Emotional Context and Predictability
in Naturalistic Reading Aloud

[author names and affiliations blinded for review]

Author Note

[author ORCiD blinded for review]

Correspondence concerning this article should be addressed to [blinded for review]

Abstract

A robust experimental literature has found that word frequency and lexical valence contribute to visual word processing at the level of the individual word. Extensions of this literature to simplified sentences have essentially corroborated single-word findings, albeit with important influences of the unfolding discourse context, which may strengthen or attenuate single-word effects. This study sought to extend current knowledge one step further, beyond standalone sentences or sentence pairs, by investigating how word frequency and lexical valence, along with their interactions, influence oral reading performance for multi-sentence stimuli in a naturalistic context. Lexical features were averaged over short passages of text, which were presented to participants on-screen simultaneously, and performance was assessed as reading speed, in words per second. Overall, we find that the same patterns emerge for multi-sentence oral reading as in the prior literature: strong frequency effects that benefit higher frequency content, a positivity bias that increases reading speed for more positive content, and an important interaction that disfavors relatively more negative (less positive), high-frequency content. We discuss these findings in light of possible interpretations based on associative connectivity in the mental lexicon, as well as oculomotor dynamics during naturalistic reading. Our data suggest that reading speed of multi-sentence texts is a viable alternative, and one that offers enhanced ecological validity, for investigations of visual word processing dynamics.

Keywords: naturalistic reading, lexical processing, word frequency, emotional valence, discourse context

Emotional Context and Predictability in Naturalistic Reading Aloud

In reading, the meaning of individual words is accessed and integrated into the developing discourse context. This process is influenced by a broad range of lexical features, including word frequency (Balota & Chumbley, 1984) and emotional valence (Larsen et al., 2006). These various properties interact during visual perception and lexical processing (Kuperman et al., 2014), and can be moderated or subsumed by discourse-level constraints such as the build-up of discourse context (Chou et al., 2020; Van Petten & Kutas, 1990). However, despite recent calls for increasing ecological validity within reading science (Schotter & Payne, 2019), much research in the field is based on highly-constrained experimental designs that use stand-alone words or simplified sentences: there remains scant evidence for how linguistic features impact reading behavior in naturalistic contexts that more closely mirror our day-to-day interactions with the written word. In the current study, we sought to close this gap in the literature by examining the influence of word frequency and lexical valence on reading speed within the context of naturalistic, passage-length stimuli. Manipulating the emotional valence of passages also allowed us to probe the impact of higher-level context processing within an ecologically-valid setting. Below, we first review prior work at the level of individual words and simplified sentences in order to motivate our hypotheses regarding how lexical features and discourse-level constraints impact reading of passage-length stimuli.

Effects of word frequency are well-attested in the visual word recognition literature (Balota & Chumbley, 1984; Balota & Spieler, 1999). Specifically, words that are encountered more often–that is, high-frequency words–are typically recognized faster than low-frequency words. This is particularly true in contexts where frequency-based expectation strategies can improve task performance, such as making a lexical decision (Barriga-Paulino et al., 2022; Kuchinke et al., 2007; Larsen et al., 2008; Scott et al., 2012) or reading single words aloud (Fischer-Baum et al., 2014). It has been argued that word frequency effects may emerge from denser and/or stronger connections within the mental lexicon's associative network (Hulme et al., 1991; Stuart &

Hulme, 2000). As a result of such connectivity, high-frequency words are thought to exhibit a higher resting activation that may serve to reduce the threshold for activation (Plaut et al., 1996). This view is supported by pupillometry studies showing lower peak dilations for high-frequency words during lexical decision (Haro et al., 2017; Kuchinke et al., 2007), as well as in eye-tracking, where high-frequency words demonstrate shorter fixation durations (Kliegl et al., 2004; Scott et al., 2012), indicating that readers process high-frequency words more rapidly.

Message-level constraints, however, can override the effects of word-level properties like frequency. For example, studies show that readers use prior context to anticipate the semantic features of upcoming words (see Federmeier, 2007 for a review), leading to attenuated frequency effects when discourse context is high. In electrophysiology, the N400 event-related potential (ERP) component has been the focus of significant research in predictive processing for visual word recognition. The N400 is believed to index access to semantic memory, such that incoming content that binds more easily into ongoing neural activity produces smaller N400 amplitudes. In single-word presentation, high-frequency words demonstrate smaller N400 amplitudes than low-frequency words (Barber et al., 2004; Fischer-Baum et al., 2014; Rugg, 1990). Within sentential frames, however, effects of word frequency on N400 amplitudes are diminished for words that are more predictable from the established semantic or syntactic context (open-class words: Payne et al., 2015; Payne & Federmeier, 2019; Van Petten & Kutas, 1990, 1991; closed-class words: Payne et al., 2015; Van Petten & Kutas, 1991).

Across a variety of experimental paradigms, word frequency interacts with other lexical properties to modulate performance. One such property is emotional valence: words range on an emotional continuum from intrinsically appetitive ("kitten") to fundamentally aversive ("puke"). There is substantial evidence that valence impacts visual word processing (Barriga-Paulino et al., 2022; Estes & Adelman, 2008; Estes & Verges, 2008; Herbert et al., 2008, 2009; Keuper et al., 2013; Kuperman et al., 2014; Larsen et al., 2008). Moreover, prior work in both lexical decision response times and single fixation duration in eye-tracking has demonstrated an

interaction between word frequency and emotional valence that selectively disfavors words that are both high-frequency and relatively more negative (Kuchinke et al., 2007; Méndez-Bértolo et al., 2011; Scott et al., 2009, 2012, 2014).

The current understanding of how word frequency and emotional valence influence reading behavior is built on a foundation of highly-constrained laboratory tasks, particularly the lexical decision paradigm. It is difficult to generalize findings to more naturalistic reading contexts, however, given the additional task demands involved in traditional lexical decision tasks (i.e., explicit lexicality decisions and execution of motor movements to register these decisions) and the limitations of processing context-less, standalone words. Beyond lexical decision paradigms, many sentence-level reading tasks present words on-screen individually, with either the participant–or, in the case of rapid serial visual presentation (RSVP), the experimenter– controlling speed of presentation. However, the onset of the N400 has been found to shift by more than 100 milliseconds in natural reading compared to RSVP designs (Kliegl et al., 2012) and single-word presentation precludes parafoveal previewing of upcoming content, which has been found to have a substantial impact on the neural time course of visual word processing. N400 effects track semantically unexpected words presented in the parafovea (Li et al., 2022) and such N400 effects for parafoveal viewing are not duplicated when target words are subsequently processed foveally (Payne et al., 2019). These modulations of N400 amplitudes imply that word form processing involves, at least in part, a fast process that can be completed in the parafovea. It therefore remains unclear whether the extant knowledge of lexical processing in highly-constrained experimental paradigms will generalize to naturalistic tasks, especially where readers are able to visually sample upcoming content parafoveally.

While prior work has focused on investigating word frequency effects on reading/processing speed at the single-word level, either in single-item presentation or embedded within carefully manipulated sentential frames, natural interactions with written text entail engaging with longer-form content; in these scenarios, the time it takes to read a passage as a whole (or a partial

excerpt of a passage) may be of greater relevance. Traditional effects of word frequency that are observable at the single-word level may or may not manifest in the aggregate (i.e., across multiple sentences within a passage). Specifically, it is unknown whether the time to read a given passage can be predicted by the average word frequency of the words comprising that passage, and, if so, whether the relation between average word frequency and oral reading speed would reflect the traditional pattern observed for words read in isolation—faster speeds for higher frequency. Relatedly, given that prior work at the single-word level demonstrates that message level constraints can attenuate the effects of word frequency on processing speeds, it also remains unclear whether message-level constraints would similarly impact any possible link between average word frequency and reading speed at the passage level.

