

ARTICLE

The influence of text segmentation on garden path processing: evidence from self-paced reading and eye-tracking

Andromachi Tsoukala¹ , Margreet Vogelzang^{2,3}  and Ianthi Maria Tsimpli³

¹Institut für Niederlandistik, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany; ²School of Psychology, Newcastle University, Newcastle upon Tyne, UK and ³Department of Theoretical and Applied Linguistics, University of Cambridge, Cambridge, UK

Corresponding author: Ianthi Maria Tsimpli; Email: imt20@cam.ac.uk

(Received 31 January 2025; Revised 12 May 2025; Accepted 13 May 2025)

Abstract

Line breaks are ubiquitous in continuous text, as in this article. Despite this prevalence, their effects on parsing and interpretation have been markedly understudied in previous research on written language processing. To shed light on these effects, we conducted a self-paced reading and an eye-tracking study in which participants read multiline texts that contained direct object–subject ambiguity, a type of temporary clause boundary ambiguity. Within these texts, we manipulated the placement of line breaks so that they either regularly coincided or clashed with clause boundaries. We hypothesised that this manipulation would cause readers to adjust their parsing strategies and interpretative commitments. Results revealed that the way in which text is segmented through line breaks can significantly affect how readers parse syntactically ambiguous structures. While coinciding line breaks and clause boundaries helped readers arrive at the correct analysis of the ambiguous structures, cases of line break and clause boundary clash led readers down the garden path during online processing, and in some cases also impacted their comprehension. Findings are discussed in terms of their implications for the importance of text segmentation in real-world settings, such as books, educational material and digital content.

Keywords: implicit prosody; line breaks; priming; reading; syntactic ambiguity; text segmentation

1. Introduction

An integral feature of continuous text is that it is segmented into distinct lines. This formal property of text and the influence it may have on linguistic analysis have remained severely understudied. This is partly because most prior research has focused on manipulations of textual content rather than form and partly because short sentential stimuli have been preferred over multiline text.

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

To address this gap, the present studies shed light on the effects of text segmentation. We manipulated whether line breaks coincided or clashed with clause boundaries within multiline texts that contained direct object–subject ambiguity, a well-studied type of local/temporary clause boundary ambiguity (e.g. Christianson et al., 2001; Frazier & Rayner, 1982; Mitchell, 1987; Staub, 2007). Our aim was to examine whether readers consider this formal textual feature to adjust their parsing strategies and interpretative commitments. We regard this work as an important step towards understanding whether nonlinguistic properties of text are irrelevant for parsing or whether they play a role in determining syntactic-prosodic analysis decisions. Throughout this article, we refer to ‘syntactic-prosodic’ analysis because we assume correspondence constraints, which often – but not always – result in alignment between syntactic and prosodic boundaries (see Cole, 2015 for a review).

1.1. Line break cue and implicit prosody

The ways in which means of text segmentation, such as line breaks, can affect parsing and interpretation have received little attention in previous research. In fact, line breaks and other visuospatial segmentation manipulations (e.g. phrase-by-phrase reading) were originally disregarded in the broader sentence processing literature and in syntactic ambiguity research more specifically. They were treated as ‘extraneous factors’ (Rayner et al., 1989, p. SI37) that bias readers against their ‘own better judgement’ (Fodor, 1989, p. SI166). Accordingly, any effects that could have been induced by line breaks or other segmentation means (e.g. as in Ferreira & Clifton, 1986; Mitchell, 1987; Rayner & Frazier, 1987) were in later work criticised as accidental and artificial (e.g. Adams et al., 1998; Trueswell et al., 1994). Since then, it has become the norm to avoid splitting text into different lines, or in cases where that has not been possible, to conduct further experiments to rule out confounding segmentation effects (e.g. Sturt et al., 2002).

