Dennis Nikolas Pauly | Guido Nottbusch

The Influence of the German Capitalization Rules on Reading

Check for updates

The Influence of the German Capitalization Rules on Reading

Dennis Nikolas Pauly* and Guido Nottbusch

Teacher Education's Department, University of Potsdam, Potsdam, Germany

German orthography systematically marks all nouns (even other nominalized word classes) by capitalizing their first letter. It is often claimed that readers benefit from the uppercase-letter syntactic and semantic information, which makes the processing of sentences easier (e.g., Bock et al., 1985, 1989). In order to test this hypothesis, we asked 54 German readers to read single sentences systematically manipulated by a target word (N). In the experimental condition (EXP), we used semantic priming (in the following example: *sick → cold*) in order to build up a strong expectation of a noun, which was actually an attribute for the following noun (N+1) (translated to English e.g., "*The sick writer had a cold* (N) *nose* (N+1) …"). The sentences in the control condition were built analogously, but word N was purposefully altered (keeping word length and frequency constant) to make its interpretation as a noun extremely unlikely (e.g., "*The sick writer had a blue* (N) *nose* (N+1) …"). In both conditions, the sentences were presented either following German standard orthography (Cap) or in lowercase spelling (NoCap). The capitalized nouns in the EXP/Cap condition should then prevent garden-path parsing, as capital letters can be recognized parafoveally. However, in the EXP/NoCap condition, we expected a garden-path effect on word N+1 affecting first-pass fixations and the number of regressions, as the reader realizes that word N is instead an adjective. As the control condition does not include a garden-path, we expected to find (small) effects of the violation of the orthographic rule in the CON/NoCap condition, but no garden-path effect. As a global result, it can be stated that reading sentences in which nouns are not marked by a majuscule slows a native German reader down significantly, but from an absolute point of view, the effect is small. Compared with other manipulations (e.g., transpositions or substitutions), a lowercase letter still represents the correct allograph in the correct position without affecting phonology. Furthermore, most German readers do have experience with other alphabetic writing systems that lack consistent noun capitalization, and in (private) digital communication lowercase nouns are quite common. Although our garden-path sentences did not show the desired effect, we found an indication of grammatical pre-processing enabled by the majuscule in the regularly spelled sentences: In the case of high noun frequency, we *post hoc* located parafovea-on-fovea effects, i.e., longer fixation durations, on the attributive adjective (word N). These benefits of capitalization could only be detected under specific circumstances. In other cases, we conclude that longer reading durations are mainly the result of disturbance in readers' habituation when the expected capitalization is missing.

Keywords: orthography, eye-tracking, reading, noun, parafoveal and foveal processing

INTRODUCTION

To our knowledge, German is (besides Luxembourgish, which is in fact a German dialect) the only remaining language which systematically marks nouns by capitalized letters (majuscules), which also includes nominalizations of all word classes, while using the Latin alphabet. Not only are all noun classes capitalized, but also all kinds of nominalization, which includes all word classes. Besides punctuation rules, capitalization rules take up most of the space in the German standard orthography regulations called "Amtliches Regelwerk" (Rat für deutsche Rechtschreibung, 2018). Due to the fact that German capitalization rules are the most fault-prone area in German orthography from 4th grade on (Moser, 1958; Menzel, 1985; Zimmermann and Heckel, 1986; Zabel, 1992; Kluge, 1995; Röber-Siekmeyer, 1999; Granzow-Emden, 2002; Günther and Nünke, 2005; Röber, 2011), discussions are held at regular intervals regarding its necessity (e.g., Mentrup, 1979a,b).

The most relevant arguments for keeping the capitalization rules should be either that they (1) clarify ambiguities or that (2) the reader benefits to the effect of faster text processing. If neither of those were true, the only arguments left would be historical or aesthetic.

Concerning (1), Augst (1980) argued, by investigating a massive text corpus, that there would be almost no sentence ambiguity and absolutely no context ambiguity when using a purely lowercase (minuscule) spelling. Augst's final conclusion that marking a noun by a majuscule is functionless for the writer and unnecessary for the reader's text comprehension leaves open the question whether there is any other benefit for the reader at all.

For this reason, an examination of the potential benefits becomes necessary. In the mid-1980s and early 1990s, several studies were conducted: Bock et al. (1985) ran different studies (for more detailed descriptions and discussion, see: Nottbusch and Jonischkait, 2011; Pauly and Nottbusch, 2016) concerning the question of the function of majuscule and minuscule for German readers. The hypothesis was that the majuscule gives a clear sign to the noun word class, which supports semantic analysis or may even make it obsolete, so that reading is supported. The methodology for validating the theory was a systematic manipulation of the beginning of words in all word classes. Subjects were supposed to read 20 short texts with five spelling varieties: (1) German capitalization rules, (2) English standard orthography, (3) all words beginning capitalized, (4) a random mixture of lower- and uppercase letters, and (5) inversion of the German capitalization rules. A total of 80 subjects read 20 texts (~125 words, 30% nouns), four in sequence from one of the varieties described above. The only significant result was the difference in reading duration between the varieties. The more the spelling varied from the regularities the slower the reading duration was, which led to the following ranking: German capitalization rules [33.1 s mean reading time, 229.6 words per minute (wpm)] < English standard orthography (34.6 s, 219.7 wpm) < All words capitalized (36.8 s, 206.4 wpm) = Random mixture (37.2 s, 204.6 wpm) < Inversion of capitalization rules (39.4 s, 192.8 wpm). The result is not astonishing and allows no conclusion regarding any benefit of the German capitalization rules, because it only confirms that reading speed decreases the more a spelling varies from the standard orthography. In Bock et al. (1989), 100 German students of English and 100 Dutch students of German took part. The participants read 20 German texts (fables, mean length: 247 words), 10 in German (all participants) and 10 in English (German participants) or 10 in Dutch (Dutch participants). The texts were presented in the five spelling varieties described above. The results of Bock et al. (1985) were replicated for the Germans reading German texts. However, when German students read English texts there was no significant difference between the texts following German standard orthography (which is irregular here) and the regular English standard orthography (212 wpm each). Bock concluded that this can be explained by two contrary effects on reading: The violation of the English orthography produces a cost, but the German capitalization rules lead to a benefit due to syntactic hints, which in sum leads to a zero difference. He further concluded that the German Capitalization System would be transferable to other languages. The Dutch participants read the German texts with the same ranking as the Germans did: The Dutch texts were read fastest when following the standard orthography (like in English, very fast: 361 wpm), the German capitalization rules ranked second (348 wpm), but led to less deceleration than the three other spellings (Bock et al., 1989, p. 48–49). In sum, for both groups it can be stated that the more spelling varies from the norm, the longer the reading time.

The first eye-tracking study so far which systematically examined the function of the German Capitalization System was conducted by Gfroerer et al. (1989), who replicated the experiment just reported: 15 Dutch participants read the same texts as in Bock et al. (1989) in Dutch and German in the five spelling modes (50 texts total, 10 per spelling mode). The authors found the Dutch texts to be read faster than the German texts in all spelling modes. However, one result is very interesting: Not only the German (182 wpm) but also the Dutch texts (311 wpm vs. 286 wpm when following standard orthography) were read fastest by the Dutch readers when they followed the German capitalization rules. Therefore, the authors concluded that the German Capitalization System is transferable to other writing systems. The results are astonishing, but should be treated carefully, because it was a pilot study and because of technical problems the authors reported.

While these studies report interesting findings by making global conclusions regarding the function of one orthographic rule, we wanted to validate these results by using the eye-tracking technique, which has become even more precise over the years.

Our overall task was to find an answer to the question of whether the German Capitalization System serves a function for the reader, i.e., whether the additional syntactic information helps the reader parse the sentence. To clarify our hypothesis, we first describe here the chronological processing of foveal word identification: As a first step, visual information is obtained before the orthography (letter identity and word length), phonology (sounds), and morphology (units of meaning, grammatical gender, etc.) of the word can be analyzed. It may be noted that the orthographic analysis takes place at a very

early stage of processing. Afterwards, the lexical representation (the abstract representation of the word form) is accessed. In the end, the semantic (word meaning) and syntactic (grammatical role) representations of the word are accessed and integrated into the meaning of the sentence (cf. Schotter et al., 2012, p. 7f). Although the syntactic processing is one of the last steps in processing, it could be argued that in the case of German nouns it somehow occurs together with orthographic processing, due to the fact that a reader implicitly knows that a noun is systematically marked by a majuscule. A similar effect could also be assumed in a more parallel and interactive model, in which orthographic information triggers syntactic processes.

In the very beginning of word identification, at least parafoveally, a majuscule can be identified. Readers of a left-to-right alphabetic writing system obtain useful information from a region extending from 3 to 4 letters to the left of the fixation to about 14–15 character spaces to the right of the fixation, although information used for word identification is obtained from an even smaller region, 7–8 letters to the right of the fixation (Rayner and Castelhano, 2008). A majuscule, however, can be recognized even further away. After having identified and processed this capital letter, a competent speaker and reader of German should possess an implicit knowledge about the syntactical function a noun serves in a German sentence, which should speed up processing. If capitalization is omitted, not only is the orthography violated, but the syntactical information is also missing. The capitalization of nouns makes parafovea-on-fovea effects (PoF) more likely, while these effects are controversial. There is a "great deal of evidence that PoF effects are driven by orthography, and not lexical status, syntactic role, or meaning" (Rayner and Schotter, 2014, p. 24). However, this could be different for noun capitalization in German, which is easy to identify parafoveally and which exclusively marks a word class.

Therefore, we chose a study design that should pinpoint the effect. The idea was to create an experimental context which ascribes differences in the measured parameters to the fact that nouns are being systematically marked, i.e., via the German Capitalization System, which makes it possible to eliminate the habit factor. That means that readers process sentences following the German standard orthography faster compared to those in which the orthography is violated not just because of habituation.

EXPERIMENT

In our study design, we systematically manipulated single sentences, which were presented in regular German spelling (Cap) and in lowercase spelling (like in English orthography; NoCap). Our hypotheses were the following:

- The German Capitalization System does serve a function for the reader, facilitating sentence processing through the systematic recognition of heads of noun phrases.
- The reader parafoveally recognizes the majuscule and is therefore able to systematically pre-process the following noun because a syntactic category becomes activated as well.