Previous studies investigating the role of emotional valence on reading/processing speed have been similarly limited and, in much the same manner, it is unknown whether effects of lexical valence, and the ways in which valence interacts with word frequency, would likewise manifest in the aggregate, across an entire passage. Prior work suggests that emotional evaluation of multi-sentence content is, to some degree, "the sum of its parts," demonstrating a linear relation between subjective evaluation of the overall emotional tone of a passage and the simple mathematical averaging of the lexical valence of its content words (Bestgen, 1994; Hsu et al., 2015). Using four-sentence snippets of Harry Potter texts, presented in an fMRI scanner and displayed consecutively, Hsu et al. (2015) further found that subjective ratings of passage-level emotional tone and the average of the lexical valence ratings of all words in the passage were comparable predictors of brain activity during reading. Thus, it is reasonable to expect that lexical valence effects may also average over the course of a passage, and any underlying facilitation of higher valenced words should be borne out by reading speeds over multiple sentences.

In the current study, we sought to investigate potential effects of average word frequency and average valence on naturalistic oral reading speed. To this end, we constructed twenty

short passages on diverse topics (one topic per passage, each 140-223 words), with each passage coded for average word frequency and valence. Critically, in the style of a literary mid-course turn, passages were constructed such that the average emotional valence switched between the first and second half of the passage. Sudden aesthetic or narrative changes, such as the dramatic peripeteia (Lucas, 1923) or the poetic volta (Theune, 2007), are common in literature, although the neurocognitive effects of such aesthetic devices are only recently starting to be studied scientifically (Jacobs, 2015). Participants read these passages aloud in a naturalistic setting, and we used the audio recordings to calculate the time elapsed during reading. Reading comprehension was assessed following each passage to ensure task engagement. Collectively, this naturalistic design allowed us to test whether traditional effects of word frequency and emotional valence, as well as potential interactions between these lexical features and interactions with higher-level discourse context, impacted oral reading speed at the multi-sentence level.

Our hypotheses were premised on the assumption that averaged lexical effects would resemble effects previously demonstrated at the single-item level. We anticipated that passage halves with higher average word frequency would be read faster than those lower in average word frequency. Based on theories of negativity bias, whereby relatively more negative stimuli preferentially capture attention and slow task-based responses (e.g. automatic vigilance: Pratto & John, 1991), as well as theories that posit a positivity bias that enhances responses to relatively more positive stimuli (e.g., the density hypothesis: Unkelbach et al., 2008), one would predict that reading speeds for more negative passage halves (i.e., those with lower average lexical valence) would be slower than reading speeds for more positive passage halves (i.e., those with higher average lexical valence). Importantly, prior work has demonstrated an interaction between word frequency and emotional valence that selectively disfavors relatively more negative high-frequency words (Kuchinke et al., 2007; Méndez-Bértolo et al., 2011; Scott et al., 2009, 2012, 2014). We therefore anticipate a similar pattern in oral reading speeds, with

slower speeds for relatively more negative passage halves of higher average word frequency compared to more positive, high-frequency passage halves.

Given that discourse context can mitigate the positive relation between word frequency and reading speed (Payne et al., 2015; Payne & Federmeier, 2019; Van Petten & Kutas, 1990, 1991), we expected that the positive relation between average word frequency and reading speed would be most pronounced in the first half of each passage, when discourse context is relatively low. However, in the second half of each passage, and following a shift in emotional valence, there are two possible hypotheses. First, if a shift in emotional valence does not disrupt the accrual of higher-level discourse context (Delaney-Busch & Kuperberg, 2013), then the relatively high discourse context available in the second half of each passage would lead us to predict a reduction in the positive link between word frequency and reading speed. However, if, due to relative differences in the positivity of the semantic content contributing to different predictive frames, the midpassage shift in emotional valence does, in fact, disrupt the accrual of higher-level discourse context, then it is reasonable to assume that oral reading speed would revert to being primarily driven by word frequency effects in the lower discourse context available following such a disruption. In this case, the positive association between word frequency and reading speed expected for the first half of the passage would be likely to manifest in the second half, as well.

Methods

Participants

Fifty-eight students from a psychology student participant pool at [blinded for review] (see Table 1 for participant demographics) participated in this experiment for course credit from January-June 2022. The initial sample size was based on power analyses performed in G*Power 3.1.9.6 (Faul et al., 2009). Inclusion criteria included normal or corrected-to-normal vision, no diagnosis of colorblindness, and no prior head injury. Participants were required to have an internet connection, webcam, and microphone, and to express willingness to record

themselves as part of the study. Additionally, participants were required to have a desktop or laptop computer on which to complete the study, as the experimental task was not designed to be compatible with a phone or tablet. Research protocols were approved by the Institutional Review Board of [blinded for review] and informed consent was obtained from all participants. Participants reported no history of communication disorders. Given that 72% of residents in [blinded for review] County, where our research was performed, speak a language other than English at home (U.S. Census Bureau, 2015), we chose to include both monolingual and multilingual participants. All participants self-reported having learned English prior to the age of six. Prior work in bilinguals who acquired English at an early age demonstrates comparable behavior in valenced lexical decision tasks to monolingual English participants (Kazanas & Altarriba, 2016). Due to a strong female bias in the [blinded for review] psychology participant pool, the participant population was heavily female (> 90%).

Table 1

Participant Demographics

	Sex			Pronouns			
	Female	Male	She/her	He/him	They/them	Other	Undisclosed
Total	54 (93.1%)	4 (6.9%)	50 (86.2%)	4 (6.9%)	1 (1.7%)	1 (1.7%)	2 (3.4%)
Speed Analyses	50 (92.6%)	4 (7.4%)	46 (85.2%)	4 (7.4%)	1 (1.9%)	1 (1.9%)	2 (3.7%)

|	Race/Ethnicity					
	Hispanic, Latino/a/x, or Spanish Origin	White	Asian	Black or African American	American Indian or Alaska Native	Undisclosed
Total	43 (74.1%)	7 (12.1%)	3 (5.2%)	3 (5.2%)	1 (1.7%)	1 (1.7%)
Speed Analyses	41 (75.9%)	7 (13.0%)	3 (5.6%)	1 (1.9%)	1 (1.9%)	1 (1.9%)

|	Age			Socioeconomic Class Affiliation			
	Mean	SD	Range	Poor	Working Class	Middle Class	Affluent
Total	22.66	4.6	18 - 40	2 (3.4%)	23 (39.7%)	31 (53.4%)	2 (3.4%)
Speed Analyses	22.67	4.77	18 - 40	2 (3.7%)	21 (39.0%)	29 (53.7%)	2 (3.7%)

Stimuli

Twenty reading passages (see supplement) were drafted with the explicit intent of serving as quasi-naturalistic stimuli. In order to investigate behavioral differences in processing and reading aloud content comprised of more positive or more negative words, ten passages were constructed to be relatively more positively valenced for the first half of the passage and relatively more negatively valenced for the second half ("positive-to-negative"); ten passages were the reverse ("negative-to-positive"). Passages ranged in length from 140 to 223 words (219 to 363 syllables) with a "switch" word positioned at the midway point; in this way, each passage half ranged in length from 70 to 113 words (101 to 196 syllables). Lexical valence for each passage half was calculated by averaging across valence ratings available in the Warriner et al. (2013) dataset; for words with no relevant entry in the corpus, the median score was imputed. In this way, our analyses avoided overly distorting the distinction between more positive passage halves, which ranged from 5.66 to 5.86 (mean = 5.75) on a 9-point scale, and relatively more negative passage halves, which ranged from 4.76 to 5.01 (mean = 4.93). Note that hypotheses for the current study conceptualize valence as existing on a continuum and do not depend on distinguishing between explicitly valenced vs. "neutral" stimuli. As such, we refer to passage halves with higher/lower mean valence ratings as "relatively more positive" and "relatively more negative," respectively (for convenience, we also refer to these passage halves in the shorthand as "positive" and "negative" within the manuscript). To create the emotional "volta", switch words were designed to be dramatic points of departure from the ongoing passage valence: positive-to-negative switch words fell below 2.5 and negative-to-positive switches landed above 7.5 on the same 9-point scale. An ANOVA model was used to confirm successful manipulation: relatively more positive/negative passage halves were significantly different from one another in average valence rating $[F(1,36) = 1638.01, p < 0.001]$ whereas there was no significant difference in valence rating as a function of passage position $[F(1,36) = 0.14, p = 0.715]$ nor any interaction $[F(1,36) = 3.57, p = 0.067]$. No explicit effort was

made to control for word frequency during the drafting of stimulus passages, resulting in content

that varied naturally in frequency and allowing us to analyze our results as a function of the

average word frequency of passage halves. For frequency analyses, we used the log-

transformed frequency counts from the SUBTLEXus (Brysbaert & New, 2009) corpus of

