ZEITSCHRIFT FÜR SPRACHWISSENSCHAFT
Journal of the Linguistic Society of Germany

Volume 44, Issue 1, Article 6

https://doi.org/10.18148/zs/2025-2006

Natascha Raue (corresponding author), Universität Osnabrück, Germany, nraue@uni-osnabrueck.de
Álvaro Cortés Rodríguez, Universität Potsdam, Germany, alvaro.cortes.rodriguez@uni-potsdam.de
Holden Härtl, Universität Kassel, Germany, holden.haertl@uni-kassel.de

The processing of quotation marks in German

Evidence from eye tracking

Abstract: In name-mentioning quotations, the name of a concept is mentioned and not used denotationally. Although there is substantial philosophical research on the notion of quotation, empirical evidence is sparse. In the current study, we use eye-tracking data to look into the nature of the processing of quotation marks in name-mentioning constructions. The results indicate that while there are no significant differences for early eye-tracking measures (e.g., first fixation duration), a significant effect was detected for the expression in the target interest area for dwell time (i.e., late measures). Target words enclosed in quotation marks take longer to be processed than target words without quotes. We argue that our findings suggest the involvement of higher cognitive processes in the comprehension of sentences with quotation marks.

Keywords: quotation, eye tracking, cognitive processes, reading processes

Cite as: Raue, Natascha, Álvaro Cortés Rodríguez & Holden Härtl. 2025. The processing of quotation marks in German: Evidence from eye tracking. *Zeitschrift für Sprachwissenschaft* 44(1). https://doi.org/10.18148/zs/2025-2006. **Received:** 30 December 2024; **accepted:** 17 April 2025; **published:** 11 July 2025
Published by: publikon / dgfs (German Linguistic Society)

1 Introduction

Language users sometimes choose to discuss specific characteristics of a certain word rather than using the word to directly refer to its denotation. For example, when someone asserts that *wobble* is a funny word, they may be referring to its whimsical sound, which can evoke a sense of playfulness and contribute to its somewhat amusing quality. One way of marking such a metalinguistic use of an expression is by employing quotation marks (or quotes for short), which signify that the enclosed word is not used in its conventional, that is, denotational sense. While there is a broad range of language-philosophical literature on quotation, experimental research on quotation is sparse. Thus, for instance, the question at which comprehension stage a quotational use of an expression is processed remains unanswered. Furthermore, in this context, the linguistic basis of quotation – whether it should be represented as part of the compositional representation or rather as part of the pragmatic domain of language – is still undetermined. More specifically, a semantic view would be related to truth-conditional effects while a pragmatic account would be supported by a violation on discourse level.

To address these issues, we present the results of an eye-tracking study conducted in German in the current paper. Specifically, we investigated processing reflexes derived from eye-tracking data collected during the comprehension of sentences containing name-mentioning predicates such as *call*, as in the sentence *This phenomenon is called "solar eclipse"*. In sentences of this type, a quotational meaning is entailed compositionally, as predicates like *call* require the mentioning of the name of a lexicalized concept (i.e., *"solar eclipse"*) as one of the predicate's arguments.[1] In particular, we are interested in processing reflexes that inform us about the temporal characteristics of comprehending a quotational construction. This study is groundbreaking in that, to our knowledge, it is the first to employ eye-tracking methodology to explore the processing of pure quotes. Specifically, we expect differences in eye movements and fixation durations between sentences with and without quotes. Given that our processing system is sensitive to punctuation marks, we hypothesize that quotes will also influence eye-tracking metrics, specifically leading to increased fixation durations, more fixations on quoted items, and a greater number of regressions to quoted words. Against the background of these research questions, we follow accounts that interpret effects in earlier time windows to reflect more automatized processes rooted, e.g., in the grammatical-compositional system. In contrast, later processing stages are assumed to be controlled by higher cognitive, reasoning processes, which compute pragmatic as well as discourse-structural information. The empirical results will contribute to locating the operation of quotation within the linguistic system, specifically, at the boundary between compositional grammar and pragmatic domains of language.

[1] Please note that name-mentioning constructions can also be used to refer to non-lexicalized items which need to be introduced within a specific speech community as a specialized term. Despite this, we are focusing entirely on lexicalized items in the following analysis.

The structure of this paper is as follows: In Section 2, we look at quotation from the perspective of the interface between semantics and pragmatics. Our claim is that quotational meanings are construed pragmatically rather than as part of the semantic compositional representation. In Section 3, we delve into empirical evidence on quotation and quotes. In the following three sections, we present and discuss the findings of an experimental study conducted to investigate the processing of quotations, using eye-tracking data. Finally, Section 7 concludes our investigation.

2 Quotation and the semantics-pragmatics interface

Quotation is a metalinguistic device used to talk about certain dimensions of language (see, e.g., Davidson 1979; Cappelen & Lepore 1997; Saka 1998).[2] In quotational constructions, expressions are mentioned rather than or in addition to being used denotationally. With an assertion like in (1a) in contrast to (1b), the syllabic setup of the word *sofa* is described and the quotation marks around *sofa* indicate this use, which means reference is made to a linguistic dimension of the quoted expression, see, e.g., Quine (1981).

(1) a. *"Sofa" has two syllables.*

b. *A sofa is a piece of furniture.*

At least five different quotation types are commonly distinguished in the literature, including direct, indirect, pure, mixed, and scare quotes (see Brendel et al. 2011 for an overview). A special subtype of quotation that is the focus of this analysis is *name-informing quotation*. Corresponding constructions typically contain predicates like *call*, *name*, *refer to*, etc., as embodied in (2) below, and are used to display the linguistic shape of a concept's conventionalized name.

(2) a. *One calls this disease "septicemia".*

b. *A function that calls itself is named "recursive function".*

c. *The purity of gold is referred to with the word "karat".*

In (2a), for instance, *call* is used to describe the naming convention that some occurrence of blood poisoning (*this disease*) is commonly referred to as *septicemia*. Thus, the verbal root of *call* involves three thematic arguments: an agent x, which is bound generically here, a theme y and a relational argument that, in this case, introduces a shape *"n"* of the name of the theme argument y, see (3).

[2] Please note that the argumentation outlined in this section is based on theoretical work discussed in Härtl (2018; 2020) and Schlechtweg & Härtl (2023).

(3) a. x call- y “n”

b. λy λn λx [CALL(x, y, NAME(“n”, y))]

c. GENx [CALL(x, this disease, NAME(“septicemia”, this disease))]

As argued in Härtl (2018), quotations of this sort are instances of pure quotation, i.e., a metalinguistic device used to demonstrate linguistic shapes in a rule-like fashion (see, e.g., Davidson 1979; Cappelen & Lepore 1997; Maier 2014). A standard case of pure quotation is represented in the example in (1a) above. As an explication of their metalinguistic status, pure quotations can be preceded by appositions like *the word*, as exemplified in (2c).

The notion of quotation relates to a mental operation that enables us to talk about language, and quotes, as a typographical means, represent a material realization of this operation. In many cases quotes are optional. Consider the examples in (4), containing expressions that are mentioned metalinguistically (i.e., *cats* and *thoracic outlet syndrome*) but used without quotes.

(4) a. *Cats is a noun.*

b. *The pediatrician has diagnosed a so-called thoracic outlet syndrome.*

The optionality of quotes is a central matter in the theoretical debates centering around the question whether quotes are an essential part of a quotation’s compositional semantic representation or not. Proponents of a semantic analysis of quotes often claim that the presence of quotes has truth-conditional effects and that they are used to produce truth-conditionally relevant content (Simchen 1999; Predelli 2003). On a semantic account, the apparent optionality of quotes can be motivated with the quoted material’s contextual embedding, which in many cases is sufficient to generate a mentioning reading of an expression (see Cappelen & Lepore 1999: 743). Under such a view, quotes materialize opaquely or on a different linguistic level such as the acoustic level.

The optionality of quotes can also be used as evidence for a pragmatic approach to quotes. Under a pragmatic view, quotes may not materialize at all. In this way, Washington (1992) argues that neither graphemic quotes nor their gestural and acoustic equivalents are an essential part of a quotational construction. On his account, quotes are no more than a punctuation device and, as such, “are neither mentioning expressions nor parts of mentioning expressions” (Washington 1992: 591). Approaches of this sort imply that quotes are not semantic in the sense that their manifestation is not part of the compositional semantic representation of a quotational construction. Instead, quotes are considered pragmatic in nature. From this perspective, quotes have been analyzed as pragmatic markers, indicating a deviation from the standard, denotational use of an expression and giving rise to a non-stereotypical interpretation instead (see, e.g., Klockow 1978; Gutzmann & Stei 2011).

We assume pragmatic operations to be related with later time windows in processing, as opposed to earlier time windows, which are associated with automatized processes, for instance,

with the processing of grammatical or lexical-semantic information. On such a perspective, pragmatically anomalous sentences have been reported to elicit a P600 component, i.e., a late ERP component, in a study by Kuperberg et al. (2003). In addition, the processing of verbal irony – a pragmatically rooted phenomenon – has been observed to elicit a P600 effect and suggested to be connected to comprehension processes at a pragmatic level (Regel et al. 2011). In contrast, grammatically and lexically rooted processing is typically associated with earlier time windows such as the N400 or the ELAN component (Friederici et al. 2004). Against this background, we relate later processing stages to be controlled by higher cognitive, reasoning processes, which compute contextual as well as discourse-structural, that is, pragmatic information. Based on acoustic data from a production study and, respectively, reaction-time data from a study on the processing of verbal irony, Schlechtweg & Härtl (2020, 2023) argued that quotation is established at the pragmatic level. In the current study, we test the *pragmatic* view on quotation using eye-tracking data.