We created garden-path sentences containing ambiguous adjectives (word N) in the second noun phrase. By means of semantic priming, which was provoked by the first noun phrase (e.g., The sick writer had a cold (N) nose (N+1), cf. **Table 1**), the adjective was manipulated in order to be interpreted as a noun (experimental condition, EXP). Actually, this adjective was acting as an attribute for an upcoming noun (word N+1; cf. **Table 1**), so that reparsing of the sentence should become necessary on N+1. The garden-path effect and thus reparsing of the sentence read so far, however, should only become necessary in the NoCap condition because, if the sentences followed the standard orthography, the reader would not misinterpret an adjective as a noun because the nouns were capitalized and the adjective was not. Orthographic familiarity should not be a problem, i.e., the fact that capitalization or non-capitalization is linked strongly to the word, because in German the so-called "morphematisches Prinzip" [morphemic principle] is very dominant in the orthography, which means that words which consist of the same morpheme are written identically or at least similarly. This principle was even strengthened in the last orthography reform (1996–2006). German readers are, since nominalization is very common in German, used to seeing the same morphemes capitalized and non-capitalized.

We also created control sentences (CON) in which word N was replaced by a non-ambiguous adjective that could not be misinterpreted as a noun. CON was also presented in both conditions (Cap and NoCap).

We expected the following effects (on N and N+1):

- EXP in Cap: Because of the capitalized nouns (N+1), processing should be possible without any confusion; parafoveal processing of the noun while fixating on the adjective should lead to shorter a fixation duration on the noun.
- EXP in NoCap: Again, due to violation of orthography, the fixation duration should be longer in general. Besides that, reparsing of the sentence due to the garden-path effect, which leads to a distinctly longer fixation duration on the noun in first-pass reading and regression to the adjective afterwards.
- CON in Cap: If possible, even less confusion; otherwise, see EXP in Cap.
- CON in NoCap: Again, due to orthography violation, the fixation duration on the noun should be longer compared to CON in Cap and shorter on the adjective due to a lack of parafoveal pre-processing.

Method
Material
A total of 40 experimental sentences (EXP) and 40 control sentences (CON) were constructed. CON was identical to EXP besides the manipulated position "N." To find matching ambiguous items for position "N," we searched through the CELEX Database (Baayen et al., 1995): We found 1,032 examples including all flections; only 68 noun-adjective pairs were left after sighting of the examples, e.g., deletion of uncommon, incorrect or inflected word forms, derivations, etc. The inflected word forms in particular reduced the number to at least a third.

TABLE 1 | Example sentence of EXP and CON in all conditions (differences between the modes are printed in bold letters).

Sentence mode	Spelling				N−1	N (adjective)	N+1 (noun)				Garden-path
		Adpositional phrase		Verb		Noun phrase		Adpositional phrase			
EXP	CAP	In	der	**G**ymnastikhalle	liegen	die	**matten**	**S**portler	auf	dem	**B**oden.	**No**
	NoCAP			**g**ymnastikhalle			**matten**	**s**portler			**b**oden.	**Yes**
CON	CAP			**G**ymnastikhalle			**steten**	**S**portler			**B**oden.	**No**
	NoCAP			**g**ymnastikhalle			**steten**	**s**portler			**b**oden.	**No**

In EXP/NoCap, the nouns are not capitalized; hence, the adjective in position N ("matten" / "tired") can be misinterpreted as a noun ("Matten" / "mat"). This is not the case for "steten" / "steady." Analogous examples in English would be: "The sick writer had a cold nose throughout the winter" vs. "The sick writer had a blue nose throughout the winter".

We then created sentences which were independently evaluated by six language experts from our department regarding the garden-path trap and their naturality, and finally chose 40 items (see **Appendix**).

The sentences had a length of 7–10 words (M: 8.6, SD: 0.74); word length within the sentences varied from 2 to 16 letters (M: 5.7, SD: 3.05). The sentences were 44–74 characters long (including spaces and final period; M: 57.6, SD: 6.56). All sentences followed a very similar syntactical structure which consisted of four units: Position one was either occupied by a Nominal Phrase (NP) or an Adpositional Phrase (AP), which always contained a noun. The following position was occupied by the verb. The third and fourth units again were composed of either a NP or an AP. The EXP sentences were systematically manipulated following the concept of semantic and partially syntactic priming (cf. **Table 1**): The first AP/NP and the verb were supposed to create a semantic context that evoked a certain reading of the following (second) AP/NP, depending on the orthography of the sentence. The second AP/NP consisted of three words: "N-1," which was either the determiner of the noun (indefinite or definite article) or a preposition. Position two ("N") always represented an attribute of the following noun ("N+1"). The noun N+1 had a word length from 3 to 13 letters (M: 7.38, SD: 2.60); syllables ranged from 1 to 4 (M: 2.33, SD: 0.88). Word N in EXP was semantically related to the first noun phrase and the sentence could syntactically and semantically end after N (i.e., the adjective would serve as a noun). The frequency of interpreting N in EXP either as an adjective (849, $\log_{10}$: 2.93; SD: 1,240, $\log_{10}$: 3.1) or a noun (1,020, $\log_{10}$: 3.0; SD: 1,486, log 10: 3.2) was comparable. As mentioned above, the semantic context should lead the reader into a garden-path trap when capitalization is missing in the whole sentence, which should make parsing as a noun more likely. EXP and CON were presented in two different spellings: (1) The German standard orthography including the capitalization rules and (2) following the English standard orthography with a majuscule (only) at the beginning of the sentence—proper names and other exceptions were avoided.

As mentioned before, the only difference between EXP and CON was position "N," which had a length of 5 to 11 letters (M: 7.58, SD: 2.11). The number of syllables in N differed slightly between EXP (Range: 2–5, M: 2.73, SD: 0.89) and CON (Range: 2–4, M: 2.9, SD: 0.77). The adjectives (i.e., N) for EXP and the corresponding CON were identical in word length and their frequency was matched as far as it was possible to construct a CON sentence which made syntactic and semantic sense. The frequency of N was determined by the two possible readings of EXP: N could be identified as either a noun or an adjective, depending on the spelling presented. Therefore, we decided to use the calculated mean of the two frequencies (adjective and noun) in EXP (M: 934, $\log_{10}$: 2.97; SD: 1,034, $\log_{10}$: 3.01) when matching the frequency of N in CON (M: 922, $\log_{10}$: 2.96; SD: 1,057, $\log_{10}$: 3.02). N+1 had a mean frequency of 7,861 ($\log_{10}$: 3.90; SD: 12,103, $\log_{10}$: 4.08). Word frequencies (i.e., frequencies of complete word forms) were based on "DLEXDB" (Corpus: DWDS [Digital Dictionary of the German Language]) with over 100 million running words (Heister et al., 2011).

The filler sentences were 6 to 10 words in length (M: 8.1, SD: 1.21); word length varied from 2 to 20 letters (M: 5.48, SD: 2.68); the sentences were 43–69 characters long (including spaces and final period; M: 52.5, SD: 6.06). Their syntactic structure differed from the EXP and CON sentences. The filler sentences were mainly extracted from the Potsdam Sentence Corpus (PSC) v. 1.0, which was "constructed with the goal of representing a large variety of grammatical structures" (Kliegl et al., 2004, 267; White et al., 2008, p. 1,271). By using a large range of syntactic constructions, we wanted to make sure that it was not very likely for the participants to see through the syntactic scheme of EXP and CON (each participant read three times more filler sentences than experimental/control ones, see below). As in the EXP and CON sentences, partial minor modifications were made, e.g., avoiding proper names within the sentence, so that no capitalization was necessary.

Participants

Fifty-four participants, mostly undergraduate students and a small number of academic staff members from Potsdam University (37 female, 17 male) with normal or corrected-to-normal vision, took part in the experiment. All participants were totally fluent in German; 52 of them were native speakers, the non-natives without any accent. The age span ranged from 19 to 46 years; the mean age was 25.8 (SD: 6.01).

The sessions lasted for about 30 min, after which the participants filled out an anonymous questionnaire asking for personal data and questions concerning spoken languages and possible problems with German. Ninty-three percent of our participants had learned at least one foreign (alphabetic) language, i.e., were familiar with at least one non-capitalized

script; 87% of the participants were familiar with English, which they had practiced for a mean duration of 13.5 years (SD: 5.4). In addition, they had to answer questions regarding their exposure to (in reading) and production of text violating German capitalization rules, e.g., in digital form. In addition to that, they were asked to evaluate how disturbed they feel by spelling violations (especially non-capitalized nouns) during reading.

Apparatus

Eye movements of the subjects were recorded monocularly in a darkened laboratory with an SR EyeLink 1000 system (500 Hz); the experiment was programmed with Experiment Builder software by SR Research. Sentences occupied only one line on the screen and were presented one at a time in the 1/3 vertical position from the top of the screen of a 22-in. TFT Monitor (1,680 × 1,050 resolution; frame rate 120 Hz). Subjects were seated 60 cm in front of the monitor with their head positioned on a chin rest; participants' dominant eye was tracked. Texts were displayed using 26 pt. bold font Courier New.

Procedure

Before the beginning of the main experiment, there was a pretest consisting of three sentences, so that the participants had a chance to ask questions to clarify ambiguity if necessary. Each subject read 160 sentences in the main experiment: all 120 filler sentences, 20 EXP (10 capitalized, 10 not capitalized) and the remaining counterpart of 20 control sentences, again 10 Cap and 10 NoCap. The order was pseudo-randomized, i.e., all 40 experimental sentences (see **Appendix**) were presented in all four conditions resulting in each sentence per condition being read by 14 participants.

Subjects' measurements were calibrated with a standard nine-point grid. After the validation of calibration accuracy, a drift correction point appeared on the center of the screen. When the dot was fixated, the fixation point re-appeared on the left side of the monitor. If the eye tracker identified a fixation on the fixation spot within 2,000 ms for at least 50 ms, the fixation point disappeared and a sentence was presented such that the center of the first letter in the sentence appeared to the right of the fixation-point position. If there was no identification, the calibrating procedure was repeated.

Subjects were instructed to read the sentences for comprehension and to signal the completion of a trial by fixating on a dot in the lower right corner which either automatically led to the next sentence or to a dual choice task regarding the content of the sentence they had read before. The reason for the question was to ensure that the participants read all the sentences attentively, supported by the fact that the questions appeared in a random order. The answer options were of similar but different structures [e.g., different word classes; example question: "How was the game? (A) Spectacular (B) Boring"]. The answer was given by a left (answer A) or right (answer B) mouse-click, which was supported by an image below the question to avoid confusion. Questions were already asked in the pretest mentioned before. A total of 40 questions were asked per subject (i.e., in 25% of the trials), and subjects correctly

answered 99.2% of questions with a maximum of two wrong answers per participant.