However, some have considered line breaks to be relevant for the parser, arguing that they ‘are interpreted by subjects as signals for potential clause endings’ (Kennedy et al., 1989, p. SI51). We can think of this as the line break cue. The core idea was that the visual information at the end of the line encourages structural closure due to the spatial discontinuity between lines during reading (also coupled with the lack of unprocessed parafoveal information to the right of the break; see Mitchell, 1987). Thus, this function of the line break cue in marking clausal endings was originally understood to be a by-product of visuospatial information encoding processes. In later work, the idea that implicit prosody could be a relevant factor was entertained.

A key impetus into subvocal prosody research was the Implicit Prosody Hypothesis (IPH) (Fodor, 2002; see also the Prosodic Constraint on Reanalysis proposal by Bader, 1998). According to the IPH, a default prosodic contour is projected onto text during silent reading, resulting in both a prosodic and syntactic structure being generated. This is because the parser postulates syntactic boundaries at the location of prosodic breaks. Crucially, in the face of ambiguity, this prosodic packaging helps identify sentence constituents, clarifies structure and facilitates interpretation.

Clearly, studying implicit prosodic phrasing presents unique challenges, particularly due to the silent nature of the phenomenon and individual differences in the use of subvocalisation during reading. Despite these difficulties, several investigations conducted over the past two decades have yielded promising results that are consistent

with the predictions of the IPH (see Breen, 2014 for a review). Some of these studies have used manipulations associated with the imposition of prosodic boundaries during reading. These include comma punctuation (Drury et al., 2016; Hirotani et al., 2006; Steinhauer, 2003; Steinhauer & Friederici, 2001; see also Jun & Bishop, 2015; Staub, 2007), segment-by-segment sentence presentation (Swets et al., 2007), frame breaks in self-paced reading (Hirotani et al., 2016) and physical line breaks in text (Hirotani et al., 2016; Traxler, 2009).

Most relevant for present purposes are the findings reported in Hirotani et al. (2016). In their first experiment, they observed that frame-break-induced boundaries affected how syntactically complex structures were parsed online, and in their second experiment, they found that line-break-induced boundaries determined what interpretations of globally ambiguous structures were preferentially reached offline. These findings were attributed to prosodic packaging caused by text segmentation means, which affected how materials placed within the same (prosodic) processing unit were analysed. Based on these results, the authors proposed an extension to the IPH, namely the Line Break Hypothesis, which states that ‘line breaks inserted into written text induce implicit prosodic boundaries’ (Hirotani et al., 2016, p. 1). Below, we consider whether line breaks can play a similar disambiguating role in object–subject garden paths.

1.2. Object–subject ambiguity

One type of local/temporary syntactic ambiguity that has been widely studied over the past decades is found in direct object–subject garden path sentences, as shown below.

- (1) While Anna dressed the baby played in the crib.
- (2) While Anna dressed the baby
played in the crib.
- (3) While Anna dressed
the baby played in the crib.

Example (1), found in Christianson et al. (2001), illustrates what is often referred to as ‘object–subject’ ambiguity. This is because the noun phrase ‘the baby’ could be analysed in two alternative ways, depending on whether the dependent clause verb ‘dressed’ is transitive or intransitive. If transitive, it takes as its object the noun phrase ‘the baby’; if intransitive, it is reflexive (that is, Anna dressed herself), while the noun phrase ‘the baby’ is the subject of the verb ‘played’ in the main clause. Out of these two options, the latter is grammatical; yet, in the absence of a comma to demarcate clause boundaries, readers are often ‘garden pathed’. That is, they tend to entertain the transitive misparse initially, only to realise later on that reanalysis is needed once they encounter the disambiguating information (‘played’ in the example above). Typically, the induced garden path effects manifest as processing difficulty (e.g. longer reading times) and poorer comprehension, as the initial misinterpretation often persists. This may be because syntactic repair is initiated but only partially completed to a ‘good enough’ standard (e.g. Christianson et al., 2001), or fully completed but susceptible to interference from the lingering memory trace of the misanalysis (Slattery et al., 2013), or even completed to varying degrees of coherence and co-existing with other local interpretations simultaneously (Ceháková & Chromý, 2023).

