< 0.01	26.00	1.03	.310
Strategy: saturation	0.16	0.12	27.68	1.30	.204
Strategy: negation	0.08	0.12	26.48	0.70	.491
Strategy: metaphor	-0.08	0.14	26.99	-0.59	.563
Significant interactions (attenuation baseline)					
Familiarity × negation	-0.03	0.01	26.00	-2.75	.010
Sincerity × negation	0.03	0.01	26.00	2.54	.017

Baselines for categorical variables: Sex = female, English L1 = yes, satirical strategy = attenuation. All other predictors were centered by subtracting the grand mean from each variable. Adjusted degrees of freedom ( $df$ ) were calculated using the *lmerTest* package. No significant interactions existed between satirical strategy type and total number of languages, prior knowledge scores, age, and crowd-sourced ratings of positivity and understandability. For the significant interactions between strategy and familiarity ratings and strategy and sincerity ratings, multiple comparisons revealed only two significant differences, reported above.

<sup>a</sup>This variable had a significant influence on reading times.

The second research question asked whether individual differences, such as prior knowledge, age, and perceptions of humor, influence the processing of satire. Results from both our first and second models suggest that age plays an important role in satire processing. In our first model there was a significant ordinal interaction between age and headline type, where the reading time difference between satirical and nonsatirical headlines increased as participants' age increased (Figure 1). Age also appeared in the second model, but this time as a significant main effect on all satirical headlines, in that all satirical headlines took longer to read as age increased. Together, the results from these two models indicate that age plays an important role in satire processing, and that differences in reading times between satirical and nonsatirical headlines may reflect the relative difficulty of reading satire as one's age increases. It may also be the case that age is reflective of relative familiarity or exposure to satirical news genres, with younger readers being more experienced with media similar to *The Onion* and other satirical outlets. This aligns with previous research demonstrating that younger participants were more engaged with satire when compared with older participants (Boukes et al., 2015). In the current study participants were all relatively young (only four participants were over 35, the upper threshold identified by Boukes et al. [2015] for satire engagement). Thus, the lack of significant differences in reading times between satirical and nonsatirical headlines when age was controlled for may be attributed to the relative youth of the participants who are potentially more familiar and engaged with satire.

In addition to the main effect of age, prior knowledge significantly interacted with satirical strategy in the second model. As Figure 2 displays, the difference in reading times between saturation and negation headlines became significant at higher levels of prior knowledge, in that negation headlines took significantly longer to read when compared with saturation headlines. Reasons for this difference can be explained by considering how each strategy achieves a satirical effect. Because saturation headlines exaggerate stereotypes of known entities using a "more of the same" approach (Simpson, 2003, p. 128), higher levels of prior knowledge could correlate with knowledge related to the saturated entities. For example, more explicit knowledge of the poverty levels in North Korea along with the desire among North Korea's leadership to develop technology on par with more technologically advanced countries may help a reader better understand the satirical message behind the headline "North Korea Successfully Harvests Wheat in Show of Growing Strength." Conversely, higher levels of prior knowledge may create a tension when reading negation headlines, as the negation strategy is one that purposefully subverts actions and ideas that may be regarded as common sense. For example, the headline "Wildlife Experts Report Not Climbing into Gorilla Enclosure Likely Prevented Man's Death" reports on an event that rarely occurs because most people do not require an expert to inform them that climbing into a gorilla enclosure could be dangerous. As a result, readers with higher levels of prior knowledge may be left unsure as to why this story is newsworthy or even if it is satirical, resulting in higher reading times when compared with saturation headlines.

The answers to the second research question help qualify the answers to the first research question. Although a simple comparison between satirical and nonsatirical headline reading times suggests that satirical headlines are read significantly slower than nonsatirical headlines, this effect is reliant on an individual's age, which may itself be a proxy for familiarity with or exposure to satire. Future research investigating differences in satirical and nonsatirical texts should consider tightly controlling participants' age (e.g., using only young or only older participants or collecting a larger sample from all age ranges) to better understand effects related to age. In addition to the satire and nonsatire distinction, there were no significant differences for reading times among the four satirical conditions until prior knowledge was included in the model, which significantly interacted with reading times and led to a significant difference between negation and saturation headline reading times. In general, these results support prior research suggesting that age and levels of prior knowledge are important components of satire processing (Becker, 2014; Boukes et al., 2015; LaMarre et al., 2009, 2014; Pfaff & Gibbs, 1997).