Data Analysis

Fixations were automatically determined by the EyeLink Data Viewer by SR Research. Outliers within the eye movement data were omitted in two steps: First, all potentially mislocated fixations (50 ms or less) were deleted. Second, fixations exceeding 2.5 SD of the mean for each eye movement measure were deleted for each participant. The following measures were calculated:

- First fixation duration (the first of multiple fixations or the only fixation on a target),
- Gaze duration (summation of the duration across all fixations of the first run within the current area of interest),
- Total fixation duration (all fixations on the target, including regressions),
- Regression duration (Total minus Gaze),
- Reading speed (the start time of the first fixation within a sentence until the end of the last fixation, divided by the number of words, minimum duration of 1,200 ms).

This left 97.3 % of the data remaining. Apart from that, the number of fixations in Gaze and Total, the number of regressions is part of our analysis. Linear-mixed models (lme) were used to analyze the eye movement data for each dependent measure in the R environment with the lme4 package version 1.1–13 (Bates et al., 2015b).

The independent measures were as follows:

- Mode of spelling ("SPELLING"), which means whether the nouns have been capitalized (Cap) or not (NoCap);
- Mode of the sentence ("SENTENCEMODE"), which means the two conditions EXP (garden-path, ambiguous in NoCap) and CON (non-ambiguous);
- Position/location of fixation ("POSITION"), which means either Adjective ("N") or Noun ("N+1");
- Word length ("LENGTH"), which means word length in letters—measures were centered;
- Frequency ("FREQUENCY"), which means the absolute token frequency of adjective N and noun N+1—measures were logarithmized and centered;
- Frequency of noun ("FREQNOUN"; only for one model), which means frequency of the noun N+1, divided into two equal-sized groups "high" (Min: 2,589, $\log_{10}$: 3.41; Max: 48,779, $\log_{10}$: 4.69; M: 14,938, $\log_{10}$: 4.17, SD: 13,863, $\log_{10}$: 4.14) and "low" (Min: 1, $\log_{10}$: 0; Max: 2,381, $\log_{10}$: 3.38; M: 785, $\log_{10}$: 2.89; SD: 771, $\log_{10}$: 2.89).

SPELLING, SENTENCEMODE, POSITION, and the noun frequency (FREQNOUN) were transformed into effect contrasts. Participants and items were treated as crossed random effects, and all fixation duration measures were log transformed. By including word length and frequency as a main effect, we ensured that the found effects would be ascribed to our experimental variations and not to linguistic determinants of the words themselves. The criterion for significance was a *t*-value of two. Statistical models were selected following the pattern of parsimonious mixed models (Bates et al., 2015a). We selected

the models by using the RePsychLing package which provides the function "rePCA" (random-effects Principal Components Analysis), which makes it possible to check models fitted with lmer for rank deficiency. Using the rePCA function, it is possible to determine parsimonious random effect structures because the "maximal model in many analyses of data from Psychology and Linguistics experiments, is almost always shown by this analysis to be degenerate" (Bates et al., 2015a, p. 3). A detailed description of model selection by using RePsychLing can be found in Bates et al. (2015a, p. 3).

RESULTS

Global Reading Speed EXP and CON

Mean reading speed was 183.8 words per minute (wpm) in the capitalized condition and 171.1 wpm in the non-capitalized condition, i.e., reading sentences which conform to German orthography is faster than reading sentences without capitalization of nouns. This difference was highly significant (*t*-value: −6.54). The sentence mode had a highly significant influence as well (*t*-value: −4.60); it becomes obvious that the experimental sentences (EXP) were read significantly faster (181.4 wpm) than the control sentences (CON: 173.4 wpm). Including the interaction of spelling and sentence mode (n.s.), the effect of spelling/capitalization is very similar in both modes; reading speed for EXP is 189.5 wpm (Cap) vs. 173.7 wpm (NoCap), and for CON 178.3 wpm (Cap) vs. 168.6 wpm (NoCap).

Global Reading Speed Filler Sentences

Comparing the results of the experimental and control sentences with the 120 filler sentences, which consist of simpler syntactic structures, it becomes obvious that they were generally read faster: 210.2 wpm in the capitalized and 200.7 wpm in the non-capitalized condition; this effect was significant (*t*-value: −7.5).

Gaze and Total Fixation Duration (Adjective and Noun)

Table 2 summarizes the model data for the main effects and interactions of Gaze and the Total fixation duration. The fixation times for the words N (Adjective) and N+1 (Noun) in Gaze duration (first pass fixations, left panel) show a significant main effect for the mode of spelling (noun capitalized M: 278.5, SD: 90.5 vs. not capitalized M: 288.3, SD: 92.4), as well as for the mode of the sentence (EXP M: 276.8, SD: 88.1 vs. CON M: 290.1, SD: 95.4), i.e., comparable to the global reading speed, the words in focus (N, N+1) were fixated on for less time when capitalization was present (as expected) in both experimental settings. Apart from that, a main effect for word length is detectable. Although there is no main effect for the location of fixation (i.e., adjective vs. noun), the interaction with capitalization is striking (*t*-value: −7.08, cf. Figure 1): In regularly capitalized sentences, the adjective is fixated on for longer than the following noun, but if capitalization is missing the noun receives longer fixations; apart

from that, adjectives show generally longer Gaze duration when capitalization is present.

An interaction between the mode of the sentence and the location of fixation did not reach significance (cf. Table 2, Figure 2), nor did, contrary to our hypothesis, the three-way interaction of all independent measures (cf. Figure 3), i.e., the sentence mode did not modulate the reading behavior for any kind of adjective or noun. Due to the complexity of our experimental design, we decided to show these plots to support the reader's understanding.

Generally, the settings EXP and CON differed to such an extent that the control sentences were read more slowly. We will discuss this issue below.

The fixation times for the word N (Adjective) and N+1 (Noun) in the Total fixation duration (all fixations for each word including regressions) again show a significant main effect for the spelling mode (noun capitalized M: 324.0, SD: 118.3 vs. not capitalized M: 340.4, SD: 130.2) as well as for the sentence mode (EXP M: 318.7, SD: 118.2 vs. CON M: 346.1, SD: 130.6), i.e., we found shorter fixation durations on N/N+1 when capitalization was present in both experimental settings. Once more, the word length effect is significant. The interaction with capitalization is again striking (*t*-value: −7.54, cf. Figure 1): In regularly capitalized sentences, the adjective is fixated on for longer than the following noun, but if capitalization is missing the noun receives longer fixations; however, this difference is not significant.

The only difference compared to Gaze is the significant interaction of the sentence mode (cf. Table 2, Figure 2) and the fixation location in such a way that the control sentences receive longer regressions on the adjective when the noun is not capitalized compared to the experimental sentences, i.e., the reader seems to clarify the meaning of this specific adjective. A possible explanation for this will be discussed below. Again, the three-way interaction of all independent measures did not reach significance.

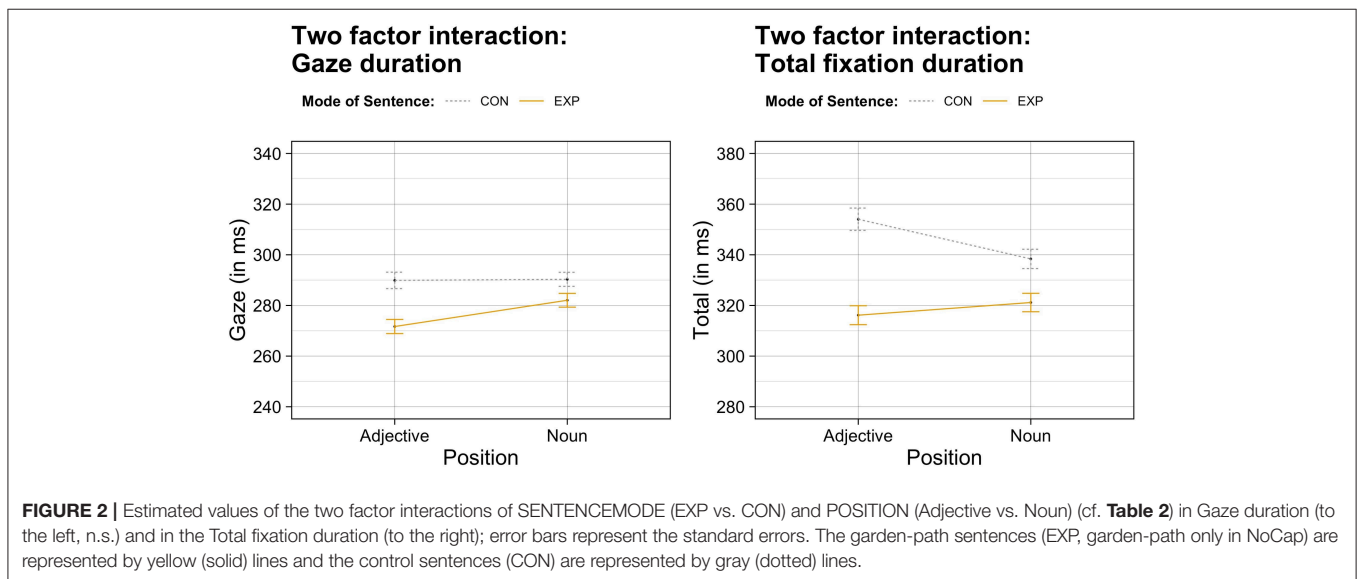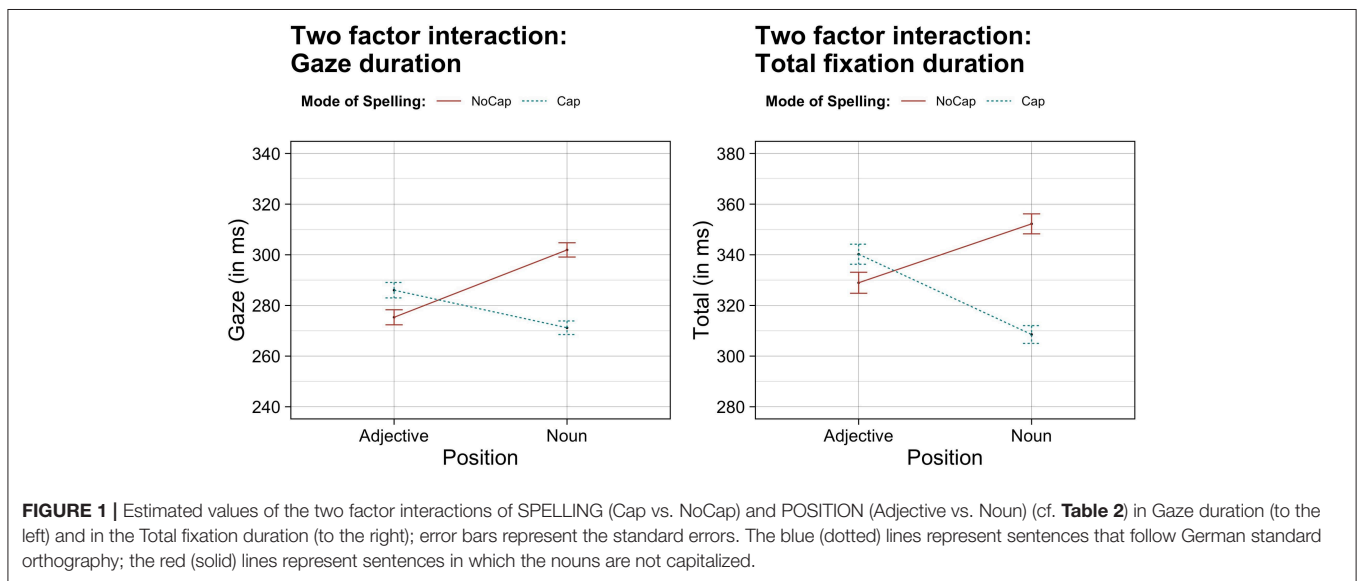### Gaze and Total Fixation Rate (Adjective and Noun) and Inter-word Regressions