In an early theoretical framework known as the garden model (Frazier, 1978; Frazier & Rayner, 1982), this misanalysis phenomenon was attributed to a syntax-based parsing bias, namely a tendency to attach incoming words inside the clause currently being processed (late closure), rather than outside it (early closure). In later work, different sentence processing models were proposed, such as the ‘good enough’ approach (Christianson et al., 2001; Ferreira et al., 2002, 2001) and constraint-based accounts (MacDonald et al., 1994; Seidenberg & MacDonald, 1999), which motivated further research to better understand factors that influence garden path effects. A key question was whether comprehenders are guided solely by syntax-based parsing strategies, or whether nonsyntactic information is available early on to influence parsing decisions. Since then, several investigations have been conducted to examine whether nonsyntactic factors, such as lexico-semantic, contextual and prosodic information, can mitigate or block garden path effects (for reviews, see Lee & Watson, 2012; Pickering & Van Gompel, 2006).

Within this literature, a question that has remained unaddressed concerns the effects of line breaks and, in particular, the extent to which they can help the reader avoid the late closure misanalysis instead of promoting it. For instance, when sentences containing object–subject ambiguity are segmented as in (2), readers display strong transitivity reflexes not only with optionally transitive verbs but also with obligatorily intransitive verbs (Mitchell, 1987; but see Adams et al., 1998; Staub, 2007). Such findings may point towards a ‘powerful influence’ of nonsyntactic cues on structural closure (Kennedy et al., 1989, p. SI70), affecting how material placed within the same (prosodic) processing unit is analysed (Hirotani et al., 2016). By extension, one could hypothesise that garden path effects are not observed when the text is segmented in a way that promotes early closure, as in (3), where the line break after the verb aids in triggering a prosodic break, which corresponds with a legitimate clause boundary.

This is the first hypothesis we tested in the present studies. Below, we present relevant literature that led us to develop our second hypothesis.

1.3. *Flouting the line break cue*

The information presented so far suggests that parsers are likely to assume structural closure at line endings, either due to visuospatial information encoding parameters (Kennedy et al., 1989) or prosodic packaging (Hirotani et al., 2016).

However, readers cannot always realise syntactic–prosodic closure at the end of a line. This is particularly relevant in cases where the discontinuity of the line conflicts with the continuity of the syntax in a conspicuous way. Consider cases in which syntactic constituents are ‘scissored’ by line endings, such as a line-final article being separated from a noun complement positioned on the next line. In that scenario, readers would know pre-break not to assume closure, as obligatory elements will need to be incorporated into the phrasal unit post-break. There is some evidence from studies using nonprosaic text (poetic fragments) that readers develop distinct processing strategies to deal with severed syntax on the line (Koops van’t Jagt et al., 2014; Schafffler et al., 2022). For instance, during silent reading, parsers recognise incompleteness in fragments with scissored syntactic constituents, which causes them to adopt a ‘fast forward’ strategy in search of closure, that is, a speed-up in pre-break regions (Koops van’t Jagt et al., 2014). Similarly, in oral recitation, when a line end

scissors a syntactic unit, readers display certain prosodic adaptations, such as a shorter pause and no F_0 reset (Schauffler et al., 2022).

Thus, while there are cases in which line breaks coincide with structural boundaries and act as ‘good’ or helpful cues to closure (Kennedy et al., 1989), there are also cases in which line endings break up syntactic-prosodic units; in such cases, they can be conceptualised as ‘bad’ or misleading cues. This ‘good-bad’ conceptual distinction is inspired by previous related work on effects of disfluencies on syntactic parsing. Specifically, Bailey and Ferreira (2003) found that when disfluencies, such as ‘uh’, occurred in positions that clashed with a clause boundary in aurally presented object–subject garden paths (late closure promoted), listeners were less likely to judge them as grammatically acceptable compared to when the disfluency closely followed the clause boundary (early closure promoted). Based on these findings, the authors argued that disfluencies can serve as both good and bad cues to sentence structure. We consider that line breaks can also be thought of in a similar way: they can act as helpful cues by demarcating linguistic structure in reading, but they can also cause reading disfluencies when they clash with structural boundaries.

