Originally published in: Quarterly Journal of Experimental Psychology vol. 71 (2018) no. 1, pp. 198-210.
DOI: *https://doi.org/10.1080/17470218.2017.1307862*

Anna Fiona Weiss[1], Franziska Kretzschmar[2], Matthias Schlesewsky[3], Ina Bornkessel-Schlesewsky[3] and Adrian Staub[4]

Comprehension demands modulate re-reading, but not first-pass reading behavior

Abstract: Several studies have examined effects of explicit task demands on eye movements in reading. However, there is relatively little prior research investigating the influence of implicit processing demands. In this study, processing demands were manipulated by means of a between-subject manipulation of comprehension question difficulty. Consistent with previous results from Wotschack and Kliegl, the question difficulty manipulation influenced the probability of regressing from late in sentences and re-reading earlier regions; readers who expected difficult comprehension questions were more likely to re-read. However, this manipulation had no reliable influence on eye movements during first-pass reading of earlier sentence regions. Moreover, for the subset of sentences that contained a plausibility manipulation, the disruption induced by implausibility was not modulated by the question manipulation. We interpret these results as suggesting that comprehension demands influence reading behavior primarily by modulating a criterion for comprehension that readers apply after completing first-pass processing.

Keywords: Reading strategy; eye movements; syntactic processing; semantic processing; regressions

It is often noted that despite its complexity, skilled reading is a fairly automatic process (e.g., LaBerge & Samuels, 1974). Under normal circumstances, the visual, cognitive and linguistic aspects of processing that are necessary for text comprehension work together seamlessly, with little need for conscious control or strategy. Consistent with this assumption, models of eye movements in reading such as E-Z Reader (Pollatsek, Reichle, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998) and SWIFT (Engbert, Longtin, & Kliegl, 2002) have regarded the speed of basic processes of word recognition as the primary determinant of the eyes' forward movement through the text and have not explicitly provided any role for variability in readers' strategies or goals. And though higher level processes of syntactic parsing and semantic interpretation also influence eye movements (see, for example, Liversedge & Rayner, 2011; Staub, 2015, for reviews), these processes are also usually assumed, at least implicitly, to operate without substantial variation based on, for example, the level of understanding that is required of the reader.

At the same time, several studies have examined effects of explicit task manipulations (e.g., proofreading vs reading for comprehension) on eye movements in reading. These studies have shown that eye movement measures like fixation duration, skipping rate and probability of regressions can be modulated by specific reading tasks (e.g., Kaakinen & Hyönä, 2010; Radach, Huestegge, & Reilly, 2008; Schotter, Bicknell, Howard, Levy, & Rayner, 2014). Arguably, however, these studies may not

[1]Department of German Linguistics, Philipps University of Marburg, Marburg, Germany
[2]CRC 1252 Prominence in Language & Department of German Language and Literature I, University of Cologne, Cologne, Germany
[3]School of Psychology, Social Work and Social Policy, University of South Australia, Adelaide, SA, Australia
[4]Department of Psychological and Brain Sciences, University of Massachusetts, Amherst, MA, USA

Corresponding author:
Anna Fiona Weiss, Department of German Linguistics, Philipps University of Marburg, Pilgrimstein 16, 35032 Marburg, Germany.
Email: fiona.weiss@staff.uni-marburg.de

be informative about how eye movements in reading are modulated by more typical variation in the demands of the reading situation. There are situations in which the reader needs to extract only the "gist" of the text, but there are also situations in which the reader must be sure to encode information in relatively fine detail. There is little prior research investigating how implicit modulations of reading strategy or depth of processing influence eye movements.

In addition to being of intrinsic interest, the question of how variability in readers' comprehension goals might influence eye movements is of substantial methodological importance. In most eye movement experiments, comprehension is tested by means of very simple two-alternative questions that appear after some or all of the sentences, once the critical sentence itself has been removed from the screen. A few studies have found that the presence of more targeted questions probing particular aspects of sentence interpretation can influence how the relevant sentences are read (Logačev & Vasishth, 2016; Swets, Desmet, Clifton, & Ferreira, 2008). To our knowledge, however, only a single study (Wotschack & Kliegl, 2013; hereafter WK) has addressed the question of whether, and how, the difficulty of comprehension questions asked after sentences might influence eye movements more generally. WK manipulated comprehension demands between subjects, by means of presenting either easy or difficult comprehension questions after sentences from the Potsdam Sentence Corpus (Kliegl, Grabner, Rolfs, & Engbert, 2004; Kliegl, Nuthmann, & Engbert, 2006). Focusing on word-based eye movement measures, WK found that among younger readers, difficult comprehension questions increased the frequency of regressive eye movements and, as a result, had a pronounced effect on total reading time. However, they did not find reliable effects on other measures such as skipping probability and first fixation duration. (For a group of older readers, with a mean age of about 70 years, the question manipulation did influence these measures.)

This study follows up on WK's results in several ways. The first question we addressed is whether more pronounced effects of question difficulty may appear in the eye movement record with a more pronounced difficulty manipulation. The limited influence of the question manipulation on young adults' reading behavior in WK's study may be due to the fact that the "difficult" comprehension questions in that study were not, in fact, very difficult. The young adult readers in the difficult question condition tested by WK achieved 95% accuracy on the comprehension questions, compared with 97% in the easy condition; the high level of accuracy in the difficult condition is especially striking given that WK used three-alternative multiple-choice questions. Second, our analyses allow an assessment of the location in sentences from which regressions are initiated when readers are confronted with difficult comprehension questions. Are they distributed throughout the sentence, or do readers simply re-read once they have reached the end of a sentence? WK's results are consistent with difficult questions generally increasing the probability of a leftward as opposed to rightward saccades but are also consistent with an increase in regressions that is restricted to near the end of sentences.