Regarding the *number of fixations* (cf. Table 3) in the first run on adjectives and nouns (i.e., Gaze), there is no main effect for the spelling mode (capitalization vs. missing capitalization of the noun), which means that in contrast to the significantly longer fixation durations reported above when capitalization is missing, the mean number of fixations does not differ significantly. In contrast to that, there is a very small effect for Total, in which significantly more fixations in NoCap are made compared to Cap (Cap: M: 1.616, SD: 0.634; NoCap: M: 1.659, SD: 0.683). Word length and frequency as main effects are detectable in Gaze and Total. The latter is exclusive for the number of fixations and does not occur in case of mean fixation durations.

Regarding the sentence mode, there is another main effect which only occurs in Total (cf. Table 3), i.e., significantly more fixations are made in CON compared to EXP. An explanation for that will be discussed below and is connected with the longer

TABLE 2 | Gaze and Total, model data for the independent variables of the fixation duration.

	GAZE			TOTAL		
	Estimate	**Std. error**	*t*-value	**Estimate**	**Std. error**	*t*-value
(Intercept)	5.596	0.021	266.26	5.74	0.028	201.89
SPELLING	−0.034	0.010	**−3.29**	−0.045	0.012	**−3.86**
SENTENCEMODE	−0.044	0.014	**−3.14**	−0.081	0.019	**−4.26**
POSITION	0.032	0.026	1.19	−0.003	0.026	−0.10
LENGTH	0.155	0.013	**12.22**	0.147	0.014	**10.30**
FREQUENCY	−0.010	0.013	−0.76	−0.029	0.015	−1.96
SPELLING:SENTENCEMODE	−0.007	0.021	−0.34	−0.005	0.023	−0.20
SPELLING:POSITION	−0.147	0.021	**−7.08**	−0.174	0.023	**−7.54**
SENTENCEMODE:POSITION	0.030	0.021	1.42	0.055	0.023	**2.36**
SPELLING:SENTENCEMODE:POSITION	−0.004	0.042	−0.090	−0.047	0.046	−1.02

Values appearing bold are significant.

FIGURE 1 | Estimated values of the two factor interactions of SPELLING (Cap vs. NoCap) and POSITION (Adjective vs. Noun) (cf. **Table 2**) in Gaze duration (to the left) and in the Total fixation duration (to the right); error bars represent the standard errors. The blue (dotted) lines represent sentences that follow German standard orthography; the red (solid) lines represent sentences in which the nouns are not capitalized.

FIGURE 2 | Estimated values of the two factor interactions of SENTENCEMODE (EXP vs. CON) and POSITION (Adjective vs. Noun) (cf. **Table 2**) in Gaze duration (to the left, n.s.) and in the Total fixation duration (to the right); error bars represent the standard errors. The garden-path sentences (EXP, garden-path only in NoCap) are represented by yellow (solid) lines and the control sentences (CON) are represented by gray (dotted) lines.
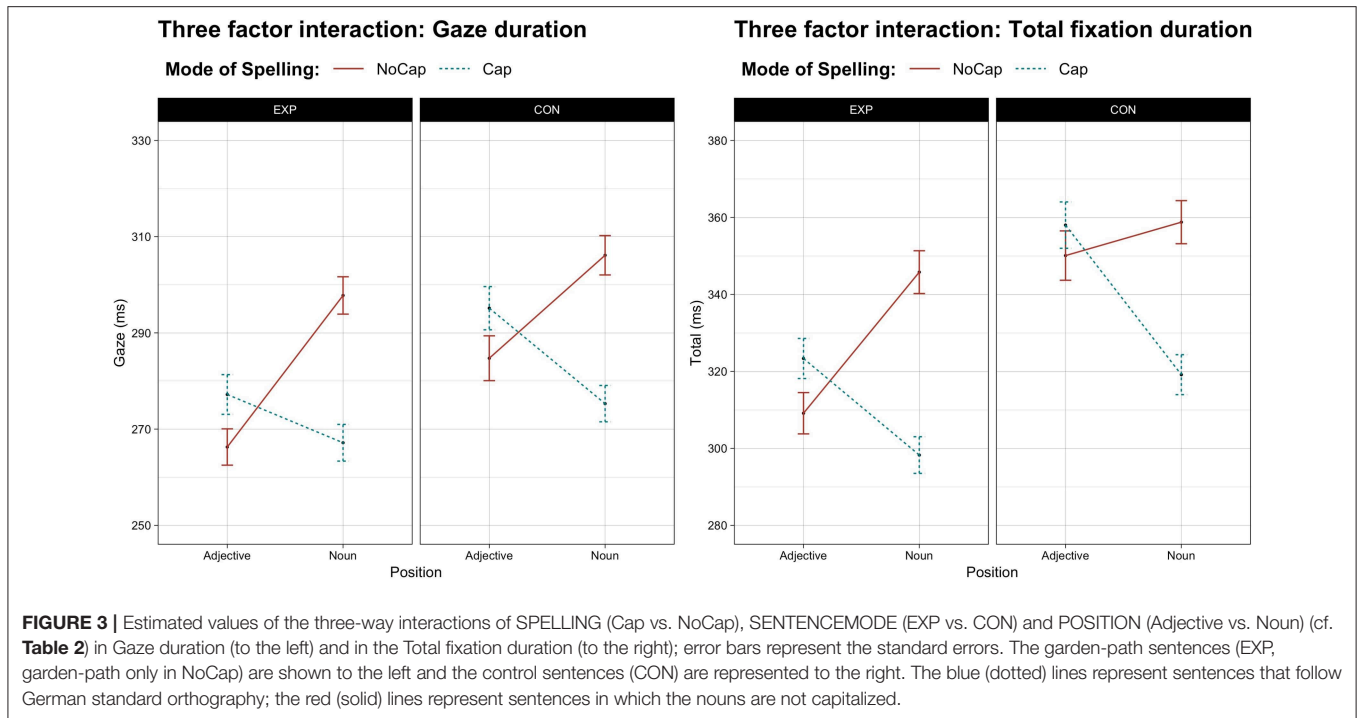
FIGURE 3 | Estimated values of the three-way interactions of SPELLING (Cap vs. NoCap), SENTENCEMODE (EXP vs. CON) and POSITION (Adjective vs. Noun) (cf. **Table 2**) in Gaze duration (to the left) and in the Total fixation duration (to the right); error bars represent the standard errors. The garden-path sentences (EXP, garden-path only in NoCap) are shown to the left and the control sentences (CON) are represented to the right. The blue (dotted) lines represent sentences that follow German standard orthography; the red (solid) lines represent sentences in which the nouns are not capitalized.

TABLE 3 | Gaze and Total, model data for the independent variables of the mean number of fixations.

	GAZE			TOTAL		
	Estimate	**Std. error**	***t*-value**	**Estimate**	**Std. error**	***t*-value**
(Intercept)	1.378	0.023	58.91	1.637	0.050	32.58
SPELLING	−0.023	0.015	−1.52	−0.043	0.021	**−2.02**
SENTENCEMODE	−0.012	0.015	−0.78	−0.074	0.027	**−2.69**
POSITION	0.014	0.036	0.38	−0.041	0.038	−1.07
LENGTH	0.244	0.017	**14.78**	0.253	0.021	**12.33**
FREQUENCY	−0.039	0.018	**−2.14**	−0.062	0.022	**−2.76**
SPELLING:SENTENCEMODE	0.009	0.030	0.28	−0.006	0.042	−0.13
SPELLING:POSITION	−0.079	0.030	**−2.64**	−0.138	0.042	**−3.27**
SENTENCEMODE:POSITION	0.032	0.030	1.06	0.083	0.042	1.95
SPELLING:SENTENCEMODE:POSITION	0.091	0.060	1.51	0.032	0.084	0.38

Values appearing bold are significant.

regressions on the adjective mentioned above (CON M: 1.674, SD 0.667; EXP M: 1.600, SD 0.651). Hand in hand with the mean fixation durations, there is also a significant—although weaker—interaction between the spelling mode and fixation position (cf. **Figure 4**) in Gaze and Total. Please note that the characteristic between Gaze and Total is slightly different: When capitalization is present, adjectives are fixated on more often compared to nouns, but this is only significant in Total. Instead, in cases where the capitalization is absent, it is vice versa: Adjectives are fixated on less often compared to nouns, but the difference between them is only significant in Gaze. The explanation for that is that adjectives receive more regressions compared to nouns (see also **Table 4**). Apart from that, there is no significant difference

between the two spelling modes. In sum, the number of fixations is less affected than the fixation durations. Regarding the inter-word regression rate (cf. **Table 4**), i.e. the rate of receiving at least one regression from a word to the right, we found values between 6.7% and 13.4%. A clear tendency becomes apparent: the adjectives receive more regressions compared to the nouns, and regressions are more likely in the NoCap condition.