American English. Similar to our handling of averaged valence values, we imputed the median

frequency rating for words with no relevant entry in the corpus. Beyond word frequency and

lexical valence, naturalistic reading stimuli can differ on various dimensions, such as syntactic

complexity and average word length, which relate directly and indirectly to characteristics known

to affect reading speed. Differences in these dimensions can be operationalized with

standardized measures of reading ease (Flesch, 1948). The passage stimuli for the current

study were constructed such they did not differ in reading ease, neither as a function of passage

position [$F(1,36) = 0.06$, p = 0.806] nor valence [$F(1,36) = 0.61$, p = 0.439] nor their interaction

[$F(1,36) = 2.84$, p = 0.100]. Additionally, we confirmed that passage halves did not differ in

number of words, neither as a function of passage position [$F(1,36) = 0.03$, p = 0.876] nor

valence [$F(1,36) = 0.04$, p = 0.839] nor their interaction [$F(1,36) = 0.14$, p = 0.711]. Likewise,

passage halves did not differ in number of syllables, neither as a function of position

[$F(1,36) = 0.01$, p = 0.935] nor valence [$F(1,36) = 0.18$, p = 0.673] nor their interaction

[$F(1,36) = 0.80$, p = 0.377]. Passage halves also did not differ in average number of syllables

per word, neither as a function of position [$F(1,36) = 0.14$, p = 0.712] nor valence

[$F(1,36) = 0.21$, p = 0.648] nor their interaction [$F(1,36) = 1.82$, p = 0.186]. Due to experimenter

error, a typographical mistake was present in the final sentence of one of the passages

("broccoli" passage); this passage was therefore dropped from all analyses.

Figure 1

Passage Characteristics
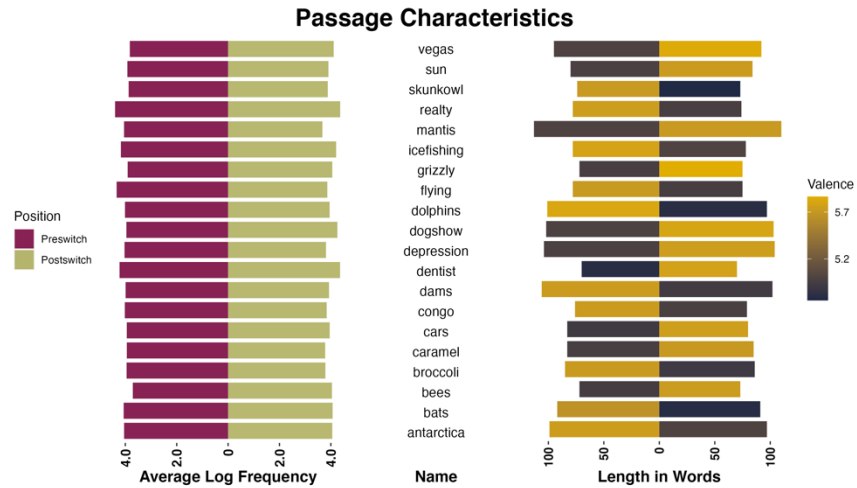
Note. On the left, bar length represents the average log frequency for each passage half (preswitch/postswitch). On the right, bar length represents the length (in words) for each passage half (preswitch/postswitch), color represents average valence for that passage half.

Figure 2

Task Procedure

Zoom setup

Instructions

Passage text

Comprehension question

Optional break
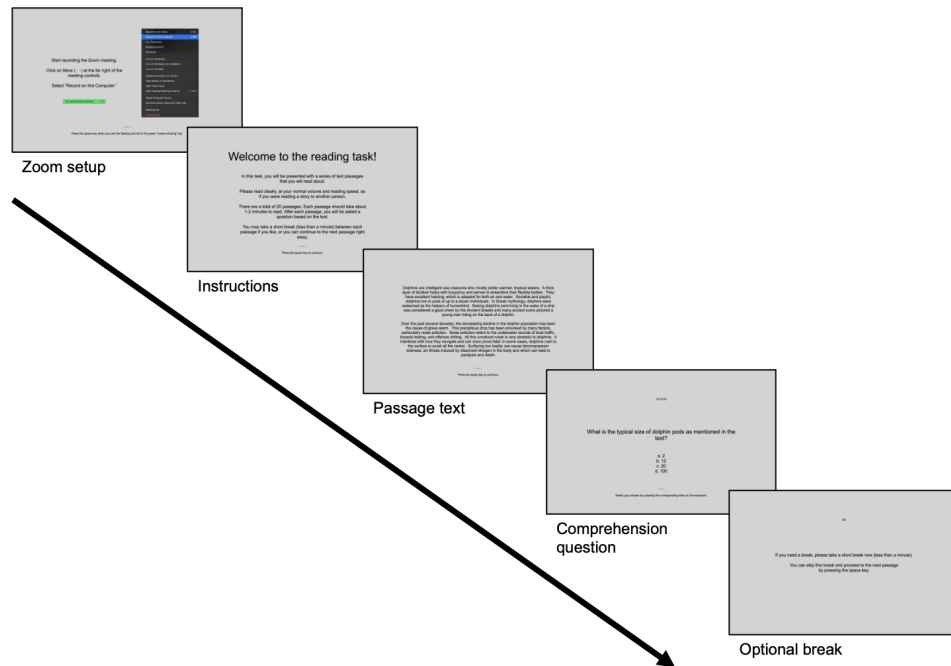
Note. The PsychoPy task, hosted by Pavlovia, led participants through setting up their Zoom recording, after which they were given task instructions. Participants proceeded through task screens by pressing the spacebar. There were no time limits for each screen. Screens with passage text were each followed by a multiple-choice comprehension question based on the prior passage, which participants answered by selecting the associated keyboard letter ('a', 'b', 'c', or 'd'). The comprehension question letter press took participants to an optional 60-second break screen, which could be skipped by pressing the spacebar to continue to the next passage text.

Procedure

Using either a desktop or laptop computer, participants completed self-paced questionnaires relating to demographic information via REDCap (Harris et al., 2019) before clicking a link to a PsychoPy (Peirce et al., 2019) task (version 2021.2.3), hosted by Pavlovia (pavlovia.org; see:

Bridges et al., 2020). Participants also completed a battery of questionnaires and additional behavioral tasks; a subset of the questionnaire data was analyzed in a series of control analyses to rule out potential confounds arising from demographic or affective state/trait variables (see supplement for details) whereas other data were beyond the scope of the current report and are not discussed further. Participants were informed that they would read twenty passages aloud, that each passage would take 1-2 minutes to read, that they would be asked to answer a comprehension question after each text, and that they would be able to take a short break between passages. Using Zoom (Zoom Video Communications, Inc., San José, California) to record their screen and microphone, they were instructed not to "pre-read" passages, but rather to begin reading aloud immediately and to read each passage at their normal volume and speed. For each passage, all text appeared on-screen at once, as black Arial text, centered on a light grey background. The experimental task used the "height" unit for font sizing (PsychoPy: Peirce et al., 2019), so that text would scale for each individual user's screen without distortion. After reading each passage, participants pressed the spacebar to proceed to a multiple-choice comprehension question, which served to confirm task engagement. For each question, four possible answers were presented (chance performance = 25%) and questions were drawn equally from the four categories of passage halves: preswitch positive, preswitch negative, postswitch positive, and postswitch negative. That is, we counterbalanced the location in the passage from which the information required to successfully answer the comprehension question was selected. There was no time limit for reading each passage nor for answering each comprehension question. Following each comprehension question, participants were given an optional 60-second break prior to proceeding to reading the next passage. Given that each reading passage was either positive-to-negative or negative-to-positive, we aimed to keep valence switches passage internal, rather than allowing a valence shift between passages. This was achieved by semi-randomization of the passage stimuli within the task setup: ten sets of passage pairs (positive-to-negative +

negative-to-positive) were created so that their presentation could be randomized across participants. A second set of passage pairs (negative-to-positive + positive-to-negative) was additionally created, and participants were randomly assigned to one of the two sets, such that half of the participants began the experiment with a positive-to-negative passage and half began with a negative-to-positive passage.