Finally, this study assessed whether a global manipulation of question difficulty would specifically impact how readers deal with anomalous content. In recent years, the notion that language processing is merely "good enough" for the task at hand has received extensive discussion (e.g., Ferreira, Bailey, & Ferraro, 2002; Ferreira & Patson, 2007; see also Christianson, 2016, which introduces a special issue of *The Quarterly Journal of Experimental Psychology* (*QJEP*) on the topic of "good enough" processing). It has been suggested that comprehenders may often impose an interpretation that is consistent with real-world knowledge rather than with the literal meaning of the text, when such an interpretation suffices for the task at hand. In this study, we ask whether readers who are presented with only easy comprehension questions would be more likely to overlook a reversal of thematic roles (e.g., *the flower is drawing the girl*), and whether they would be especially likely to do so when the subject and verb are highly lexically associated (e.g., *the flower is picking the girl*). Specifically, we test the hypothesis that the anomaly effect on first-pass reading of the verb and object in these sentences may be reduced when the reader receives only easy questions and that this effect may be reduced especially when there is a high degree of association between the subject and the verb. Processing of these so-called "semantic reversal anomalies" (SRAs) has been investigated using event-related potentials (ERPs; for example, Kim & Osterhout, 2005; Kolk, Chwilla, van Herten, & Oor, 2003; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003, among many others) but has not been previously investigated using eye movements in reading.

Method

Participants

Ninety-two undergraduate students at the University of Massachusetts, Amherst, participated in the study (45 in the easy condition), in exchange for course credit. Two additional participants were excluded due to poor accuracy on comprehension questions in the easy condition (<80%, which is >2.5 *SD* from group mean), and three were excluded due to very slow average total sentence reading time (>2.5 *SD* from group mean). All participants had normal or corrected-to-normal vision and were native speakers of American English. All were naive concerning the purpose of the experiment.

Materials

The materials in the experiment as a whole consisted of three sets of sentences, which we refer to as SRAs

Table 1. Sentence types used as stimuli in the experiment.

1. Semantic reversal anomalies (SRAs)	
(a) On a sunny afternoon | the girl | is picking | the flower | for the dining table.	*non-anomalous, high associated*
(b) On a sunny afternoon the girl is drawing the flower on a little sketchpad.	*non-anomalous, low associated*
(c) On a sunny afternoon the flower is picking the girl for the dining table.	*anomalous, high associated*
(d) On a sunny afternoon the flower is drawing the girl on a little sketchpad.	*anomalous, low associated*
2. Relative clause sentences (RC)	
(a) The chef | that distracted the waiter | sifted the flour onto the counter.	*subject relative (SRC)*
(I) Did a chef do something?	*easy*
(II) Did the waiter distract the chef?	*difficult*
(b) The executives | that the lawyers sued | roused themselves from slumber.	*object relative (ORC)*
(I) Did a policeman do something?	*easy*
(II) Was it the executives who roused themselves?	*difficult*
3. Garden path sentences (GP)	
John borrowed | the rake or the shovel | turned out to be sufficient.	
(I) Is there a shovel?	*easy*
(II) Might the rake have been borrowed?	*difficult*

(36 sentences); relative clause (RC) sentences (39 sentences), divided between sentences that contained object relative clause (ORC) and subject relative clause (SRC), as described below; and garden path (GP) sentences (24 sentences). However, only the SRA sentences were manipulated within subjects and within items; all subjects read the same versions of the RC and GP sentences. All RC and GP sentences were followed by comprehension questions, but because half of the SRA sentences were anomalous, and because different questions would be necessary for the different versions of each item, we did not include comprehension questions for these items. In total, questions followed 63 of the 99 sentences in the experiment. The SRA sentences were allocated to four experimental lists according to a Latin Square design and randomly intermixed with the RC and GP items. Question difficulty was manipulated between subjects, and participants were randomly assigned to one of the four lists and one of the two question difficulty conditions. Thirty-three of the difficult questions and 32 of the easy questions required a YES answer.

An example SRA stimulus set is given in Table 1. We manipulated the factors ANOMALY (non-anomalous vs anomalous) and ASSOCIATION (high vs low), in a 2×2 design. Non-anomalous sentences had an animate subject and an inanimate object. In anomalous sentences, the subject and object were reversed. The verbs required an animate subject, so that the reversed sentences were anomalous at the point of reaching the verb. The verb used in the high association condition (e.g., *picking*) was judged to be strongly associated with the inanimate object (e.g., *flower*) based on a pre-test, described below. Note that verbs were used in the present progressive tense, as opposed to the past tense, to avoid the possibility that subjects would initially adopt a reduced RC reading at the verb (e.g., *The flower picked by the girl was . . .*).

The high and low associated verbs in the sentences did not differ significantly in mean length in characters (high: 7.4, low: 7.5; $t(35) = -0.36$, $p = 0.72$) or mean frequency (log Subltex; Brysbaert & New, 2009: 2.6 for high, 2.5 for low; $t(35) = 0.27$, $p = 0.79$). The inanimate nouns were slightly shorter on average than animate nouns (animate: 6.6, inanimate: 5.5; $t(35) = 2.48$, $p < 0.05$) and were also slightly more frequent (animate: 2.8, inanimate: 3.2; $t(35) = -2.25$, $p < 0.05$).

In an online pre-test on Amazon Mechanical Turk, verb and noun pairs were assigned to four lists according to a Latin Square design. For each pair of words, the participants were asked to give a rating on a 5-point scale of how strongly associated the meanings of the words are (1 = *low associated*, 5 = *high associated*). Each list contained 55 word pairs in random order and was rated by 13 participants who self-reported as native speakers of American English. The questionnaire took between 5 and 10 min and participants received $1. Twenty-four additional word pairs were rated by eight further subjects in a paper-and-pencil questionnaire.

Based on the results from the pre-test, we selected 36 stimuli for which the association ratings between the verb and the two inanimate nouns were very different. The difference between was highly significant for both animate (high associated: 3.95, low associated: 2.25; $t(35) = 7.53$, $p < 0.001$) and inanimate nouns (high associated: 4.49, low associated: 1.92; $t(35) = 24.99$, $p < 0.001$).

In a second questionnaire on Mechanical Turk, we asked participants to rate the naturalness of the entire sentence (1 = *very unnatural*, 5 = *very natural*). The critical sentences were presented as they would appear in the main experiment and were again allocated to four lists. Each list contained 55 sentences in random order and was rated by 13 further participants. Each questionnaire took about 10 min. Participants received the same remuneration for their participation. Again, 24 additional sentences were

rated by eight further subjects in a paper questionnaire. The ratings of the 36 experimental stimuli that were selected based on the association criteria described above suggest no influence of ASSOCIATION within either the anomalous (high associated: 1.25, low associated: 1.29) or the non-anomalous sentences (high associated: 4.27, low associated: 4.12) but a clear effect of ANOMALY. This pattern was confirmed by a repeated measures by-items ANOVA with ANOMALY and ASSOCIATION as fixed effects that showed a significant main effect of ANOMALY ($F_2(1, 35) = 758.20$, $p < 0.001$), but neither a main effect of ASSOCIATION ($p = 0.59$) nor an interaction ($p = 0.11$).