Furthermore, the findings of Koops van’t Jagt et al. (2014) and Schauffler et al. (2022) suggest that parsers adapt their syntactic–prosodic processing decisions in response to the structural incompleteness of textual lines. What has not been investigated yet is whether incompleteness can be ‘regularised’ and anticipated in parsing. To explicate, consider a scenario where readers are repeatedly exposed to syntactically incomplete lines at the beginning of a text. After this exposure, they encounter a line that could be perceived as structurally complete or incomplete (e.g. a line ending featuring an optionally transitive verb, as in (3)). Under these conditions, it is hypothesised that parsers would not be quick to assume that the end of that line signals the end of a clause (line break cue). Rather, they may hold off on syntactic-prosodic closure. In other words, repeated exposure to syntactically incomplete lines would render readers likely to flout the line break cue which signals closure, instead expecting upcoming material to be accommodated in the parse post-break (e.g. an object argument of the verb). Essentially, this manipulation of exposure frequency resembles a priming manipulation.

Priming is observed when prior exposure to a stimulus (prime item or structure) influences the processing of subsequent information (target item or structure) due to lexical repetition or featural overlap at semantic, structural or other representational levels (for a review of lexical priming, see Jones & Estes, 2012; for structural priming, see Pickering & Ferreira, 2008; Tooley & Traxler, 2010). Previous research has shown that exposure to particular syntactic frames – even in the absence of lexical repetition – can modulate the pre-activation of a given parse, facilitating processing in subsequent trials (e.g., Fine et al., 2013; Pickering et al., 2013; Prasad & Linzen, 2021; Tooley & Bock, 2014; Traxler, 2008). We consider that the manipulation of exposure frequency to line-clause coterminality or clash can be thought of as a similar priming manipulation. Specifically, we assume that early lines within a given item can shape participants’ expectations regarding subsequent line breaks, thereby modulating the activation of competing syntactic–prosodic analyses as the item unfolds. Crucially, these within-item effects occur in conjunction with broader between-item adaptation mechanisms, whereby readers adjust to the experimental task itself as well as to the structural patterns they encounter repeatedly over the course of the experiment, a phenomenon well-documented in the literature (e.g., Chromý & Tomaschek, 2024).

1.4. *The present studies*

The present studies had two main aims. First, we wanted to test whether line breaks can disambiguate object–subject garden paths by triggering early closure and blocking the late closure misanalysis. To that end, we designed multiline texts in which a locally ambiguous verb that could be analysed as transitive or intransitive was positioned at a line end. We hypothesised that readers will be less likely to analyse the verb as transitive and more likely to analyse it as intransitive; this is due to the function of the line break in triggering early closure (Hirotani et al., 2016; Kennedy et al., 1989). This hypothesis (H1) applies in texts that contain no repeated line break and syntactic boundary clashes (hence readers have no reason to flout the line break cue).

Our second aim was to test how strong this line break cue to closure is and whether it can be overridden via repeated exposure to line break and syntactic boundary clashes. Thus, we manipulated the aforementioned texts so that prior to the critical line ending with the ambiguous verb, all lines would be syntactically incomplete. We hypothesised that readers will be more likely to analyse the verb as transitive and less likely to analyse it as intransitive; this is due to repeated exposure to line break and syntactic boundary clashes, making them flout the line break cue and refrain from early closure. This hypothesis (H2) applies in texts containing repeated line break and syntactic boundary clashes.

To investigate these hypotheses, adult participants completed a self-paced reading (Study 1) and an eye-tracking experiment (Study 2). The eye-tracking experiment allowed us to follow up on the results of the self-paced reading experiment and gain data of higher spatiotemporal precision.