Fixation Duration Depending on the Frequency of the Noun

The following analysis of noun frequency effects was integrated *post-hoc*, and the division of the nouns into high- and low-frequency items cuts the sample size per condition by half.
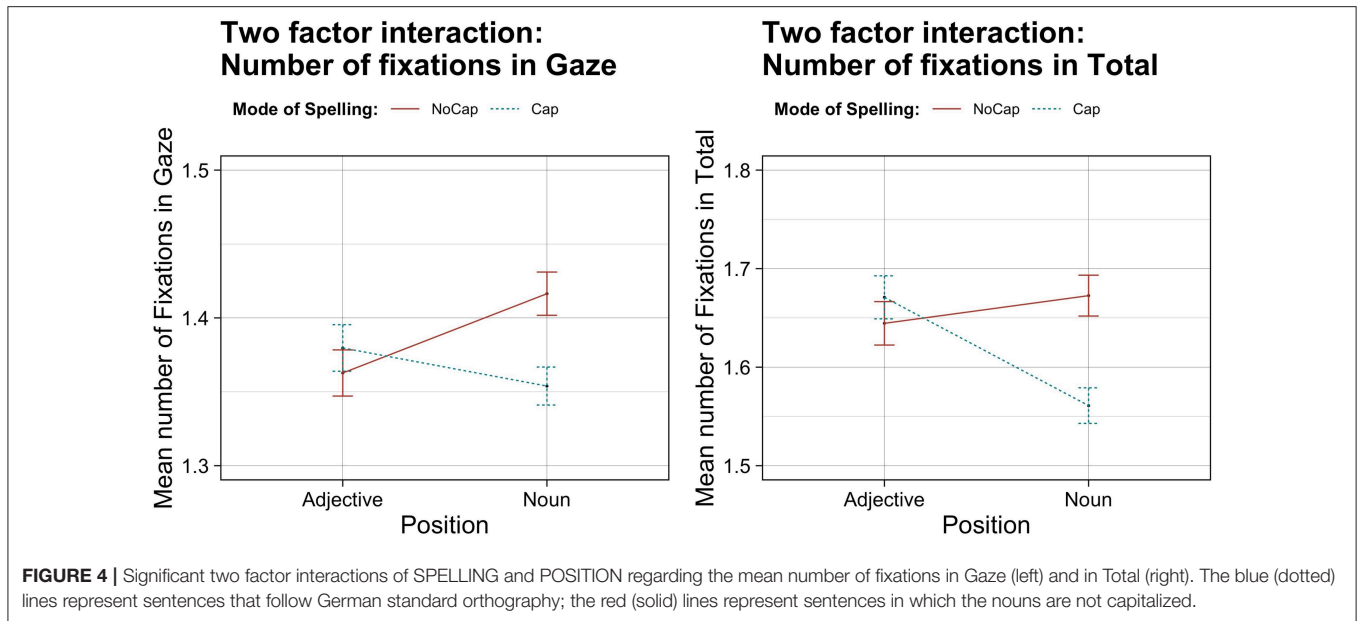
FIGURE 4 | Significant two factor interactions of SPELLING and POSITION regarding the mean number of fixations in Gaze (left) and in Total (right). The blue (dotted) lines represent sentences that follow German standard orthography; the red (solid) lines represent sentences in which the nouns are not capitalized.

TABLE 4 | Descriptive statistics: probability of receiving at least one regression from a higher Interest Area (word from the right); colors are shading from dark green (low value) to yellow (middle value) to red (high value).

	Cap (%)	NoCap (%)	Difference (%)
all (Noun & Adj. in EXP & CON)	9.53	10.62	1.09
CON all	9.66	10.94	1.28
EXP all	9.40	10.31	0.91
all Adjective	11.92	12.28	0.36
CON adjective	11.85	13.35	1.50
EXP adjective	11.99	11.24	−0.75
all Noun	7.10	9.00	1.90
CON noun	7.51	8.62	1.11
EXP noun	6.68	9.39	2.71

TABLE 5 | Gaze duration depending on the Frequency of the Noun, model data for the independent variables of the Fixation Duration.

	Estimate	Std. error	t-value
(Intercept)	5.477	0.023	240.56
SPELLING	−0.034	0.01	**−3.29**
SENTENCEMODE	−0.044	0.014	**−3.16**
POSITION	0.024	0.025	0.96
FREQNOUN	−0.027	0.024	−1.13
LENGTH	0.204	0.015	**13.89**
SPELLING:SENTENCEMODE	−0.008	0.021	−0.39
SPELLING:POSITION	−0.148	0.021	**−7.16**
SENTENCEMODE:POSITION	0.028	0.021	1.34
SPELLING:FREQNOUN	−0.001	0.021	−0.04
SENTENCEMODE:FREQNOUN	0.025	0.028	0.90
POSITION:FREQNOUN	−0.034	0.044	−0.78
SPELLING:SENTENCEMODE:POSITION	−0.004	0.041	−0.10
SPELLING:SENTENCEMODE:FREQNOUN	0.018	0.042	0.42
SPELLING:POSITION:FREQNOUN	−0.065	0.042	−1.55
SENTENCEMODE:POSITION:FREQNOUN	−0.114	0.042	**−2.71**
SPELLING:SENTENCEMODE:POSITION: FREQNOUN	−0.222	0.084	**−2.66**

Values appearing bold are significant.

Therefore, we present these results with considerable caution. Relating the Gaze duration (cf. **Table 5**) of the adjective and the noun to the frequency of the noun (FREQNOUN, low vs. high frequency), there are once again notable effects caused by SPELLING (Cap vs. NoCap), i.e., shorter fixation durations on N/N+1 in cases where the spelling of the noun is correct, and caused by the SENTENCEMODE (EXP vs. CON), i.e., shorter fixation durations in the EXP condition. Furthermore, a notable effect caused by word length is demonstrable (LENGTH). Moreover, interaction for SPELLING:POSITION (Cap vs. NoCap: Adjective vs. Noun) is highly significant. The three-way (SENTENCEMODE:POSITION:FREQNOUN) and four-way interactions (SPELLING:SENTENCEMODE:POSITION: FREQNOUN) are significant as well, notwithstanding the small sample size.

Among the capitalized conditions of the control sentences, the duration of the Gaze on the nouns is not affected by the frequency of the nouns but rather the adjectives (cf. **Figure 5**,

upper-left and upper-right graph, gray dotted line; see "Problems with Garden-Path and Control Sentences" below and discussion). The Gaze durations on the adjective in the correctly spelled EXP non-garden-path sentences, however, show a highly interesting finding: If the following noun is of low frequency, the Gaze durations on the adjective are significantly shorter compared to a high-frequency noun with a difference of more than 21 ms. Over the course of the exercise, the processing time of the noun is influenced. Apart from that, the ratios of adjective and noun
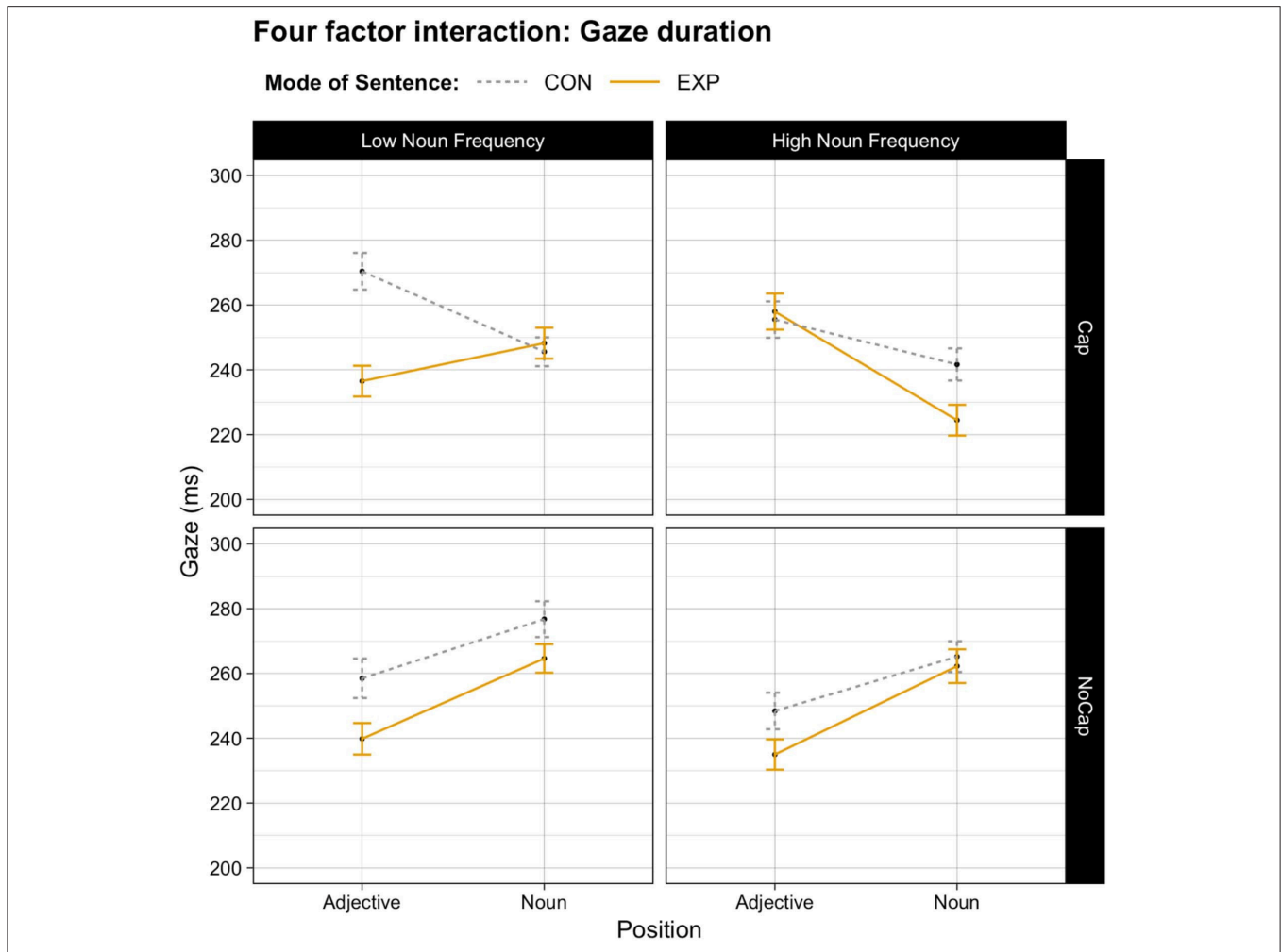
FIGURE 5 | Gaze duration on Adjective and Noun depending on the noun frequency. Estimated values of the four-way interactions of SPELLING (Cap vs. NoCap), SENTENCEMODE (EXP vs. CON), POSITION (Adjective vs. Noun) and FREQNOUN, i.e., the influence of the noun frequency on the fixation durations on the adjective and the noun themselves, divided into two groups: Adjective vs. Noun in Gaze duration (to the left); error bars represent the standard errors. Within both panels, "Low Noun Frequency" is shown to the left, "High Noun Frequency" to the right. Sentences following German standard orthography (Cap) are shown in the upper part of the graph, while the lower part represents sentences, in which the standard orthography is violated, i.e., noun with lowercase letter (NoCap). "Low" and "high" show the fixation duration depending on the level of noun frequency. The garden-path sentences (EXP, garden-path only in NoCap) are represented by solid yellow lines and the control sentences (CON) are represented by dotted gray lines.

mean durations in EXP depend on the frequency (cf. **Figure 5** upper-left and upper-right graphs, yellow solid line): While in the context of a low-frequency noun the adjective is fixated on shorter compared to the noun, this is vice versa when the noun frequency is high.