The 39 RC sentences were followed by either easy or difficult comprehension questions, depending on task condition. Two examples are shown in Table 1. Seventeen of these sentences contained a subject-modifying SRC (2a), and 22 contained a subject-modifying ORC (2b). Object RCs in particular are known to induce measurable processing difficulty in the eye movement record (e.g., Staub, 2010). In all cases, the difficult question required the reader to correctly assign thematic roles to the noun phrases in the sentence, while the easy question did not.

Finally, an example of the 24 GP sentences is also shown in Table 1, with the corresponding easy and difficult comprehension questions. The GP sentences always involved a temporary ambiguity between a noun phrase coordination structure and a clausal coordination structure, which was ultimately resolved toward the (initially dispreferred) clausal coordination analysis (e.g., Staub, 2007; Staub & Clifton, 2006). A comma after the initial clause (e.g., after *rake* in the example in Table 1) would avert the GP, but the comma was never present. The difficult question always required the reader to successfully interpret the thematic roles of the noun phrases in the sentence, while the easy question did not.

Procedure

Eye movements were recorded using an SR Research EyeLink 1000 tracker with a sampling rate of 1000 Hz. The viewing was binocular, but only one eye's movement was monitored. In most cases, this was the right eye, but due to technical problems in some cases the left eye was tracked. All sentences were displayed on one line on a cathode ray tube (CRT) monitor 55 cm from the participant, in 12-point Monaco font. At this distance, three characters corresponded to approximately 1° of visual angle; the resolution of the eyetracker was less than one character.

Participants were asked to read for comprehension and were told that after some of the sentences they would be required to respond to a comprehension question presented on the screen by pressing one of two buttons on a gamepad. They were also told that some of the items might be "a little weird." Once the participant was seated at the eyetracker, the tracker was aligned and calibrated in a single line calibration. The experiment began with six practice trials and took about 25 min in total.

Because the verb region of the SRA sentences was of critical interest as the location in at which an implausibility arises, we deleted all trials with first-pass blinks or track losses in this region, which resulted in excluding about 11% of these sentences. No RC or GP sentences were excluded due to blink or track loss, as there was no pre-identified critical region of these sentences. The initial regions for analysis were defined as given in Table 1; below we discuss results with an alternate regioning scheme.

Eye movement analysis focused on first-pass time (the sum of all fixation durations on a region before leaving it to the left or right) and go-past time (the sum of all fixation durations from first entering a region until leaving it to the right, including any regressive re-reading), as well as the probability of a first-pass regressive eye movement from a region. These three measures have the ability to clearly capture potential effects of the task difficulty manipulation on incremental processing of the sentence, and together to distinguish first-pass reading effects from re-reading effects. We calculated linear mixed-effects models (for first-pass time and go-past time) and logistic regression models (for probability of regressions out) for each region, for each sentence type, using the lme4 package (Version 1.1-5; Bates, Mächler, Bolker, & Walker, 2014) and the NLopt nonlinear-optimization package (http://ab-initio.mit.edu/nlopt) for R (Version 2.15.1). Following Barr, Levy, Scheepers, and Tily (2013), we used the maximal random effect structure justified by our design. These models included random intercepts for subjects and items, and random by-item slopes for the question difficulty manipulation. For the SRA models, random by-subject and by-item slopes were also included for the effects of anomaly and association, and their interaction. Following convention, we treat $|t|$ or $|z| > 2$ as significant.

Results

Question Accuracy

On average, question accuracy in the easy condition was 93.57% (after exclusion of two subjects, as noted above; range: 80.30%-100%), whereas for the difficult group, average accuracy dropped to 82.70% (range: 63.63%-98.48%; see Figure 1). The difference between conditions was significant ($t(69.99) = -7.71$, $p < 0.001$) and was apparent for both kinds of RC sentences (SRC: easy: 93.99%, difficult: 82.48%, $t(84.68) = -5.22$, $p < 0.001$; ORC: easy: 92.53%, difficult: 77.72%, $t(60.21) = -6.59$, $p < 0.001$) as well as the GP sentences (easy: 96.11%, difficult: 87.94%, $t(95.37) = -4.78$, $p < 0.001$).

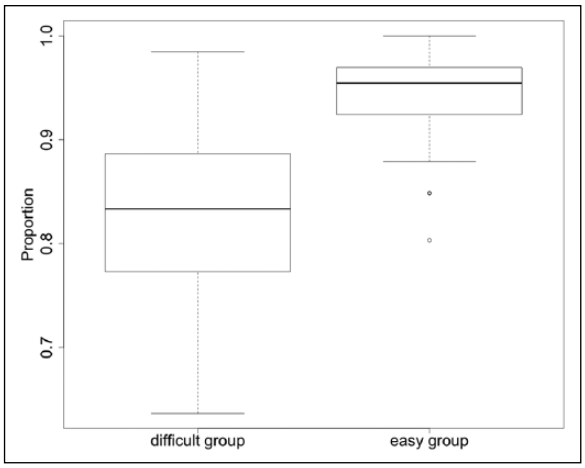

Figure 1. Comprehension accuracy, by question difficulty condition.

Semantic reversal anomalies

Tables 2 and 3 provide mean reading times for the SRA sentences and statistical results, respectively, for each of the sentence regions. In line with previous research indicating that anomaly has immediate effects in the eye movement record (Rayner, Warren, Juhasz, & Liversedge, 2004), anomaly increased first-pass time and go-past time on the verb and object regions (Regions 3 and 4), and the probability of a regression from each of these regions. The effect of anomaly also reached significance in go-past time and regressions out for the sentence-final Region 5. The main effect of association was significant only in first-pass time on the object and in go-past time on the final region, but there was a significant Anomaly × Association interaction in first-pass and go-past time on Regions 3 and 4. There was a reduced anomaly effect in the associated conditions in Region 3, but an increased effect of anomaly in the associated conditions in Region 4. We discuss this pattern in more detail in the "Discussion" section.

