CrossMark

The role of look-backs in the processing of written sarcasm

Henri Olkoniemi[1] · Eerika Johander[1] · Johanna K. Kaakinen[1,2]

Abstract

Previous eye-tracking studies suggest that when resolving the meaning of sarcastic utterances in a text, readers often initiate fixations that return to the sarcastic utterance from subsequent parts of the text. We used a modified trailing mask paradigm to examine both the role of these look-back fixations in sarcasm comprehension and whether there are individual differences in how readers resolve sarcasm. Sixty-two adult participants read short paragraphs containing either a literal or a sarcastic utterance while their eye movements were recorded. The texts were presented using a modified trailing mask paradigm: sentences were initially masked with a string of x's and were revealed to the reader one at a time. In the normal reading condition, sentences remained visible on the screen when the reader moved on to the next sentence; in the masked condition, the sentences were replaced with a mask. Individual differences in working memory capacity (WMC) and the processing of emotional information were also measured. The results showed that readers adjusted their reading behavior when a mask prevented them from re-examining the text content. Interestingly, the readers' compensatory strategies depended on spatial WMC. Moreover, the results showed that the ability to process emotional information was related to less processing effort invested in resolving sarcasm. The present study suggests that look-backs are driven by a need to re-examine the text contents but that they are not necessary for the successful comprehension of sarcasm. The strategies used to resolve sarcasm are mediated by individual differences.

Keywords Sarcasm · Working memory · Emotion · Eye-tracking · Individual differences

Introduction

Sarcasm is a form of verbal irony that is often used to criticize someone (Attardo, 2000; Kreuz & Glucksberg, 1989; cf. positive sarcasm, e.g., Clark & Gerrig, 1984; Gibbs, 2000). Along with ironic language in general, it is an integral part of everyday communication. Ironic language is found in 7.4% of e-mails sent between friends and 72.8% of personal blog entries (Whalen, Pexman, & Gill, 2009; Whalen, Pexman, Gill, & Nowson, 2013), and it is used eight times per hour in American television shows (Schwoebel, Dews, Winner, & Srinivas, 2000). Although it carries a higher risk of miscommunication, it has been suggested that ironic language such as sarcasm is used more frequently in written, computer-mediated communication than in face-to-face conversations

(Hancock, 2004). Misunderstanding sarcasm may have dramatic consequences: misinterpreted sarcastic and ironic messages on Twitter have caused people to be fired from their jobs and to be arrested (Fallows, 2016; Robbins, 2010; Ronson, 2015). In the present study, we examined how readers resolve the meaning of written sarcasm and explored individual differences in how readers process sarcastic utterances.

Resolving the meaning of sarcastic utterances

According to current theories of sarcasm comprehension, it is particularly difficult to process and comprehend sarcasm when the utterance is not typically used in sarcastic meaning and the context does not provide advance cues for a sarcastic interpretation (Gibbs, 1994; Giora, 2003; Grice, 1975; Pexman, 2008). In this type of setting, a reader must search for an alternative interpretation of the utterance and incorporate this into the memory representation of the text. In support of these theoretical views, recent eye-tracking studies have shown that written sarcastic utterances take more time to process and are harder to comprehend than their literal counterparts (Au-Yeung, Kaakinen, Liversedge, & Benson, 2015; Filik, Leuthold, Wallington, & Page, 2014; Filik & Moxey, 2010; Kaakinen, Olkoniemi, Kinnari, &

✉ Henri Olkoniemi
henri.olkoniemi@utu.fi

[1] Department of Psychology, University of Turku, 20014 Turku, Finland

[2] Turku Institute for Advanced Studies, University of Turku, 20014 Turku, Finland

Hyönä, 2014; Olkoniemi, Ranta, & Kaakinen, 2016; Olkoniemi, Strömberg, & Kaakinen, in press; Turcan & Filik, 2016).

The exact time-course of resolving sarcasm can be studied with eye tracking, which is particularly useful for examining how readers comprehend connected discourse (Hyönä, Lorch, & Rinck, 2003; Raney, Campbell, & Bovee, 2014). Eye fixations can be categorized into those done during the *first-pass reading* of a sentence and those done during later *look-backs* initiated by subsequent parts of the text. First-pass sentence reading can be further divided into *progressive fixations*, which land on unread parts of the sentence, and *first-pass re-readings*, which are fixations that return to earlier parts of the sentence (Liversedge, Paterson, & Pickering, 1998). Sentence-level analyses are informative when the "area of interest" is not a single word (cf. Rayner, 1998) but instead consists of a phrase or a sentence (Hyönä et al., 2003). This is often the case with sarcastic utterances.

What previous eye-tracking studies show about the processing of sarcasm is that readers tend to do more re-reading of sarcastic than literal utterances during first-pass reading (Kaakinen et al., 2014; Olkoniemi et al., 2016, Olkoniemi et al., in press). Moreover, readers tend to look back to the sarcastic utterance from subsequent parts of the text (Kaakinen et al., 2014; Olkoniemi et al., 2016; Olkoniemi et al., in press) and to initiate a look-back from the sarcastic utterance to the preceding context (Olkoniemi et al., 2016; Olkoniemi et al., in press). From a theoretical point of view, these look-backs may reflect that readers are trying to integrate the sarcastic utterance with the developing text representation, which requires a reanalysis of the utterance (e.g., Grice, 1975). The findings resemble those of studies on the processing of syntactically ambiguous sentences (i.e., garden-path sentences), showing that readers typically regress back to the syntactically ambiguous sentence region from subsequent areas of the text (e.g., Frazier & Rayner, 1982; Meseguer, Carreiras, & Clifton, 2002; Mitchell, Shen, Green, & Hodgson, 2008) and sometimes re-read the preceding sentence context (von der Malsburg & Vasishth, 2013). However, it should be noted that these garden-path studies have examined the processing of complex single sentences, not entire passages.

Even though previous studies suggest that readers tend to do more first-pass re-reading and later looking back to the sarcastic utterance and, occasionally, to the preceding context (Kaakinen et al., 2014; Olkoniemi et al., 2016; Olkoniemi et al., in press), the time-course of resolving sarcasm is still unclear. If readers already tend to do more re-reading of the sarcastic utterance during the first-pass, why do they still look back to it from subsequent parts of the text?

The role of re-inspections in text comprehension

An obvious reason for re-inspections is that readers want to refresh the text information in their memory, as this supports comprehension (e.g., Booth and Weger, 2013; Schotter, Tran, & Rayner, 2014; Hyönä & Nurminen, 2006; Hyönä, Lorch, & Kaakinen, 2002; Raney et al., 2014; Rayner, Chace, Slattery, & Ashby, 2006; White, Lantz, & Paterson, 2016). Studies that have examined the function of regressive eye movements (i.e., a backward-directed eye movement) have found that fixations made during first-pass re-reading are crucial for comprehending the sentence (Booth & Weger, 2013; Schotter et al., 2014; White et al., 2016). For example, readers made more comprehension errors when re-reading of the already fixated words was eliminated due to masking the words after the reader had moved on in the sentence (Booth & Weger, 2013; Schotter et al., 2014; White et al., 2016, Exp1). Moreover, Booth and Weger (2013) showed that readers return to previous words in a sentence to re-access their meanings. In their study, participants read single sentences containing a predefined target word. In cases where the reader made a regression back to the target word, a display change was performed during the returning saccade, and the target word was changed to another word that influenced the interpretation of the sentence. After reading, participants' comprehension of the sentence was measured. The results showed that when the target word was changed, readers were likely to assign a new meaning to the sentence. This suggests that information extracted during re-reading overwrites the information gathered during initial reading (Booth & Weger, 2013, Exp. 3).

Previous research suggests that look-backs (i.e., returning to previous sentences) reflect a conscious effort (Hyönä & Nurminen, 2006) to build a comprehensive mental representation of the text contents (Hyönä et al., 2002). First of all, readers are able to report after reading whether they looked back in the text, and where (Hyönä & Nurminen, 2006). Second, readers who initiate look-backs to important parts of the text gain better comprehension than readers who look back more randomly (e.g., Hyönä et al., 2002).

Re-inspecting parts of sentences or a text means that the reader makes a regression to a previously read word or sentence. Previous research suggests that readers are accurate in making saccades to the part of the text that caused them comprehension difficulty (e.g., Frazier & Rayner, 1982; Meseguer et al., 2002), thus indicating that they build a representation of the text's spatial layout while reading (e.g., Baccino & Pynte, 1994, 1998; Kennedy, Brooks, Flynn, & Prophet, 2003; Murray & Kennedy, 1988). The representation of the spatial layout allows readers to selectively re-inspect the text, which is more efficient than simply re-reading the whole text again (Kennedy et al., 2003).

An interesting suggestion is that looking back in a text does not necessarily indicate a need to re-inspect the information itself, but rather a need to focus attention on what appeared there (Meseguer et al., 2002). Eye movements to previously viewed locations are sometimes triggered by reactivation of the memory representation, and this enhances subsequent

memory retrieval (Ferreira, Apel, & Henderson, 2008). Therefore, look-backs may support the refreshing of text information in memory in two ways: by focusing attention on a certain spatial location – which helps in retrieving the contents of the text from memory – and/or by providing a review of the text content itself.

Even though previous studies suggest that look-back fixations play an important role in text comprehension (Hyönä & Nurminen, 2006; Hyönä et al., 2002; Raney et al., 2006, 2014), very little is known about their actual function in reading comprehension. Studies that experimentally manipulated the availability of previously read information to examine the role of regressive eye movements (e.g., Booth & Weger, 2013; Schotter et al., 2014; White et al., 2016) used only single sentences, thus providing information about the function of first-pass re-readings in sentence comprehension. In the present study, we modified the "trailing mask" paradigm used in these previous studies (Schotter et al., 2014) to be suitable for examining the reading of passages. Our particular interest was the role of look-back fixations in processing sarcastic utterances embedded in passages.

Individual differences in the processing of sarcasm

One recent theory on sarcasm comprehension, the parallel constraint-satisfaction framework, explicitly states that there are individual differences in sarcasm comprehension (Pexman, 2008). According to this framework, reader-related factors – such as how frequently a person uses sarcasm – influence the likelihood that different interpretations (literal, sarcasm, or white lie) are active in the reader's mind. In line with this, recent eye-tracking studies have suggested that individual differences in *working memory capacity* (WMC) and the ability to process emotional information impact the likelihood of re-reading sarcastic texts (Kaakinen et al., 2014; Olkoniemi et al., 2016; Olkoniemi et al., in press).