Finally, the third research question asked whether participants' humor ratings would differ based on satirical strategy. The results demonstrated that humor ratings did not differ significantly based on headline strategy, but we did find that prior knowledge scores significantly predicted humor ratings for all strategies (with no interaction), such that higher levels of prior knowledge led to higher humor ratings (Table 5). We interpret these findings to further indicate the importance of prior knowledge on satire comprehension as addressed in the second research question. Although humor comprehension is not necessary for a reader to understand a satirical message (Johnson et al., 2010), finding a satirical message humorous can still serve as a proxy of satire comprehension (Simpson, 2003). Additionally, while previous studies based their findings on background knowledge measures related specifically to the types of satire participants read or watched (Boukes et al., 2015; LaMarre et al., 2009; Lee & Kwak, 2014; Pfaff & Gibbs, 1997; Simpson, 2003), our measure of prior knowledge was much more general and related to nonspecific knowledge (i.e., knowledge of literature, science, and world history). Thus, these results also provide evidence suggesting that general knowledge, in addition to specific knowledge, may facilitate satire comprehension when measured through perceptions of humor.

It should not be surprising that generalized, prior knowledge is an important component of satire processing because levels of background knowledge are important for text processing in general. Readers draw on background knowledge when constructing a mental representation of a text, and different levels of background knowledge partially determine whether a reader alternates between passive and reader-initiated processes during reading (van den Broek & Helder, 2017). With satire, a reader that does not engage in reader-initiated processes may never make the inferences necessary to comprehend the satirical intent of the author. Conversely, a reader with sufficient general background knowledge may better understand satirical intent using passive and reader-initiated processes. Furthermore, the goals of a reader strongly influence how he or she reads a text (van den Broek & Helder, 2017). In this study, the participants' goal was to read the headlines for the purposes of answering true or false comprehension questions. Because these comprehension questions did not quiz the participants on the satirical meaning of the headline (but rather the surface meaning), the participants' standard of coherence may have been set to a relatively low threshold that focused on answering the subsequent question, even if the participant was aware of the satirical intent. In addition, although we used perceptions of humor as a proxy of satire comprehension, this is an imperfect measure because participants do not need to find satire funny to understand the satire (Johnson et al., 2010). Therefore, developing measures more capable of recognizing whether a participant understands the satirical intent of a satirical text is an important step for future research.

In this study we also included a wide battery of control variables that affect text processing in general. Many of these variables exerted significant effects on participants' headline processing. Specifically, variables measuring participants' language background, number of content words in the headlines, and crowd-sourced perceptions of the headlines all served to facilitate or inhibit reading times in some manner. Most of these variables represented main effects with no interaction among headline conditions, suggesting they are representative of general text processing effects. However, two control variables significantly interacted with satirical strategy and warrant brief mention. First, higher levels of familiarity (i.e., how familiar online raters were with the content of the headlines) resulted in significantly higher humor ratings for attenuation headlines when compared with negation headlines. Second, although higher levels of sincerity (i.e., how sincere online raters believed the headline was) significantly reduced humor ratings for all satire strategies (intercept = 2.12, coefficient =  $-.01$ ,  $SE < .01$ ,  $t = -3.64$ ,  $p < .001$ ), there was a significant interaction between sincerity and headline type in that higher levels of sincerity resulted in higher humor ratings for negation compared with attenuation headlines. However, because the measures of familiarity and sincerity were obtained from online raters, they do not reflect the participants' perceptions of familiarity or sincerity, so it is difficult to draw strong conclusions from these interactions. That being said, these findings do highlight the potential importance of measuring additional perceptual features from participants themselves.

## Conclusion

In this study we found that higher age and levels of prior knowledge influence the processing of satire as well as different satirical strategies. When comparing satire with nonsatire, our results, like previous research (Boukes et al., 2015), suggest that age plays an important role in satire processing, with higher age resulting in significantly slower reading times for satire compared with nonsatire. We also found that prior knowledge played a significant role during satire comprehension, with specific satirical strategies being more strongly affected by prior knowledge than others. This provides further evidence that age and prior knowledge are important components of satire comprehension, but more research is needed to determine the true effect of age (i.e., whether it represents relative familiarity or exposure to satirical genres). Finally, we also found that perceptions of humor in the satirical headlines, which can be taken as a proxy of satirical understanding, were affected by prior knowledge and outside ratings of familiarity and sincerity in the headlines.

In general, these findings suggest that the processing of satire interacts with several different variables related to both the reader (i.e., age and prior knowledge) and the language itself (i.e., satirical strategies). Future investigations into satire should explore other methods of assessing satire comprehension while also using more fine-grained behavioral techniques (e.g., eye-tracking) to better assess the generalizability of these findings across registers (e.g., spoken and written satire) and genres (e.g., comics, fiction, political commentary).