3 Empirical evidence: quotation and quotes

This section provides an overview of previous research on quotation and quotes, with a focus on semantic-pragmatic issues and studies examining punctuation. Generally speaking, there is limited research on the influence of punctuation marks on the reading process using eye-tracking methodology. The focus of empirical evidence on punctuation marks using eye-tracking data has mostly centered around the effects of commas on processing texts (Hill & Murray 2000; Hirotani et al. 2006; Pynte & Kennedy 2007).

One exception is found in two eye-tracking studies presented in Yao & Scheepers (2011) on the reading of written stories using direct and indirect quotations. In these studies, the context was modulated using appropriate adverbs giving rise either to a fast or slow-speaking context (for instance *She said slowly*/*quickly*). In both oral reading and silent reading experiments, the context affected the reading rates. Direct speech was, in contrast to indirect speech, read significantly faster in fast-speaking quoted contexts in both oral and silent reading settings. These findings were interpreted as reflecting the vivid representation, i.e., the speed simulation, of a person's inner speech. Stites et al. (2013) empirically support these findings, presenting evidence of a study in which the speaking rate was modulated using a single adverb in direct and indirect quotations. They observed a tendency that speed-related adverbs like *quickly* and *slowly* impacted the reading rates so that those with *quickly* were read faster and vice versa. Comparing direct to indirect quotes, these effects were interestingly not evident in a nearly identical section of text when presented as an indirect quote.

Previous investigations into the nature of name-mentioning constructions (NMCs) have utilized various methodologies, including analysis of corpus data (see Härtl 2020; Raue & Cortés Rodríguez 2022; Raue 2024), acceptability judgements (Cortés Rodríguez et al. 2022), and reading studies implementing the self-paced reading paradigm (Schlechtweg & Härtl 2023; Raue 2024).

Empirical evidence from a large-scale corpus study (Raue 2024) found that NMCs in German involving the predicate *nennen* 'call', as in the sentence *Man nennt dieses Phänomen einen Sonnenring* '*One calls this phenomenon a sun halo*', have a significant preference for not being accompanied by quotes. In addition, this study found that NMCs involving *nennen* 'call' appeared significantly more often with nominals of low lexical frequency as opposed to expressions with a high lexical frequency, indicating that NMCs are predominantly used to introduced unknown terminology to the listeners/readers. In addition, a recent study using the self-paced reading paradigm revealed that the processing of irony in NMCs is facilitated if the nominal is enclosed in quotes (Schlechtweg & Härtl 2023). Quotes have been argued to have an effect on language processing, creating a processing burden first but facilitating reading times at the end of quotation-marked target words, helping the reader to recognize ironic content.

Despite this, the nature of the reading process involved in the processing of metalinguistic sentences and of the special type of pure quotation has not been investigated using eye-tracking methodology. We regard this approach to NMCs as particularly beneficial to gain insights into the moment-to-moment processing with millisecond accuracy and possible differences in early and late processing time windows.

4 Methodology

The present study aims at investigating the processing of quotes in the written mode. More specifically, the study aims at answering the research question when and how quotes affect the eye movements in NMCs during reading.

4.1 Eye-tracking in reading

This section provides an overview of key findings in eye-tracking research on reading, focusing on the basic eye movements relevant to this study. Since initial observations of eye movements during reading in 1879 by Hering and Lamare (Wade 2010), eye tracking has become a powerful tool for studying cognitive processes related to language processing and reading.

The primary eye movements measured include saccades and fixations. Saccades are rapid eye movements between fixations, while fixations involve the stabilization of the eyes on a single point. It is generally necessary to fixate on a word to extract its meaning and morphological information (Staub & Rayner 2007). Moreover, fixations are categorized into progressive fixations, which involve forward-looking movements, and regressive fixations, which are backward movements. For left-to-right reading languages, skilled readers typically make forward saccades averaging 7–9 letters (Rayner 1978), while regressions account for approximately 10 % of all saccades (Engbert & Kliegl 2011). Importantly, high-frequency or short words often appear as skipped – i.e., not fixated – in reading data (Engbert & Kliegl 2011).

Additionally, three lexical influences affect eye movements at the word level: the word length effect (longer words take longer to process), the predictability effect (more predictable words are processed faster), and the frequency effect (frequently occurring words are processed quickly) (Ehrlich & Rayner 1981; Inhoff & Rayner 1986; Rayner & Duffy 1986; Rayner et al. 2011). Other factors include word familiarity, acquisition age, orthographic transparency, and semantic properties (Clifton et al. 2016; Attardo & Pickering 2023).

Appropriate eye measures for analyzing individual words include, among others, *first fixation duration* (FFD) and *dwell time* (Conklin et al. 2019). For the purpose of this study, we are particularly interested in fixations and the duration of fixations as these measures are more responsive to linguistic factors (Staub & Rayner 2007; Conklin et al. 2019) as opposed to saccades. The reading measures employed in this study are commonly used reading measures related to language processing. As demonstrated by Holmqvist et al. (2011), there is a great variety in the range of terminology used by researchers for the different kinds of measures. To avoid terminological confusion, all relevant measures are defined here. For the target reading region, the following reading time measures are considered: (i) the FFD, (ii) the *first run dwell time* (FRD), (iii) total number of fixations (*total fixation count*), and (iv) *total regression count*. Both *fixations* and *dwell time* are position measures. In the period spent on fixating an item, the foveal vision is focused on a small area, so that the position of the fovea is measured by the eye tracker. Thus, fixations are the time spent between saccades. The dwell time, in contrast, is defined as the time spent within a certain *interest area* (IA) which includes not only the first fixation on a word but also all subsequent fixations including returns to that area of interest. Consequently, the position measure for dwells is the respective area of interest. While the FRD sums up the duration spent in one IA before entering either to the right or left, the *dwell time* takes into consideration the whole duration spent in one IA, including re-fixations.

Different position measures for fixations, i.e., first fixation and later fixations, have been chosen for this study because the different types of fixations can reflect different cognitive processes. It is expected that a combination of investigating both early and late measures is necessary to gain a comprehensive understanding of how quotes influence reading. Generally speaking, measurements of fixations are classified into either early or late measures, reflecting different stages of reading processing. As the name indicates, early measures focus on the initial stages of processing, and "are seen primarily as a reflection of highly automatic word recognition and lexical access processes while later measures tend to reflect more conscious, controlled, strategic processes" (Conklin et al. 2019: 66). While FFD and FRD clearly reflect early measures, regressions can be regarded as a late measure (Rayner 2009). For the purpose of this study, early measures can provide insight into how quotes capture the readers' initial attention while late measures give insight into high-level cognitive processes. More specifically, the early measurement FFD, which describes only the first fixation on a visual target, activates the lexical process. The dwell time measurement, in turn, has been employed as a late measurement (Carter & Luke

2020; Siyanova-Chanturia & Elgort 2023) and has been discussed with respect to a higher informativeness and difficulties regarding the extraction of word information (Holmqvist et al. 2011).

Crucial for our investigation is the fact that different types of fixations correlate with varying cognitive processes. Lexical activation occurs with the first fixation on a word (Holmqvist et al. 2011), while later fixations may involve discourse integration. Research also suggests that "the longer the fixation, the 'deeper' the processing" (Holmqvistet al. 2011: 382). This study adopts a cognitive-linguistic perspective, positing that the duration of word fixation is influenced by contextual factors and word properties, reflecting the time needed to access and incorporate meaning into the text (Rayner & Liversedge 2012: 752).

4.2 Research question and hypotheses

To investigate the effect of quotes in NMCs, we created minimal pairs in German following the sentential pattern illustrated in (5). We expect a difference to arise in eye movements and fixation durations between the conditions *Quotes* (5a) and *noQuotes* (5b) in NMC sentences.

(5) a. *Ella erfährt, dass man dieses Gefäß „Vase" nennt, und erzählt Niko davon.* [*Quotes*]
'Ella realizes that this jar is called "vase" and tells Niko about it.'

b. *Ella erfährt, dass man dieses Gefäß Vase nennt, und erzählt Niko davon.* [*noQuotes*]
'Ella realizes that this jar is called vase and tells Niko about it.'

As we have explained above, the processing system is sensitive to punctuation marks like commas. Consistent with these findings, we hypothesize (HA) the processing system to be sensitive to the processing of quotes, reflected in extended fixation durations on the word embedded in quotes as well as a higher number of fixations on the quoted item. Based on the pragmatic approach to quotes as outlined in Schlechtweg & Härtl (2023), a difference is expected to manifest between early and late eye measurements (HB). We hypothesize that the FFD should not be sensitive to this effect as the first fixation is commonly associated with lexical access and quotes should not have an impact in this early measurement. However, we expect a difference to be manifested in later measures like the dwell time of the target IA. The last hypothesis is derived from regressive saccades argued to involve higher order linguistic information and may be a result of a disruptive processing (Rayner & Liversedge 2012). We therefore hypothesize a higher amount of regressions to quoted items (HC). The hypotheses are stated out below.

HA: More fixations and a longer fixation duration in the target IA
HB: A significant effect in the dwell time with longer duration on quoted item
HC: Higher amount of saccadic regressions from n+1 to target word n

To test for these hypotheses, we conducted the eye-tracking experiment presented below.