Working memory refers to the mental process of maintaining information in an active state for later recall and manipulating information during the execution of ongoing tasks (e.g., Cowan, 2010; Daneman & Carpenter, 1980; Engle, 2010). There are individual differences in the capacity of working memory, and these are related to the ability to control attentional resources so that relevant information is quickly activated and irrelevant information is inhibited (e.g., Engle, 2010). WMC thus plays a crucial role in text comprehension, as it is needed to direct attention to, and maintain attention on, relevant information (see also Gernsbacher, 1993).

Recent eye-tracking studies have shown that individual differences in WMC are related to the processing of sarcasm (Kaakinen et al., 2014; Olkoniemi et al., 2016). Readers with high WMC demonstrate an increase in first-pass re-reading of sarcastic sentences (Kaakinen et al., 2014; Olkoniemi et al., 2016), whereas low WMC readers are more likely to initiate

look-backs to the sarcastic sentences (Olkoniemi et al., 2016). Hence, the time-course of resolving sarcasm seems to depend on WMC, such that high WMC readers detect sarcasm faster and/or resolve it earlier than low WMC readers, who show mainly delayed effects. One possible explanation for these results is that low WMC readers may have trouble in inhibiting the initial literal interpretation of the utterance, which is why they need to engage in later reprocessing to validate the sarcastic meaning (e.g., Giora, 1999). For low WMC readers, then, look-backs should be important in forming sarcastic interpretations.

In the present study, we examined the roles of both verbal (Daneman & Carpenter, 1980) and visuo-spatial (Redick et al., 2012) WMC in sarcasm comprehension. There is an ongoing theoretical debate on whether WMC is domain specific or domain general (e.g., Oswald, McAbee, Redick, and Hambrick, 2015; Fougnie, Zughni, Godwin, & Marois, 2015). The present study does not aim to resolve this debate for good, but we thought it was important to include both types of measures, as look-backs may be driven by a memory representation of the text's spatial layout. If look-backs are related to episodic memory processes that launch a reader's eyes toward a spatial location where relevant information was previously identified, readers with higher spatial WMC may be better able to use this strategy to resolve sarcasm than low spatial WMC readers. Using more than one measure of WMC also provides a more reliable view of the role of working memory in comprehending sarcasm (e.g., Conway et al., 2005; Walczyk & Taylor, 1996).

Another important factor in sarcasm comprehension is *the ability to process emotional information*. Sarcastic statements are typically meant to sting and are usually perceived as insulting (e.g., Akimoto et al., 2014; Bowes & Katz, 2011; Olkoniemi et al., in press). Sarcasm can also be used as a form of humor, and sarcastic comments might be perceived as funny (e.g., Akimoto et al., 2014; Olkoniemi et al., in press). Previous studies have shown that sensitivity to the emotional message delivered by the speaker or by the sarcastic protagonist in a story is important in sarcasm comprehension (Amenta, Noël, Verbanck, & Campanella, 2013; Nicholson, Whalen, & Pexman, 2013; Olkoniemi et al., 2016; Shamay-Tsoory, Tomer, & Aharon-Peretz, 2005). Readers who can rapidly assign an emotional tone to the utterance should thus be able to interpret sarcasm more quickly (Jacob, Kreifelts, Nizielski, Shütz & Wildgruber, 2016; Olkoniemi et al., 2016; Olkoniemi et al., in press). In their eye-tracking study, Olkoniemi et al. (2016) showed that a poor ability to make use of emotional information (as measured with the Iowa Gambling Task; Bechara, Damasio, Damasio, & Anderson, 1994) was reflected in eye-movement records as an increased probability of looking back from the sarcastic target utterance to earlier parts of the text. In another study, readers with a poor ability to recognize emotions (as measured by the 20-item

Toronto Alexithymia Scale; Bagby, Parker, & Taylor, 1994) showed increased first-pass reading time of the utterance following a sarcastic statement that validated the sarcastic interpretation (Olkoniemi et al., in press). These findings suggest that a poor ability to process emotional information is related to greater confusion when encountering sarcastic utterances and that this is reflected as an increased need to reprocess the sarcastic utterance and the passage context.

In previous studies, the Iowa Gambling Task (IGT; Bechara et al., 1994) and the Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994) have been used to study individual differences in the processing of written sarcasm (Olkoniemi et al., 2016; Olkoniemi et al., in press). Although they should gage different abilities, both measures have been related to similar effects on the processing of sarcasm. IGT is thought to measure an individual's sensitivity to emotional responses to reward and/or punishment in a decision-making task. The effect of the emotional response measured by IGT is thought to be, at least to some extent, unconscious (e.g., Buelow & Suhr, 2009). On the other hand, TAS-20 measures individuals' ability to recognize and name emotions (i.e., alexithymix traits). TAS-20 should thus reflect more conscious emotional processing ability. In terms of written sarcasm comprehension, IGT should be related to the immediate emotional response evoked after a reader has processed the sarcastic utterance and starts to build an inference. TAS-20 should be related to a reader's ability to name and recognize that the protagonist in the story intends to insult another person, and this should help in inferring the meaning of the utterance. In the present study, both measures were used to explore the role of emotion processing ability in the comprehension of written sarcasm.

Overview of the present experiment

The goal of the present study was to examine the role of look-backs in resolving the meaning of sarcastic utterances embedded in story context. We modified the "trailing mask" paradigm used in previous sentence reading studies (Schotter et al., 2014) to be suitable for passage reading. In the mask condition, readers revealed the text one sentence at the time, and the previously read sentence was always replaced with x's as soon as the reader moved to the next sentence. This type of masking allowed readers to perform normal first-pass re-reading of the sentence but prevented re-examination of the text content during later look-backs (see Fig. 1). In the no-mask condition, readers revealed initially masked text one sentence at the time, and the sentences remained visible after the reader had moved on.

Participants read short passages that contained either sarcastic or literal utterances embedded in sentences (e.g., "*What a clown you are!*" *Zachary shouted.*). After each passage, participants responded to a question about their interpretation

No-Mask Condition

Paavo ja Sakari työskentelevät sirkuksessa. Esitys on ollut loppuunmyyty ja Paavon ollessa lavalla yleisö nauraa ratketakseen. Paavon poistuessa näyttämöltä hänelle annetaan raikuvat aplodit. "Sinä se vasta olet sirkuspelle!" Sakari huudahtaa. xxx xxxxxxxx xxxxxxxx, xxxx xx xxxxxxxxx xxxxxxxxx xxxxxxx xxxxxxxxx.

Mask Condition

xxxxx xx xxxxxx xxxxxxxxxxxxx xxxxxxxxxx. xxxxxx xx xxxxx xxxxxxxxxxx xx xxxxxx xxxxxxx xxxxxxx xxxxxx xxxxxx xxxxxxxxxx. xxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxx xxxxxxxx xxxxxxxx xxxxxxx. "Sinä se vasta olet sirkuspelle!" Sakari huudahtaa. xxx xxxxxxxx xxxxxxxx, xxxx xx xxxxxxxxx xxxxxxxxx xxxxxxx xxxxxxxxx.

Fig. 1 Examples of the no-mask and mask conditions

of the target utterance and a question that tested their memory about the passage content. To get a detailed picture of the time-course of processing, we computed sentence-level fixation measures (Hyönä et al., 2003) separately for the target utterance ("What a clown you are!") and the spillover region ("Zachary shouted.").

Our primary interest was the impact of masking manipulation on the processing and comprehension of sarcastic versus literal utterances. We assumed that if an increase in look-backs to sarcastic utterances and passage context reflects a need to re-examine the text content to resolve the meaning of the utterance (e.g., Grice, 1975; Olkoniemi et al., 2016), we should observe a decrease in look-backs to and from the sarcastic utterances in the mask condition. It is also possible that readers develop compensatory strategies to deal with the masking of the previous sentences and start spending longer first-pass reading times, and especially re-reading times (Walczyk & Taylor, 1996; White et al., 2016), on sarcastic utterances. As the spillover region is within the same sentence as the target utterance, increased first-pass reading times and look-backs initiated from the spillover region to the target utterance may also occur. However, if such compensatory strategies are not used, we should observe poorer comprehension of the sarcastic utterances in the mask condition.

We also explored individual differences in how readers process sarcastic utterances. Previous studies suggest that WMC (Kaakinen et al., 2014; Olkoniemi et al., 2016) and the ability to process emotional information (e.g., Amenta et al., 2013; Nicholson et al., 2013; Olkoniemi et al., 2016) may be crucial in determining how quickly readers can resolve sarcasm. Assuming that look-backs reflect the need to re-examine the text content, and thus to resolve the meaning of the utterance, we expected that readers with low WMC and readers with relatively poor emotion processing abilities are more influenced by the mask than readers with relatively high WMC and good emotion processing abilities. Thus, the eye-movement measures of these readers should reflect a greater impact of masking. Both verbal and spatial WMC tasks were

used to examine the contributions of the different aspects of WMC to resolving sarcasm. Moreover, a test of the ability to name and recognize emotions and a test of the ability to make use of emotional information in higher-order cognitive tasks were used to explore how different aspects of emotion processing relate to the comprehension of sarcasm.

Method

Participants

Sixty-two University of Turku (Finland) students (53 women, M_{age} = 23.57 years, SD_{Age} = 6.06) participated in the study to fulfill a course requirement. All were native speakers of Finnish (the language studied here) and had normal or corrected-to-normal vision. All participants provided written informed consent before the experiment.

Based on effect size measures from the Olkoniemi et al. (2016) eye-tracking study, a power analysis was conducted using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007) to estimate the sample size needed to detect an interaction between text type and an individual differences measure. The aimed statistical power was set to .80 and the significance level to α = .05. This analysis showed that at least 49 participants were needed.

Apparatus

Eye movements were recorded monocularly using EyeLink 1000 (SR Research Ltd., Ontario, Canada) at 1,000 Hz sampling frequency. The stimuli were presented on a 21-in. CRT screen (resolution: 1,024 × 768, refresh rate: 120 Hz). Participants were seated 70 cm from the screen, and a chin rest was used to stabilize the head.

Text materials

Each participant read a total of 42 text paragraphs on a computer screen (font: Courier New, font size: 14 pt, line height: 3 pt) while their eye movements were recorded. Fourteen of the paragraphs included a sarcastic utterance, 14 included a literal utterance, and there were 14 filler paragraphs. There were two versions of each experimental paragraph: a sarcastic and a literal (28 paragraphs × 2 text types, resulting in 56 experimental paragraphs). The paragraphs used were the same as in the Olkoniemi et al. (2016) study. They were pre-tested for the utterances' familiarity as sarcasm, naturality in the story context, whether the emotional state of the speaker was apparent in the sarcastic statement, and whether the context supported a sarcastic interpretation. All sarcastic utterances were rated as unfamiliar; literal and sarcastic target utterances were evaluated to be natural in the story context; sarcastic utterances were

evaluated to be less positive and more negative than their literal counterparts; and the story contexts did not provide cues about sarcastic interpretation for the utterance. English translations of paragraphs used in the experiment are available via the Open Science Framework (https://osf.io/9maz8). The filler passages were similar to the experimental items, but they did not contain indirect language (Kaakinen et al., 2014).