2. Study 1: Self-paced reading

2.1. *Participants*

Out of the originally recruited 42 participants, 39 native English speakers (21 females, $M_{AGE} = 21.3$, $SD_{AGE} = 2.08$) formed this study's sample (see [Data Analysis](#) for exclusions). All were university students in the UK. They were recruited through the Prolific platform (<https://www.prolific.com/>) and mailing lists at the University of Cambridge. Participants had no history of dyslexia, neurological or psychiatric disorders. They all provided informed consent and received £15 as payment. Both Study 1 and 2 have received ethical approval by the ethics committee of the Modern and Medieval Languages and Linguistics faculty at the University of Cambridge.

2.2. *Materials*

The items were 32 poem-like texts consisting of five lines. In a 2×2 experimental design, we manipulated Transitivity (transitive or intransitive third line verb) and Line Completeness (complete or incomplete lines preceding the third line). The resulting four conditions are shown in [Table 1](#). For the full list of items, see the [Supplementary Materials](#).

In the conditions labelled as Complete, the first and second line contained syntactically complete clauses, namely 'Alice attended a talk' and 'that the speaker named James gave'. In Incomplete conditions, the clause-final words 'talk' and 'gave' were transposed on the immediately subsequent line so that these clauses would be

Table 1. Example of an item showing the experimental conditions

Condition		Example
Line completeness	Transitivity	
Complete	Intransitive	Alice attended a talk that the speaker named James gave and because Alice heckled James who was sickly and frail was a little mortified
Complete	Transitive	Alice attended a talk that the speaker named James gave and because Alice heckled [I] ^a James who was sickly and frail she was truly mortified
Incomplete	Intransitive	Alice once attended a [I] talk that the speaker named James [I] gave and when Alice heckled James who was sickly and quite frail was a bit mortified
Incomplete	Transitive	Alice once attended a [I] talk that the speaker named James [I] gave and when Alice heckled [I] James who was sickly and quite frail she was so mortified
Question: Options:		Who was it that was mortified? (a) Alice (b) James (c) Other(s)

^a[I] marks all lines which preceded or contained the verb region that were structurally incomplete; this was not shown to participants.

syntactically incomplete. Given this displacement, minor lexical substitutions or insertions were considered for Incomplete conditions to control for line length (e.g. addition of the adverb ‘once’ on the first line, replacement of the conjunction ‘because’ with ‘when’ on the third as shown in Table 1). Accordingly, syllable count on each line remained constant ($N = 7$) across items and within-item conditions. Importantly, the conspicuous ‘scissoring’ of clauses at line endings from the beginning of texts was designed to create an expectation that subsequent lines would also be syntactically incomplete and that upcoming material would need to be incorporated into the parse after each line break.

This Line Completeness manipulation is important to consider as we now turn to the dependent clause ‘because/when Alice heckled’ on line 3. Therein, a verb that could be analysed as transitive or intransitive was placed in line-final position, as in ‘heckle’. In Transitive conditions, the verb would take as its object the proper name found at the beginning of line 4 (‘James’), whereas in Intransitive conditions, this proper name would be the subject of a main clause on the final line (‘was a little/bit mortified’). Texts remained temporarily ambiguous until line 5 where the presence or absence of a new subject for the main clause, namely an anaphoric pronoun (e.g. she), would determine whether the verb was transitive or intransitive, respectively. Note that if a pronoun was present – thus rendering the verb transitive – global referential ambiguity was avoided, as only one pronoun antecedent met anaphoric binding criteria (e.g., ‘she’ refers to ‘Alice’, not ‘James’). Importantly, the inclusion of structurally incomplete lines prior to the presentation of the third line ambiguous verb was intended to influence parsing decisions on its transitivity status.

For all stimuli, including fillers, a multiple choice comprehension question was generated of the form ‘Who did what’. The question for critical items enquired about the subject of the main clause verb on line 5, e.g. ‘Who was it that was mortified?’ The options were third-line candidate (‘Alice’), fourth-line candidate (‘James’) or ‘Other (s)’ as a fallback option.