4.3 Participants

Twenty-four native speakers of German (20 female, 4 male) participated in the eye-tracking experiment. Subjects were university students with an academic background and were paid €8 for their participation. All participants reported to have normal or corrected-to-normal vision. The mean age of the participants was 25.12 years (SD = 4.93 years). Thirteen people had eye dominance on the right eye and 10 on the left eye. A consent form was signed by all participants prior to commencing the study. The experiment was approved by the Research Ethics Committee of the University of Kassel.

4.4 Apparatus

All eye movements were collected using the SR-Research EyeLink 1000 Plus (SR Research Ltd., Toronto, Ontario, Canada). The eye tracker was running at 500 Hz sampling rate. Viewing for the participants was binocular but only the dominant eye was tracked.

For the eye-tracking setup, we used a desktop mount combined with a chinrest and forehead rest. A 35 mm lens was used, and the camera and illuminator were placed at a distance of approx. 57 cm between the eyes and the monitor as recommended by SR Research (User Manual 2009: 52). The monocular tracking mode was used based on the participant's dominant eye, also referred to as "ocular dominance". As Hooge et al. (2019) argue, tracking one eye should in some cases be preferred over binocular tracking. Prior to the experiment, the *Porta test* (also known as the *Porta dot test*) was implemented to test for eye dominance (Bossi et al. 2018). For this test, the participants were instructed to extend their arms to form a rectangular shape with their hands through which they were able to see a dot glued to the wall the lab. When opening and closing the eyes in succession, the dominant eye remains aligned with the dot.[3]

SR Research report a high accuracy of approximately 0.5 degrees of visual angle for the EyeLink 1000 tracker, which relates to the recorded gaze position and the actual position of the eye.[4] A high accuracy of the eye tracker is especially important for reading studies because of the small stimuli, i.e., the sequence of letters presented to the participants. Regarding precision, EyeLink 1000 reports the value of about 0.01 degrees of visual angle, which describes the consistency of the measurement. This precision has been argued to be reliable at detecting small changes in the participant's gaze position.

[3] Please note that closing the eyes independently from each other caused difficulties for some participants. Therefore, in cases of doubt or unclear eye dominance, the right eye was tracked as the vast majority of people have ocular dominance on the right eye, see, e.g., Roth et al. (2002), who report 65–70 % of right eye dominance.
[4] https://www.sr-research.com/support/thread-214.html, 27 May 2025

The experimental items were presented in black letters 20-pt Times New Roman font against a white background with single line spacing for the target items. The text resource was horizontally aligned to the left and vertically aligned to the top. The sampling rate for each trial was set to 500 Hz. The EyeLink 1000 Plus Host computer was used, which has vertical refresh rates of at least 100 Hz with a microsaccade resolution of 0.05° for a head supported mode.

4.5 Materials and experimental design

The software used for creating and presenting the stimuli was *Experiment Builder* (SR Research, 2023). The experimental items consisted of 80 target items out of which 60 were fillers. After randomly presenting 40 items, there was an approx. 5-minute break followed by a recalibration of the eye tracker.[5] The stimuli consisted of a sentence, either a target sentence or filler item, followed by a comprehension question. The comprehension question was implemented to ensure that participants paid attention throughout the experiment.

All experimental items were divided into four groups and the design was a between-subjects design. The position of the target stimuli was not varied across the target sentences. The order of the items was randomized for each participant. The experimental items were NMCs in German. The target and filler items were created for the purpose of the study. The target items consisted of 20 minimal pairs that were manipulated by varying the presence or absence of quotes around the nominal in NMCs. Therefore, the factor we manipulated is QUOTES, with the level *Quotes* and *noQuotes*. All participants were exposed to a total of 20 critical items. As a consequence, participants read ten target items that contained no quotes, i.e., condition *noQuotes*, and ten target items with quotes, condition *Quotes*. Crucially, none of the filler items contained quotes. The comprehension of the sentence was controlled for by asking a yes-no comprehension question after each sentence, which required a button press for confirmation.

All target nominals were controlled for lexical frequency using the corpus of the German language provided by the Leipzig Corpora Collection (LCC)[6] so that the frequency classes ranged between 11 and 22,[7] here considered as a low lexical frequency[8] (see Table 7 in Appendix A for

[5] The recalibration was used as the participants were encouraged to move around in the lab and get some refreshments to avoid fatigue effects. Further, recalibrations have the advantage of avoiding "drifts", i.e., a deteriorating accuracy of the eye tracker (see Attardo & Pickering 2023).
[6] https://www.wortschatz.uni-leipzig.de/de, 3 June 2025
[7] The frequency in this corpus is clustered into frequency classes from 0 with the highest frequency and 24 for words with the lowest frequency. Previous studies (see, e.g., Raue 2024) have clustered words into high and low frequencies based on the frequency classes provided by this corpus which has been taken for our classification of frequencies, meaning that all items considered in this analysis belong to the category of low frequency items.
[8] We would like to thank an anonymous reviewer for pointing out that in addition to name-informing readings, a scare-quotational interpretation, i.e., a non-literal reading, of the quoted nominal may also arise, which are frequently

an overview of the target nominals with their word length and lexical frequency). Regarding word length, the number of syllables ranged between 2–4 syllables per word.[9] The average WORD LENGTH was 8.35 letters (SD = 2.39). The mean lexical frequency ranged at 15.55 (SD = 2.85), and the mean number of syllables was 2.6 (SD = 0.58). Target nominals were grouped into the following categories: *short* (4–6 letters), *medium* (7–9 letters), and *long* (10–13 letters).[10]

The location of the target nominal was consistent across all critical items and placed sentence-medially as illustrated in the following minimal pair, with the *Quotes* condition shown in (6a) and the *noQuotes* condition in (6b).

(6) a. *Kim weiß, dass man dieses Verfahren „Neuwahl" nennt, und belehrt Anna darüber.*
'Kim knows that this procedure is called "re-election" and informs Anna about this.'

b. *Kim weiß, dass man dieses Verfahren Neuwahl nennt, und belehrt Anna darüber.*
'Kim knows that this procedure is called re-election and informs Anna about this.'

In order to investigate how the reading of NMCs is modulated by the factor QUOTES, we implemented an eye-tracking paradigm. Eye-tracking during reading has been proven to be a useful method for investigating both sentence-level and word-level phenomena.

4.6 Procedure

The experiment lasted between 30 to 45 minutes.[11] Prior to the experiment, participants were asked to read and sign consent form. All participants wearing glasses cleaned their glasses in advance to avoid tracking issues. The blinds of the windows in the room were closed to ensure that no sunlight would interfere with the infra-red light of the eye tracker and the lights were turned off at least one minute before the experiment started. Then, it was ensured that the subjects were comfortably seated in front of a computer. In order to minimize head movements, the experiment was conducted using solely the so-called *head-fixed mode*, which is characterized by using a chinrest and forehead rest. Prior to the experiment, it was ensured that there were no distractions in the peripheral field. All phones were silenced beforehand, and it was

combined with a high lexical frequency. Despite this, we do not expect a scare-quotational reading to arise as the contextual embedding consistently allows for a literal interpretation.

[9] Please note that instead of using syllables, other researchers might have used (average) word length based on letters of the respective word(s), see Eskenazi (2024). This calculation is based on the word frequency effect reported in Rayner et al. (2011), in which target words were clustered into 3 different types of length (short: 4–5 letters, medium: 7–9 letters, and long: 10–12 letters). Word length has been reported as influencing the fixation duration and word skipping probability (Rayner et al. 2011).

[10] The categories are used following Rayner et al. (2011) with the only exception that long words ranged between 10–12 letters in their study.

[11] Please note that the time is highly individual and depends on (i) the participant setup prior to the experiment, (ii) the duration of the calibration and validation process, and (iii) the length of the mid-experiment break.

ensured that the participants were feeling comfortable and not distracted before starting the experiment. Participants were encouraged to read the sentences at their normal speed, and they were told to read all sentences silently.

The experiment started with a welcome slide. After pressing the enter key, participants read a slide with instructions concerning the experiment, in which they were informed that the study's aim was to investigate the understanding of sentences (see Table 8, Appendix B). The experiment started with questions regarding their gender, age, handedness, and dominant eye to let them settle in and feel relaxed during the experiment. Afterwards, a 9-point-calibration, which is recommended for a head-fixed mode, was performed followed by a 9-point-validation. In case that the calibration was asymmetrical, i.e., a poor calibration, we either used a manual calibration instead of an automatic one or the calibration was re-done to ensure data quality. The validation was only accepted once it was marked as a "good" validation, which means that the worst point error was < 1.5° with an average error of < 1.0°.[12] Another calibration and validation procedure were used after participants read half of the items as multiple calibrations are necessary to ensure data quality.

In the next step, participants continued with a practice trial which consisted of 4 sentences including feedback on the accuracy of their answer to the comprehension question. The first four sentences were presented to familiarize the participants with the procedure and were not related to the target items. Participants were allowed to ask questions, if necessary. After presenting another slide announcing the start of the experiment and reminding them that no questions could be answered in between, the experiment started with a key press.

A trial always started with a fixation cross on the screen, which was aligned to the beginning of the sentence and presented for 2 seconds. Implementing a fixation cross is especially important for short trials in reading experiments as these may have an impact on the average FFD, and the last fixation of the previously presented trial can spill over to the first fixation in the following trial (see Holmqvist et al. 2011).

In order to analyze the data, IAs were created around each word of the target sentence, as displayed in (7).

(7) Sample sentence with IAs marked with squares

Tom	*erfährt,*	*dass*	*man*	*diese*	*Krankheit*	*"Sepsis"*	*nennt,*	*und*	*erzählt*	*Anna*	*davon.*
Tom	learns	that	one	this	disease	"sepsis"	calls	and	tells	Anna	about it

'Tom learns that this disease is called "sepsis", and tells Anna about it.'