Acoustic Preprocessing

In order to extract reading speed, timestamps were obtained using Praat, version 6.2.14 (Boersma & Weenink, 2001). For each passage, for each reader, three time points were recorded: the onset of the first syllable in the first word of the passage, the onset of the first syllable in the switch word (denoting the border between the first and second half of the passage), and the end of the coda of the last syllable in the passage. Two coders trained by the first author annotated 57% of the recordings. The first author annotated the remaining recordings and cross-annotated 20% of the timestamps for each participant that were annotated by each of the two coders. The psych package, version 2.2.5 (Revelle, 2022) in R, version 4.1.1 (R Core Team, 2021), was used to calculate inter-rater reliability for the 20% of cross-annotated recordings from a two-way agreement model; this was done individually for each coder. Single measure intra-class correlation coefficients were found to be very high (coder 1: 192 timestamps (16 participants × 4 passages × 3 timestamps each), kappa > 0.999; coder 2: 276 timestamps (23 participants × 4 passages × 3 timestamps each), kappa > 0.999).

Reading speed per participant was calculated as the total number of words in the passage half divided by the total number of seconds spent reading, such that higher values for reading speed correspond to faster rates of oral reading. We chose to operationalize reading speed as words per second given that our planned predictors, valence and frequency, were both lexical level features. Additionally, prior work has used words per time-unit to analyze standardized

reading rates of passage-length stimuli (Lewandowski et al., 2003; Trauzettel-Klosinski & Dietz, 2012), including assessments of screen-based reading (Wallace et al., 2022).

Statistical Analyses

Reading comprehension questions after each stimulus passage were included to confirm task engagement. We removed four participants from further analysis due to low overall accuracy (≤ 50%; chance performance = 25%) across passage comprehension questions. Overall accuracy for the remaining participants was 79.7% (SD = 11.3%).

Passages for which the reading speed of either the first or second passage half could not be calculated (e.g., if the participant failed to read the full passage text aloud in the recording) were removed prior to analysis. Given that participants recorded themselves outside a laboratory environment, infrequent instances of participant interruption (for example, by family members) were observed during the task. For this reason, we also removed passages where the difference in reading speed between the first and second passage halves was ±3 standard deviations from the individual delta of each participant. In combination, this eliminated 9 (0.9% of all passages) from further analysis.

Only passages for which the correct response was provided to the reading comprehension question were considered for the analysis of reading speed. In addition, passages whose reading speed was ±3 standard deviations from the mean reading speed were removed prior to analysis. These trimming procedures resulted in the removal of a further 218 passages (21.2%).

To analyze the effects of stimulus characteristics on reading speed, lme4, version 1.1-28 (Bates et al., 2015), and the lmerTest wrapper, version 3.1-3 (Kuznetsova et al., 2017), in R, version 4.1.1 (R Core Team, 2021), were used to construct a mixed effects model via restricted maximum likelihood estimation with reading speed, calculated in words per second, as the dependent measure. For each passage half, position (preswitch/postswitch), average valence, average frequency, and their interactions were entered into the model as fixed effects, with

random intercepts per participant and per passage. The position variable was contrast coded (preswitch: -1, postswitch: +1) and, following outlier removal, the two continuous variables were mean centered across all data points. Effect sizes are reported as unstandardized β coefficients for ease of interpretation and 95% bootstrapped confidence intervals were computed via the confint() function in lme4. Johnson-Neyman intervals were calculated using the sim_slopes() function from the interactions package, version 1.1.5 (Long, 2019).

Transparency and Openness

All data exclusions and manipulations, along with all software employed for analysis, are reported above. Results plots were created using: ggplot2, version 3.4.0 (Wickham, 2016); interactions, version 1.1.5 (Long, 2019); and gridExtra, version 2.3 (Baptiste, 2017). This study's design and its analysis were not pre-registered. Study stimuli are provided as supplementary information, and additionally available at [blinded for review]. Data and analysis code have been made publicly available at [blinded for review].

Results

As described above, reading speed (in words per second) for each passage half was analyzed via a linear mixed effects model with position (preswitch/postswitch), average valence, average frequency, and their interactions as fixed effects, and random intercepts per participant and per passage. The model identified a significant main effect of frequency ($\beta = 0.71$, SE = 0.06, 95% parametric bootstrapped CI [0.59, 0.83], $p < 0.001$) such that, on average, reading speeds were faster for passage halves of higher average word frequency. A main effect of valence was also identified ($\beta = 0.06$, SE = 0.02, 95% parametric bootstrapped CI [0.02, 0.09], $p < 0.001$), such that, on average, relatively more positive passage halves were also read at faster speeds. Crucially, these main effects of frequency and valence were qualified by a significant frequency × valence interaction ($\beta = 0.29$, SE = 0.1, 95% parametric

bootstrapped CI [0.07, 0.47], p = 0.003). The nature of this interaction was such that higher average word frequency was associated with faster reading speeds, with this effect being stronger in relatively more positive passage halves (see Figure 3). In order to assess at which values of each predictor (frequency, valence) the slope of the other was significant, we calculated the Johnson-Neyman intervals (Johnson & Fay, 1950). For all observed values of lexical valence, the frequency slope was significant. On the other hand, the effect of valence on reading speed was significant only for passages with high average frequency (frequency values greater than -0.08 following mean-centering). No main effects or interactions involving passage position were identified.

Individual differences have the potential to influence processing of affective content (Carstensen & DeLiema, 2018; Foland-Ross & Gotlib, 2012; Lang & Cuthbert, 1984); to rule out this potential confound, we re-ran our primary statistical model (described above) while also controlling for a series of individual differences variables, including: age, sex, trait depression/anxiety symptoms, and affective state (see supplement for details). Briefly, in each additional model, we added the individual difference measure of interest, as well as its interaction with average valence. Across all additional statistical models, the results of our primary model remained qualitatively unchanged; see supplement for complete results.