Each of the participants saw half of the paragraphs in the mask condition and the other half in the no-mask condition. In the no-mask condition, text was first masked with x's and readers revealed the text one sentence at a time by pressing the spacebar. Once the sentences were read, they remained visible on the screen as the reader progressed. In the mask condition, the text was also initially masked and the readers revealed sentences one at a time by pressing a spacebar. However, after the spacebar was pressed to reveal a new sentence, the previous sentence was masked with x's (cf. moving window paradigm, McConkie & Rayner, 1975). An example text is shown in Table 1.

After reading, participants responded to two questions that required a yes/no response. The first tapped into the meaning of the target utterance (e.g., Did Zachary think that Paul's show was a success?), and the second tested memory for the text information (e.g., Does Paul work at the circus as a clown?). Participants answered the questions by pressing "yes" or "no" buttons on the keyboard. A correct answer was rewarded with one point. The percentage of correct answers was computed separately for different question types. Reaction times to the inference questions were also recorded.

Table 1 An example of experimental text (translated from Finnish)

Text region	Text type	
	Literal	Sarcasm
Introduction	Paul and Zachary work at a circus. Tonight's show is sold out	Paul and Zachary work at a circus. Tonight's show is sold out
Critical context	When Paul is at stage the audience laughs loudly. When Paul leaves the stage he gets standing applause	When Paul is at stage the audience looks bored and some people yawn. When Paul leaves the stage he gets small applause
Target utterance	"What a clown you are!"	"What a clown you are!"
Spillover region	Zachary shouts	Zachary shouts
Ending	Later Zachary asks Paul if he would like to go to the movies	Later Zachary asks Paul if he would like to go to the movies

Note. There were two different versions (literal and sarcastic) of each text; each participant read only one of the versions. Text versions were counterbalanced across participants

Individual differences measures

Toronto Alexithymia Scale (TAS-20) The ability to recognize emotions was measured with the TAS-20 (Bagby et al., 1994; Joukamaa et al., 2001). TAS-20 is a paper-and-pencil self-report scale that includes short claims (e.g., "I am often confused about what emotion I am feeling."). Participants responded to the items on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). Five of the items are negatively keyed. The scale was scored by summing up all responses. The scores can vary between 20 and 100 points, with higher scores indicating a poorer ability to recognize emotions (i.e., higher alexithymia). The internal reliability (Cronbach's alpha) of the TAS total score was $\alpha = .74$.

Iowa Gambling Task (IGT) A computerized version (Psychology Experiment Building Language Version 0.13; Mueller, 2012) of the IGT (Bechara et al., 1994) was used to measure the ability to use emotional information in decision-making. In the task, participants are given four decks of cards (referred to here as A, B, C, and D) and an imaginary capital of $2,000. They are instructed to use the capital in a game of cards so that they win more money than they lose. In the game, turning some cards (by clicking on the deck with a mouse) leads to a reward ($100 from decks A and B, $50 from decks C and D) and turning others leads to a penalty (net outcome of -$250/10 cards in decks A and B and net outcome of +$250/10 cards in decks C and D). Playing mostly from decks A and B leads to an overall loss, and playing mostly from decks C and D results in an overall gain.

Because the cards in the decks are shuffled, the occurrence of the penalty cannot be predicted. In the long run, half of the decks (decks A and B) cost the most. This losing trend should activate a negative emotional signal that these disadvantageous decks should be avoided. If emotional markers are absent or cannot be used properly in the decision-making process, then disadvantageous decks are not avoided. This causes a net loss or minimal gains in the task. The task ends when 100 cards have been drawn, though this is not revealed to the participants beforehand. After a few draws, people tend to choose the advantageous decks (C and D; Bechara et al., 1994). The task is scored by summing the number of draws from the advantageous decks, and scores can vary between 0 and 100. To estimate the internal consistency of the task, the split half correlation was calculated for the choice of advantageous decks (between the odd- and even-numbered trials to avoid a learning effect). The correlation was $r = .88$, p <.001, 95% CI [.87–.89].

Reading span task (RSPAN) The reading span task (Daneman & Carpenter, 1980; Kaakinen & Hyönä, 2007) was used to measure verbal WMC. In this task, participants read aloud sets of unrelated sentences presented on a computer screen. After every set, they are asked to recall the last word of each sentence in the set. The task begins with sets of two sentences. The set size increases as long as the participant is able to recall the final words of the sentences. Each set size is repeated three times. The task ends when the participant fails to recall the final words of a sentence of a particular set size for its three repetitions. The task was preceded by a practice session that included reading three sets of two sentences. It was scored for the total number of correctly recalled final words (Friedman & Miyake, 2005). The task scores can vary between 0 and 81 points. The internal reliability (Cronbach's alpha) of the RSPAN varies between .70 and .90 (see, e.g., Conway et al., 2005).

Symmetry span task (SSPAN) The symmetry span task (Redick et al., 2012) was used to measure spatial WMC. In the task, participants view an 8×8 matrix of white and black squares and determine whether the pattern is symmetrical along its vertical axis. After the judgment, participants are presented with a 4×4 matrix of squares in which one cell is filled in red. After the series of matrix presentations, participants must recall the serial order of the positions of the red cells (2–5 spatial locations). The task was scored for the total number of items recalled in the memory trials (the partial span score; Conway et al., 2005). Scores can vary between 0 and 42 points. The internal reliability (Cronbach's alpha) of the SSPAN varies between .76 and .81 when partial scoring is used (see, e.g., Redick et al., 2012).

Procedure

Participants were tested individually. They were naïve to the purpose of the experiment: upon arrival, they were only informed that the experiment was about reading. The specific nature of the task was explained to the participants only after the experiment.

First, each participant signed an informed consent form and the eye-tracking system was introduced. The experimental procedure was then explained. The eye-tracker was set up and calibrated using a nine-point calibration screen. Participants were instructed to read each paragraph at their own pace. Each paragraph was presented on one screen. Participants were told to press the spacebar on the keyboard when they wanted to move from one sentence to another within a paragraph. After every paragraph, two questions concerning the previously read paragraph were presented, one at a time. After the participant had answered the second question, the next paragraph was presented.

The reading task was followed by SSPAN, IGT, RSPAN, and TAS-20. Each experimental session lasted for about 90 min.

Results

Data preparation

Fixations shorter than 50 ms were either merged with a nearby fixation (if the distance between the fixations was < 1°) or removed from the data. Sentence-level measures for the target utterance and the spillover region were computed from the eye-movement data (see Hyönä et al., 2003). The reading time measures were divided into first-pass fixation times and later look-backs. *First-pass reading time* was the summed duration of fixations falling within the sentence during first-pass reading. First-pass reading time was further divided into *forward-fixation time* and *re-reading time*. *Forward-fixation time* is the summed duration of fixations that land on unread parts of the sentence during first-pass reading, and *first-pass re-reading time* is the summed duration of fixations made during re-inspection of the target utterance before moving on in the text. The *look-back fixation time* is the summed duration of fixations returning to the sentence from other parts of the text after the first-pass reading. *Look-froms* are look-back fixations that were initiated from the sentence.

For the re-reading and look-back measures (first-pass re-reading, look-back fixations, and look-from fixations), we computed separately the probability of re-reading or a look-back (binomial measure). The summed fixation times were computed on the condition that re-reading or a look-back was done. Because readers saw the target utterance and the spillover region at the same time, look-back times to the target utterance were computed separately for look-backs made from the spillover region and look-backs made from other text regions to the target utterance.

The reading time measures were skewed and were consequently transformed. The best-fitting transformation was selected to normalize the measures. The first-pass reading times on the target utterance and the spillover region and the forward-fixation time on the target utterance were square root transformed. The first-pass re-reading time measures were logarithmically transformed before the analyses. All measures were analyzed for the target utterances; first-pass reading time, probability to look back to the target utterance, and look-back time to the target utterance were analyzed for the spillover region. Observed means and standard deviations of the different eye-movement measures are presented in Table 2.

Statistical analyses

Data were analyzed with linear mixed-effects models (LMM) using the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015) in the R statistical software (Version 3.4.3; R Core Team, 2017). The models were estimated using Maximum Likelihood estimation. Separate models were built for each eye-movement measure for the target utterance and the spillover region. In addition, separate models were built for reaction times and correct answers to text memory and inference questions (descriptive statistics are presented in Table 3).

Participants and items were entered into the models as crossed random factors (Baayen, Davidson, & Bates, 2008). The influences of the text type and mask were tested by fitting models with sum coded fixed effects variables. The individual differences variables (RSPAN, SSPAN, IGT, and TAS-20; see Table 4 for descriptive statistics) were added to the models as centered fixed effects variables. The correlations between the individual differences measures were low or modest (see Table 4). The RSPAN and TAS-20 scores were skewed; consequently, the RSPAN score was square root transformed and the TAS-20 was logarithmically transformed. To examine the potential effects of presentation order on the observed effects (e.g., Olkoniemi et al., 2016), a trial order was added as a sum coded fixed effect to the models (first half of the experiment = -1, end half of the experiment = 1).

Model fitting was performed in a step-wise fashion, starting with the most complex model and including text type, mask condition, trial order, and the individual differences measures and their interactions as fixed effects. At this point, only intercepts for participants and texts were fitted as random effects. The fixed effect associated with the smallest *t*-value was removed from the model, beginning with the interaction terms, and the reduced model was compared to the former model using the *anova* function in the *lme4* package (Bates et al., 2015) to compare the fit of the models. Fixed effects were removed one at a time until nothing else could be removed without significantly reducing the fit of the model. The fixed effects of the primary interests (text type, mask condition, and their interaction term) were always retained in the model, as they were crucial with respect to our hypotheses. Finally, a full random structure was fitted to the model (Barr, Levy, Scheepers, & Tily, 2013), and fixed and random effects were removed if further changes did not significantly reduce the fit of the model. If the model failed to converge after fitting the full random structure, the random structure was trimmed top-down, starting with correlations between factors.

The exact degrees of freedom are difficult to determine for the *t*-statistics estimated by LMMs, leading to problems in determining exact *p*-values (Baayen et al., 2008). Statistical significance at the .05 level is indicated by the values of |*t* and *z*| > 1.96.