To create IAs, the default parser was used in *Experiment Builder*. A total of 12 IAs were created for each target sentence. The IA templates with the segmentation information were constant

[12] https://www.sr-research.com/support/showthread.php?tid=244, 14 August 2023

for each condition. Importantly, the variable of interest, i.e., the target nominal, was placed sentence-medially in order to be able to take delayed effects of the processing of quotes into consideration (see also Schlechtweg & Härtl 2023) and to avoid the wrap-up effects (see Hirotani et al. 2006; Conklin et al. 2019).

4.7 Data cleaning

Before moving on to reporting the results of the eye-tracking experiment, the data processing will be described, using data cleaning methods and details regarding outlier removal. *Data Viewer* (SR Research, 2023) was used for visual inspection and pre-processing of the data as a preparation for the data analysis. For the analysis of the critical items, i.e., the conditions *Quotes* and *noQuotes*, all trials with a wrong answer to the comprehension question were removed. Filtering for correct answers only resulted in a total of 18 incorrect answers out of the 480 items (3.75 % of data loss). In the next step of the analysis, the 462 data points were visually inspected trial-by-trial and instances with no or only few fixations and saccades being removed.[13] A total of 27 answers were removed, out of which 19 belonged to a single recording session. As a consequence, a total of 435 observations are left for the statistical analysis.

After applying templates across all conditions, fixation adjustments were made using the drift correction in *Data Viewer* where possible.[14] Vertically misaligned fixations are in principle correctable for single-line texts with minor adjustments on the y-axis, allowing for taking into consideration more trials in the analysis.[15] A total of 26 tokens showed vertically misaligned fixations out of which 20 were corrected and 6 discarded for further analysis because of severe vertical drift that could not be corrected in *Data Viewer*.

Data cleaning methods for fixations in reading research usually include fixation merging methods, temporal cut-offs, and outlier removal. The data files were analyzed following the most commonly used data cleaning method, which has been described as the baseline approach by Eskenazi (2024). This approach is characterized by using no fixation merging techniques, temporal cut-offs or outlier treatments.

In order to reduce the impact of idiosyncratic factors and explain the participants'[16] and items' effect and contribution to the variance, a multilevel modelling was required as the statistical approach. For each eye-tracking measure reported below, separate models were built. The eye-

[13] When the inspector window in Data Viewer did not show any fixations or only two fixations per sentence, known as track loss, the data was removed.
[14] Manually correcting fixations in reading experiments is a common practice (see Cohen 2013) but should never be done horizontally.
[15] A total of 26 individual trials showed misaligned fixations, which would have resulted in about 6.42 % of data loss.
[16] Idiosyncrasy is reflected in the variety of average fixation duration across participants (Holmqvist & Nyström 2011).

tracking measures constituted the dependent variable, and the factor QUOTES with the two levels *Quotes* & *noQuotes* as well as the factor WORDLENGTH were the independent variables. Participant and items were entered as random intercepts and/or slopes to the model.

5 Results

Overall, the mean count of fixations for the target IA was 2.35 for the *noQuotes* condition and 2.78 for the *Quotes* condition, revealing a slightly higher count for the quoted instances. The distribution of the duration of fixations for *Quotes* and *noQuotes* is presented in Figure 1.

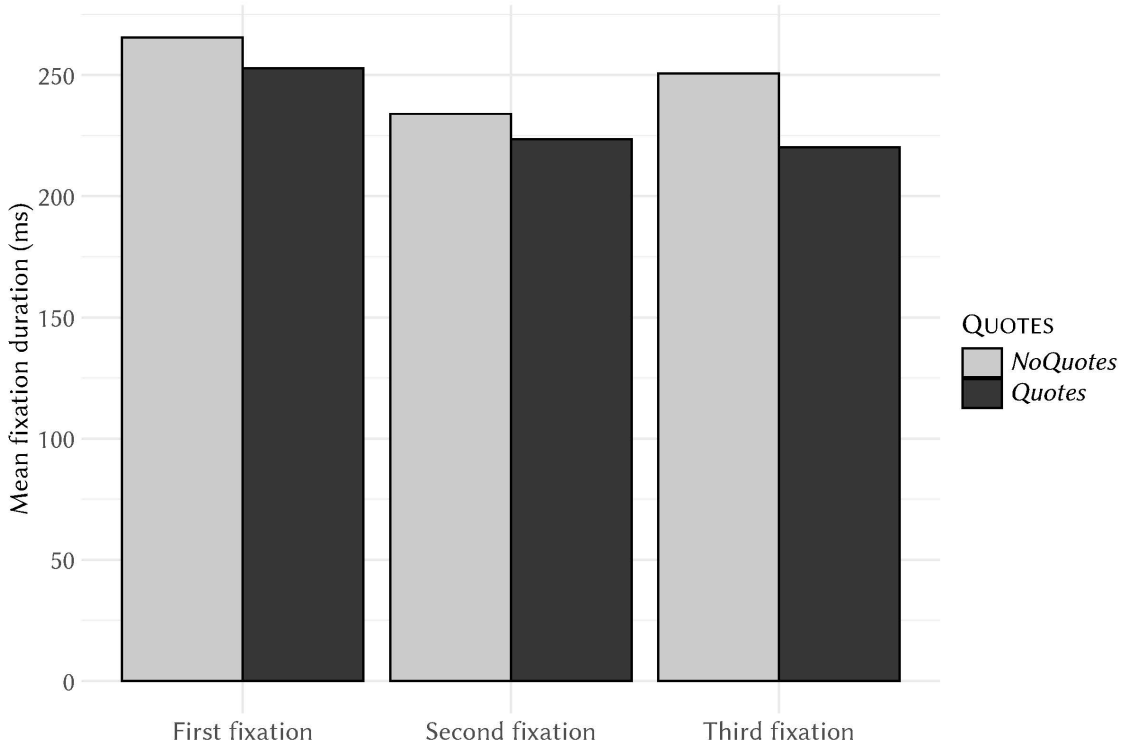

Figure 1: Mean fixation duration on the target item

In addition, regressions from a higher IA into the target IA, i.e., regressions from right to left, were analyzed. For the target sentences, these regressions were performed from the word *nennt* 'calls' back to the target nominal. The regression-in count into the target IA showed that the maximum number of regressions observed was three. Notably, all words receiving a third regression were enclosed in quotes. In 152 cases, the target IA received a regression, while 241 received no regression. When considering solely the instances with regressions, there were 88 regressions for the *Quotes* and 64 regressions for the *noQuotes* conditions. The analysis showed

a statistically significant effect between the two conditions ($\chi^2(1) = 8.6879$, $p = .003$) with more regressions to nominals enclosed in quotes. This indicates that conditions with quotes lead to a higher amount of regressions compared to conditions where nominals are presented without quotes.

For the statistical analysis of the fixation durations, i.e., the dependent variables FFD, FRD, and *IA dwell time*, linear mixed-effects models (LMEM) were implemented using the *lmer* function of the *lme4* package (Bates et al. 2015) in the statistics software *R* (R Core Team 2024). For each of the continuous variables, we used nested models to obtain the best fitting. The variables QUOTES and WORDLENGTH were entered in the model as fixed effects. For subjects and items, random effects and random slopes were also entered into the model. Models were selected based on the lowest AIC value (Akaike 1974), which indicates a good balance between model fit and complexity (see Venables & Ripley 2002). The best fitting model was determined following Matuschek et al.'s (2017) approach, whereby optimal random effect structure was included only when they enhanced the model's goodness of fit. If the model did not converge, we followed Barr et al. (2013) by systematically eliminating random slopes by participant or item that accounted for the least. Only the models with the best fit were selected and are reported here. In the following, the eye movement measures are reported in the order of the time of processing during reading, starting with the descriptive statistics and the model results.

First fixation duration: For the early measure of the FFD, there was a skipping rate of 2.76 % for the target item (see column *Missing values* in Table 1): six skips for short words, three skips for words with medium length, and three skips for long words. Crucially, all skips happened exclusively in the *noQuotes* condition.

Table 1: Descriptive statistics for first fixation duration

Condition	N	Mean	Variance	SD	Missing values
Quotes	213	252.8592	7679.829	87.63463	0
noQuotes	222	265.5429	8303.379	91.12288	12

All missing values were excluded prior to the statistical analysis. By definition, skipped items have a first fixation of 0 ms which would be problematic if included in further analyses (see Conklin et al. 2019). The mean duration of the FFD measure obtained for each corresponding word length is illustrated in Figure 2. The inferential statistics with the best fitting LMEM and its corresponding formula are given in Table 2.

Even though descriptively shorter FFDs were obtained for target items with quotes, the result of the models show a non-significant difference between the conditions *Quotes* and *noQuotes*. The significant effects in WORDLENGTH in Table 2 reveal that FFDs were significantly longer for *medium* and *long* words compared to short words. In order to obtain the difference between

the levels *medium* and *long*, we relevel this factor and set the level *long* as the baseline. This model shows a non-significant difference between WORDLENGTH *medium* and *long* ($t = -0.22$; $p = .828$). No interaction was obtained between the experimental factors.