Figure 3

Frequency × Valence Interaction Effects in Reading Speed
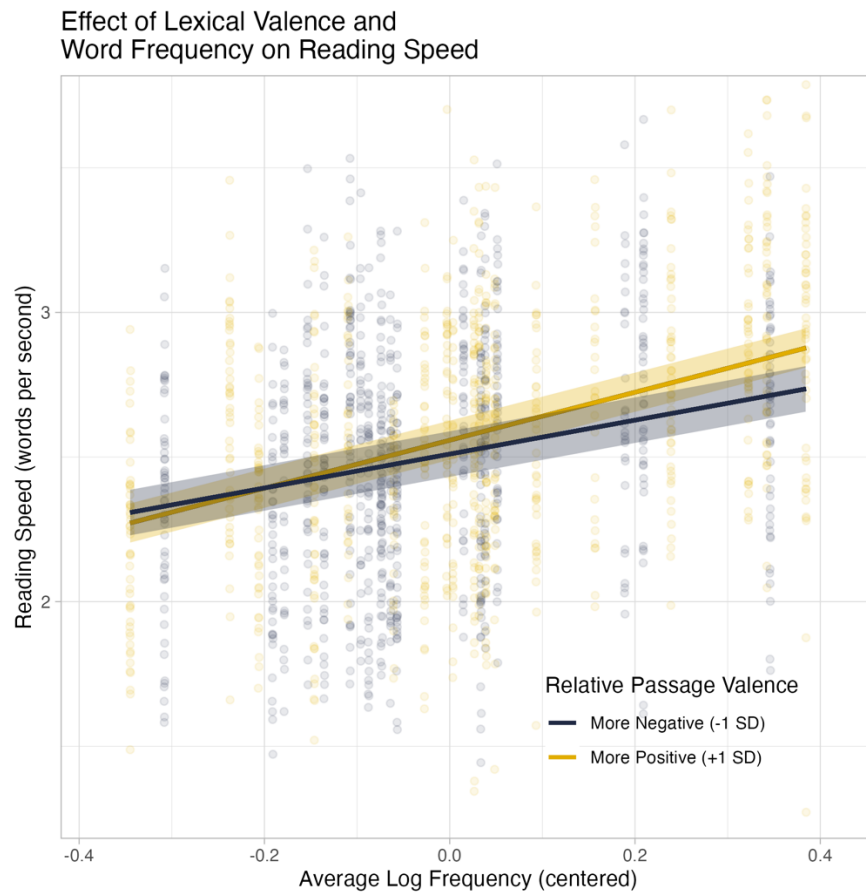
Note. Shape of the frequency × valence interaction (p = 0.003) on reading speed. Points represent individual performance on each passage half. More negative passage valence values represent performance on passage halves whose centered valence rating was less than one standard deviation (SD) below the mean; more positive values represent performance on passage halves whose centered valence rating was greater than one standard deviation (SD) above the mean.

Discussion

We examined the influence of word frequency and lexical valence on reading speed in a naturalistic oral reading task. Effects of word frequency, lexical valence, and their interaction have been found at the single-word level (e.g., Balota & Chumbley, 1984; Kuperman et al.,

2014; Larsen et al., 2006) and within the context of simplified sentences (e.g., Holt et al., 2009; Scott et al., 2012; Van Petten & Kutas, 1990), but little is known about whether these effects accumulate over the course of multi-sentence passages as in everyday reading. As a corollary, we also manipulated emotional valence within each passage to investigate whether a sudden shift in valence during the reading of a passage would disrupt the processing of the higher-level discourse context, further impacting effects of word frequency on reading speed.

Overall, we found that effects of lexical frequency on oral reading speed, measured in words per second over a naturalistic passage, are generally consistent with the pattern of results reported in more traditional studies that employ highly-constrained experimental designs and use stand-alone words or simplified sentences. On average, reading speed displayed traditional frequency effects, with faster speeds for passage halves with higher average word frequency. Likewise, reading speed demonstrated, on average, a positivity bias, with faster reading speeds for passage halves that were relatively more positive in valence. Crucially, the effects of frequency and valence were qualified by an interaction, providing a generalization of previously reported interactions between frequency and valence whereby the processing of high-frequency negative words is slowed (Kuchinke et al., 2007; Méndez-Bértolo et al., 2011; Scott et al., 2009, 2012, 2014). With respect to modulations of frequency effects based on higher level discourse, we did not observe a significant position × frequency interaction. This null result suggests that the midpassage shift in valence may have disrupted the accrual of higher-level discourse context, resulting in reading speeds being driven primarily by average word frequency in both the first and second halves of each passage. We discuss each of these findings below.

Word frequency effects extend to passage-length stimuli

In single word reading aloud, response latencies for high-frequency words are shorter than those for low-frequency words (Balota & Spieler, 1999; Fischer-Baum et al., 2014). That is, participants require less time to process and produce words that are more frequently

encountered. A similar pattern emerges in eye-tracking research with single sentence stimuli, in which high-frequency words are fixated for a shorter period of time than low-frequency words (Kliegl et al., 2004; Scott et al., 2012). In this study, we find that the speed advantage for high-frequency stimuli extends beyond single words and short sentences to multi-sentence texts, with passages of higher average word frequency being read faster in words per second.

Word frequency effects are thought to be related to higher connectivity in the associative network of the mental lexicon (Hulme et al., 1991; Stuart & Hulme, 2000). Speeded naming of standalone words, which is similar in nature to the oral reading in our task, has previously been shown to display frequency effects, with lower frequency words produced more slowly (Balota & Chumbley, 1984; Balota & Spieler, 1999; Larsen et al., 2008). Longer-form oral reading, however, benefits not only from parafoveal preview, but also from higher levels of semantic and syntactic context that can be used to predict upcoming words. Similar to production times in speeded naming, the amplitude of the N400 ERP component is attenuated by word frequency during single-word presentation. However, when words are combined into sentential frames, the accumulating discourse context serves to modulate frequency effects on the N400. For example, Van Petten & Kutas (1990) found larger N400 amplitudes for low frequency words, but only in early sentence positions. One might therefore anticipate that potential frequency effects would be obscured in passage-length stimuli. Instead, we find robust frequency effects across passage halves that are 70-113 words in length.

We posit that longer passage lengths offer a larger window for underlying word frequency effects to become salient, due, in part, to an interaction between oculomotor control and lexical access. During the ongoing dynamics of oral reading, the reader has simultaneous visual access to past, current, and upcoming words, and the distance that separates the word currently being produced from the word on which the eyes are fixated is called the eye-voice span (Laubrock & Kliegl, 2015). In naturalistic paradigms, low frequency words are less likely to be visually skipped (Kliegl et al., 2004) and generally have longer fixation times (Kliegl et al.,

2006). That is, low frequency words can cause the eye-voice span to shrink, presumably due to the need to allocate more processing resources to the fixated word (Laubrock & Kliegl, 2015). In this way, higher average word frequency may allow readers to maximize the eye-voice span closer to the threshold of the buffer in which the visual code is translated into phonological working memory in anticipation of articulatory output (Wagner & Torgesen, 1987). At maximal span, reading speed would be (theoretically) bounded only by the articulatory limitations of the vocal tract. We do not propose that such a boundary is reached in our study, but rather suggest that increased eye-voice span during oral reading of content with higher average frequency might increase the speed of articulatory output compared to speeds when the span is smaller.

Positivity promotes faster reading

In speeded naming, response times are faster for relatively more positive words (Estes & Adelman, 2008; Kuperman et al., 2014; Larsen et al., 2008). Similarly, more positive words demonstrate stronger affective priming effects (Kazanas & Altarriba, 2015, 2016; Lüdtke & Jacobs, 2015; Sass et al., 2012; Unkelbach et al., 2008). Many researchers (Hofmann & Jacobs, 2014; Lüdtke & Jacobs, 2015; Unkelbach et al., 2008) have interpreted this positivity bias in affective priming as the result of greater semantic-associative clustering among more positive words. In this way, higher valence words bolster spreading activation across a densely-connected associative network while, conversely, the semantic distinctiveness of lower valence words slows their evaluation and integration into the unfolding discourse context. Importantly, such affective priming effects are strongly moderated by stimulus onset asynchrony, with effects most salient when the lag between prime and target is very short (Hermans et al., 2001). Similar to the discussion of frequency effects above, longer form naturalistic reading may create larger windows in which a positivity bias in the processing of individual words can become compounded and lead to overall faster reading speeds for relatively more positive content. Nonetheless, our findings introduce an important caveat: relatively more positive passage

halves were only read faster at higher levels of average frequency. Phrased differently, the speed distinction between passage halves that were relatively more or less positive in their average valence disappeared when average word frequency was low. We discuss this interaction between frequency and valence next.