The items included locally ambiguous verbs (e.g. optionally transitive, reflexives) used in Adams et al. (1998) and Mitchell et al. (2008). We also normed the stimuli with an independent group of native English speakers ($N = 20$, 10 females, $M_{AGE} = 21.4$, $SD_{AGE} = 3.05$). Our aim was to ensure that the chosen verbs would lead to the activation of the two alternative analyses. Results revealed that although both analyses were activated (about one-third of raters found the transitive and intransitive versions equally comprehensible), raters favoured the transitive interpretation (44.3%) over the intransitive one (20.3%), a difference that was significant ($p = 0.003$). For further details and discussion of the norming results, see the [Supplementary Materials](#).

Apart from the 32 critical items, participants also read 80 five-line texts, including distractors containing global ambiguity and unambiguous fillers. Four counterbalanced lists were prepared. Participants saw eight items for each one of the four conditions, and each item was seen in only one of its four versions.

2.3. Procedure

This web-based reading study employed the self-paced (line-by-line) moving-window paradigm, programmed in JsPsych (de Leeuw, 2015).

To approximate lab-based experimental conditions, a remote testing method was used, which involved participants completing the study while being on a live call with the experimenter. Participants were informed that they would view short texts made up of a few lines, which they had to read at their normal pace, following which they would be presented with a comprehension question. We did not mention anything more specific about the structure of the stimuli in order to avoid biasing participants or drawing their attention to the poetic particularities of the stimuli.

During testing sessions, participants started the main reading task, after going through three practice items. The main reading task was split in five blocks. The first four blocks contained 22 texts each, while the last one consisted of 24 texts. In between blocks, participants could take a short break and would then proceed to complete a cognitive task (results on these tasks are beyond the scope of the present report).

2.4. Data analysis

Prior to analyses, three participants who gave more than 20% incorrect responses to questions following unambiguous fillers were excluded, yielding a sample size of thirty-nine. Then, trials in which participants had responded with ‘Other(s)’ to the question for the critical items were excluded (0.8% data loss). Next, reading time data were checked for outliers. Inspection of line reading times through histograms indicated that most values ranged between 500 and 10000 ms, so we trimmed data falling outside this range (7.1% data loss) and then log-transformed data to reduce skew. Finally, to account for differences in line character count between items, we

regressed logged reading times per line on the character count of the respective line and extracted the residuals of these models, which we then used as the dependent variables in the main analyses. We analysed residual reading times for line 5, which contains the disambiguating region and is central to our hypotheses. Additionally, we analysed data for lines 3 and 4, since they contain information relevant for disambiguation, namely the transitive/intransitive verb and the object/subject noun phrase, both commonly examined in the garden path literature. In contrast, lines 1 and 2 were not relevant for our analyses.

Analyses were performed in R using the *lme4* package (Bates et al., 2015). We used linear mixed-effects models (LMEMs) to analyse reading times and generalised LMEMs for responses to the comprehension questions. We took the following modelling steps. Firstly, we started with an empty model and used the Akaike Information Criterion (AIC) to identify the random effects structure that best fitted the data (Matuschek et al., 2017). We built and compared (a) a model with only by-participant and by-item intercepts, (b) a model that also included by-participant slopes for Line Completeness and Transitivity, (c) a model with by-item slopes for these factors instead and (d) a ‘maximal’ model with both by-participant and by-item intercepts and slopes for Line Completeness and Transitivity. The maximal model often led to non-convergence issues, so we selected among one of the other options based on AIC. Subsequently, Line Completeness (negative level: Complete) and Transitivity (negative level: Intransitive) were deviation-coded and entered in all models as fixed effects along with their interaction. For interactions, we performed post-hoc tests and corrected for multiple comparisons with false discovery rate adjustments. As effect size indices, we report odds ratio (OR) or Cohen’s *d* (*d*).