The question difficulty manipulation had a significant effect on go-past time and regressions out for the sentence-final region, with longer reading times and more regressions with difficult questions. There were no significant effects of this manipulation on the earlier regions. In particular, question difficulty did not modulate the size of the anomaly effect on the Critical Regions 3 and 4.

A follow-up analysis addressed the possibility that participants adopted a fairly superficial reading strategy in the SRA sentences, regardless of task condition, because these sentences were never followed by comprehension questions. This hypothesis predicts that the anomaly effect should decline over the course of the experiment, as the lack of comprehension questions becomes apparent. We computed additional statistical models of first-pass and go-past time on the Critical Regions 3 and 4 that included (centered and scaled) trial order and its interactions with the other variables. Random slopes for trial order and its interactions were also included. No main effects of trial order approached significance, and the only significant interaction was between trial order and anomaly in go-past time on Region 4 ($\beta = -0.05$, $SE = 0.02$, $t = 2.48$); the corresponding interactions for first-pass time, and the go-past interaction for Region 3, all had $|t| < 0.8$. To test whether the significant interaction in go-past time on Region 4 resulted in elimination of the anomaly effect late in the experiment, we computed a model identical to our original model, but restricted to the second half of the experiment. The anomaly effect was still highly significant ($\beta = 0.28$, $SE = 0.04$, $t = 6.78$). Thus, the effect of anomaly was present even late in the experiment on both Regions 3 and 4, despite the lack of comprehension questions with the SRA sentences.

RC sentences

Mean reading times and statistical results for the RC sentences are given in Tables 4 and 5. Though the analysis does not distinguish between the RC types, as these sentences were not matched on factors such as lexical frequency and length, we present descriptive statistics separately to illustrate that the effect of task condition is in fact similar for both types. The results for the RC sentences show the same pattern as the SRA sentences. Question difficulty influenced re-reading, indicated by significant effects in go-past time and regressions out in the last region. But again, there were no first-pass effects for question difficulty in any region.

GP sentences

Tables 4 and 5 show mean reading times and statistical results for the GP sentences. As for the SRA and RC sentences, there was no effect of the question difficulty manipulation on first-pass reading, but this factor did influence the probability of making regressions out of the sentence-final region, as indicated by a significant effect on go-past time as well as regressions out.

Alternate regioning of sentences

In the analyses above, the task manipulation reliably influenced reading behavior only once readers reached the final region of sentences. We performed an additional analysis focusing on go-past time (the measure in which task effects most clearly emerged) to further clarify this pattern. One possibility is that task effects on material prior to the final region are weak and that these effects are not clearly seen when this material is divided into multiple regions. To test this possibility, we combined the original pre-final regions (Regions 1-4 for the SRA sentences, and 1-2 for the RC and GP sentences) into a single region. In addition, to test

Table 2. By-subject means and standard deviations (in parentheses) of first-pass time and go-past time, and probability of regressions out for semantic reversal anomalies (SRAs).

	Region 1				Region 2				Region 3				Region 4				Region 5			
	On the sunny afternoon				*the flower / the girl*				*is picking / is drawing*				*the girl / the flower*				*for the dining table*			
	easy		difficult		easy		difficult		easy		difficult		easy		difficult		easy		difficult	
	low	high	low	high	low	high	low	high	low	high	low	high	low	high	low	high	low	high	low	high
First-pass time																				
ano	526	514	525	520	380	379	390	356	430	400	422	398	402	402	395	401	897	836	860	856
	(190)	(151)	(128)	(173)	(114)	(105)	(110)	(90)	(122)	(90)	(112)	(101)	(95)	(100)	(94)	(94)	(269)	(205)	(274)	(307)
non-ano	507	497	530	547	382	393	384	385	372	371	360	368	354	324	373	345	903	892	893	881
	(157)	(128)	(183)	(172)	(122)	(105)	(114)	(88)	(97)	(82)	(81)	(101)	(87)	(73)	(79)	(77)	(288)	(247)	(272)	(238)
Go-past time																				
ano	549	518	534	528	457	443	475	453	516	492	526	499	556	597	593	599	1830	1655	2269	2323
	(201)	(152)	(127)	(171)	(168)	(125)	(147)	(155)	(160)	(145)	(131)	(158)	(174)	(229)	(172)	(185)	(620)	(647)	(737)	(1144)
non-ano	519	503	540	555	494	487	483	486	437	426	424	444	419	359	430	409	1491	1377	1886	1753
	(160)	(137)	(189)	(167)	(147)	(165)	(170)	(146)	(114)	(118)	(146)	(133)	(138)	(94)	(113)	(162)	(486)	(497)	(731)	(605)
Regressions out																				
ano					11.11	12.62	13.27	15.32	12.97	12.72	16.65	16.70	22.31	25.47	23.54	27.24	58.21	54.88	74.56	64.81
					(12.70)	(13.14)	(15.61)	(17.88)	(11.49)	(11.21)	(14.77)	(16.43)	(16.60)	(17.93)	(16.18)	(18.49)	(27.31)	(26.77)	(21.81)	(28.05)
non-ano					17.16	12.92	14.97	16.76	10.41	8.34	9.15	10.19	10.34	6.67	9.85	10.00	49.98	47.43	60.69	63.10
					(19.18)	(14.10)	(15.23)	(15.61)	(11.66)	(9.73)	(10.28)	(10.17)	(11.50)	(10.02)	(10.72)	(11.89)	(30.59)	(31.44)	(26.30)	(27.25)

ano, anomaly; easy, easy comprehension questions; difficult, difficult comprehension questions; low, low associated; high, high associated.

Table 3. Results of the mixed models analysis (first-pass time and go-past time) and the logistic regression models (regressions out) for SRA.