Higher-frequency, lower-valence content is disadvantaged

When interactions between word frequency and lexical valence are considered in the prior literature, it is specifically high-frequency negative words that underperform, demonstrating slowed response times in lexical decision (Kuchinke et al., 2007; Méndez-Bértolo et al., 2011; Scott et al., 2009, 2014) and longer fixation durations in eye-tracking (Scott et al., 2012). A similar pattern emerges in our results. Thus, interactive effects of valence and word frequency previously demonstrated in traditional, highly-constrained experimental designs appear to generalize to naturalistic oral reading of multi-sentence passages.

Yap & Seow (2014) argue that words with greater familiarity and/or meaningfulness, including emotional words, have richer semantic representations that may serve to provide stronger feedback to lexical representations, thereby promoting faster access. Indeed, semantic neighborhood density, one measure used to operationalize semantic richness, correlates strongly with word frequency (M. J. Yap et al., 2012), and more frequent words are more likely to be nodes in word association networks (De Deyne & Storms, 2008). Similarly, higher valence words have demonstrated greater semantic density in the associative network of the mental lexicon (Unkelbach et al., 2008) while lower valence words, both alone and in n-grams up to n = 4, have been shown to contain more information in an information theoretic sense (Garcia et al., 2012), suggesting greater distinctiveness. That is, relatively more positive words ("sweet," "kind") are more alike than words that are more negative ("cruel," "rude"), and therefore more densely associated in the mental lexicon. In this case, both high average frequency and positive valence may benefit from the same underlying mechanism, namely tighter relations in a small-

world associative network, which facilitates construction of an ongoing model of the discourse context during oral reading and reduces word-to-word processing time.

From another perspective, we might consider that words which require additional processing, either because they are infrequently encountered or higher in information content, may operate to reduce the eye-voice span during reading aloud. Above, we suggested that the eye-voice span has a theoretical ceiling at which point reading speed is limited only by articulatory motor control. It is likewise reasonable to propose that, for a given reader of a given text, the eye-voice span has a hard floor; namely, when the eyes are fixated on the same word that is being verbally produced (for example, see figure 1 in Laubrock & Kliegl, 2015). In the context of reading aloud a multi-sentence passage, the additional processing required for low frequency words may have similar time dynamics regardless of valence, leading to the pattern of results observed in the current study: similar average reading speeds for passage halves with lower average word frequency, regardless of emotional valence.

It is important to note that it remains speculative as to whether the pattern of results observed here are best explained by lexical access, oculomotor control, or both. The primary goal of the present study was to confirm that traditional lexical effects of frequency and valence are visible in longer-form, naturalistic reading aloud. As such, our design does not allow us to definitively adjudicate between competing interpretations of our results. Further research is necessary to shed light on the extent to which lexical access and oculomotor control contribute, independently or in tandem, to slower oral reading speeds for higher-frequency, lower-valence content.

Shifts in valence may disrupt the discourse context

Given the body of literature supporting traditional word frequency effects (Balota & Chumbley, 1984; Balota & Spieler, 1999; Barriga-Paulino et al., 2022; Fischer-Baum et al., 2014; Kuchinke et al., 2007; Larsen et al., 2008; Scott et al., 2012), we expected a pronounced

positive relation between average word frequency and reading speed in the first half of each passage, when discourse context is relatively low and word frequency effects are more likely to drive reading performance. Following the midpassage shift in emotional valence, however, any reduction in the positive association between word frequency and reading speed (i.e., the presence of a significant position by frequency interaction) would depend on the degree to which the higher-level discourse context were disrupted by this shift in valence. That is, if midcourse turn in valence did, indeed, disrupt the accrual of higher-level discourse context, then reading speed would presumably revert to being driven primarily by word frequency effects, in which case the positive association between word frequency and reading speed expected in the first half of the passage would be more likely to manifest in the second half, as well; no interaction between position and frequency would therefore be observed. In contrast, if the valence switch did not disrupt the unfolding discourse context, and given prior work demonstrating that discourse context can mitigate the positive relation between word frequency and processing speed (Payne et al., 2015; Payne & Federmeier, 2019; Van Petten & Kutas, 1990, 1991), we would expect a reduction in the positive link between word frequency and reading speed in the second half of passages, manifesting as an interaction between position and frequency.

In the preswitch passage halves of our study, and similar to the traditional frequency effects displayed on the N400 in early open-class words within sentential frames (Van Petten & Kutas, 1990), we found an accumulation of word-level frequency effects that caused preswitch passage halves with higher average word frequency to be read at faster speeds. The postswitch passage halves, following the shift in valence, displayed the same pattern, with no significant interaction between position and frequency observed. These results therefore suggest that the midpassage shift in emotional valence may have disrupted the accrual of discourse context used to predict upcoming content, causing processing speeds in the second half of the passage to again be driven primarily by lexical frequency effects. However, caution is warranted when

attempting to draw inferences from a null effect, particularly when the sample size is relatively small. As such, future work is needed to confirm that the patterns reported here replicate.

Constraints on generality and future directions

A key strength of the current study is the use of an experimental protocol offering enhanced ecologically validity and naturalistic, paragraph-length stimuli. Additionally, we incorporated a mixed-effects analytic approach in all analyses. Nonetheless, this study is not without limitations. The sample is modest in size and, given the novelty of the paradigm employed, the current results should be replicated within a larger group of participants. Our sample is also predominantly female due to the gender bias of our student participant pool, and prior work has found that word valence ratings can be influenced by gender (Warriner et al., 2013); future work with more symmetrical gender statistics should therefore leverage designs that explore whether gender or biological sex moderate the effects reported in the current study. Furthermore, in order to balance competing demands between data volume and time burden of participation sessions, we limited our stimuli to twenty passages, which is comparatively few data points for analysis. As described above, our investigation of whether shifts in valence disrupt the development of discourse context relies on interpreting the null; as such, it is particularly important that future work seek to replicate our results within the context of both a larger sample of participants and a larger set of passage-length stimuli. Finally, our naturalistic design of the stimuli precluded the creation of passages with extreme values of average valence. Given our findings, future work should assess the effect of extreme manipulations of valence on oral reading speed and how such manipulations influence the interaction with average word frequency.

Conclusions

Current research on lexical processing is heavily focused on word-level behaviors, and much is now known about many of the features that influence visual word processing, as well as how this influence unfolds over time. However, conventional experimental paradigms are distinctly unlike naturalistic reading, not only in their presentation (standalone words or RSVP designs) but also in their construction: 95% of open-class words are less contextually constrained than those used in typical language studies (Luke & Christianson, 2016). The current study takes an initial step toward understanding how word-level features affect reading in more ecologically-valid task contexts, and whether such word-level features reliably map onto reading behavior across naturalistic, multi-sentence frames. To these ends, we find that previously studied lexical processing effects do, indeed, map onto oral reading speed when passage-length texts are presented on-screen as a whole. Across passage halves and average levels of emotional valence, we observed traditional word frequency effects, with faster reading speeds when average word frequency was high. We also observed a positivity effect across passage halves, with faster reading speeds for passage halves with higher (more positive) average valence. In addition to these two main effects, we observed an interaction between frequency and valence. Here again, our findings mirror the existing literature: similar reading speeds for low-frequency content regardless of valence, but faster reading speeds for positive content when average frequency is high. Overall, we demonstrate that oral reading speed is a useful proxy for classic measures used in reading research.

References

Balota, D. A., & Chumbley, J. I. (1984). Are Lexical Decisions a Good Measure of Lexical Access? The Role of Word Frequency in the Neglected Decision Stage. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 340–357. https://doi.org/10.1037//0096-1523.10.3.340

Balota, D. A., & Spieler, D. H. (1999). Word Frequency, Repetition, and Lexicality Effects in Word Recognition Tasks: Beyond Measures of Central Tendency. *Journal of Experimental Psychology: General*, *128*(1), 32–55. https://doi.org/10.1037//0096-3445.128.1.32

Baptiste, A. (2017). *gridExtra: Miscellaneous Functions for "Grid" Graphics* (R package version 2.3). https://cran.r-project.org/package=gridExtra

Barber, H., Vergara, M., & Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: Evidence from ERPs. *Cognitive Neuroscience and Neuropsychology*, *15*(3), 545–548.