## Acknowledgements

We thank the editor and the anonymous reviewers for their extremely helpful suggestions and advice during the revision of this manuscript. We also thank the attendees of the 2017 annual meeting of the Society for Text and Discourse for their feedback and suggestions during a presentation of this project.

## References

- 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, 255–278. doi:10.1016/j.jml.2012.11.001
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. doi:10.18637/jss.v067.i01
- Becker, A. B. (2014). Playing with politics: Online political parody, affinity for political humor, anxiety reduction, and implications for political efficacy. *Mass Communication and Society*, 17, 424–445. doi:10.1080/15205436.2014.891134
- Boukes, M., Boomgaarden, H. G., Moorman, M., & de Vreese, C. H. (2015). At odds: laughing and thinking? The appreciation, processing, and persuasiveness of political satire. *Journal of Communication*, 65, 721–744. doi:10.1111/jcom.12173
- Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). Sentiment analysis and social cognition engine (SEANCE): An automatic tool for sentiment, social cognition, and social-order analysis. *Behavior Research Methods*, 49, 803–821. doi:10.3758/s13428-016-0743-z
- Field, A. (2009). *Discovering statistics using SPSS* (4th ed.). Thousand Oaks, CA: Sage.
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50, 346–363. doi:10.1002/bimj.200810425
- Johnson, A., Del Rio, E., & Kemmitt, A. (2010). Missing the joke: A reception analysis of satirical texts. *Communication, Culture & Critique*, 3, 396–415. doi:10.1111/j.1753-9137.2010.01077.x
- Kuznetsova, A., Brockhoff, B., & Christensen, H. B. (2016). lmerTest: Tests in linear mixed effects models. *R package version 2.0-32*. Retrieved from <https://CRAN.R-project.org/package=lmerTest>
- Kyle, K. (2016). Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication. Doctoral dissertation, Georgia State University. Retrieved from [http://scholarworks.gsu.edu/alesl\\_diss/35/](http://scholarworks.gsu.edu/alesl_diss/35/)
- Kyle, K., & Crossley, S. A. (2015). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly*, 49, 757–786. doi:10.1002/tesq.194



- LaMarre, H. L., Landreville, K. D., & Beam, M. A. (2009). The irony of satire: Political ideology and the motivation to see what you want to see in *The Colbert Report*. *The International Journal of Press/Politics*, 14, 212–231. doi:[10.1177/1940161208330904](https://doi.org/10.1177/1940161208330904)
- LaMarre, H. L., Landreville, K. D., Young, D., & Gilkerson, N. (2014). Humor works in funny ways: Examining satirical tone as a key determinant in political humor message processing. *Mass Communication and Society*, 17, 400–423. doi:[10.1080/15205436.2014.891137](https://doi.org/10.1080/15205436.2014.891137)
- Lee, H., & Kwak, N. (2014). The affect effect of political satire: Sarcastic humor, negative emotions, and political participation. *Mass Communication and Society*, 17, 307–328. doi:[10.1080/15205436.2014.891133](https://doi.org/10.1080/15205436.2014.891133)
- McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1–43. doi:[10.1207/s1532690xci1401\\_1](https://doi.org/10.1207/s1532690xci1401_1)
- McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, 22, 247–288. doi:[10.1080/01638539609544975](https://doi.org/10.1080/01638539609544975)
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R<sup>2</sup> from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133–142. doi:[10.1111/j.2041-210x.2012.00261.x](https://doi.org/10.1111/j.2041-210x.2012.00261.x)
- Nilsen, A., & Nilsen, D. (2008). Literature and humor. In V. Raskin (Ed.), *The primer of humor research* (pp. 243–280). New York, NY: Mouton de Gruyter.
- Pfaff, K. L., & Gibbs, R. W. (1997). Authorial intentions in understanding satirical texts. *Poetics*, 25, 45–70. doi:[10.1016/S0304-422X\(97\)00006-5](https://doi.org/10.1016/S0304-422X(97)00006-5)
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114, 510–532. doi:[10.1037/0033-2909.114.3.510](https://doi.org/10.1037/0033-2909.114.3.510)
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburg, PA: Psychology Software Tools Inc.
- Simpson, P. (2003). *On the discourse of satire: Towards a stylistic model of satirical humour*. Philadelphia, PA: John Benjamins Publishing.
- van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes*, 54, 360–372. doi:[10.1080/0163853X.2017.1306677](https://doi.org/10.1080/0163853X.2017.1306677)
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162–185. doi:[10.1037/0033-2909.123.2.162](https://doi.org/10.1037/0033-2909.123.2.162)