In the text, we always report the main effects of text type and masking condition and their interactions. As for the other interaction effects, only significant effects involving text type are reported for the sake of brevity. Interactions were further examined by computing the simple effects of text type; the estimates and their 95% CIs are illustrated in the figures. All of the final models are reported in Appendix A.

Table 2 Descriptive statistics of the reading time measures

Text region	Measure	Text type							
		Literal				Sarcasm			
		No-Mask		Mask		No-Mask		Mask	
		M	SD	M	SD	M	SD	M	SD
Target utterance	First-pass reading time	598	330	597	336	621	334	603	334
	Forward-fixation time	504	234	497	251	521	249	504	242
	First-pass re-reading time	296	263	305	237	273	206	314	207
	Probability of re-reading	.32	.47	.33	.47	.37	.48	.31	.46
	Look-back time	387	299	277	236	460	395	298	227
	Probability to look-back	.30	.46	.19	.40	.37	.48	.23	.42
	Probability to look-from	.27	.44	.20	.40	.35	.48	.22	.42
Spillover region	First-pass reading time	564	384	614	427	592	405	659	448
	Look-back time to target utterance	462	350	544	344	588	434	605	557

Reading of the target utterance

First-pass reading The analysis of the *first-pass reading times, the forward-fixation times*, or the *probability of first-pass re-reading* of the target utterance did not show effects of text type, mask condition, or their interaction.

The analysis of the *first-pass re-reading times on the target utterance* did not show main effects of text type, mask condition, or their interaction. However, there was a three-way interaction between text type, mask, and SSPAN. This indicates that the effect of masking was different in literal and sarcastic texts – and that this difference depended on the level of SSPAN. Estimates of the effects of text type at low (-1 *SD*) and high (+ 1 *SD*) levels of SSPAN and their 95% CIs are illustrated in Fig. 2. As is apparent in Fig. 2, readers scoring relatively low in SSPAN (i.e., lower spatial working memory capacity) showed shorter first-pass re-reading times for sarcastic than for literal target utterances in the no-mask condition, whereas in the mask condition, low SSPAN readers showed *longer* first-pass re-reading times on sarcastic target utterances

when compared to literal. In other words, low SSPAN readers' first-pass re-reading times on sarcastic target utterances got longer in the masked condition. Readers scoring relatively high in SSPAN, on the other hand, did not show reliable sarcasm effects in either condition. The analysis also showed an interaction between text type and TAS-20 (see Fig. 3). Readers scoring relatively low in TAS-20 (i.e., fewer alexithymic traits) showed shorter re-reading times for sarcastic than for literal target utterances, whereas readers scoring relatively high showed higher re-reading times for sarcastic than for literal target utterances.

Look-backs The analysis of the *probability to initiate a look-back to the target utterance* showed a main effect of text type, indicating that readers were more likely to initiate look-backs to sarcastic than to literal target utterances. There also was a main effect of mask condition, indicating that masking reduced the likelihood of look-backs to the target utterance in general. There was no indication of a Text type × Mask condition interaction. However, the analysis revealed a three-way

Table 3 Reaction times (RTs) and percentage of correct answers (%) to text memory and inference questions

Measure	Literal				Sarcasm			
	No-mask		Mask		No-mask		Mask	
	M	SD	M	SD	M	SD	M	SD
Text memory (%)	94	24	92	28	93	25	95	23
Text memory (RT)	2,628	1,291	2,684	1,278	2,616	1,275	2,752	1,788
Inference (%)	96	20	97	17	85	36	84	37
Inference (RT)	2,771	1,183	2,781	1,273	3,302	1,939	3,299	1,889

interaction between text type, mask, and SSPAN, indicating that the effect of masking was different in sarcastic and literal texts – and that this difference depended on the level of SSPAN (see Fig. 4). Readers scoring relatively low in SSPAN were more likely to initiate a look-back to sarcastic than to literal target utterances, especially in the no-mask condition. In the masked condition, the difference between text types was reduced. On the other hand, readers scoring relatively high in SSPAN showed an opposite pattern: they were more likely to look back to sarcastic than to literal target utterances in the mask condition, and the difference between text types was reduced in the no-mask condition.

The analysis of the *look-back time* on the target utterance showed a main effect of masking, indicating that overall, look-back times were shorter in the mask than in the no-mask condition. There was no evidence for a main effect of text type or for a Text type × Mask condition interaction.

As for the *probability to initiate a look-from from the target utterance*, there was a main effect of text type. This indicated that readers were more likely to look to the other parts of the text from the sarcastic than from the literal target utterance. There also was a main effect of mask condition, indicating that masking reduced the likelihood of looking back in the text in general. There was no indication of an interaction between text type and mask condition.

Reading of the spillover region

The analysis of *first-pass reading times on the spillover region* showed a main effect of mask condition, indicating that the first-pass reading time in the spillover region was longer in the mask than in the no-mask condition. There was no main effect of text type or an interaction between text type and mask condition.

The analysis of the *probability to initiate a look-back to the target utterance from the spillover region* did not show effects of text type, mask condition, or their interaction. Furthermore,

Fig. 2 Model estimates for the first-pass re-reading time on the target utterance. Y-axis represents the sarcasm effect, which is the difference in the reading times between sarcastic and literal texts. For illustration purposes SSPAN score is divided to High and Low (± 1 *SD*), and the model means and confidence intervals are back-transformed from log values. Error bars represent 95% CI

the analysis of the *look-back time to the target utterance from the spillover region* did not show main effects of text type, mask condition, or their interaction. There was only an interaction between text type and trial order (see Fig. 5). The result indicates that look-back times were longer for sarcastic than for literal texts in the beginning of the experiment. However, look-back times to sarcastic target utterances decreased toward the end of the experiment – during the second half of the experiment, look-back times were shorter for sarcastic than for literal target utterances.

Text memory and inference questions

The analysis of the *correct answers to the text memory questions* revealed a main effect of text type, indicating that readers were better at responding to the text memory questions after sarcastic paragraphs than after literal paragraphs. There were no effects of masking or an interaction between text type
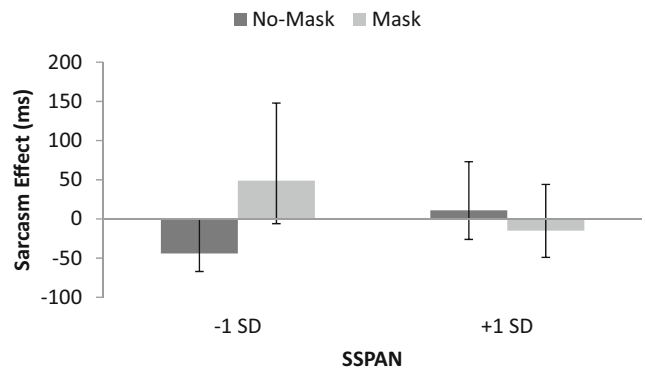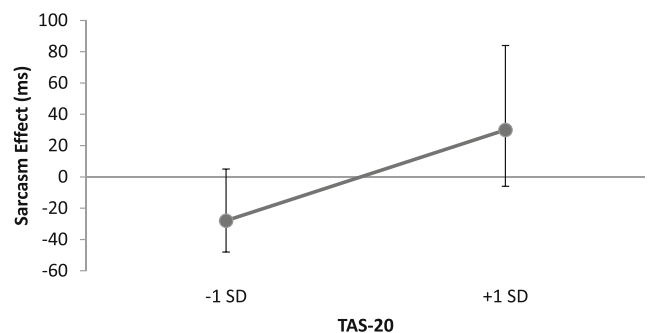

Fig. 3 Model estimates for the first-pass re-reading time on the target utterance. Y-axis represents the sarcasm effect, which is the difference in the reading times between sarcastic and literal texts. For illustration purposes TAS-20 score is divided to High and Low (± 1 *SD*), and the model means and confidence intervals are back-transformed from log values. Error bars represent 95% CI

Table 4 Descriptive statistics of the individual differences measures and correlations between the measures

Measure	M	SD	Min	Max	Correlation		
					RSPAN	SSPAN	TAS-20
RSPAN	26.03	12.74	10	73			
SSPAN	28.13	7.87	5	42	.26		
TAS-20	41.97	7.35	24	62	-.31	-.09	
IGT	54.27	11.76	2	81	-.02	.03	-.06

Note. RSPAN Reading Span task score, *SSPAN* Symmetry Span task score, *IGT* Iowa Gambling Task score, *TAS-20* 20-item Toronto Alexithymia Scale score

Fig. 4 Model estimates for the probability to initiate a look-back to the target utterance from the other parts of text than spillover region. Y-axis represents the sarcasm effect, which is the difference in the reading times between sarcastic and literal texts. For illustration purposes SSPAN score is divided to High and Low (± 1 *SD*), and the model means and confidence intervals are back-transformed from log values. Error bars represent 95% CI
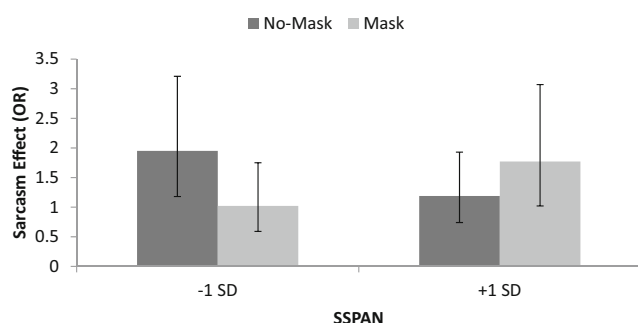

Fig. 6 Model estimates for the reaction times to the inference questions. Y-axis represents the sarcasm effect, which is the difference in the reaction times between sarcastic and literal texts. The model means and confidence intervals are back-transformed from log values. Error bars represent 95% CI

and mask condition. The analysis of the *reaction times to the text memory questions* did not show effects of text type, mask, or their interaction.

The analysis of the *correct answers to the inference questions* showed a main effect of text type, indicating that readers made more correct responses after literal than after sarcastic passages. There were no effects of masking or an interaction between text type and mask condition. The analysis of *reaction times to the inference questions* showed a main effect of text type, indicating that it took longer to answer inference questions after sarcastic than after literal passages. There was also an interaction between text type and trial order, indicating that readers were slower in answering inference questions after sarcastic paragraphs than after literal paragraphs in the first half of the experiment. This effect wore off toward the end of the experiment, however (see Fig. 6). There was no indication of a main effect of masking or a Text type × Mask condition interaction.