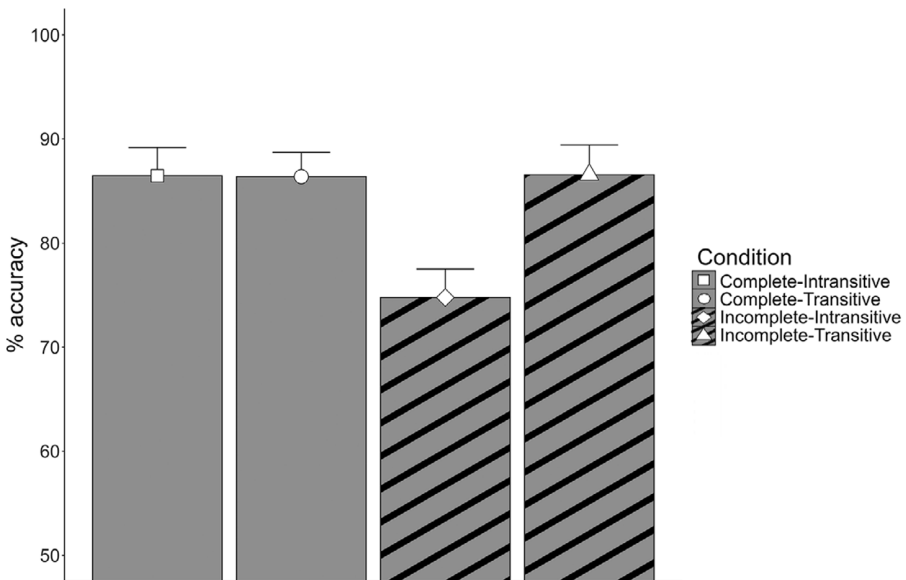


Figure 1. Mean percent correct responses to the comprehension question by condition in Study 1 (SE error bars).

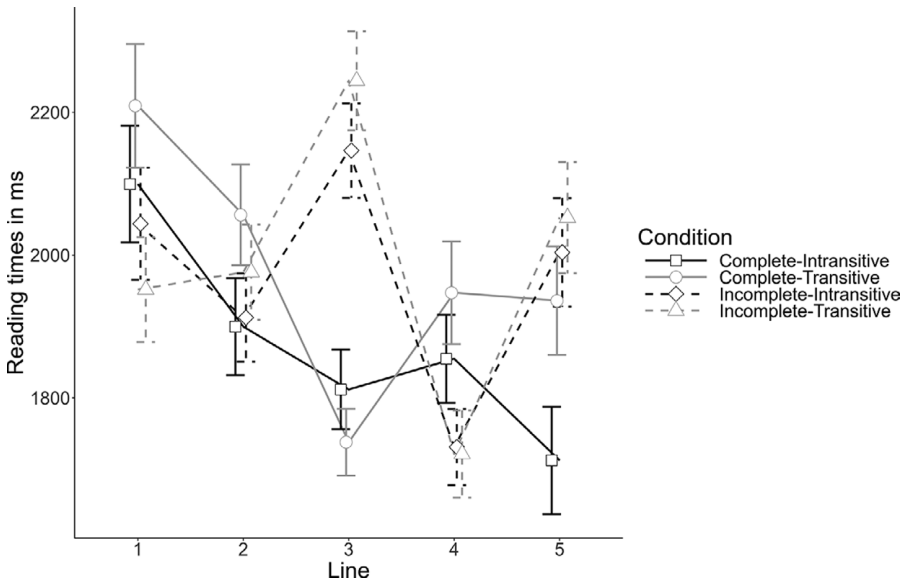


Figure 2. Mean reading times by condition in Study 1 (SE error bars).

Table 2. Summary of the statistical results in Study 1.