Barriga-Paulino, C. I., Guerreiro, M., Faísca, L., & Reis, A. (2022). Does emotional valence modulate word recognition? A behavioral study manipulating frequency and arousal. *Acta Psychologica*, *223*, 103484. https://doi.org/10.1016/j.actpsy.2021.103484

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bestgen, Y. (1994). Can emotional valence in stories be determined from words? *Cognition & Emotion*, *8*(1), 21–36. https://doi.org/10.1080/02699939408408926

Boersma, P., & Weenink, D. (2001). PRAAT, a system for doing phonetics by computer. *Glot International*, *5*, 341–345.

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414. https://doi.org/10.7717/peerj.9414

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of

current word frequency norms and the introduction of a new and improved word

frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990.

https://doi.org/10.3758/BRM.41.4.977

Carstensen, L. L., & DeLiema, M. (2018). The positivity effect: A negativity bias in youth fades

with age. *Current Opinion in Behavioral Sciences*, *19*, 7–12.

https://doi.org/10.1016/j.cobeha.2017.07.009

Chou, L.-C., Pan, Y.-L., & Lee, C. (2020). Emotion anticipation induces emotion effects in

neutral words during sentence reading: Evidence from event-related potentials.

Cognitive, Affective, & Behavioral Neuroscience, *20*(6), 1294–1308.

https://doi.org/10.3758/s13415-020-00835-z

De Deyne, S., & Storms, G. (2008). Word associations: Network and semantic properties.

Behavior Research Methods, *40*(1), 213–231. https://doi.org/10.3758/BRM.40.1.213

Delaney-Busch, N., & Kuperberg, G. (2013). Friendly drug-dealers and terrifying puppies:

Affective primacy can attenuate the N400 effect in emotional discourse contexts.

Cognitive, Affective, & Behavioral Neuroscience, *13*(3), 473–490.

https://doi.org/10.3758/s13415-013-0159-5

Estes, Z., & Adelman, J. S. (2008). Automatic vigilance for negative words in lexical decision

and naming: Comment on Larsen, Mercer, and Balota (2006). *Emotion*, *8*(4), 441–444.

https://doi.org/10.1037/1528-3542.8.4.441

Estes, Z., & Verges, M. (2008). Freeze or flee? Negative stimuli elicit selective responding.

Cognition, *108*(2), 557–565. https://doi.org/10.1016/j.cognition.2008.03.003

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using

G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research

Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language

    comprehension. *Psychophysiology*, *44*(4), 491–505. https://doi.org/10.1111/j.1469-

    8986.2007.00531.x

Fischer-Baum, S., Dickson, D. S., & Federmeier, K. D. (2014). Frequency and regularity effects

    in reading are task dependent: Evidence from ERPs. *Language, Cognition and*

    Neuroscience, *29*(10), 1342–1355. https://doi.org/10.1080/23273798.2014.927067

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*(3), 221–233.

    https://doi.org/10.1037/h0057532

Foland-Ross, L. C., & Gotlib, I. H. (2012). Cognitive and Neural Aspects of Information

    Processing in Major Depressive Disorder: An Integrative Perspective. *Frontiers in*

    Psychology, *3*. https://doi.org/10.3389/fpsyg.2012.00489

Garcia, D., Garas, A., & Schweitzer, F. (2012). Positive words carry less information than

    negative words. *EPJ Data Science*, *1*(1), 3. https://doi.org/10.1140/epjds3

Haro, J., Guasch, M., Vallès, B., & Ferré, P. (2017). Is pupillary response a reliable index of

    word recognition? Evidence from a delayed lexical decision task. *Behavior Research*

    Methods, *49*(5), 1930–1938. https://doi.org/10.3758/s13428-016-0835-9

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L.,

    Delacqua, G., Delacqua, F., Kirby, J., & Duda, S. N. (2019). The REDCap consortium:

    Building an international community of software platform partners. *Journal of Biomedical*

    Informatics, *95*, 103208. https://doi.org/10.1016/j.jbi.2019.103208

Herbert, C., Ethofer, T., Anders, S., Junghofer, M., Wildgruber, D., Grodd, W., & Kissler, J.

    (2009). Amygdala activation during reading of emotional adjectives—An advantage for

    pleasant content. *Social Cognitive and Affective Neuroscience*, *4*(1), 35–49.

    https://doi.org/10.1093/scan/nsn027

Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology*, *45*(3), 487–498. https://doi.org/10.1111/j.1469-8986.2007.00638.x

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition & Emotion*, *15*(2), 143–165. https://doi.org/10.1080/02699930125768

Hofmann, M. J., & Jacobs, A. M. (2014). Interactive activation and competition models and semantic context: From behavioral to brain data. *Neuroscience & Biobehavioral Reviews*, *46*, 85–104. https://doi.org/10.1016/j.neubiorev.2014.06.011

Holt, D. J., Lynn, S. K., & Kuperberg, G. R. (2009). Neurophysiological correlates of comprehending emotional meaning in context. *Journal of Cognitive Neuroscience*, *21*(11), 2245–2262. https://doi.org/10.1162/jocn.2008.21151

Hsu, C.-T., Jacobs, A. M., Citron, F. M. M., & Conrad, M. (2015). The emotion potential of words and passages in reading Harry Potter – An fMRI study. *Brain and Language*, *142*, 96–114. https://doi.org/10.1016/j.bandl.2015.01.011

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*(6), 685–701. https://doi.org/10.1016/0749-596X(91)90032-F

Jacobs, A. M. (2015). Towards a neurocognitive poetics model of literary reading. In R. M. Willems (Ed.), *Cognitive Neuroscience of Natural Language Use* (1st ed., pp. 135–159). Cambridge University Press. https://doi.org/10.1017/CBO9781107323667.007

Johnson, P. O., & Fay, L. C. (1950). The Johnson-Neyman technique, its theory and application. *Psychometrika*, *15*(4), 349–367. https://doi.org/10.1007/BF02288864

Kazanas, S., & Altarriba, J. (2015). Emotion Word Type and Affective Valence Priming at a Long Stimulus Onset Asynchrony. *Language and Speech*, *59*. https://doi.org/10.1177/0023830915590677

Kazanas, S., & Altarriba, J. (2016). Emotion Word Processing: Effects of Word Type and Valence in Spanish–English Bilinguals. *Journal of Psycholinguistic Research*, *45*, 395–406. https://doi.org/10.1007/s10936-015-9357-3

Keuper, K., Zwitserlood, P., Rehbein, M. A., Eden, A. S., Laeger, I., Junghöfer, M., Zwanzger, P., & Dobel, C. (2013). Early Prefrontal Brain Responses to the Hedonic Quality of Emotional Words – A Simultaneous EEG and MEG Study. *PLoS ONE*, *8*(8), e70788. https://doi.org/10.1371/journal.pone.0070788

Kliegl, R., Dambacher, M., Dimigen, O., Jacobs, A. M., & Sommer, W. (2012). Eye movements and brain electric potentials during reading. *Psychological Research*, *76*(2), 145–158. https://doi.org/10.1007/s00426-011-0376-x

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1–2), 262–284. https://doi.org/10.1080/09541440340000213

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, *135*(1), 12–35. https://doi.org/10.1037/0096-3445.135.1.12

Kuchinke, L., Vo, M., Hofmann, M., & Jacobs, A. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, *65*(2), 132–140. https://doi.org/10.1016/j.ijpsycho.2007.04.004

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. (2014). Emotion and Language: Valence and Arousal Affect Word Recognition. *Journal of Experimental Psychology: General*, *143*(3), 1065–1081. https://doi.org/10.1037/a0035669

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lang, P. J., & Cuthbert, B. N. (1984). Affective information processing and the assessment of

anxiety. *Journal of Behavioral Assessment*, *6*(4), 369–395.