Relationship between reading times and sarcasm comprehension

Based on the suggestions of the reviewers, a new set of analyses was conducted to explore the correlation between reading times and sarcasm comprehension. Separate models were built for each eye-movement measure shown to be related to the processing of sarcasm in the previous analyses (first-pass re-reading time, probability to initiate a look-back, look-back time from the spillover region, and the probability to initiate a look-from). Because literal items showed a roof effect with limited variance, the analyses were conducted only for the sarcasm comprehension scores. The reading time measures were fitted to the models as centered fixed effects variables, and probability measures were fitted as sum coded fixed effects variables. To explore the possible effects of mask manipulation, separate models were built for mask versus no-mask conditions. Model coefficients for mask and no-mask conditions did not differ from each other across the models; consequently, mask was not fitted as a predictor to the final models. All of the final models are reported in Appendix B.

The analysis showed that *first-pass re-reading time on sarcastic utterance, probability to initiate a look-back to the sarcastic utterance* (from parts of the text other than the spillover region), *or look-back time to the sarcastic utterance from the spillover region* did not have a statistically significant effect on correct responses to comprehension questions after sarcastic passages. However, the analysis of the *probability to initiate a look-from the sarcastic utterance* showed that look-froms were associated with poorer comprehension of the sarcastic utterance.
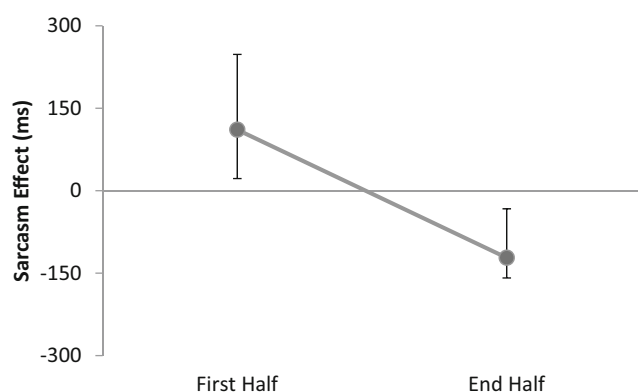

Fig. 5 Model estimates for the look-back time from the spillover region to the target utterance. Y-axis represents the sarcasm effect, which is the difference in the reading times between sarcastic and literal texts. The model means and confidence intervals are back-transformed from log values. Error bars represent 95% CI
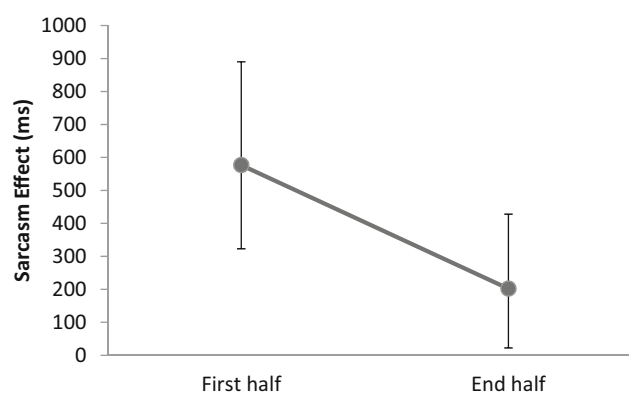
Discussion

In the present experiment, we examined the role of look-backs in the processing of written sarcasm by using a version of the

trailing mask paradigm (Schotter et al., 2014) modified to be suitable for studying the reading of texts. The results showed some effects of the masking manipulation. Readers were less likely to look back to and from the target utterance, showed shorter look-back times, and demonstrated increased first-pass reading time in the spillover region in the mask versus the no-mask condition. The increased first-pass reading time in the spillover region can be taken as an indicator of increased wrap-up processing at the sentence end (e.g., Rayner, Kambe, & Duffy, 2000), suggesting that when look-backs do not provide an opportunity to re-examine the text content, readers invest extra effort into sentence integration before they move on to the next sentence. In other words, look-backs do seem to provide an opportunity to re-examine the text contents. Even though look-backs are likely to be important in forming a comprehensive memory representation of the text contents, masking did not hamper memory or comprehension of the target utterances in the present study. This is probably because readers compensated by increasing sentence wrap-up processing during first-pass reading.

We had expected that if resolving the meaning of a sarcastic utterance requires re-examining the text content (see, e.g., Olkoniemi et al., 2016; Olkoniemi et al., in press), masking should have greater effects on the processing of passages containing sarcastic utterances. However, we obtained no evidence of a critical interaction between text type and mask condition in any of the dependent measures. Interactions only appeared in combination with individual differences measures – specifically, the spatial working memory task.

Processing of sarcasm and sarcasm comprehension

We replicated the previous findings that sarcastic utterances are more likely to attract look-backs and that readers tend to initiate more look-backs to other parts of text from sarcastic than from literal utterances (Au-Yeung et al., 2015; Filik et al., 2014; Filik & Moxey, 2010; Kaakinen et al., 2014; Olkoniemi et al., 2016; Olkoniemi et al., in press; Turcan & Filik, 2016). In the beginning of the experiment, sarcastic utterances also received longer look-backs from the spillover region, implying increased sentence wrap-up processing (Rayner et al., 2000).

The finding that the sentence wrap-up effect (i.e., look-back time to the target utterance from the spillover region) observed for sarcastic utterances diminished across the experiment is in line with previous studies that have reported trial effects (Olkoniemi et al., 2016; Olkoniemi et al., in press). It seems that after encountering several sarcastic utterances during the course of the experiment, readers formed an expectation for sarcasm (Olkoniemi et al., 2016). This, in turn, facilitated the processing of sarcastic utterances. This interpretation is supported by the finding that response times to inference questions about sarcastic utterances decreased across the experiment. The results are in line with theoretical views that assume that context may bias the interpretation of an utterance toward the non-salient indirect meaning (Gibbs, 1994; Giora, 2003; Pexman, 2008).

Sarcastic utterances were also more difficult to comprehend than literal utterances, as evidenced by lower accuracy scores and slower response times to the inference questions (Au-Yeung et al., 2015; Kaakinen et al., 2014; Olkoniemi et al., 2016; Olkoniemi et al., in press). These results show that despite the extra processing effort readers invest in reading sarcastic utterances, they do not always understand the intended meaning. Moreover, the results show that the extra processing effort invested in the processing of sarcasm sometimes reflects problems in comprehension: look-backs initiated from the sarcastic utterance to previous text parts were related to poorer comprehension of the utterance. The result indicates that readers' need to return to previous text parts after reading the sarcastic utterance may reflect confusion about a possible interpretation of the utterance. The result is in line with Gibbs and Colston's (2012) suggestion that a failure to integrate the utterance with the context is one reason why a reader might not understand a sarcastic comment.

However, readers were more accurate in answering text memory questions related to sarcastic than to literal texts (Olkoniemi et al., in press). The extra processing effort invested in the processing of sarcasm may help readers to recall the text content.

Individual differences in the processing of sarcasm

The results showed that SSPAN score was related to compensatory strategies utilized during the reading of sarcastic passages in the mask condition. A lower SSPAN score correlated to longer first-pass re-reading time of sarcastic versus literal utterances in the mask condition. Readers with a relatively low SSPAN score also showed a lower likelihood of initiating look-backs to the sarcastic utterances in the mask condition. These results suggest that spatial working memory is associated with the processing of sarcasm, such that look-backs provide an opportunity for lower spatial WMC readers to re-examine the target utterance (Walczyk & Taylor, 1996). It seems that for higher visuospatial WMC readers, look-backs work as cues to the text content (e.g., Ferreira et al., 2008; Meseguer et al., 2002), as these readers still made look-backs to the sarcastic target utterance even in the mask condition. It is possible that high-SSPAN readers are better at maintaining an episodic memory representation that contains both the text content and the spatial information and that look-backs to previous parts of the text help them to retrieve the content information from memory.

We did not find any relation between RSPAN and the processing of sarcasm, as in our previous studies (Kaakinen et al., 2014; Olkoniemi et al., 2016). A possible explanation is that

even though the correlation between RSPAN and SSPAN scores was relatively low, SSPAN captured most of the variance in WMC, while RSPAN had no explanatory power over and above it. This would be understandable if SSPAN captured the task-specific variance in spatial WMC as well as domain-general variations in WMC.

In accordance with previous research (Olkoniemi et al., 2016), the present results suggest that the ability to recognize emotions is related to the processing of sarcastic utterances. Readers who were better able to recognize emotions (i.e., low TAS-20 score) showed shorter first-pass re-reading times of sarcastic utterances. Sarcasm conveys an emotional message that differs from the meaning of a literal interpretation (Akimoto et al., 2014). Sarcasm is typically used to criticize someone (e.g., Bowes & Katz, 2011), and it may be used as a form of humor and experienced as funny (e.g., Gibbs, Bryant, & Colston, 2014). The intended emotional message was apparent in the sarcastic utterances used in the present experiment: they were evaluated as more negative and less positive than their literal counterparts (Olkoniemi et al., 2016). The result implies that those who are better able to recognize this emotional component are able to resolve sarcasm more quickly. The finding is in line with theoretical views stating that sarcasm comprehension is, to some extent, dependent on emotional inferences (Pexman, 2008). In contrast to our hypothesis, we failed to observe interactions between the effects of masking and the ability to recognize emotions. One possible explanation is that readers slowed down their reading on the spillover region after the sarcastic statements in the mask condition, indicating that readers with a poorer ability to recognize emotions might have had enough time to do extra processing that they would normally need.

The lack of correlation between IGT and processing times for sarcastic texts (cf. Olkoniemi et al., 2016) suggests that the ability to name and recognize the emotions the sarcastic protagonist wants to deliver may be a stronger predictor of sarcasm comprehension than the automatic activation of emotion. Future studies are needed to examine in more detail the role of emotional processing in the comprehension of emotionally complex language such as sarcasm.

Limitations of the study

Naturally, there are some limitations to the present study. First, we failed to observe positive correlations between processing time and the comprehension of sarcastic utterances. We would like to note that exploring correlations between these measures is tricky, as the comprehension scores were relatively high, even for sarcastic passages. The analyses are further complicated by the experimental manipulation employed in the present study (i.e., masking), which had an overall impact on the number and duration of re-inspections. Even though

these results should be considered with some caution, the analyses do complement the planned analyses of the data.

Moreover, the gender distribution of the present sample was skewed (53 of the 62 participants were women). Previous studies suggest that men might be more likely to use sarcasm than women (Colston & Lee, 2004; Gibbs, 2000; Rockwell & Theriot, 2001; cf. Taylor, 2017). However, only a few studies have examined gender differences in the comprehension of sarcasm, and most have failed to find evidence for gender differences (Baptista, Macedo, & Boggio, 2015; see also Holtgraves, 1997). However, it has been suggested that the processing strategies used to understand irony might differ between men and women (Baptista et al., 2015) and that women might be better in recognizing sarcasm (Rothermich & Pell, 2015). Thus, the potential gender effects should be taken into account in the future.