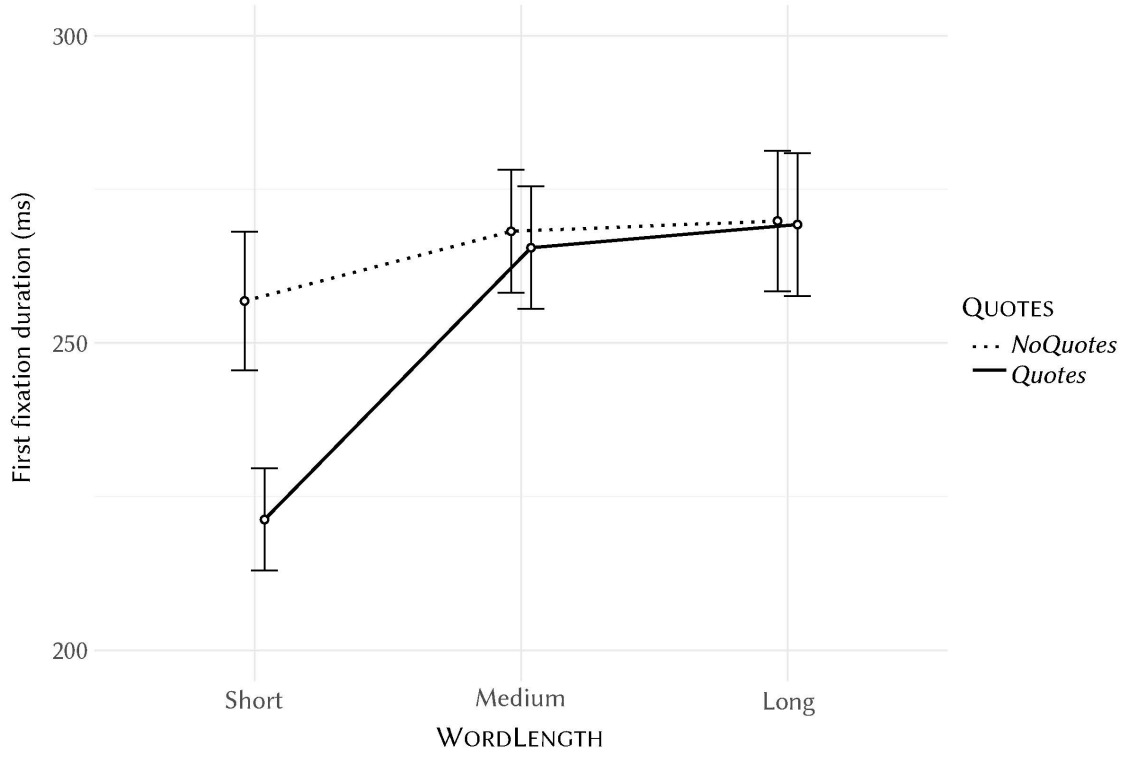

Figure 2: Mean first fixation durations and standard error per condition

Table 2: Summary of statistical analysis for FFD[a],
Formula: lmer(FFD ~ 1+ QUOTES + WORDLENGTH + (1 | Subject) + (1 | Item), REML = "FALSE")

	Estimate	Std. error	df	*t* value	Pr(>|*t*|)
(Intercept)	243.344	12.346	37.580	19.710	<2e−16***
QUOTES*Quotes*	−10.977	8.009	385.015	−1.371	0.1713
WORDLENGTH*medium*	29.709	13.668	20.007	2.174	0.0419*
WORDLENGTH*long*	32.600	13.743	20.416	2.372	0.0276*

[a] The significance levels reported here are the following: p < 0.05 = *; p < 0.01 = **; p < 0.001 = ***.

First run dwell time: The descriptive statistics for the conditions of the FRD in the target IA are given in Table 3 and the mean values per condition for FRD are plotted in Figure 3.

Table 3: Descriptive statistics for first run dwell time

Condition	N	Mean	Variance	SD	Missing values
Quotes	213	377.6761	40399.84	200.9971	0
noQuotes	222	393.5524	77405.22	278.2179	12

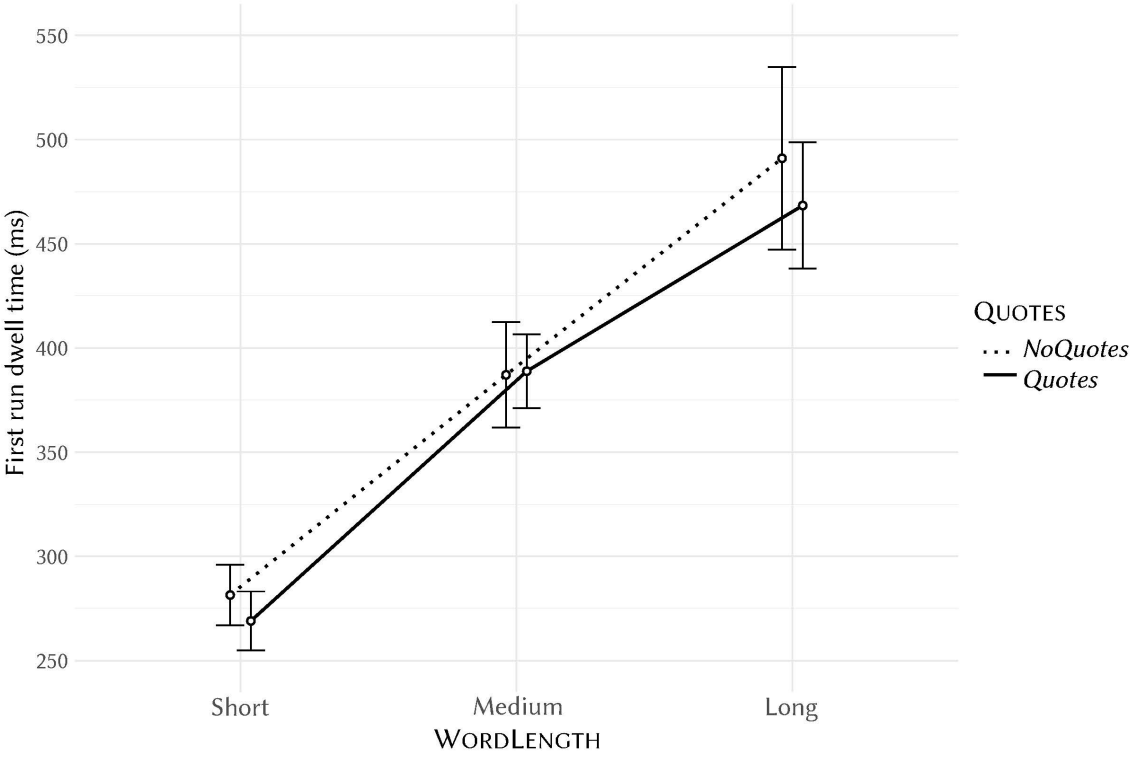

Figure 3: Mean first run dwell times and standard error per condition

The results of the LMEM with the best fit for the variable FRD are presented in Table 4.

These results reveal a significant effect for WordLength between *long* and *short* words and a marginal effect between *medium* and *short* words. To compare the levels *medium* and *long*, we relevel the factor, setting *long* as the baseline. The model results indicate no significant difference between *medium* and *long* words ($t = -1.377$; $p = 0.184$). The *noQuotes* condition receives longer mean fixation times, as can be seen in Figure 3; however, for the factor Quotes, the LMEM did not reveal any significant differences in the duration of the FRD in the target IA. Lastly, no interaction was obtained between the experimental factors.

Table 4: Summary of statistical analysis for FRD[a],
Formula: lmer(FRD ~ 1 + QUOTES + WORDLENGTH + (1 | Subject) + (1 + | Item), REML = "FALSE")

	Estimate	Std. error	df	*t* value	Pr(>|*t*|)
(Intercept)	274.82	51.29	26.45	5.358	1.24e−05***
QUOTES*Quotes*	−5.28	17.92	382.65	−0.295	0.76848
WORDLENGTH*medium*	118.75	64.52	19.75	1.840	0.08079
WORDLENGTH*long*	204.20	64.61	19.85	3.161	0.00495**

[a] The significance levels reported here are the following: p < 0.05 = *; p < 0.01 = **; p < 0.001 = ***.

Dwell time: The descriptive statistics for the IA dwell time are given in Table 5 and the mean dwell times are presented in Figure 4.