SRA	Region 1			Region 2			Region 3			Region 4			Region 5		
	On the sunny afternoon			*the flower / the girl*			*is picking / is drawing*			*the girl / the flower*			*for the dining table*		
	Estimate	*SE*	|*t*| value / |*z*| value	Estimate	*SE*	|*t*| value / |*z*| value	Estimate	*SE*	|*t*| value / |*z*| value	Estimate	*SE*	|*t*| value / |*z*| value	Estimate	*SE*	|*t*| value / |*z*| value
First-pass time															
Intercept	6.06	0.07	88.74	5.82	0.03	205.84	5.86	0.02	239.38	5.82	0.02	272.46	6.59	0.04	160.82
Anomaly	0.01	0.02	0.31	–0.02	0.04	0.64	0.10	0.02	**4.63**	0.11	0.03	**4.37**	–0.04	0.03	1.23
Association	0.01	0.02	0.60	0.00	0.02	0.08	–0.02	0.02	0.86	–0.04	0.02	**2.28**	–0.02	0.05	0.44
Difficulty	0.05	0.05	0.89	0.00	0.04	0.09	–0.02	0.04	0.43	0.02	0.04	0.46	–0.02	0.05	0.31
Anomaly × Association	–0.04	0.04	0.89	–0.03	0.04	0.75	–0.09	0.03	**2.78**	0.08	0.03	**2.36**	–0.02	0.05	0.37
Anomaly × Difficulty	–0.06	0.04	1.29	–0.01	0.04	0.36	–0.02	0.03	0.45	–0.04	0.04	1.24	–0.04	0.05	0.83
Association × Difficulty	0.02	0.05	0.48	–0.05	0.04	1.39	0.03	0.04	0.70	0.00	0.03	0.08	0.05	0.05	0.92
Anomaly × Association × Difficulty	–0.02	0.09	0.23	–0.07	0.08	0.95	–0.02	0.07	0.35	0.01	0.06	0.23	0.11	0.10	1.08
Go-past time															
Intercept	6.11	0.07	90.13	6.01	0.03	182.89	6.01	0.03	201.66	6.03	0.03	224.30	7.31	0.04	171.87
Anomaly	0.00	0.02	0.17	–0.06	0.04	1.31	0.14	0.03	**4.45**	0.32	0.04	**9.03**	0.18	0.03	**5.95**
Association	–0.01	0.02	0.53	–0.01	0.02	0.28	–0.02	0.02	1.14	–0.04	0.03	1.31	–0.07	0.03	**2.28**
Difficulty	0.03	0.05	0.64	0.00	0.05	0.04	0.00	0.05	0.03	0.03	0.05	0.71	0.21	0.07	**2.96**
Anomaly × Association	–0.02	0.04	0.64	–0.01	0.04	0.39	–0.07	0.04	**2.04**	0.12	0.05	**2.71**	–0.04	0.05	0.80
Anomaly × Difficulty	–0.05	0.03	1.47	0.03	0.03	0.82	0.02	0.03	0.54	–0.03	0.04	0.64	0.02	0.06	0.31
Association × Difficulty	0.06	0.04	1.30	–0.02	0.03	0.68	0.04	0.04	0.94	0.01	0.04	0.16	0.05	0.04	1.34
Anomaly × Association × Difficulty	–0.02	0.07	0.31	–0.09	0.07	1.30	–0.06	0.07	0.91	–0.08	0.08	1.02	–0.03	0.08	0.37
Regressions out															
Intercept				–2.13	0.13	16.57	–2.34	0.12	19.01	–2.00	0.09	21.79	0.52	0.14	3.65
Anomaly				–0.06	0.12	0.56	0.59	0.17	**3.39**	1.54	0.14	**10.90**	0.44	0.14	**3.05**
Association				0.00	0.12	0.02	–0.17	0.13	1.30	0.02	0.17	0.10	–0.17	0.12	1.47
Difficulty				0.23	0.21	1.08	0.12	0.18	0.68	0.26	0.17	1.51	0.73	0.28	**2.62**
Anomaly × Association				0.28	0.24	1.18	0.19	0.29	0.67	0.34	0.34	0.98	–0.39	0.20	1.91
Anomaly × Difficulty				0.11	0.26	0.43	0.28	0.27	1.02	–0.31	0.26	1.17	–0.01	0.25	0.03
Association × Difficulty				0.20	0.23	0.89	0.28	0.30	0.94	0.34	0.33	1.04	–0.04	0.21	0.20
Anomaly × Association × Difficulty				–0.77	0.47	1.63	–0.30	0.53	0.57	–0.66	0.56	1.18	–0.78	0.41	1.91

SRAs, semantic reversal anomalies; *SE*, standard error.

Table 4. By-subject means and standard deviations (in parentheses) of first-pass time, go-past time and probability of regressions out for relative clause sentences (RC) and garden path sentences (GP).

	Region 1		Region 2		Region 3	
	easy	difficult	easy	difficult	easy	difficult
First-pass time						
Garden path	728 (165)	776 (195)	927 (216)	923 (206)	877 (216)	848 (219)
Subject relative	413 (90)	423 (94)	885 (244)	854 (170)	1326 (283)	1304 (295)
Object relative	415 (98)	441 (106)	915 (261)	905 (189)	1375 (281)	1390 (353)
Go-past time						
Garden path	747 (172)	793 (205)	1099 (286)	1128 (274)	1634 (451)	2049 (694)
Subject relative	420 (89)	434 (93)	1027 (270)	1045 (323)	2188 (618)	2808 (904)
Object relative	424 (100)	448 (109)	1072 (299)	1123 (295)	2440 (713)	3071 (1054)
Regressions out						
Garden path			13.98 (9.66)	16.43 (10.90)	54.17 (23.60)	70.62 (21.90)
Subject relative			11.63 (10.90)	13.64 (10.05)	50.20 (27.95)	71.21 (24.66)
Object relative			12.83 (10.58)	17.43 (14.49)	55.25 (25.55)	70.29 (23.13)

easy, easy comprehension questions; difficult, difficult comprehension questions.

Table 5. Results of the mixed models analysis (first-pass time and go-past time) and the logistic regression models (regressions out) for relative clause sentences (RC) and garden path sentences (GP).