Conclusions

Even though look-backs provide an opportunity to re-examine the text contents when resolving sarcasm, the present results indicate that they are not necessary for the successful comprehension of sarcasm. Readers seem to compensate for their inability to retrieve text content with look-backs by investing extra effort during first-pass reading (see also White et al., 2016). This result suggests that readers are already aware of sarcasm during first-pass reading and that re-inspecting the sarcastic utterance, either during first-pass reading or during look-backs, is important in forming the sarcastic interpretation. Furthermore, the results showed that the need to re-access the sarcastic utterance was mediated by spatial WMC. Readers with lower spatial WMC were more bound to re-accessing the text content, whereas high spatial WMC readers seemed able to use look-backs to the utterance location as cues about the text content. These finding are in line with *the Compensatory-Encoding Model* (Walczyk & Taylor, 1996), suggesting that low WMC readers use text as external memory and employ compensatory strategies (i.e., looking back or slowing down their reading). However, as this is the first study to experimentally examine the role of look-backs in the processing of sarcasm, it is clear that further studies are needed to fully understand how different readers react to and process sarcasm.

The results also showed that the ability to recognize emotions is related to the efficiency of resolving sarcasm. Readers who are better able to recognize emotions invest less processing effort to form a sarcastic interpretation. As sarcasm is typically used to criticize someone (Attardo, 2000; Kreuz & Glucksberg, 1989), it delivers an emotional message (i.e., negative emotions; Olkoniemi et al., 2016). The finding suggests that readers who are better able to recognize emotions are faster to categorize the utterance as sarcastic.

Finally, the results demonstrated that longer sentence wrap-up processing time and slower responses to inference questions observed for sarcastic texts wore off toward the end of the experiment. This suggests that after encountering several sarcastic utterances during the course of the experiment, readers form an expectation of sarcasm, which facilitates its processing.

Of the current theories of sarcasm comprehension, the direct access view (Gibbs, 1994) and the parallel constraint-satisfaction framework (Pexman, 2008) can best accommodate the present results, as they make the general assumption that various reader-, text-, and context-related factors constrain the accessibility of different interpretations during reading. The present study shows that while comprehending unfamiliar sarcastic utterances is more difficult and takes more time than comprehending literal utterances (Giora, 2003; Grice, 1975), the effort required and the strategies used to resolve the meaning of the sarcastic utterances depends on the reading context and reader characteristics.

Appendix A Final models for each eye movement, text memory, and inference measure

Table 5 Final model for the first-pass reading time on the target utterance

Random effects	n	Variance	SD
Participant (Intercept)	62	10.23	3.20
Participant (Literal vs. Sarcasm)		0.14	0.37
Participant (Mask)		0.60	0.78
Participant (Literal vs. Sarcasm × Mask)		0.17	0.41
Text (Intercept)	28	5.97	2.44
Text (Literal vs. Sarcasm)		0.12	0.35
Text (Mask)		0.07	0.27
Text (Order)		0.13	0.36
Text (Literal vs. Sarcasm × Mask)		0.01	0.12
Residual		23.59	4.86
Fixed effects	b	SE	t
Intercept	23.76	0.63	37.93
Literal vs. Sarcasm	-0.14	0.14	-1.01
Mask	0.12	0.16	0.74
Order	-0.55	0.17	**-3.29**
Literal vs. Sarcasm × Mask	-0.09	0.13	-0.72

Note. The *t*-values > 1.96 are in boldface. Order = Trial Order

Table 6 Final model for the forward-fixation time on the target utterance

Random effects	n	Variance	SD
Participant (Intercept)	62	7.23	2.69
Participant (Literal vs. Sarcasm)		0.08	0.28
Participant (Mask)		0.32	0.57
Text (Intercept)	28	4.04	2.01
Text (Literal vs. Sarcasm)		0.19	0.43
Text (Mask)		0.07	0.26
Text (Order)		0.08	0.28
Text (Literal vs. Sarcasm × Mask)		0.004	0.06
Residual		16.26	4.03
Fixed effects	b	SE	t
Intercept	21.88	0.52	42.08
Literal vs. Sarcasm	-0.15	0.13	-1.16
Mask	0.15	0.13	1.18
Order	-0.43	0.13	**-3.26**
Literal vs. Sarcasm × Mask	-0.03	0.10	-0.30

Note. The *t*-values > 1.96 are in boldface. Order = Trial Order; IGT = Iowa Gambling Task score

Table 7 Final model for the probability of first-pass rereading on the target utterance

Random effects	n	Variance	SD
Participant (Intercept)	62	0.59	0.77
Text (Intercept)	28	0.21	0.46
Text (Literal vs. Sarcasm)		0.01	0.12
Text (Mask)		0.01	0.09
Text (Literal vs. Sarcasm × Mask)		0.01	0.10
Fixed effects	b	SE	z
Intercept	-0.83	0.14	-5.82
Literal vs. Sarcasm	-0.04	0.06	-0.72
Mask	0.04	0.06	0.70
Literal vs. Sarcasm × Mask	-0.09	0.06	-1.46

Note. The *z*-values > 1.96 are in boldface

Table 8 Final model for the first-pass rereading time on the target utterance

Random effect	n	Variance	SD
Participants (Intercept)	61	0.03	0.18
Text items (Literal)	28	0.02	0.14
Text items (Sarcasm)		0.04	0.20
Text items (Mask)		0.01	0.12
Text items (SSPAN)		2.30E-5	0.005
Text items (TAS-20)		0.68	0.83
Residual		0.33	0.58

Fixed effect	b	SE	t
Intercept	5.42	0.05	114.83
Literal vs. Sarcasm	-3.21E-4	0.03	0.01
Mask	-0.04	0.03	-1.28
SSPAN	0.002	0.005	0.51
TAS-20	0.22	0.26	0.83
Literal vs. Sarcasm × Mask	0.04	0.03	1.41
Literal vs. Sarcasm × SSPAN	2.70E-4	0.004	0.08
SSPAN × Mask	-0.002	0.003	-0.58
Literal vs. Sarcasm × TAS-20	-0.37	0.15	**-2.46**
Literal vs. Sarcasm × SSPAN × Mask	-0.01	0.004	**-2.33**

Note. The t-*values* > 1.96 are in boldface. SSPAN = Symmetry Span score; TAS-20 = 20 item Toronto Alexithymia Scale score. In the analysis n = 61, because one participant did not do any first-pass rereading

Table 10 Final model for the look-back time on the target utterance from the other text regions than spillover region

Random effect	n	Variance	SD
Participants (Intercept)	60	0.03	0.18
Participant (Literal vs. Sarcasm)		0.001	0.03
Participant (Mask)		0.01	0.08
Text items (Intercept)	28	0.01	0.09
Text items (Literal vs. Sarcasm)		0.02	0.13
Text items (Mask)		4.34E-4	0.02
Text items (Order)		0.002	0.05
Residual		0.50	0.70

Fixed effect	b	SE	t
Intercept	5.54	0.05	120.03
Literal vs. Sarcasm	-0.05	0.04	-1.18
Mask	0.14	0.04	**3.68**
Order	-0.08	0.04	**-2.20**
Literal vs. Sarcasm × Mask	-0.005	0.03	-0.14

Note. The t-*values* > 1.96 are in boldface. Order = Trial Order. In the analysis n = 60, because two participants didn't do any look-backs to target utterance

Table 9 Final model for the probability to initiate a look-back to the target utterance from the other text regions than spillover region

Random effect	n	Variance	SD
Participants (Intercept)	62	0.68	0.83
Participant (Literal vs. Sarcasm)		0.06	0.24
Participant (Mask)		0.13	0.37
Text items (Intercept)	28	0.15	0.38
Text items (Literal vs. Sarcasm)		0.003	0.05
Text items (SSPAN)		4.22E-5	0.01

Fixed effect	b	SE	z
Intercept	-1.24	0.15	-8.53
Literal vs. Sarcasm	-0.18	0.07	**-2.41**
Mask	0.33	0.08	**3.99**
Order	-0.36	0.08	**-4.49**
SSPAN	0.01	0.02	0.36
Literal vs. Sarcasm × Mask	-0.03	0.06	-0.50
Literal vs. Sarcasm × SSPAN	-0.001	0.01	-0.11
SSPAN × Mask	0.02	0.01	**2.03**
Literal vs. Sarcasm × SSPAN × Mask	0.02	0.01	**2.17**

Note. The z-*values* > 1.96 are in boldface. Order = Trial Order; SSPAN = Symmetry Span score

Table 11 Final model for the probability to initiate a look-from from the target utterance

Random effect	n	Variance	SD
Participants (Intercept)	62	0.89	0.94
Text items (Intercept)		0.17	0.41
Text items (Literal vs. Sarcasm)		0.03	0.16
Text items (Mask)	28	0.02	0.13
Text items (Order)		0.02	0.13
Text items (SSPAN)		3.44E-5	0.01

Fixed effect	b	SE	z
Intercept	-1.39	0.16	-8.65
Literal vs. Sarcasm	-0.16	0.07	**-2.20**
Mask	0.30	0.07	**4.29**
Order	-0.53	0.07	**-7.37**
SSPAN	0.03	0.02	1.55
Literal vs. Sarcasm × Mask	-0.08	0.06	-1.25
Order × SSPAN	0.03	0.01	**3.05**

Note. The z-*values* > 1.96 are in boldface. Order = Trial Order; SSPAN = Symmetry Span score

Table 12 Final model for the first-pass reading time on the spillover region

Random effects	v	Variance	SD
Participant (Intercept)	62	9.53	3.09
Text (Intercept)	28	6.76	2.60
Text (Literal vs. Sarcasm)		0.22	0.47
Text (Mask)		0.02	0.14
Text (Order)		0.06	0.24
Residual		37.59	6.13
Fixed effects	b	SE	t
Intercept	23.38	0.65	36.18
Literal vs. Sarcasm	0.32	0.17	1.84
Mask	-0.55	0.15	**-3.59**
RSPAN	1.85	1.04	1.78
Order	-0.52	0.16	**-3.33**
IGT	0.11	0.04	**3.02**
Literal vs. Sarcasm × Mask	-0.03	0.15	-0.18
Mask × RSPAN	-1.37	0.38	**-3.63**

Note. The *t*-values > 1.96 are in boldface. Order = Trial Order; RSPAN = Reading Span score