The Gaze durations on the noun—in cases where it is correctly spelled—follow a typical pattern of frequency effects, which are well-known (e.g., Broadbent, 1967; Becker, 1979): The mean fixation durations are significantly higher when the frequency is low compared to a higher frequency. In our experiment, this effect is only clearly visible in the capitalized EXP condition. When comparing capitalized with non-capitalized nouns, the durations are significantly shorter for the former. The mode of the sentence has no significant influence on the noun as long as it is not capitalized or, if it is capitalized, is of low frequency.

However, in the case of a capitalized high-frequency noun (cf. **Figure 5**, upper-right graph), there is a significant difference in the mean Gaze duration in EXP compared to CON. The slope of the EXP line on the noun and adjective (cf. **Figure 5**, upper-right graph, solid yellow line) is the steepest of all lines. We interpret this as a sign of parafoveal pre-processing in cases where the noun is spelled correctly and of high frequency.

Problems With Garden-Path and Control Sentences

The results reported so far do not support the hypothesis of garden-path parsing by the participants in the EXP/NoCap condition. A clear demonstration of this outcome is provided by the fact that the interaction of spelling mode, sentence mode and position has no significant effect on either Gaze or Total

fixation duration, i.e., the mode of sentence (EXP vs. CON) does not influence the processing in a statistically significant way. Therefore, we provide a *post hoc* analytic explanation: As described above, the garden-path type was always semantic, i.e., the first noun phrase initiated a semantic field that was supposed to lead the reader onto the garden path, i.e., parsing the adjective in the second NP as a noun. However, it is possible to divide the sentences (see **Appendix**) into two subtypes that differ according to the semantic relation between the adjective and the following noun:

- Subtype I: The adjective and the noun are semantically similar or tautological. The full sentence would remain intact if the noun were omitted: "jugendliche Kinder" (*juvenile children*), "kriminellen Dieb" (*criminal thief*), "neugeborene Baby" (*newborn baby*); 24 sentences (60 %) belong to this category;
- Subtype II: The adjective is semantically dissimilar: "Der Kapitän kommandiert die flotte [engl. *brisk*] Mannschaft…" (*The captain commands the fleet crew…*), "In der Gymnastikhalle liegen die matten Sportler… (*In the gym hall lie the [exercise] mats* [garden-path parsing] / *faint* [actual meaning] *sportsmen…*"); 16 sentences (40 %) belong to this category.

Concerning sentences of Subtype I, the garden-path word is semantically in a similar field as the following noun. If capitalization is missing in the sentence *"Die mutter hält das neugeborene baby im arm"* (English: The mother holds the newborn baby in her arm"), "newborn" is likely to be interpreted as a noun, but a complete reparsing of the sentence will not be necessary when continuing reading, because a newborn baby is in fact a "newborn" and a "baby"—it is tautological. The adjectives in sentences of Subtype I have different shades of meaning (Juhasz and Pollatsek, 2011, p. 883). The only thing that fundamentally changes is the word class. After having analyzed the data, we came to the conclusion that, especially in these cases, the garden-path effect is unable to fully unfold, because a complete reparsing is not necessary.

Unfortunately, the number of items of Subtype II is too small to report inferential statistics here. Nevertheless, we report some tendencies that will be of interest in future studies. The two subtypes differ substantially concerning Gaze and Total in both conditions (EXP / CON), showing longer fixations on the noun than on the adjective in Gaze except for CON/Cap. However, there should be no differences between Subtypes I and II in CON. This may be largely due to non-balanced frequencies between the subtypes. It is interesting, though, that there are more regressions toward the adjective in EXP / Subtype II, even in the capitalized condition.

Regardless, not all items of Subtype II lead to a reparsing of the sentence as clearly as the example sentences above indicate. One reason for that can be found in the German language: The garden-path words were all adjectives which acted as an adjunct of a following noun. Compared to English, in which attributive adjectives do not have flectional endings (e.g., "blue": *the blue nose*; *the blue noses),* adjectives as adjuncts in German always consist of the adjective stem and a flectional ending ("*blau*": *die blau Nase*; *die blau Nasen*). The flectional ending itself cannot

be predicted easily. On the one hand, it depends on the case, number and gender of the following noun to which the adjective adapts itself (case-number-gender-congruency). On the other hand, the flection type is defined by the determiner before the adjective, which leads to three different types of flection: the so-called strong flection, the weak flection, and the mixed flection (Eisenberg, 2013). Since nouns in German also belong to the inflectable word classes, which consist of flectional endings that are added to the stem even though the flection type of nouns and adjectives differ (Eisenberg, 2013), this proves to be another indicator for a hard-to-predict flectional ending.

In sum, it seems as though semantic processing is stronger than a minor orthographic spelling deviation, both foveally (garden-path parsing on the adjective in the NoCap condition) and parafoveally (garden-path parsing on the adjective in the Cap condition). In other words, the ambiguity of the adjective in EXP is not dissolved by the parafoveal perception of the majuscule.

Additionally, the control sentences generally showed consistently longer fixation durations without delivering other results regarding the direction of the effects. We think that this is primarily explainable by the complexity of the experimental design: As described above, when designing the experimental sentences many properties had to be taken into account in addition to the semantic requirements of ambiguity. Also as mentioned above, only 68 potential EXP items remained after CELEX sighting, from which we constructed 40 sentences. The CON items were created at the very end: They had to be of the same length, matching semantically and matching a specific frequency corridor, and it had to be impossible to misinterpret them as a noun of any kind. All of the CONs are proper German sentences in theory, but at least some of them must have appeared a little peculiar to the participants. This is also an explanation for the significant interaction of the mode of the sentence (EXP vs. CON) and the position (Adjective vs. Noun) in the Total fixation duration, because the adjectives in CON are regressed longer, we assume, due to their oddity.

DISCUSSION

In the present study, it was our goal to find out whether the German capitalization rules for nouns present a benefit to the reader. A benefit could result from the fact that it is possible to perceive the majuscule parafoveally, which could lead to grammatical pre-processing due to the syntactic function of a noun as the core of a nominal phrase.

Reaching disambiguation through capitalization in garden-path sentences and thus singularizing an effect of the German noun capitalization did not work out, mainly due to language constraints and the design-specific reasons outlined above. This conclusion is supported by the non-significant interaction of SENTENCEMODE:POSITION in Gaze duration (cf. **Table 2**, **Figure 2** to the left); i.e., the fact that processing in the Gaze duration is not affected by the sentence modes (garden-path or control sentence) regarding a capitalized or a non-capitalized noun. The significance of this interaction in Total (cf. **Table 2**,

Figure 2 to the right) stems from longer and more regressions in the control sentence condition on the adjective when the upcoming noun is not capitalized. We explain that based on the peculiarity of the adjectives in the CON sentences due to the reasons mentioned above.

As a global result, it can be stated that reading sentences in which nouns are not marked by a majuscule slows down a native German reader significantly, but from an absolute point of view, the effect is small. These results coincide with several previous findings (e.g., Bock et al., 1985), but differ from a result by Hohenstein and Kliegl (2013, p. 2,556), in which the global reading speed was higher compared to our study and reading speed was almost identical for both conditions: 208 words per minute in the capitalized condition vs. 207 words per minute in the non-capitalized condition (184 vs. 171 words in our study, but more comparable regarding our filler sentences: 210 vs. 200).

Furthermore, it can be noted that the mean fixation duration on a non-capitalized noun in Gaze is always significantly higher compared to a capitalized noun (by 30 ms on average). The effect is stable and independent of the sentence mode (EXP vs. CON) in our experiment. The result is expected but, compared to other deviations from the standard orthography, its impact is less strong: In an experiment by Rayner and Kaiser (1975), where internal or ending letters were substituted by visually similar letters (*pncblem/problnc* for *problem*), reading time doubled; when initial letters (*qroblem*) were substituted, reading time became even 2.5 times longer than normal. Our results are more in line with those from an experiment by White et al. (2008), in which letters (among other conditions) were transposed in the beginning of the word (e.g., *rpoblem* for *problem*). The mean difference in Gaze between a correct and a transposed word was 45 ms (White et al., 2008, p. 1,271). The authors state that a transposition in the beginning of a word is most critical for foveal and parafoveal processing (White et al., 2008, p. 1,274).

The difference between a capitalized and a non-capitalized noun for the Total fixation duration increases to about 43 ms (plus 13 ms) in our experiment. Compared to the experiment of White et al. (2008), this is a moderate increase; they report a difference of 120 ms (Correct: M: 336, SD: 169 vs. Transposed: M: 456, SD: 390), which means an absolute increase of 75 ms. While the experiment of White et al. (2008) did not examine capitalization, it is therefore worth noting that capitalization differs from substitution or transposition of letters because the underlying allograph remains constant; the letter is just substituted by its own lower- or uppercase form and the phonological level is not affected. Therefore, capitalization can be conceived as a lesser violation of orthography resulting in less strong effects on the number and duration of fixations. In addition, our participants are used to seeing non-capitalized nouns in English and other alphabetic writing systems (see participants section) and, finally, in modern communication channels the concept of non-capitalized nouns in German is quite familiar.