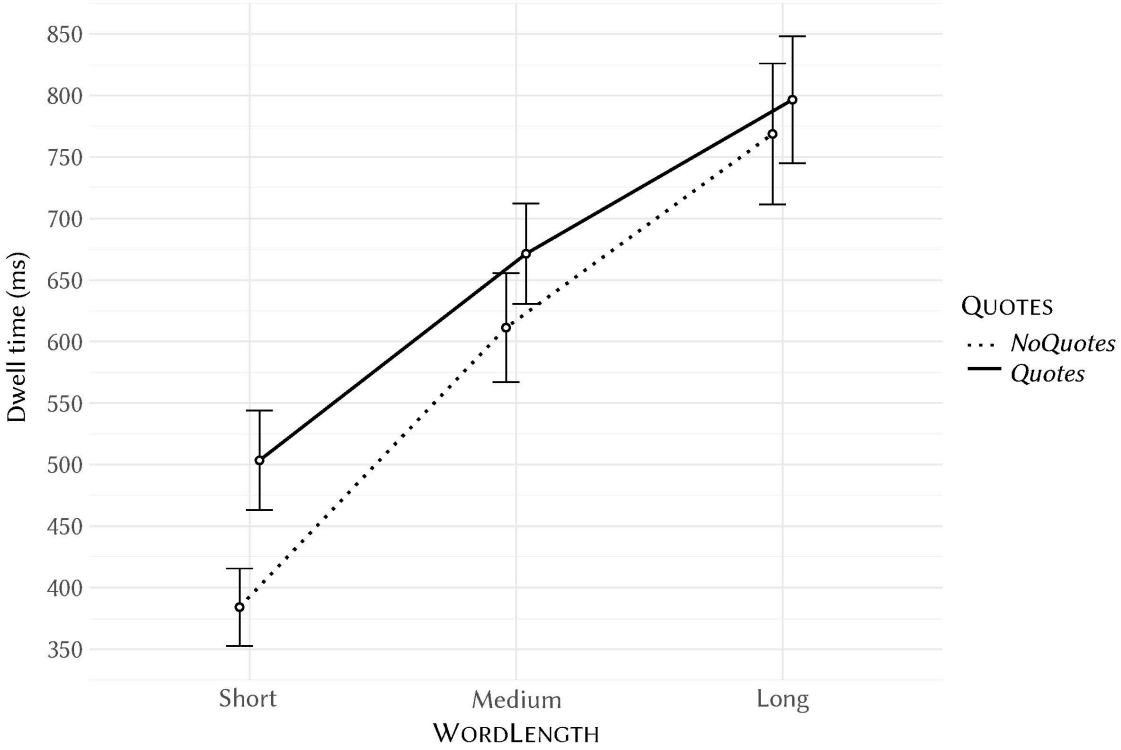

Figure 4: Mean dwell times and standard error per condition

Table 5: Descriptive statistics for dwell time at the IA

Condition	N	Mean	Variance	SD

Quotes	213	660.2676	154144.5	392.6124
noQuotes	222	598.3468	184644.5	429.7028

In contrast to the measures explored beforehand, sentences in the *Quotes* condition have a longer mean duration in the dwell time measure. The model with the best fit for the dwell time in the IA is given in Table 6.

The statistical model for the dwell time reveals a significant effect for QUOTES, whereby the *Quotes* condition received significantly longer dwell time than the *noQuotes* condition (see Figure 4). Similar to the results reported in previous models, the main effects obtained for WORDLENGTH showed a significant difference between *long* and *short* words as well as between *medium* and *short* words, whereby items with less lengthy words received shorter dwell times. To examine the difference between *medium* and *long* levels, we releveled the factor by setting *long* as the baseline. The model results show no significant difference between *medium* and *long* words ($t = -1.426$; $p = .170$). The model showed no interaction between the factors.

Table 6: Summary of statistical analysis for dwell time[a],
Formula: Dwell_Time ~ 1 + QUOTES + WORDLENGTH + (1 + | Item) + (1 + Subject)

	Estimate	Std. error	df	*t* value	Pr(>|*t*|)
(Intercept)	402.75	82.72	33.76	4.869	2.58e−05***
QUOTES*Quotes*	70.90	28.15	393.39	2.518	0.01220*
WORDLENGTH*medium*	210.79	94.12	19.20	2.239	0.03715*
WORDLENGTH*long*	340.00	94.26	19.32	3.607	0.00184**

[a] The significance levels reported here are the following: p < 0.05 = *; p < 0.01 = **; p < 0.001 = ***.

The results for the two early measurements (i.e., FFD and FRD) will be discussed together with the results obtained for the late measurement (i.e., total dwell time at the IA) in the following section.

6 Discussion

The aim of this investigation was to provide empirical evidence for reading differences between sentences with and without quotes. In the present eye-tracking study on the reading of NMCs, we have observed that in the first fixation only 12 out of 435 target IAs are not fixated and that crucially all the skipped items were of the *noQuotes* condition. The skipping rate is usually influenced by visual factors and linguistic information like length, frequency, and lexical status (Conklin et al. 2019). Although words with only 4 letters were included as experimental items,

which are frequently skipped in reading – more specifically not fixated but only included in the parafoveal view – they were never skipped when surrounded by quotes. We therefore accept HA in claiming that the target IA receives more fixations. This indicates that quotes make a substantial contribution to the understanding of the sentence.

Moreover, a significant effect materialized in the dwell time measurement. In contrast to the FFD and FRD, where the *noQuotes* condition had longer, though not statistically significant, duration, the dwell time measure reveals that quoted nominals are fixated on significantly longer than non-quoted nominals. This finding is supported by a reading-time experiment reported in Schlechtweg (2022), who found longer reading times for quoted nominals as opposed to non-quoted, yet the effect obtained in the present experiment is not statistically significant. The significantly longer fixation duration for the *Quotes* condition in the dwell time leads us to accept HB. Quotes activate processing in later stages, which becomes evident from the contrast between early and late measures. We interpret this effect as reflecting a top-down process typically associated with higher cognitive discourse-based functions (e.g., Rayner & Pollatsek 1989; Orquin & Mueller Loose 2013; Orquin & Holmqvist 2018). Top-down processes mean that the interpretation of the sentence relies on higher cognitive functions, such as discourse processing or contextual understanding. This interpretation is compatible with a pragmatic account of pure quotation (Schlechtweg & Härtl 2023). In addition, the significant difference for the lexical factor word length found in all three measurements with varying significant effects is reflected in the word length effect, which implies that longer words generally take longer to process than shorter words. While this effect is commonly attributed to several cognitive mechanisms, including visual and lexical processing demands, we argue that the familiarity of the word and its semantic complexity have further impacted the variability across measures in the data. We consider the word length effect to be independent of NMC constructions but to be present in other types of quotation as well.

The higher percentage of re-fixations on the target nominal corresponds with the increased number of regressions. This leads us to accept hypothesis HC due to the higher number of saccadic regressions from n+1 to the target word n in the *Quotes* condition as opposed to *noQuotes*. We argue that this finding is consistent with proposals within the EZ-Reader framework, based on the assumption that post-lexical integration failures can trigger regressions, given that between-word regressions, i.e., short, regressions are considered to be a consequence of higher-level language processing (Reichle et al. 2009). The obtained results suggest that the presence of quotation marks introduces additional cognitive demands during reading, disrupting the post-lexical integration process and giving rise to saccadic regressions.

7 Conclusion

In summary, the data obtained from our eye-tracking study confirm the effects associated with quotes in fixation count, dwell time, and regressions. In comparison to FFDs, regressions and

dwell time are reflective of later, i.e., post-lexical, stages of the reading process. The variance observed in dwell time implies that quotation marks introduce processing effects in the later stages of reading, indicative of higher cognitive processes. During earlier stages, no effects attributable to the presence of quotes were discerned. We take the lack of early-stage effects as evidence that participants tended to overlook the influence of quotes during their initial processing. We believe that our findings lend support for the view that quotes are represented and processed as part of the pragmatic domain of language and not as part of the compositional representation (see Section 2). The present investigation therefore aligns with previous evidence from Schlechtweg & Härtl (2020, 2023) which also supports a pragmatic approach to the study of quotation. We conclude that the quotational meanings are construed pragmatically and are processed in the discourse-integrative stages during the reading process.

To the best of our knowledge, this study represents the first of its kind in this field, providing insights into the cognitive processing of pure quotations. Future research should address the processing of a variety of quotation types, contrasting pure quotes with scare and mixed quotes, and investigate the processing of quotation (marks) from a cross-linguistic perspective.

Declaration of interest statement: The authors report there are no competing interests to declare. The data that supports the findings and the R scripts of this study are openly available in OSF at http://doi.org/10.17605/OSF.IO/2M76X.
Funding: This research was supported by funding from the University of Kassel as part of a doctoral fellowship awarded to Natascha Raue.

Abbreviations

AIC	Akaike information criterion
FFD	first fixation duration
FRD	first run dwell time
IA	interest area
LMEM	linear mixed-effects models
NMC	name-mentioning construction
SD	standard deviation

Appendix A

Table 7: Target stimuli of the eye-tracking study[a]

ID	Target nominal	Letter count	Lexical frequency
1	*Neuwahl* 're-election'	7	13
2	*Sepsis* 'sepsis'	6	15
3	*Kernfusion* 'nuclear fusion'	10	14
4	*Supermond* 'supermoon'	9	14
5	*Bürgergeld* 'citizens' funds'	10	11
6	*Ultraschall* 'ultrasound'	11	15
7	*Radio* 'radio'	5	11
8	*Stockrose* 'hollyhock'	9	20
9	*Bahnradsport* 'track cycling'	12	18
10	*Verfilzen* 'felting'	9	22
11	*Demenz* 'dementia'	6	12
12	*Blutspende* 'blood donation'	10	14
13	*Harfe* 'harp'	5	15
14	*Vase* 'vase'	4	15
15	*Polarlichter* 'northern lights'	12	16
16	*Bleichen* 'bleaching'	8	17
17	*Wandregal* 'shelf'	9	19
18	*Platzregen* 'cloudburst'	10	16
19	*Sitzsack* 'beanbag'	8	19
20	*Klippe* 'cliff'	6	15

[a] Lexical frequency calculated based on the German News Corpus 2022 (*Deutsches Nachrichten-Korpus 2022*), Wortschatz Universität Leipzig.