https://doi.org/10.1007/BF01321326

Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical Characteristics of Words used in

Emotional Stroop Studies. *Emotion*, *6*(1), 62–72. https://doi.org/10.1037/1528-

3542.6.1.62

Larsen, R. J., Mercer, K. A., Balota, D. A., & Strube, M. J. (2008). Not all negative words slow

down lexical decision and naming speed: Importance of word arousal. *Emotion*, *8*(4),

445–452. https://doi.org/10.1037/1528-3542.8.4.445

Laubrock, J., & Kliegl, R. (2015). The eye-voice span during reading aloud. *Frontiers in

Psychology*, *0*. https://doi.org/10.3389/fpsyg.2015.01432

Lewandowski, L. J., Codding, R. S., Kleinmann, A. E., & Tucker, K. L. (2003). Assessment of

Reading Rate in Postsecondary Students. *Journal of Psychoeducational Assessment*,

21(2), 134–144. https://doi.org/10.1177/073428290302100202

Li, C., Midgley, K. J., & Holcomb, P. J. (2022). ERPs reveal how semantic and syntactic

processing unfold across parafoveal and foveal vision during sentence comprehension.

Language, Cognition and Neuroscience, 1–17.

https://doi.org/10.1080/23273798.2022.2091150

Long, J. A. (2019). *interactions: Comprehensive, User-Friendly Toolkit for Probing Interactions*

(R package version 1.1.5). https://cran.r-project.org/package=interactions

Lucas, F. L. (1923). The Reverse of Aristotle. *The Classical Review*, *37*(5–6), 98–104.

https://doi.org/10.1017/S0009840X00079634

Lüdtke, J., & Jacobs, A. M. (2015). The emotion potential of simple sentences: Additive or

interactive effects of nouns and adjectives? *Frontiers in Psychology*, *6*.

https://doi.org/10.3389/fpsyg.2015.01137

Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology*, *88*, 22–60. https://doi.org/10.1016/j.cogpsych.2016.06.002

Méndez-Bértolo, C., Pozo, M. A., & Hinojosa, J. A. (2011). Word frequency modulates the processing of emotional words: Convergent behavioral and electrophysiological data. *Neuroscience Letters*, *494*(3), 250–254. https://doi.org/10.1016/j.neulet.2011.03.026

Payne, B. R., & Federmeier, K. D. (2019). Individual Differences in Reading Speed are Linked to Variability in the Processing of Lexical and Contextual Information: Evidence from Single-trial Event-related Brain Potentials. *Word*, *65*(4), 252–272. https://doi.org/10.1080/00437956.2019.1678826

Payne, B. R., Lee, C.-L., & Federmeier, K. D. (2015). Revisiting the Incremental Effects of Context on Word Processing: Evidence from Single-Word Event-Related Brain Potentials. *Psychophysiology*, *52*(11), 1456–1469. https://doi.org/10.1111/psyp.12515

Payne, B. R., Stites, M. C., & Federmeier, K. D. (2019). Event-related brain potentials reveal how multiple aspects of semantic processing unfold across parafoveal and foveal vision during sentence reading. *Psychophysiology*, *56*(10), e13432. https://doi.org/10.1111/psyp.13432

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Plaut, D. C., McClelland, J. L., & Seidenberg, M. S. (1996). Understanding Normal and Impaired Word Reading: Computational Principles in Quasi-Regular Domains. *Psychological Review*, *103*(1), 56–115. https://doi.org/10.1037/0033-295x.103.1.56

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037/0022-3514.61.3.380

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation

for Statistical Computing. https://www.R-project.org/.

Revelle, W. (2022). *psych: Procedures for Psychological, Psychometric, and Personality*

Research (R package version 2.2.5). Northwestern University. https://cran.r-

project.org/package=psych

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high-and low-

frequency words. *Memory & Cognition*, *18*(4), 367–379.

https://doi.org/10.3758/BF03197126

Sass, K., Habel, U., Sachs, O., Huber, W., Gauggel, S., & Kircher, T. (2012). The influence of

emotional associations on the neural correlates of semantic priming. *Human Brain*

Mapping, *33*(3), 676–694. https://doi.org/10.1002/hbm.21241

Schotter, E. R., & Payne, B. R. (2019). Eye Movements and Comprehension Are Important to

Reading. *Trends in Cognitive Sciences*, *23*(10), 811–812.

https://doi.org/10.1016/j.tics.2019.06.005

Scott, G. G., O'Donnell, P. J., Leuthold, H., & Sereno, S. C. (2009). Early emotion word

processing: Evidence from event-related potentials. *Biological Psychology*, *80*(1), 95–

104. https://doi.org/10.1016/j.biopsycho.2008.03.010

Scott, G. G., O'Donnell, P. J., & Sereno, S. C. (2014). Emotion words and categories: Evidence

from lexical decision. *Cognitive Processing*, *15*(2), 209–215.

https://doi.org/10.1007/s10339-013-0589-6

Scott, G. G., O'Donnell, P., & Sereno, S. (2012). Emotion Words Affect Eye Fixations During

Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(3),

783–792. https://doi.org/10.1037/a0027209

Stuart, G., & Hulme, C. (2000). The effects of word co-occurance on short-term memory:

Associative links in long-term memory affect short-term memory performance. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *26*(3), 796–802.

https://doi.org/10.1037/0278-7393.26.3.796

Theune, M. (2007). *Structure & Surprise: Engaging Poetic Turns*. Teachers & Writers

Collaborative.

Trauzettel-Klosinski, S., & Dietz, K. (2012). Standardized Assessment of Reading Performance:

The New International Reading Speed Texts IReST. *Investigative Opthalmology &*

Visual Science, *53*(9), 5452. https://doi.org/10.1167/iovs.11-8284

Unkelbach, C., Fiedler, K., Bayer, M., Stegmüller, M., & Danner, D. (2008). Why positive

information is processed faster: The density hypothesis. *Journal of Personality and*

Social Psychology, *95*(1), 36–49. https://doi.org/10.1037/0022-3514.95.1.36

U.S. Census Bureau. (2015). *Detailed Languages Spoken at Home and Ability to Speak English*

for the Population 5 Years and Over for Counties: 2009-2013.

https://www2.census.gov/library/data/tables/2008/demo/language-use/2009-2013-acs-

lang-tables-county.xls

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency

in event-related brain potentials. *Memory & Cognition*, *18*(4), 380–393.

https://doi.org/10.3758/BF03197127

Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and

closed-class words. *Memory & Cognition*, *19*(1), 95–112.

https://doi.org/10.3758/BF03198500

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal

role in the acquisition of reading skills. *Psychological Bulletin*, *101*(2), 192–212.

https://doi.org/10.1037/0033-2909.101.2.192

Wallace, S., Bylinskii, Z., Dobres, J., Kerr, B., Berlow, S., Treitman, R., Kumawat, N., Arpin, K.,

Miller, D. B., Huang, J., & Sawyer, B. D. (2022). Towards Individuated Reading

Experiences: Different Fonts Increase Reading Speed for Different Individuals. *ACM*

Transactions on Computer-Human Interaction, *29*(4), 1–56.

https://doi.org/10.1145/3502222

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and

dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207.

https://doi.org/10.3758/s13428-012-0314-x

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (2nd ed. 2016). Springer

International Publishing : Imprint: Springer. https://doi.org/10.1007/978-3-319-24277-4

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. J. (2012). An Abundance

of Riches: Cross-Task Comparisons of Semantic Richness Effects in Visual Word

Recognition. *Frontiers in Human Neuroscience*, *6*.

https://doi.org/10.3389/fnhum.2012.00072

Yap, M., & Seow, C. (2014). The influence of emotion on lexical processing: Insights from RT

distributional analysis. *Psychonomic Bulletin & Review*, *21*, 526–533.

https://doi.org/10.3758/s13423-013-0525-x