	Region 1			Region 2			Region 3		
	Estimate	*SE*	|*t*| value / |*z*| value	Estimate	*SE*	|*t*| value / |*z*| value	Estimate	*SE*	|*t*| value / |*z*| value
First-pass time									
Garden path sentences									
Intercept	6.47	0.05	129.21	6.67	0.05	138.92	6.56	0.06	106.33
Difficulty	0.07	0.05	1.37	0.00	0.06	0.01	–0.03	0.05	0.57
Relative clause sentences									
Intercept	5.90	0.04	132.14	6.64	0.03	203.06	7.00	0.03	227.69
Difficulty	0.04	0.05	0.91	–0.01	0.05	0.28	–0.01	0.05	0.27
Go-past time									
Garden path sentences									
Intercept	6.53	0.05	125.64	6.93	0.04	177.57	7.33	0.05	151.22
Difficulty	0.06	0.05	1.26	0.03	0.05	0.52	0.20	0.07	**2.96**
Relative clause sentences									
Intercept	5.95	0.05	131.75	6.88	0.04	195.44	7.75	0.04	205.20
Difficulty	0.04	0.05	0.84	0.02	0.05	0.38	0.21	0.06	**3.30**
Regressions out									
Garden path sentences									
Intercept				–1.99	0.16	12.30	0.66	0.15	4.33
Difficulty				0.23	0.19	1.24	0.92	0.27	**3.36**
Relative clause sentences									
Intercept				–2.04	0.11	19.19	0.69	0.15	4.53
Difficulty				0.28	0.19	1.46	0.98	0.29	**3.38**

SE, standard error.

whether task effects specifically emerge only very near the end of sentences, we divided the original final region into two regions. This regioning included all but the last two words of the sentence in one region, and the last two words in another; if this new final region was less than 10 characters in length, we included three words. In sum, this alternate regioning divided the sentences into three regions: a long initial region that consisted of the combined pre-final regions from the original analyses, a short second region that consisted of the first part of the final region in the original analyses, and a new final region that consisted of only the last 2 to 3 words of the sentence (see Table 6).

Table 6. Examples of alternate regioning of the different sentence types.

	Initial region	Prefinal region	Final region
Semantic reversal anomalies	*On a sunny afternoon the flower is picking the girl*	*for the*	*dining table.*
Relative clause sentences	*The chef that distracted the waiter*	*sifted the flour onto*	*the counter.*
Garden path sentences	*John borrowed the rake or the shovel*	*turned out to*	*be sufficient.*

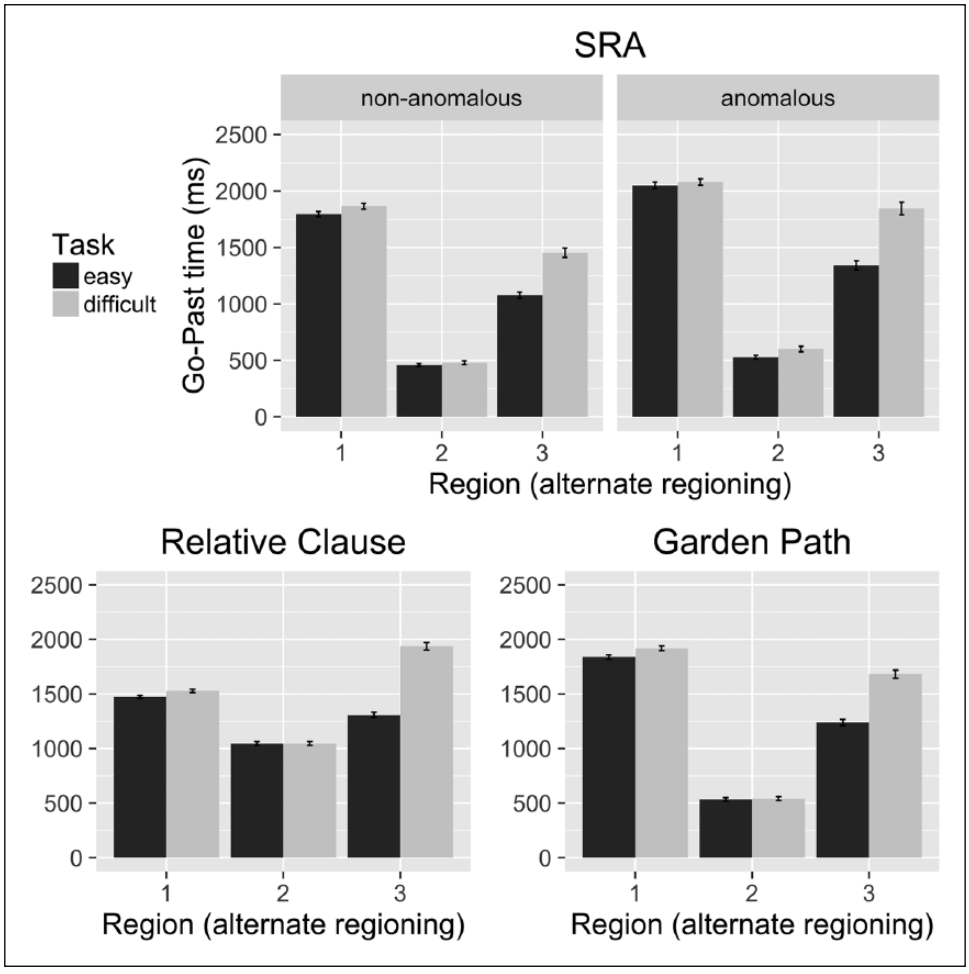

Figure 2. Go-past time means and standard errors by region, based on alternate regioning described in the text, for each sentence type, by task condition (and for SRA sentences, by anomaly condition).

Figure 2 shows go-past means and standard errors by region and sentence type. For the initial combined region in the SRA sentences, there was again a large anomaly effect on go-past time (β=0.12, *SE*=0.02, *t*=6.61), but no other significant effects or interactions. The effect of the task manipulation did not approach significance (β=0.02, *SE*=0.05, *t*=0.54). For the RC and GP sentences, there was also no task effect on go-past time on the first region (RC: β=0.03, *SE*=0.05, *t*=0.58; GP: β=0.05, *SE*=0.05, *t*=1.09). For the new second region, in SRA sentences there was again an anomaly effect on go-past time (β=0.12, *SE*=0.02, *t*=5.20), and again no task effect (β=0.03, *SE*=0.05, *t*=0.70) and no other significant effects. There was also no task effect in this region for the other sentence types (RC: β=–0.01, *SE*=0.05, *t*=0.24; GP: β=–0.04, *SE*=0.05, *t*=0.69). In the new, two- or three-word final region of the SRA sentences, however, there were significant effects of both anomaly (β=0.18, *SE*=0.04, *t*=4.75) and task (β=0.25, *SE*=0.09, *t*=2.96), as well as a significant main effect of association (β = −0.10, *SE*=0.03, *t*=3.27). There were no significant interaction effects. There was also a significant effect of task in the new final region of the other sentence types (RC: β=0.37, *SE*=0.09, *t*=4.04; GP: β=0.28, *SE*=0.08, *t*=3.45). In sum, effects of task did not reliably appear until readers reached the last 2-3 words of sentences of all three types. Note that the

absolute size of the task effect on go-past time from this final region was similar for all three sentence types, with parameter estimates ranging from 0.25 to 0.37 in log units.