Appendix B

Table 8: Overview of target sentences

Condition	Experimental Items
Quotes	*Kim weiß, dass man dieses Verfahren „Neuwahl“ nennt, und belehrt Anna darüber.* ‘Kim knows that this procedure is called “re-election”, and she teaches Anna about it.’
noQuotes	*Kim weiß, dass man dieses Verfahren Neuwahl nennt, und belehrt Anna darüber.* ‘Kim knows that this procedure is called re-election, and she teaches Anna about it.’
Quotes	*Tom erfährt, dass man diese Krankheit „Sepsis“ nennt, und erzählt Anna davon.* ‘Tom learns that this disease is called “sepsis”, and he tells Anna about it.’
noQuotes	*Tom erfährt, dass man diese Krankheit Sepsis nennt, und erzählt Anna davon.* ‘Tom learns that this disease is called sepsis, and he tells Anna about it.’
Quotes	*Luisa erfährt, dass man diesen Prozess „Kernfusion“ nennt, und berichtet Luis davon.* ‘Luisa learns that this process is called “nuclear fusion”, and she tells Luis about it.’
noQuotes	*Luisa erfährt, dass man diesen Prozess Kernfusion nennt, und berichtet Luis davon.* ‘Luisa learns that this process is called nuclear fusion, and she tells Luis about it.’
Quotes	*Martin erfährt, dass man dieses Phänomen „Supermond“ nennt, und informiert Lisa darüber.* ‘Martin learns that this phenomenon is called “super moon”, and he informs Lisa about it.’
noQuotes	*Martin erfährt, dass man dieses Phänomen Supermond nennt, und informiert Lisa darüber.* ‘Martin learns that this phenomenon is called super moon, and he informs Lisa about it.’
Quotes	*Angelina lernt, dass man diese Zahlung „Bürgergeld“ nennt, und berichtet Hanno davon.* ‘Angelina learns that this payment is called “citizen’s benefit”, and she tells Hanno about it.’
noQuotes	*Angelina lernt, dass man diese Zahlung Bürgergeld nennt, und berichtet Hanno davon.* ‘Angelina learns that this payment is called citizen’s benefit, and she tells Hanno about it.’
Quotes	*Nele erfährt, dass man dieses Verfahren „Ultraschall“ nennt, und informiert Holger darüber.* ‘Nele learns that this procedure is called “ultrasound”, and she informs Holger about it.’
noQuotes	*Nele erfährt, dass man dieses Verfahren Ultraschall nennt, und informiert Holger darüber.* ‘Nele learns that this procedure is called ultrasound, and she informs Holger about it.’
Quotes	*Charlotta erfährt, dass man dieses Gerät „Radio“ nennt, und informiert Paul darüber.* ‘Charlotta learns that this device is called “radio”, and she informs Paul about it.’
noQuotes	*Charlotta erfährt, dass man dieses Gerät Radio nennt, und informiert Paul darüber.* ‘Charlotta learns that this device is called radio, and she informs Paul about it.’
Quotes	*Maria weiß, dass man diese Pflanze „Stockrose“ nennt, und erzählt Paul davon.* ‘Mary knows that this plant is called a “hollyhock”, and she tells Paul about it.’
noQuotes	*Maria weiß, dass man diese Pflanze Stockrose nennt, und erzählt Paul davon.* ‘Mary knows that this plant is called a hollyhock, and she tells Paul about it.’
Quotes	*Christoph weiß, dass man diese Sportart „Bahnradsport“ nennt, und berichtet Tina darüber.* ‘Christoph knows that this sport is called “track cycling”, and he tells Tina about it.’
noQuotes	*Christoph weiß, dass man diese Sportart Bahnradsport nennt, und berichtet Tina darüber.* ‘Christoph knows that this sport is called track cycling, and he tells Tina about it.’
Quotes	*Annika weiß, dass man diesen Prozess „Verfilzen“ nennt, und berichtet Toni davon.* ‘Annika knows that this process is called “felting”, and she tells Toni about it.’
noQuotes	*Annika weiß, dass man diesen Prozess Verfilzen nennt, und berichtet Toni davon.* ‘Annika knows that this process is called felting, and she tells Toni about it.’
Quotes	*Emanuel lernt, dass man diese Krankheit „Demenz“ nennt, und berichtet Helena davon.* ‘Emanuel learns that this disease is called “dementia”, and he tells Helena about it.’

noQuotes	*Emanuel lernt, dass man diese Krankheit Demenz nennt, und berichtet Helena davon.* 'Emanuel learns that this disease is called dementia, and he tells Helena about it.'
Quotes	*Leonie weiß, dass man diese Verfahren „Blutspende" nennt, und belehrt Manuel darüber.* 'Leonie knows that this procedure is called "blood donation", and she teaches Manuel about it.'
noQuotes	*Leonie weiß, dass man diese Verfahren Blutspende nennt, und belehrt Manuel darüber.* 'Leonie knows that this procedure is called blood donation, and she teaches Manuel about it.'
Quotes	*Marcel erfährt, dass man dieses Instrument „Harfe" nennt, und berichtet Carla davon.* 'Marcel learns that this instrument is called a "harp", and he tells Carla about it.'
noQuotes	*Marcel erfährt, dass man dieses Instrument Harfe nennt, und berichtet Carla davon.* 'Marcel learns that this instrument is called a harp, and he tells Carla about it.'
Quotes	*Ella erfährt, dass man dieses Gefäß „Vase" nennt, und erzählt Niko davon.* 'Ella learns that this piece is called a "vase", and she tells Niko about it.'
noQuotes	*Ella erfährt, dass man dieses Gefäß Vase nennt, und erzählt Niko davon.* 'Ella learns that this piece is called a vase, and she tells Niko about it.'
Quotes	*Maria weiß, dass man diese Erscheinung „Polarlichter" nennt, und erzählt Markus davon.* 'Mary knows that this phenomenon is called "northern lights", and she tells Mark about it.'
noQuotes	*Maria weiß, dass man diese Erscheinung Polarlichter nennt, und erzählt Markus davon.* 'Mary knows that this phenomenon is called northern lights, and she tells Mark about it.'
Quotes	*Oliver erfährt, dass man dieses Verfahren „Bleichen" nennt, und informiert Beate darüber.* 'Oliver learns that this procedure is called "bleaching", and he informs Beate about it.'
noQuotes	*Oliver erfährt, dass man dieses Verfahren Bleichen nennt, und informiert Beate darüber.* 'Oliver learns that this procedure is called bleaching, and he informs Beate about it.'
Quotes	*Helena weiß, dass man diese Konstruktion „Wandregal" nennt, und erzählt Kim davon.* 'Helena knows that this construction is called a "shelf", and she tells Kim about it.'
noQuotes	*Helena weiß, dass man diese Konstruktion Wandregal nennt, und erzählt Kim davon.* 'Helena knows that this construction is called a shelf, and she tells Kim about it.'
Quotes	*Anton weiß, dass man dieses Phänomen einen „Platzregen" nennt, und erzählt Miriam davon.* 'Anton knows that this phenomenon is called a "sudden downpour", and he tells Miriam about it.'
noQuotes	*Anton weiß, dass man dieses Phänomen einen Platzregen nennt, und erzählt Miriam davon.* 'Anton knows that this phenomenon is called a sudden downpour, and he tells Miriam about it.'
Quotes	*Felicia erfährt, dass man diese Sitzgelegenheit „Sitzsack" nennt, und erzählt Johannes davon.* 'Felicia learns that this seat is called a "bean bag", and she tells Johannes about it.'
noQuotes	*Felicia erfährt, dass man diese Sitzgelegenheit Sitzsack nennt, und erzählt Johannes davon.* 'Felicia learns that this seat is called a bean bag, and she tells Johannes about it.'
Quotes	*Andreas weiß, dass man diese Felsformation „Klippe" nennt, und berichtet Daniela davon.* 'Andreas knows that this rock formation is called a "cliff", and he tells Daniela about it.'
noQuotes	*Andreas weiß, dass man diese Felsformation Klippe nennt, und berichtet Daniela davon.* 'Andreas knows that this rock formation is called a cliff, and he tells Daniela about it.'

References

Akaike, Hirotugu. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6). 716–723. https://doi.org/10.1109/TAC.1974.1100705.

Attardo, Salvatore & Lucy Pickering. 2023. *Eye tracking in linguistics*. London: Bloomsbury Academic.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. https://doi.org/10.1016/j.jml.2012.11.001.

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

Bossi, Manuela, Lisa M. Hamm, Annegret Dahlmann-Noor & Steven C. Dakin. 2018. A comparison of tests for quantifying sensory eye dominance. *Vision Research* 153. 60–69. https://doi.org/10.1016/j.visres.2018.09.006.

Brendel, Elke, Jörg Meibauer & Markus Steinbach. 2011. Exploring the meaning of quotation. In Elke Brendel (ed.), *Understanding quotation*, 1–34. Berlin: De Gruyter. https://doi.org/10.1515/9783110240085.1.

Cappelen, Herman & Ernest Lepore. 1997. Varieties of quotation. *Mind* 106(423). 429–450. https://doi.org/10.1093/mind/106.423.429.

Cappelen, Herman & Ernest Lepore. 1999. Using, mentioning and quoting: A reply to Saka. *Mind* 108(432). 741–750. https://doi.org/10.1093/mind/108.432.741.

Carter, Benjamin T. & Steven G. Luke. 2020. Best practices in eye tracking research. *International Journal of Psychophysiology* 155. 49–62. https://doi.org/10.1016/j.ijpsycho.2020.05.010.

Clifton, Charles, Fernanda Ferreira, John M. Henderson, Albrecht W. Inhoff, Simon P. Liversedge, Erik D. Reichle, Elizabeth R. Schotter. 2016. Eye movements in reading and information processing: Keith Rayner's 40 year legacy. *Journal of Memory and Language* 86. 1–19. https://doi.org/10.1016/j.jml.2015.07.004.

Cohen, Andrew L. 2013. Software for the automatic correction of recorded eye fixation locations in reading experiments. *Behavior Research Methods* 45(3). 679–683. https://doi.org/10.3758/s13428-012-0280-3.

Conklin, Kathy, Ana Pellicer-Sánchez & Gareth Carrol. 2019. *Eye-tracking: A guide for applied linguistics research*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108233279.