Concerning the advantages capitalization may have for the reader, this study provides indications that there is a (small) benefit under specific circumstances. As reported above, we also found a significant interaction between SPELLING and

POSITION in Gaze and Total, which is a stable effect and independent of SENTENCEMODE (cf. **Table 2**, **Figure 1**). This includes a significant difference for the adjective, resulting in longer fixation durations throughout the capitalization condition; the actual number of fixations on the adjective does not differ significantly (cf. **Figure 4**). The difference in the mean duration is about 10 ms in Gaze and in the Total duration. These results can be compared with those by Hohenstein and Kliegl (Hohenstein, 2013, p. 112; cf. Hohenstein and Kliegl, 2013, p. 2,557; 2014). They conducted several boundary paradigm experiments regarding the benefits of semantic preview and also included the influence of noun capitalization as a factor in one experiment. They examined whether a related or unrelated semantic preview of the target word and the target itself (always a noun) is modulated by capitalization, i.e., whether or not the noun is capitalized (example: "Beim Ausgraben waren *K/knochen* zum Vorschein gekommen" [During the excavation, bones appeared.]; related preview: *S/schädel* [skulls] unrelated: *S/stiefel* [boots]; target and preview word in italics). Fixation durations were longer for the pre-target word when the following noun was capitalized, independently of the preview mode [Capitalization related/unrelated preview (in ms): M: 260/270, SD: 101/119; Lowercase: M: 241/237, SD: 114/97 (Hohenstein and Kliegl, 2014, p. 175)]. This is similar to our finding, with the difference that in their case the pre-target word could represent all word classes (but nouns), while in our study it was consistently an adjective.

It can be seen that "fixation durations in sentences presented completely in lowercase are less modulated by properties of the fixated and surrounding text" (Hohenstein and Kliegl, 2013, p. 2,557). Below we refer to this phenomenon when discussing the influence of the frequencies of an upcoming noun. The authors reanalyzed this experiment in the form of an eye-movement corpus analysis (Hohenstein and Kliegl, 2013), in which astonishingly no main effect of capitalization was demonstrable, which is in line with the same global reading speed in both conditions. The authors explained that by referring to the proficiency of reading text without capitalization due to internet-based communication by the participants (young adults). In the present study, we did not find any systematic effects regarding the self-assessment of the participants concerning the disturbance intensity of non-capitalized nouns and/or their own violation of noun capitalization, although our participants were mainly young adults as well. The authors also report a significant interaction between word class and capitalization. They divided the word class into two groups, nouns and non-nouns, and examined the Gaze duration on the current word as a function of capitalization and word class of the previous, the current, and the next word. Significance of this interaction was found for the current and the next word. This means that in the capitalized condition a currently fixated noun is fixated on less compared to a non-noun; for the next word this effect is reversed.

Similar to our findings, the reading time for the current word increases when the next word is a noun and decreases when the current word is a noun, provided the orthography is correct.

Hohenstein and Kliegl (2013) argue that the word before a noun is highly associated with the noun and that therefore both

words are processed before the noun is actually fixated upon (word group hypothesis). The evidentiary status of the word group(ing) hypothesis—processing two words as one whole—(e.g., Radach, 1996; Kliegl, 2007; Drieghe et al., 2008), though, is unclear. Radach (1996) found this kind of processing for two-word groups (in German), in which the first one was short and the second one a five- to seven-letter noun. Drieghe et al. (2008) were able to replicate the effect in English; they showed that it was restricted to cases in which the first word was an article, but processing was different for highly frequent three- or five-letter non-articles as a first word. Based on the fact that the word before the noun was not controlled and that there is only evidence for three-letter words so far, we are unsure about the conclusion, but we agree that their and our findings are providing evidence for parafoveal processing in reading.

In another boundary paradigm experiment, Tiffin-Richards and Schroeder (2015) examined the preview effect of nouns in German. They compared regularly presented nouns with those presented without a majuscule before crossing an invisible boundary. The preview effect was not significant, but the 10 ms difference in Gaze duration is comparable with our findings reported above: There seems to be a slight parafoveal pre-processing of about 10 ms of the word before a capitalized noun which leads to a lower processing duration of the noun itself.

Parafovea-on-fovea effects (PoF) of word frequency are highly controversial (for an overview: Drieghe, 2011; Hyönä, 2011). While in some corpus studies (Pynte et al., 2005; Pynte and Kennedy, 2006) and reading-like tasks (e.g., Kennedy, 1998, 2000) the effect was reported, in experimentally manipulated settings the effect was rarely found (e.g., Carpenter and Just, 1983; Henderson and Ferreira, 1993; Angele et al., 2008); Hyönä and Bertram (2004) found an effect in one out of five of their reported experiments.

In our study, the processing of an adjective-noun word group with or without a capitalized letter also seems to interact with the frequency of the noun (cf. **Table 5**, **Figure 5**). In the capitalized EXP condition, the mean difference in Gaze duration between a high-frequency noun (224 ms) and a low-frequency noun (248 ms) is noteworthy. For the adjective, a reverse effect can be observed: An adjective followed by a high-frequency noun is fixated on 21.5 ms longer compared to an adjective followed by a low-frequency noun. Quite the opposite can be observed for the adjective in the NoCap condition; the mean fixation durations are almost identical (difference < 5 ms) and thus independent of the frequency of the upcoming noun. Interestingly, the mean fixation durations on adjectives are quite similar to the durations on the following nouns in NoCap (independent of the noun frequency) and in Cap / low-frequency nouns. Only in case of capitalized high-frequency nouns are the adjectives fixated on longer than the following noun. Besides that, no frequency effects are detectable on the noun itself when the orthographic regular capitalization is missing—a low-frequency noun is fixated on only 2.4 ms longer compared to a high-frequency one. The observations reported above are similar (but shifted) for the non-capitalized control sentences.

TABLE 6 | Mean Gaze durations (in ms) on the word group (Adjective + Noun) in both spelling modes split by Noun frequency in the EXP condition.

	Low	High
Cap	484.7	482.4
NoCap	504.5	497.3

This leads us to the following conclusions on frequency effects, which should be observed with considerable caution due to sample-size and *post hoc* analyses:

1. Orthographic violation, in the form of missing noun capitalization, levels the frequency effects on the noun itself.
2. Parafoveal pre-processing of a noun does not occur when the following noun is not capitalized, independent of the frequency of the upcoming noun.
3. Parafoveal pre-processing of a capitalized noun does not occur when the following noun has a low frequency.

However, there is reason to assume that parafoveal pre-processing occurs, but only under optimal circumstances (cf. **Table 6**): The frequency of the upcoming noun needs to be high and the current word (adjective) must not be peculiar (like in some of our CON sentences). Even then, the advantage is rather small or even non-existent: When the mean durations for adjective and noun are added together, fixation durations in the Cap condition are only 2.3 ms shorter with a high-frequency noun. The pre-processing of the following noun while fixating on the adjective seems to cost almost as much as the benefits it provides while processing the noun itself. An explanation for this could be that these effects are only present in the results of the experimental sentences and that noun frequencies in this *post hoc* analyses were not controlled in advance.

Together with the findings of Hohenstein and Kliegl (2013, 2014) and Tiffin-Richards and Schroeder (2015), our findings indicate that the capitalization of German nouns supports parafoveal pre-processing and, in cases where the capitalization is missing, there is some cost to the inhibition of proper pre-processing. However, it is unclear whether this is due to the violation of the familiar orthographic system or whether in some contexts (e.g., high frequency of the upcoming noun) the parafoveal preview of the majuscule is prevented to enable another approach to (faster) processing. In favor of the latter are results by Rayner and Schotter (2014), who did not find evidence of semantic preview benefit to readers of English for semantically related previews when following the English standard orthography, but they did find evidence for semantic preview benefit when the target/preview words were capitalized.

German capitalization of nouns supports parafoveal pre-processing, but the question whether this leads to an actual overall processing benefit is still not satisfyingly answered. Besides a replication of our study to put the results on a more solid footing with more stimulus materials and participants (cf. Brysbaert and Stevens, 2018; Brysbaert, 2019), further research could envisage two paths to a better understanding of these effects. One path should include language comparison: Are

German native speakers able to transfer the application of German orthographic rules regarding noun capitalization to another writing system, e.g., English? Flectional endings for adjectives regarding case, number and gender do not exist in English, which means that one problem of our study would be eliminated. The second path we would like to follow is a boundary paradigm experiment in which the interaction of the frequency and spelling of a noun on the preceding word is examined to clarify whether the pre-processing does actually lead to a processing benefit or whether it is just a different form of processing.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

At the time of conducting the experiment, the approval of an ethics committee was not mandatory at the University of Potsdam neither was it by national regulations. Nevertheless, the data collection was performed according to the latest guidelines of the ethics committee of the University of Potsdam. The

participants received full enlightenment regarding the tasks of the experiment and had the opportunity to ask questions. Full anonymity regarding participants' data was guaranteed. Written informed consent was obtained from all participants. The participants had the right of withdrawal without giving reasons at any time.

AUTHOR CONTRIBUTIONS

DP and GN designed the study. DP performed the experiment, analyzed the data, and wrote the paper with input from all authors.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2020.00015/full#supplementary-material

REFERENCES

Angele, B., Slattery, T. J., Yang, J., Kliegl, R., and Rayner, K. (2008). Parafoveal processing in reading: manipulating n + 1 and n + 2 previews simultaneously. *Vis. Cogn.* 16, 697–707. doi: 10.1080/13506280802009704

Augst, G. (1980). Kleinschreibung oder Großschreibung – weitere Bausteine zu einem selbständigen Urteil. *Wirkendes Wort* 30, 22–29.

Baayen, R., Piepenbrock, R., and Gulikers, L. (1995). *The {CELEX} Lexical Data Base on {CD-ROM}.* Philadelphia, PA: University of Pennsylvania; Linguistic Data Consortium.

Bates, D., Kliegl, R., Vasishth, S., and Baayen, H. (2015a). Parsimonious mixed models. *arXiv:1506.04967v1.* 1–21.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015b). Fitting linear mixed-effects models using lme4. *J. Stat. Softw. Vol* 67:12015. doi: 10.18637/jss.v067.i01

Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 5, 252–259. doi: 10.1037/0096-1523.5.2.252

Bock, M., Augst, G., and Wegner, I. (1985). Groß oder klein? Zur Funktion des Wortanfangs fuer den gegenwärtigen Leser. *Z. Für Entwicklungspsychologie Pädagog. Psychol.* 17, 191–209.