Table 14 Final model for the look-back time to the target utterance from the spillover region

Random effects	n	Variance	SD
Participant (Intercept)	62	.07	.26
Text (Intercept)	28	.01	.12
Text (Literal vs. Sarcasm)		.003	.06
Text (Mask)		.004	.06
Text (Order)		.01	.08
Text (Literal vs. Sarcasm × Mask)		.003	.05
Text (Literal vs. Sarcasm × Order)		.003	.05
Residual		.44	.66
Fixed effects	b	SE	t
Intercept	5.99	0.06	101.41
Literal vs. Sarcasm	0.001	0.04	0.02
Mask	-0.06	0.04	-1.52
Order	-0.08	0.04	-1.72
Literal vs. Sarcasm × Mask	-0.04	0.04	-0.97
Literal vs. Sarcasm × Order	0.14	0.04	**-3.49**

Note. The *t*-values > 1.96 are in boldface. Order = Trial Order

Table 13 Final model for the probability to initiate a look-back to the target utterance from the spillover region

Random effects	n	Variance	SD
Participant (Intercept)	62	0.91	0.95
Participant (Literal vs. Sarcasm)		0.08	0.28
Participant (Mask)		0.07	0.26
Participant (Literal vs. Sarcasm × Mask)		0.02	0.13
Text (Intercept)	28	0.21	0.46
Text (Literal vs. Sarcasm)		0.09	0.30
Text (Mask)		0.06	0.24
Text (Order)		0.03	0.18
Text (Literal vs. Sarcasm × Mask)		0.04	0.20
Fixed effects	b	SE	z
Intercept	1.75	0.17	10.18
Literal vs. Sarcasm	0.06	0.10	0.60
Mask	0.01	0.10	0.13
Order	0.54	0.09	**6.36**
Literal vs. Sarcasm × Mask	0.16	0.09	1.76

Note. The *z*-values > 1.96 are in boldface. Order = Trial Order

Table 15 Final model for the correct answers to the text memory questions

Random effect	n	Variance	SD
Participants (Intercept)	62	0.39	0.62
Text items (Intercept)	28	3.76	1.94
Text items (Literal vs. Sarcasm)		0.57	0.76
Participants (Mask)		0.04	0.20
Participants (Literal vs. Sarcasm × Mask)		0.04	0.21
Fixed effect	b	SE	z
Intercept	4.17	0.54	7.68
Literal vs. Sarcasm	-0.78	0.38	**-2.06**
Mask	0.04	0.30	0.14
Literal vs. Sarcasm × Mask	0.31	0.30	1.08

Note. The z-*values* > 1.96 are in boldface

Table 16 Final model for the reaction time to the text memory questions

Random effects	n	Variance	SD
Participant (Intercept)	62	0.04	0.19
Text (Intercept)	28	0.02	0.14
Text (Literal vs. Sarcasm)		0.003	0.06
Text (Mask)		0.001	0.03
Text (Order)		2.24E-4	0.01
Text (IGT)		1.98E-6	0.001
Residual		0.08	0.28
Fixed effects	b	SE	t
Intercept	7.80	0.04	209.94
Literal vs. Sarcasm	-0.002	0.01	-0.18
Mask	-0.01	0.01	-1.27
Order	-0.08	0.01	**-10.55**
IGT	0.005	0.002	**2.07**
Literal vs. Sarcasm × Mask	0.003	0.01	0.44

Note. The *t*-values > 1.96 are in boldface. Order = Trial Order; IGT = Iowa Gambling Task score

Table 17 Final model for the correct answers to the inference questions

Random effect	n	Variance	SD
Participants (Intercept)	62	1.14	1.07
Text items (Intercept)		0.17	0.41
Text items (Literal vs. Sarcasm)	28	0.56	0.75
Text items (Mask)		0.04	0.21
Text items (Order)		0.08	0.29
Text items (SSPAN)		5.87E-5	0.01
Fixed effect	b	SE	z
Intercept	3.43	0.25	13.76
Literal vs. Sarcasm	1.11	0.21	**5.25**
Mask	0.10	0.13	0.75
Order	0.54	0.12	**4.46**
SSPAN	0.05	0.02	**2.44**
Literal vs. Sarcasm × Mask	0.01	0.13	0.06

Note. The *z*-values > 1.96 are in boldface. Order = Trial Order, SSPAN = Symmetry Span score

Table 18 Final model for the reaction time to the inference questions

Random effects	n	Variance	SD
Participant (Intercept)	62	0.04	0.19
Text (Intercept)		0.02	0.15
Text (Literal vs. Sarcasm)	28	0.01	0.08
Text (Mask)		0.001	0.03
Text (Order)		0.001	0.03
Text (IGT)		3.04E-6	0.001
Residual		0.10	0.32
Fixed effects	b	SE	t
Intercept	7.91	0.04	207.81
Literal vs. Sarcasm	-0.07	0.02	**-4.16**
Mask	0.002	0.01	0.29
Order	-0.09	0.01	**-10.19**
IGT	0.01	0.002	**2.88**
Literal vs. Sarcasm × Mask	0.001	0.01	0.10
Literal vs. Sarcasm × Order	0.03	0.01	**3.51**

Note. The *t* > 1.96 are in boldface. Order = Trial Order; IGT = Iowa Gambling Task score

Table 19 Final model for the relation between sarcasm comprehension and first-pass rereading time on the sarcastic target utterance

Random effects	n	Variance	SD
Participant (Intercept)	62	2.19	1.48
Participant (F-PRR)		0.05	0.22
Text (Intercept)		0.38	0.62
Text (F-PRR)		1.20	1.10
Fixed effects	b	SE	z
Intercept	2.73	0.50	5.42
F-PRR	-0.45	0.58	-0.77

Note. The *z* > 1.96 are in boldface

F-PRR = First-Pass Rereading time

Table 20 Final model for the relation between sarcasm comprehension and probability to initiate a look-back to the sarcastic target utterance

Random effects	n	Variance	SD
Participant (Intercept)	62	2.01	1.42
Participant (LB)		0.01	0.09
Text (Intercept)		0.48	0.70
Text (LB)		0.02	0.15
Fixed effects	b	SE	z
Intercept	2.29	0.28	8.12
LB	0.34	0.19	1.82

Note. The *z* > 1.96 are in boldface

LB = Probability to Initiate a Look-Back to the Target Utterance

Appendix B Final models for the analysis on relationship between reading times and sarcasm comprehension

Table 21 Final model for the relation between sarcasm comprehension and look-back time from the spillover region to the sarcastic target utterance

Random effects	n	Variance	SD
Participant (Intercept)	62	0.41	0.64
Participant (Spill)		1.07	1.03
Text (Intercept)		0.16	0.40
Text (Spill)		0.06	0.24
Fixed effects	b	SE	z
Intercept	2.28	0.30	7.54
Spill	-0.40	0.38	-1.06

Note. The $z > 1.96$ are in boldface

Spill = Look-Back Time From the Spillover Region to the Target Utterance

Table 22 Final model for relation between sarcasm comprehension and probability to initiate a look-from the sarcastic target utterance

Random effects	n	Variance	SD
Participant (Intercept)	62	2.03	1.42
Participant (From)		0.01	0.09
Text (Intercept)		0.45	0.67
Text (From)		0.11	0.33
Fixed effects	b	SE	z
Intercept	2.27	0.28	8.09
From	0.48	0.19	**2.55**

Note. The $z > 1.96$ are in boldface

From = Probability to Initiate a Look-From the Target Utterance

References

Akimoto, Y., Sugiura, M., Yomogida, Y., Miyauchi, C. M., Miyazawa, S., & Kwashima, R. (2014). Irony comprehension: Social conceptual knowledge and emotional response. *Human Brain Mapping, 35*, 1167–1178. https://doi.org/10.1002/hbm.22242

Amenta, S., Noël, X., Verbanck, P., & Campanella, S. (2013). Decoding emotional components in complex communicative situations (irony) and its relation to empathic abilities in male chronic alcoholics: An issue for treatment. *Alcoholism: Clinical & Experimental Research, 37*, 339–347. https://doi.org/10.1111/j.1530-0277.2012.01909.x

Attardo, S. (2000). Irony as relevant inappropriateness. *Journal of Pragmatics, 32*, 793–826. https://doi.org/10.1016/S0378-2166(99)00070-3

Au-Yeung, S. K., Kaakinen, J. K., Liversedge, S. P., & Benson, V. (2015). Processing of written irony in autism spectrum disorder: An eye-movement study. *Autism Research, 8*, 749–760. https://doi.org/10.1002/aur.1490

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baccino, T., & Pynte, J. (1994). Spatial coding and discourse models during text reading. *Language and Cognitive Processes, 9*, 143–155. https://doi.org/10.1080/01690969408402114

Baccino, T., & Pynte, J. (1998). Spatial encoding and referential processing during reading. *European Psychologist, 3*, 51–61. https://doi.org/10.1027/1016-9040.3.1.51

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale-I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38*, 23–32. https://doi.org/10.1016/0022-3999(94)90005-1

Baptista, N. I., Macedo, E. C., & Boggio, P. S. (2015). Looking more and at different things: Differential gender eye-tracking patterns on an irony comprehension task. *Psychology & Neuroscience, 8*, 157–167. https://doi.org/10.1037/h0101061

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. https://doi.org/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. https://doi.org/10.18637/jss.v067.i01

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50*, 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Booth, R. W., & Weger, U. W. (2013). The function of regression in reading: Backward eye movements allow re-reading. *Memory & Cognition, 41*, 82–97. https://doi.org/10.3758/s13421-012-0244-y

Bowes, A., & Katz, A. (2011). When sarcasm stings. *Discourse Processes, 48*, 215–236. https://doi.org/10.1080/0163853X.2010.532757

Buelow, M. T., & Suhr, J. A. (2009). Construct validity of the Iowa Gambling Task. *Neuropsychology Review, 19*, 102–114. https://doi.org/10.1007/s11065-009-9083-4

Clark, H. H., & Gerrig, R. J. (1984). On the pretense theory of irony. *Journal of Experimental Psychology: General, 113*, 121–126. https://doi.org/10.1037/0096-3445.113.1.121

Colston, H. L., & Lee, S. Y. (2004). Gender differences in verbal irony use. *Metaphor and Symbol, 19*, 289–306. https://doi.org/10.1207/s15327868ms1904_3

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*, 769–786. https://doi.org/10.3758/BF03196772

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions of Psychological Science, 19*, 51–57. https://doi.org/10.1177/0963721409359277

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Engle, R. W. (2010). Role of working-memory capacity in cognitive control. *Current Anthropology, 51*, 17–26. https://doi.org/10.1086/650572

Fallows, J. (2016). Why Twitter doesn't work with sarcasm, chap. 823. *The Atlantic*. Retrieved from: https://www.theatlantic.com/notes/2016/01/why-twitter-doesn't-work-with-sarcasm-chap-823/424761/.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191. https://doi.org/10.3758/BF03193146

Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences, 12*, 405–410. https://doi.org/10.1016/j.tics.2008.07.007

Filik, R., Leuthold, H., Wallington, K., & Page, J. (2014). Testing theories of irony processing using eye-tracking and ERPs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 811–828. https://doi.org/10.1037/a0035658

Filik, R., & Moxey, L. M. (2010). The on-line processing of written irony. *Cognition, 116*, 421–436. https://doi.org/10.1016/j.cognition.2010.06.005

Fougnie, D., Zughni, S., Godwin, D., & Marois, R. (2015). Working memory storage is intrinsically domain specific. Journal of Experimental Psychology: General, 144, 30–47. https://doi.org/10.1037/a0038211

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178–210. https://doi.org/10.1016/0010-0285(82)90008-1

Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior Research Methods, 37*, 581–590. https://doi.org/10.3758/BF03192728

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanism. *Psychological Science, 4*, 294–298. https://doi.org/10.1111/j.1467-9280.1993.tb00567.x

Gibbs, R. W. (1994). The poetics of mind: Figurative thought, language, and understanding. Cambridge: Cambridge University Press.