Cortés Rodríguez, Álvaro, Holden Härtl, Natascha Raue & Kristina Weissbecker. 2022. *This text is called (an) article*: Referring nouns in name-informing quotation. *Linguistic Research* 39(2). 327–354. https://doi.org/10.17250/khisli.39.2.202206.004.

Davidson, Donald. 1979. Quotation. *Theory and Decision* 11. 27–40. https://doi.org/10.1007/BF00126690.

Ehrlich, Susan F. & Keith Rayner. 1981. Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior* 20(6). 641–655. https://doi.org/10.1016/S0022-5371(81)90220-6.

Engbert, Ralf & Reinhold Kliegl. 2011. A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research* 42(5). 621– 636. https://doi.org/10.1016/S0042-6989(01)00301-7.

Eskenazi, Michael A. 2024. Best practices for cleaning eye movement data in reading research. *Behavior Research Methods* 56. 2083–2093. https://doi.org/10.3758/s13428-023-02137-x.

Friederici, Angela D., Thomas C. Gunter, Anja Hahne & Kerstin Mauth. 2004. The relative timing of syntactic and semantic processes in sentence comprehension. *Neuroreport* 15(1). 165–169. https://doi.org/10.1097/00001756-200401190-00032.

Gutzmann, Daniel & Erik Stei. 2011. How quotation marks what people do with words. *Journal of Pragmatics* 43(10). 2650–2663. https://doi.org/10.1016/j.pragma.2011.03.010.

Härtl, Holden. 2018. Name-informing and distancing *sogenannt* 'so-called': Name mentioning and the lexicon-pragmatics interface. *Zeitschrift für Sprachwissenschaft* 37(2). 139–169. https://doi.org/10.1515/zfs-2018-0008.

Härtl, Holden. 2020. Referring nouns in name-informing quotation: A copula-based approach. *Sinn und Bedeutung* (SuB) 24(1). 291–304. https://doi.org/10.18148/sub/2020.v24i1.868.

Hill, Robin L. & Wayne S. Murray. 2000. Commas and spaces: Effects of punctuation on eye movements and sentence parsing. In: Alan Kennedy, Ralph Radach, Dieter Heller & Joël Pynte (eds.), *Reading as a perceptual process*, 565–589. Oxford: Elsevier. https://doi.org/10.1016/B978-008043642-5/50027-9.

Hirotani, Masako, Lyn Frazier & Keith Rayner. 2006. Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language* 54(3). 425–443. https://doi.org/10.1016/j.jml.2005.12.001.

Holmqvist, Kenneth, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka & Joost Van de Weijer. 2011. *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Hooge, Ignace T. C., Gijs A. Holleman, Nina C. Haukes & Roy S. Hessels. 2019. Gaze tracking accuracy in humans: One eye is sometimes better than two. *Behavior Research Methods* 51. 2712–2721. https://doi.org/10.3758/s13428-018-1135-3.

Inhoff, Albrecht W. & Keith Rayner. 1986. Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics* 40(6). 431–439. https://doi.org/10.3758/bf03208203.

Klockow, Reinhard. 1978. Anführungszeichen, Norm und Abweichung. *Linguistische Berichte* 57. 14–24.

Kuperberg, Gina R., Phillip J. Holcomb, Tatiana Sitnikova, Douglas Greve, Anders M. Dale & David Caplan. 2003. Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. *Journal of Cognitive Neuroscience* 15(2), 272–293. https://doi.org/10.1162/089892903321208204.

Maier, Emar. 2014. Pure quotation. *Philosophy Compass* 9(9). 615–630. https://doi.org/10.1111/phc3.12149.

Matuschek, Hannes, Reinhold Kliegl, Shravan Vasishth, Harald Baayen & Douglas Bates. 2017. Balancing Type I error and power in linear mixed models. *Journal of Memory and Language* 94. 305–315. https://doi.org/10.1016/j.jml.2017.01.001.

Orquin, Jacob L. & Kenneth Holmqvist. 2018. Threats to the validity of eye-movement research in psychology. *Behavior Research Methods* 50(4). 1645–1656. https://doi.org/10.3758/s13428-017-0998-z.

Orquin, Jacob L. & Simone Mueller Loose. 2013. Attention and choice: A review on eye movements in decision making. *Acta Psychologica* 144(1). 190–206. https://doi.org/10.1016/j.actpsy.2013.06.003.

Predelli, Stefano. 2003. Scare quotes and their relation to other semantic issues. *Linguistics and Philosophy* 26(1). 1–28. https://doi.org/10.1023/A:1022278209949.

Pynte, Joël & Alan Kennedy. 2007. The influence of punctuation and word class on distributed processing in normal reading. *Vision Research* 47(9), 1215–1227. https://doi.org/10.1016/j.visres.2006.12.006.

Quine, Willard Van Orman. 1981. *Mathematical logic*. Revised edn. Cambridge: Harvard University Press.

R Core Team. 2024. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/, 8 June 2025.

Raue, Natascha & Álvaro Cortés Rodríguez. 2022. The interplay between quotation and referentiality: An empirical investigation into name-mentioning constructions. Paper presented at Linguistic Evidence 11, Laboratoire de Linguistique Formelle (Université de Paris – CNRS) & Laboratoire d'Excellence « Empirical Foundations of Linguistics », 6–8 October.

Raue, Natascha. 2024. Reference to names: An empirically based study on metalinguistic uses of names of lexicalized concepts. Kassel: Universität Kassel. https://doi.org/10.17170/kobra-2024122010782.

Rayner, Keith. 1978. Eye movement latencies for parafoveally presented words. *Bulletin of the Psychonomic Society* 11(1). 13–16. https://doi.org/10.3758/bf03336753.

Rayner, Keith. 2009. The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology* 62(8). 1457–1506. https://doi.org/10.1080/17470210902816461.

Rayner, Keith & Susan A. Duffy. 1986. Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition* 14(3). 191–201. https://doi.org/10.3758/BF03197692.

Rayner, Keith & Simon P. Liversedge. 2012. Linguistic and cognitive influences on eye movements during reading. In Simon P. Liversedge, Iain Gilchrist & Stefan Everling (eds.), *The Oxford handbook of eye movements*, 752–766. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199539789.001.0001.

Rayner, Keith & Alexander Pollatsek. 1989. *The psychology of reading*. New York: Routledge.

Rayner, Keith, Timothy J. Slattery, Denis Drieghe & Simon P. Liversedge. 2011. Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human perception and performance* 37(2), 514–528. https://doi.org/10.1037/a0020990.

Regel, Stefanie, Thomas C. Gunter & Angela D. Friederici. 2011. Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience* 23(2). 277–293. https://doi.org/10.1162/jocn.2010.21411.

Reichle, Erik D., Tessa Warren & Kerry McConnell. 2009. Using E-Z Reader to model the effects of higher-level language processing on eye movements during reading. *Psychonomic Bulletin Review* 16(1), 1–21. https://doi.org/10.3758/PBR.16.1.1.

Roth, Heidi L., Andrea N. Lora & Kenneth M. Heilman. 2002. Effects of monocular viewing on eye dominance on spatial attention. *Brain: A Journal of Neurology* 125(9). 2023–2035. https://doi.org/10.1093/brain/awf210.

Saka, Paul. 1998. Quotation and the use-mention distinction. *Mind* 107(425). 113–135. https://doi.org/10.1093/mind/107.425.113.

Schlechtweg, Marcel. 2022. Quotation marks in advanced language acquisition: A reading time experiment on English pure quotation. *Linguistik online* 118(6). 113–128. https://doi.org/10.13092/lo.118.9047.

Schlechtweg, Marcel & Holden Härtl. 2020. Do we pronounce quotation? An analysis of name-informing and non-name-informing contexts. *Language and Speech* 63(4). 769–798. https://doi.org/10.1177/0023830919893393.

Schlechtweg, Marcel & Holden Härtl. 2023. Quotation marks and the processing of irony in English: Evidence from a reading time study. *Linguistics* 61(2). 355–390. https://doi.org/10.1515/ling-2021-0079.

Simchen, Ori. 1999. Quotational mixing of use and mention. *Philosophical Quarterly* 49(196), 325–336. https://doi.org/10.1111/1467-9213.00145.

Siyanova-Chanturia, Anna & Irina Elgort. 2023. Analysing reading with eye tracking. In Sandrine Zufferey & Pascal Gygax (eds.), *The Routledge handbook of experimental linguistics*, 250–267. https://doi.org/10.4324/9781003392972-19.

Staub, Arian & Keith Rayner. 2007. Eye movements and on-line comprehension processes. In M. Gareth Gaskell (ed.), *The Oxford handbook of psycholinguistics*, 327–342. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198568971.013.0019.

Stites, Mallory C., Steven G. Luke & Kiel Christianson. 2013. The psychologist said quickly, "Dialogue descriptions modulate reading speed!" *Memory & Cognition* 41(1). 137–151. https://doi.org/10.3758/s13421-012-0248-7.

Venables, William N. & Brian D. Ripley. 2002. *Modern applied statistics with S*. New York: Springer. https://doi.org/10.1007/978-0-387-21706-2.

Wade, Nicholas J. 2010. Pioneers of eye movement research. i-*Perception* 1(2). 33–68. https://doi.org/10.1068/i0389.

Washington, Corey. 1992. The identity theory of quotation. *The Journal of Philosophy* 89(11). 582–605. https://doi.org/10.2307/2941057.

Yao, Bo & Christoph Scheepers. 2011. Contextual modulation of reading rate for direct versus indirect speech quotations. *Cognition* 121(3). 447–453. https://doi.org/10.1016/j.cognition.2011.08.007.