Bock, M., Hagenschneider, K., and Schweer, A. (1989). "Zur Funktion der Groß- und Kleinschreibung beim Lesen deutscher, englischer und niederländischer Texte," in *Schriftsystem und Orthographie* Reihe Germanistische Linguistik, eds P. Eisenberg and H. Günther (Tübingen: Niemeyer), 23–55.

Broadbent, D. E. (1967). Word-frequency effect and response bias. *Psychol. Rev.* 74, 1–15. doi: 10.1037/h0024206

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *J. Cogn.* 2:16. doi: 10.5334/joc.72

Brysbaert, M., and Stevens, M. (2018). Power analysis and effect size in mixed effects models: a tutorial. *J. Cogn.* 1:9. doi: 10.5334/joc.10

Carpenter, P. A., and Just, M. A. (1983). "What your eyes do while your mind is reading," in *Eye Movements in Reading: Perceptual and Language Processes*, ed. K. Rayner (San Diego: Academic Press), 275–307. doi: 10.1016/B978-0-12-583680-7.50022-9

Drieghe, D. (2011). "Parafoveal-on-foveal effects in eye movements during reading," in *The Oxford Handbook of Eye Movements*, eds S. P. Liversedge, I. Gilchrist, S. Everling, B. J. Juhasz, and A. Pollatsek (Oxford: Oxford University Press), 839–855.

Drieghe, D., Pollatsek, A., Staub, A., and Rayner, K. (2008). The word grouping hypothesis and eye movements during reading. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 1552–1560. doi: 10.1037/a0013017

Eisenberg, P. (2013). *Das Wort. 4., Aktualisierte Und Überarb. Aufl.* Stuttgart: Metzler.

Gfroerer, S., Günther, H., and Bock, M. (1989). "Augenbewegungen und Substantivgroßschreibung - Eine Pilotstudie," in *Schriftsystem und Orthographie* Reihe Germanistische Linguistik. eds P. Eisenberg and H. Günther (Tübingen: Niemeyer), 111–135.

Granzow-Emden, M. (2002). *Zeigen und Nennen: Sprachwissenschaftliche Impulse zur Revision der Schulgrammatik am Beispiel der "Nominalgruppe".* Tübingen: Stauffenburg.

Günther, H., and Nünke, E. (2005). *Warum das Kleine groß geschrieben wird, wie man das lernt und wie man das lehrt.* Duisburg: Gilles und Franke. Available online at: http://www.koebes.uni-koeln.de/guenther_nuenke.pdf

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., et al. (2011). dlexDB–eine lexikalische datenbank für die psychologische und linguistische forschung. *Psychol. Rundsch.* 62, 10–20. doi: 10.1026/0033-3042/a000029

Henderson, J. M., and Ferreira, F. (1993). Eye movement control during reading: fixation measures reflect foveal but not parafoveal processing difficulty. *Can. J. Exp. Psychol.* 47, 201–221. doi: 10.1037/h0078814

Hohenstein, S. (2013). *Eye Movements and Processing of Semantic Information in the Parafovea During Reading.* Available online at: https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/6798

Hohenstein, S., and Kliegl, R. (2013). "Eye movements reveal interplay between noun capitalization and word class during reading," in *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, eds M. Knauff, M. Pauen, N. Sebanz, and I. Wachsmuth (Austin, TX: Cognitive Science Society), 2254–2259.

Hohenstein, S., and Kliegl, R. (2014). Semantic preview benefit during reading. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 166–190. doi: 10.1037/a0033670

Hyönä, J. (2011). "Foveal and parafoveal processing during reading," in *The Oxford Handbook of Eye Movements*, eds S. P. Liversedge, I. Gilchrist, S. Everling, B. J. Juhasz, and A. Pollatsek (Oxford: Oxford University Press), 822–838. Available online at: http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199539789.001.0001/oxfordhb-9780199539789-e-048

Hyönä, J., and Bertram, R. (2004). Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of nonfixated words influence the processing of fixated words during reading? *Eur. J. Cogn. Psychol.* 16, 104–127. doi: 10.1080/09541440340000132

Juhasz, B. J., and Pollatsek, A. (2011). "Lexical influences on eye movements in reading," in *The Oxford Handbook of Eye Movements*, eds S. P. Liversedge, I. Gilchrist, S. Everling, B. J. Juhasz, and A. Pollatsek (Oxford: Oxford University Press), 873–893. Available online at: http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199539789.001.0001/oxfordhb-9780199539789-e-048

Kennedy, A. (1998). "The influence of parafoveal words on foveal inspection time: evidence for a processing trade-off," in *Eye guidance in reading and scene perception*, ed G. Underwood (Oxford: Elsevier), 149–179.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Q. J. Exp. Psychol.* 429–455. doi: 10.1080/713755901

Kliegl, R. (2007). Toward a perceptual-span theory of distributed processing in reading: a reply to rayner, pollatsek, drieghe, slattery, and reichle (2007). *J. Exp. Psychol. Gen.* 136, 530–537. doi: 10.1037/0096-3445.136.3.530

Kliegl, R., Grabner, E., Rolfs, M., and Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *Eur. J. Cogn. Psychol.* 16, 262–284. doi: 10.1080/09541440340000213

Kluge, W. (1995). Vermutungen über ein rechtschreibliches Zaunkönigswissen. *Grundschulunterricht* 42, 4–7.

Mentrup, W. (1979a). *Die gemäßigte Kleinschreibung: Diskussion einiger Vorschläge zu ihrer Regelung und Folgerungen.* Mannheim: Bibliographisches Institut.

Mentrup, W. (1979b). *Die Groß- und Kleinschreibung im Deutschen und ihre Regeln: Historische Entwicklung und Vorschlag zur Neuregelung.* Tübingen: Narr.

Menzel, W. (1985). *Rechtschreibunterricht: Praxis und Theorie.* Prax. Dtsch.

Moser, H. (1958). "Groß- oder Kleinschreibung?: Ein Hauptproblem der Rechtschreibreform," in *Duden-Beiträge zu Fragen der Rechtschreibung, der Grammatik und des Stils*, ed. Dudenreaktion (Mannheim: Bibliographisches Institut), 9–48.

Nottbusch, G., and Jonischkait, J. (2011). "Einzeluntersuchungen zur GKS, GZS und Interpunktion," in *Weiterführender Orthographieerwerb Deutschunterricht in Theorie und Praxis*, eds U. Bredel and T. Reißig (Baltmannsweiler: Schneider-Verlag Hohengehren), 164–187.

Pauly, D. N., and Nottbusch, G. (2016). "Die Groß- und Kleinschreibung (beim Lesen) – kognitive Prozesse und didaktische Implikationen," in *System, Norm und Gebrauch - drei Seiten derselben Medaille? Orthographische Kompetenz und Performanz im Spannungsfeld zwischen System, Norm und Empirie, Thema Sprache – Wissenschaft für den Unterricht,* eds B. Mesch and C. Noack (Baltmannsweiler: Schneider Verlag Hohengehren GmbH), 122–145.

Pynte, J., and Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: evidence from reading English and French. *Vision Res.* 46, 3786–3801. doi: 10.1016/j.visres.2006.07.004

Pynte, J., Kennedy, A., and Ducrot, S. (2005). The influence of parafoveal typographical errors in eye movements in reading. *Eur. J. Cogn. Psychol.* 16, 178–202. doi: 10.1080/09541440340000169

Radach, R. (1996). *Blickbewegungen beim Lesen: Psychologische Aspekte der Determination von Fixationspositionen.* Münster/New York, NY: Waxmann.

Rat für deutsche Rechtschreibung (2018). *Deutsche Rechtschreibung: Regeln und Wörterverzeichnis. Aktualisierte Fassung des amtlichen Regelwerks entsprechend den Empfehlungen des Rats für deutsche Rechtschreibung 2016.* München und Mannheim. Available online at: https://www.rechtschreibrat.com/DOX/rfdr_Regeln_2016_redigiert_2018.pdf

Rayner, K., and Castelhano, M. S. (2008). "Eye movements during reading, scene perception, visual search, and while looking at print advertisements," in *Visual Marketing: From Attention to Action*, eds M. Wedel and R. Pieters (New York, NY: Taylor & Francis Group/Lawrence Erlbaum Associates), 9–42.

Rayner, K., and Kaiser, J. S. (1975). Reading mutilated text. *J. Educ. Psychol.* 2, 301–306. doi: 10.1037/h0077015

Rayner, K., and Schotter, E. R. (2014). Semantic preview benefit in reading English: the effect of initial letter capitalization. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1617–1628. doi: 10.1037/a0036763

Röber, C. (2011). "Konzepte des Erwerbs der Groß-/Kleinschreibung," in *Weiterführender Orthographieerwerb, Deutschunterricht in Theorie und Praxis*, eds U. Bredel and T. Reißig (Baltmannsweiler: Schneider-Verlag Hohengehren), 296–317.

Röber-Siekmeyer, C. (1999). *Ein anderer Weg zur Groß- und Kleinschreibung.* Leipzig: Klett-Grundschulverl.

Schotter, E. R., Angele, B., and Rayner, K. (2012). Parafoveal processing in reading. *Atten. Percept. Psychophys.* 74, 5–35. doi: 10.3758/s13414-011-0219-2

Tiffin-Richards, S. P., and Schroeder, S. (2015). Children's and adults' parafoveal processes in German: phonological and orthographic effects. *J. Cogn. Psychol.* 27, 531–548. doi: 10.1080/20445911.2014.999076

White, S. J., Johnson, R. L., Liversedge, S. P., and Rayner, K. (2008). Eye movements when reading transposed text: the importance of word-beginning letters. *J. Exp. Psychol.* 5, 1261–1276. doi: 10.1037/0096-1523.34.5.1261

Zabel, H. (1992). Groß- und Kleinschreibung im Deutschen: Möglichkeiten und Grenzen der sogenannten Artikel-Probe. *Muttersprache* 102, 60–85.

Zimmermann, F., and Heckel, B. (1986). "Untersuchungen zu Normverstößen bei der Groß- und Kleinschreibung," in *Forschungsinformation. Analysen Orthographischer Leistungen von Schülern der POS der DDR*, ed Pädagogische Hochschule Ernst Schneller Zwickau (Zwickau), 21–35.