Relation between re-reading and question accuracy

A final post hoc analysis assessed the relationship between regressions and go-past time on the final region of RC and GP sentences, in the difficult condition, and accuracy on the comprehension questions (see, for example, Schotter, Tran, & Rayner, 2014, for similar analyses). The previous analyses determined that the difficult task condition influenced reading behavior primarily by increasing re-reading from the ends of sentences; this analysis addressed whether such re-reading did actually improve comprehension (we did not attempt a similar analysis for the easy condition, as comprehension was near ceiling for the majority of subjects; see Figure 1). We computed logistic regression models of accuracy that included either regression from the sentence-final region (two to three final words) or go-past time on this region as a fixed effect, as well as random intercepts for subjects and items and random slopes for the fixed effect. Neither the effect of regression ($\beta = -0.23$, $SE = 0.17$, $z = 1.35$) nor the effect of go-past time ($\beta = -0.000066$, $SE = 0.000050$, $z = 1.33$) reached significance.

Discussion

This study investigated how implicit modulations of reading strategy may influence eye movements during sentence reading. We manipulated the difficulty of comprehension questions between subjects who read sentences containing different types of anomalous and non-anomalous material. The results are easily summarized: The difficulty of the comprehension questions did not significantly affect first-pass reading of sentences; instead, participants were more likely to re-read the sentence after regressing from near the end, and/or take longer in re-reading, when faced with difficult comprehension questions. Importantly, this pattern was replicated for all three sentence types in our study and held for sentences with and without SRAs. Indeed, the task manipulation did not influence readers' eye movement behavior until they reached the two or three final words of the sentence.

These results extend WK's findings in several ways. First, our results confirm, with a more pronounced question difficulty manipulation, WK's finding that such a manipulation primarily modulates reading behavior by influencing the probability of re-reading, at least with younger adults. Second, while WK showed that regressions were more common when readers were faced with difficult questions, their word-level analyses did not investigate the point in sentences from which these regressions were launched; our data clearly show that question difficulty influenced the subject's decision to re-read only when he or she reached the final words of the sentence.

Finally, we tested whether comprehension demands have a particular impact on reading of anomalous sentences, and we found that they did not: Anomaly effects during incremental processing were as pronounced with easy questions as with difficult questions. Thus, the results do not support the idea that comprehension demands determine whether readers engage in syntactically licensed assignment of thematic roles. A clear limitation is that we do not have direct evidence about the final interpretations of SRA sentences, because we did not include comprehension questions for these sentences; it is possible that readers may have sometimes misunderstood these sentences and that this was more common with easy comprehension questions. However, the on-line effect of anomaly was not contingent on comprehension demands. We acknowledge that to some extent, this issue may depend on whether sentences involve canonical (in this case, Subject-Verb-Object) word order or a non-canonical order such as in passive or RC structures. Previous studies (e.g., Christianson, Luke, & Ferreira, 2010; Ferreira, 2003) have found that comprehenders are most likely to misinterpret thematic roles when a sentence involves non-canonical word order. Thus, it is possible that comprehension demands would have a clearer influence on incremental thematic role assignment with non-canonical word order.

In addition, it remains an open question what function is played by the regressions that are specifically initiated during incremental processing of anomalous material, which have also appeared in other eyetracking studies (e.g., Rayner et al., 2004). It is possible that these regressions serve to confirm that the sentence is indeed anomalous; if so, a regression might lead to an increased probability of correctly recognizing that the sentence is anomalous. However, in the absence of explicit comprehension probes of anomalous sentences (which have also been absent from other eyetracking studies investigating processing of implausible or anomalous sentences, for example, Rayner et al., 2004), this cannot be confirmed.

Our results from sentences with SRAs do have implications for the functional architecture of the language processing system. Though the processing of these structures was not affected by task condition, processing was influenced by the lexical association between subject and verb. The anomaly effect on reading times on the verb was reduced when the subject was a highly plausible theme for the verb (*the flower is picking* vs *the flower is drawing*). On the object, by contrast, this effect was reversed, with a larger anomaly effect in the associated conditions. On its surface, this pattern suggests a very slight delay in the detection of a semantic anomaly when the subject is highly associated with the verb, and is a likely theme. It is important to note, however, that a sizable anomaly effect was

still present at the verb even when the subject and verb were highly associated: about 30 ms in first-pass time (averaging across task conditions) and about 60 ms in go-past time. Thus, the present results do not directly support the claim from the ERP literature (Kim & Osterhout, 2005) put forward to account for "semantic P600" effects in the absence of N400 modulations, namely that readers initially assign thematic roles in SRA sentences based on plausibility, rather than based on the actual syntactic structure. Our results may therefore be regarded as consistent with more recent ERP studies (e.g., Bornkessel-Schlesewsky & Schlesewsky, 2008; Brouwer, Fitz, & Hoeks, 2012; Stroud & Phillips, 2012) that emphasize the role of lexical pre-activation in explaining the lack of N400 amplitude modulation by SRAs in languages such as English (note that across languages, the situation is more complex; cf. Bornkessel-Schlesewsky et al., 2011; Tune et al., 2014). The present results demonstrate that in reading, the anomaly effect is present at the verb in both association conditions but is reduced when lexical association between subject and verb is high.

Our results complement prior findings on the role of regressions in reading comprehension. An important prior study is by Schotter, Tran, and Rayner (2014), who had participants read GP and unambiguous sentences for comprehension. In half of the sentences, an x-mask replaced words that the subject had just read, thereby inhibiting the uptake of useful information from regressions. Comprehension question accuracy, both for difficult questions after GP sentences and easy questions after unambiguous sentences, dropped significantly in the x-mask condition. However, for the GP sentences, there were no differences in accuracy between fully legible trials in which readers made a regression from the disambiguating region and those where they did not. Thus, it appears that the opportunity to make regressions is beneficial, and our data confirm that readers do use regressive re-reading when faced with difficult comprehension. But at the same time, Schotter et al.'s findings suggest that comprehension may not actually be higher when readers regress, compared with when they don't. Our data also failed to show evidence of improved comprehension when readers regressed from the end of the sentence (see also Christianson, Luke, Hussey, & Wochna, 2017, for a similar finding). Schotter et al. interpret their null finding as suggesting that regressions may be compensatory, improving comprehension up to, but not beyond, the level of comprehension in cases where the reader does not feel the need to re-read. We think this is a very plausible interpretation. The overall improvement in comprehension when readers have the opportunity to regress, coupled with the lack of comprehension benefit when they actually do regress, is expected if there are some cases in which first-pass reading yields an incomplete or defective representation of sentence meaning, and regressions are selectively initiated on these trials in order to improve comprehension.