Gibbs, R. W. (2000). Irony in talk among friends. *Metaphor and Symbol, 15*, 5–27. https://doi.org/10.1080/10926488.2000.9678862

Gibbs, R. W., Bryant, G. A., & Colston, H. L. (2014). Where is the humor in verbal irony? *Humor, 27*, 575–595. https://doi.org/10.1515/humor-2014-0106

Gibbs, R. W., & Colston, H. L. (2012). Interpreting figurative meaning. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139168779.008

Giora, R. (1999). On the priority of salient meanings: Studies of literal and figurative language. *Journal of Pragmatics, 31*, 919–929. https://doi.org/10.1016/S0378-2166(98)00100-3

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

Grice, H. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), Syntax and semantics 3: Speech acts (pp. 41–58). New York, NY: Academic Press.

Hancock, J. T. (2004). Verbal irony use in face-to-face and computer-mediated conversations. *Journal of Language and Social Psychology, 23*, 447–463. https://doi.org/10.1177/0261927X04269587

Holtgraves, T. (1997). Styles of language use: Individual and cultural variability in conversational indirectness. *Journal of Personality and Social Psychology, 73*, 624–637. https://doi.org/10.1037/0022-3514.73.3.624

Hyönä, J., Lorch, R. F., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94*, 44–55. https://doi.org/10.1037//0022-0663.94.1.44

Hyönä, J., Lorch, R. F., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), The mind's eye: Cognitive and applied aspects of eye movement research (pp. 313–334). Amsterdam: Elsevier Science. https://doi.org/10.1016/B978-044451020-4/50018-9

Hyönä, J., & Nurminen, A. M. (2006). Do adult readers know how they read? Evidence from eye movement patterns and verbal reports. *British Journal of Psychology, 97*, 31–50. https://doi.org/10.1348/000712605X53678

Jacob, H., Kreifelts, B., Nizielski, S., Schütz, A., & Wildgruber, D. (2016). Effects of emotional intelligence on the impression of irony created by the mismatch between verbal and nonverbal cues. *PLoS ONE, 11*, 1–17. https://doi.org/10.1371/journal.pone.0163211

Joukamaa, M., Miettunen, J., Kokkonen, P., Koskinen, M., Julkunen, J., Kauhanen, J.(2001). Psychometric properties of the Finnish 20-item Toronto Alexithymia Scale. *Nordic Journal of Psychiatry, 55*, 123–127. https://doi.org/10.1080/08039480116694

Kaakinen, J. K., & Hyönä, J. (2007). Strategy use in the reading span test: An analysis of eye movements and reported encoding strategies. *Memory, 15*, 634–645. https://doi.org/10.1080/09658210701457096

Kaakinen, J. K., Olkoniemi, H., Kinnari, T., & Hyönä, J. (2014). Processing of written irony: An eye movement study. *Discourse Processes, 51*, 287–311. https://doi.org/10.1080/0163853X.2013.870024

Kennedy, A., Brooks, R., Flynn, L.-A., & Prophet, C. (2003). The reader's spatial code. In J. Hyönä, R. Radach, & H. Deubel (Eds.), The mind's eye: Cognitive and applied aspects of eye movement research (pp. 193–212). Amsterdam: Elsevier Science. https://doi.org/10.1016/B978-044451020-4/50012-8

Kreuz, R. J., & Glucksberg, S. (1989). How to be sarcastic: The echoic reminder theory of verbal irony. *Journal of Experimental Psychology: General, 118*, 374–386. https://doi.org/10.1037/0096-3445.118.4.374

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), Eye guidance in reading and scene perception (pp. 55–100). Oxford: Elsevier Science. https://doi.org/10.1016/B978-008043361-5/50004-3

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*, 578–586. https://doi.org/10.3758/BF03203972

Meseguer, E., Carreiras, M., & Clifton, C. (2002). Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory & Cognition, 30*, 551–561. https://doi.org/10.3758/BF03194956

Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. L. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the Selective Reanalysis hypothesis. *Journal of Memory and Language, 59*, 266–293. https://doi.org/10.1016/j.jml.2008.06.002

Mueller, S. (2012). The Psychology Experiment Building Language (Version 0.13.) [Computer software]. Retrieved from http://pebl.sourceforge.net.

Murray, W. S., & Kennedy, A. (1988) Spatial coding in the processing of anaphor by good and poor readers: Evidence from eye movement analyses. *The Quarterly Journal of Experimental Psychology, 40*, 693–718. https://doi.org/10.1080/14640748808402294

Nicholson, A., Whalen, J. M., & Pexman, P. M. (2013). Children's processing of emotion in ironic language. *Frontiers in Psychology, 4*, 1–10. https://doi.org/10.3389/fpsyg.2013.00691

Olkoniemi, H., Ranta, H., & Kaakinen, J. K. (2016). Individual differences in the processing of written sarcasm and metaphor: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 433–450. https://doi.org/10.1037/xlm0000176

Olkoniemi, H., Strömberg, V., & Kaakinen, J.K. (in press). The ability to recognizeognise emotions predicts the time-course of sarcasm

processing: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*.

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods, 47*, 1343–1355. https://doi.org/10.3758/s13428-014-0543-2

Pexman, P. M. (2008). It's fascinating research: The cognition of verbal irony. *Current Directions in Psychological Science, 17*, 286–290. https://doi.org/10.1111/j.1467-8721.2008.00591.x

R Core Team (2017). R: A language and environment for statistical computing. Retrieved from http://www.R-project.org.

Raney, G. E., Campbell, S. J., & Bovee, J. C. (2014). Using eye movements to evaluate the cognitive processes involved in text comprehension. *Journal of Visualized Experiments, 83*, e50780. https://doi.org/10.3791/50780

Rayner, K. (1998) Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflection of comprehension processes in reading. *Scientific Studies of Reading, 10*, 241–255. https://doi.org/10.1207/s1532799xssr1003_3

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology Section A, 53*, 1061–1080. https://doi.org/10.1080/713755934

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28*, 164–171. https://doi.org/10.1027/1015-5759/a000123

Robbins, M. (2010). The war on irony (on the Twitter joke trial). *The Guardian.* Retrieved from: https://www.theguardian.com/science/the-lay-scientist/2010/nov/11/2

Rockwell, P., & Theriot, E. M. (2001). Culture, gender, and gender mix in encoders of sarcasm: A self-assessment analysis. *Communication Research Report, 18*, 44–52. https://doi.org/10.1080/08824090109384781

Rothermich, K., & Pell, M. D. (2015). Introducing RISC: A new video inventory for testing social perception. *PLoS ONE, 10*, e0133902. https://doi.org/10.1371/journal.pone.0133902

Ronson, J. (2015) How one stupid tweet blew up Justine Sacco's life. Retrieved from: https://www.nytimes.com/2015/02/15/magazine/how-one-stupid-tweet-ruined-justine-saccos-life.html.

Schotter, E. R., Tran, R., & Rayner, K. (2014'. Don't believe what you read (only once): Comprehension is supported by regressions during reading. *Psychological Science, 25*, 1218–1226. https://doi.org/10.1177/0956797614531148

Schwoebel, J., Dews, S., Winner, E., & Srinivas, K. (2000). Obligatory processing of the literal meaning of ironic utterances: Further evidence. *Metaphor and Symbol, 15*, 47–61. https://doi.org/10.1080/10926488.2000.9678864

Shamay-Tsoory, S. G., Tomer, R., & Aharon-Peretz, J. (2005). The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology, 19*, 288–300. https://doi.org/10.1037/0894-4105.19.3.288

Taylor, C. (2017). Women are bitchy but men are sarcastic? Investigating gender and sarcasm. *Gender and Language, 11*, 415–445. https://doi.org/10.1558/genl.27906

Turcan, A., & Filik, R. (2016). An eye-tracking investigation of written sarcasm comprehension: The role of familiarity and context. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 1867–1893. https://doi.org/10.1037/xlm0000285

Walczyk, J. J., & Taylor, R. W. (1996). How do the efficiencies of reading subcomponents relate to looking back in text? *Journal of Educational Psychology, 88*, 537–545. https://doi.org/10.1037/0022-0663.88.3.537

Whalen, J. M., Pexman, P. M., & Gill, A. J. (2009). "Should be fun—Not!" Incidence and marking of nonliteral language in e-mail. *Journal of Language and Social Psychology, 28,* 263–280. https://doi.org/10.1177/0261927X09335253

Whalen, J. M., Pexman, P. M., Gill, A. J., & Nowson, S. (2013). Verbal irony use in personal blogs. *Behaviour & Information Technology, 32*, 560–569. https://doi.org/10.1080/0144929X.2011.630418

White, S. J., Lantz, L. M., & Paterson, K. B. (2016). Spontaneous re-reading within sentences: Eye movement control and visual sampling. *Journal of Experimental Psychology: Human Perception and Performance, 43*, 395–413. https://doi.org/10.1037/xhp0000307

von der Malsburg, T., & Vasishth, S. (2013). Scanpaths reveal syntactic underspecification and reanalysis strategies, *Language and Cognitive Processes, 28*, 1545–1578. https://doi.org/10.1080/01690965.2012.728232