Finally, it is worth considering the interpretation of the fact that task effects on regressive eye movements were reliably present only at the end of the sentence. We suggest that this pattern may reflect the operation of a checking mechanism, whereby, after completing an initial pass through the sentence, the reader assesses whether his or her understanding of the sentence is sufficiently clear that it suffices for the task at hand. In the present case, this is the decision about whether the reader's comprehension will suffice to answer the kind of comprehension question that the reader has come to expect. The answer to this question will more often be "no" in the difficult task condition, leading to an increased probability of regression. If this checking mechanism is employed only at the end of each trial, the lack of interaction between task demands and sentence-internal difficulty-inducing manipulations (e.g., implausibility) may be expected. The present experiment cannot, however, answer the question of whether such a checking mechanism is specifically deployed at the end of each sentence, perhaps as part of a sentence "wrap-up" mechanism (Hirotani, Frazier, & Rayner, 2006; Rayner, Kambe, & Duffy, 2000), or if instead its deployment depends on the task-specific schedule with which comprehension probes are provided.

Acknowledgements

The authors thank Joshua Levy, Nikhil Garg, Wesley Albright, Atreyi Mukherji, Chhing Tiv and Hannah Simes for assistance with data acquisition.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a German Academic Exchange Service scholarship to Anna Fiona Weiss. Ina Bornkessel-Schlesewsky is the recipient of an Australian Research Council Future Fellowship (Project Number: FT160100437). The funders had no influence on study design, data collection and analysis or manuscript preparation.

Note

1. The number of semantic reversal anomaly (SRA) items was motivated by the 2 × 2 factorial design of that subexperiment, with the 36 items allowing each subject to read nine sentences in each condition. However, the number of relative clause (RC) and green park (GP) sentences was determined simply by the ease of constructing these sentences and the general desire to avoid too many repetitions of the same structure; for these sentence types, there was no need to, for example, use an even number of items.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). *Fitting linear mixed-effects models using lme4* (arXiv Preprint arXiv:1406.*5823)*. Retrieved from http://arxiv.org/abs/*1406*.5823

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., . . . Schlesewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain & Language*, *117*, 133–152.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews*, *59*, 55–73. doi:10.1016/j.brainresrev.2008.05.003

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143. doi:10.1016/j.brainres.2012.01.055

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*, 977–990.

Christianson, K. (2016). When language comprehension goes wrong for the right reasons: Good-enough, underspecified, or shallow language processing. *The Quarterly Journal of Experimental Psychology*, *69*, 817–828. doi:10.1080/17470218.2015.1134603

Christianson, K., Luke, S. G., & Ferreira, F. (2010). Effects of plausibility on structural priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 538–544.

Christianson, K., Luke, S. G., Hussey, E. K., & Wochna, K. L. (2017). Why reread? Evidence from garden-path and local coherence structures. *The Quarterly Journal of Experimental Psychology*, *70*, 1380–1405.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, *42*, 621–636.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164–203.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.

Ferreira, F., & Patson, N. D. (2007). The "good enough" approach to language comprehension. *Language & Linguistics Compass*, *1*, 71–83.

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language*, *54*, 425–443.

Kaakinen, J. K., & Hyönä, J. (2010). Task effects on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1561–1566. doi:10.1037/a0020693

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52*, 205–225. doi:10.1016/j.jml.2004.10.002

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*, 262–284. doi: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*, 12–35. doi:10.1037/0096-3445.135.1.12

Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain & Language*, *8*, 51–36. doi:10.1016/S0093-934X(02)00548-5

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*, 117–129.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323.

Liversedge, S. P., & Rayner, K. (2011). Linguistic and cognitive influences on eye movements during reading. In I. Gilchrist & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 751–766). Oxford, UK, Oxford University Press.

Logačev, P., & Vasishth, S. (2016). A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension. *Cognitive Science*, *40*, 266–298. doi:10.1111/cogs.12228

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *5*, 21–56. doi:10.1016/j.cogpsych.2005.06.001

Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye-movement control in reading. *Psychological Research*, *72*, 675–688. doi:10.1007/s00426-008-0173-3

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

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1290–1301. doi:10.1037/0278-7393.30.6.1290

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *10*, 125–157.

Schotter, E. R., Bicknell, K., Howard, I., Levy, R., & Rayner, K. (2014). Task effects reveal cognitive flexibility responding to frequency and predictability: Evidence from eye movements in reading and proofreading. *Cognition*, *13*, 11–27. doi:10.1016/j.cognition.2013.11.018

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

Staub, A. (2007). The return of the repressed: Abandoned parses facilitate syntactic reanalysis. *Journal of Memory and Language*, *57*, 299–323.

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, *11*, 71–86. doi:10.1016/j.cognition.2010.04.002

Staub, A. (2015). Reading sentences: Syntactic parsing and semantic interpretation. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 202–216). Oxford, UK: Oxford University Press.

Staub, A., & Clifton, C., Jr. (2006). Syntactic prediction in language comprehension: Evidence from *either. . . or*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 425–436.

Stroud, C., & Phillips, C. (2012). Examining the evidence for an independent semantic analyzer: An ERP study in Spanish. *Brain & Language*, *120*, 108–126.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, *36*, 201–216. doi:10.3758/MC.36.1.201

Tune, S., Schlesewsky, M., Small, S. L., Sanford, A. J., Bohan, J., Sassenhagen, J., & Bornkessel-Schlesewsky, I. (2014). Cross-linguistic variation in the neurophysiological response to semantic processing: Evidence from anomalies at the borderline of awareness. *Neuropsychologia*, *56*, 147–166.

Wotschack, C., & Kliegl, R. (2013). Reading strategy modulates parafoveal-on-foveal effects in sentence reading. *The Quarterly Journal of Experimental Psychology*, *66*, 548–562. doi:10.1080/17470218.2011.625094