Measure	Predictors	Model estimate (CI)	p-value	OR/d effect size (CI)
Accuracy	Line completeness	−0.41 (−0.82, −0.002)	0.049*	OR = 0.66 (0.44, 1.00)
	Transitivity	0.62 (0.11, 1.13)	0.016*	OR = 1.87 (1.12, 3.10)
	Line completeness × Transitivity	1.07 (0.29, 1.84)	0.007**	
Reading Times Line 3	Line completeness	0.14 (0.10, 0.18)	< 0.001***	d = 0.40 (0.28, 0.52)
	Transitivity	0.02 (−0.02, 0.06)	0.420	d = 0.04 (−0.06, 0.16)
	Line completeness × Transitivity	0.05 (−0.03, 0.14)	0.213	
Reading Times Line 4	Line completeness	−0.05 (−0.09, −0.005)	0.027*	d = −0.13 (−0.25, −0.01)
	Transitivity	0.004 (−0.04, 0.05)	0.814	d = 0.01 (−0.10, 0.13)
	Line completeness × Transitivity	−0.03 (−0.12, 0.05)	0.455	
Reading Times Line 5	Line completeness	0.08 (0.02, 0.15)	0.012*	d = 0.18 (0.03, 0.33)
	Transitivity	0.08 (0.003, 0.17)	0.076	d = 0.17 (−0.03, 0.38)
	Line completeness × Transitivity	−0.12 (−0.23, −0.006)	0.038*	

Note: To derive effect sizes for significant interactions we performed post-hoc tests which we present in the text. Significant p values are highlighted in bold (* denotes $p < 0.05$; ** denotes $p < 0.01$; *** denotes $p < 0.001$).

2.5. Results

Mean accuracy and reading time results are shown in Figures 1 and 2, respectively. A summary of the statistical results can be found in Table 2. The full model outputs and descriptive statistics are provided in the Supplementary Materials.

Comprehension accuracy was well above chance in all conditions ($>74\%$). Analyses revealed an effect of Line Completeness, as lower odds of providing an accurate response were estimated in Incomplete conditions. Transitivity was also significant, as Transitive conditions were associated with greater odds of an accurate response. A significant interaction between Transitivity and Line Completeness was detected. Post-hoc tests revealed that accuracy in the Incomplete-Intransitive condition was significantly lower compared to all other conditions, namely Complete-Intransitive ($\beta = -0.94$, $p < 0.001$; OR = 0.38, 95% CI [0.19, 0.76]), Complete-Transitive ($\beta = -1.03$, $p < 0.001$; OR = 0.35, 95% CI [0.16, 0.74]) and Incomplete-Transitive ($\beta = -1.15$, $p < 0.001$; OR = 0.31, 95% CI [0.14, 0.70]). No other significant results were obtained through post-hoc comparisons.

Regarding reading time data, the model for line 3 revealed an effect of Line Completeness; participants took longer to process the region with the ambiguous verb in Incomplete conditions. Line Completeness also had an effect on line 4 reading times; in this case, participants read this region faster in Incomplete conditions.

In line 5, there was again an effect of Line Completeness; participants were slower to process the disambiguating region in Incomplete conditions. Importantly, a significant interaction between Line Completeness and Transitivity was detected. Post-hoc tests revealed that the interaction was driven by the Complete-Intransitive condition, which was read faster compared to all other conditions, namely Incomplete-Intransitive ($\beta = -0.14$, $p = 0.011$; $d = -0.31$, 95% CI [-0.50, -0.11]), Incomplete-Transitive ($\beta = -0.16$, $p = 0.026$; $d = -0.35$, 95% CI [-0.62, -0.09]) and Complete-Transitive ($\beta = -0.14$, $p = 0.026$; $d = -0.30$, 95% CI [-0.54, -0.06]). No other significant differences were detected through post-hoc comparisons, and there were no other significant effects in any of the models reported above.

2.6. Discussion

These results provide partial support for our hypotheses, as we detected the effects we expected in either online or offline measures, but not in both.

Regarding H1, which concerns Complete conditions, we expected the line break after the critical verb on line 3 to trigger early closure, thus promoting the intransitive analysis. Consistent with H1, post-hoc tests revealed that line 5 (disambiguating region) in the Complete-Intransitive condition was read significantly faster than the Complete-Transitive one. However, the comprehension results do not suggest similar facilitation for the intransitive analysis, as accuracy was similar in the Complete-Intransitive condition and the Complete-Transitive one.

Regarding H2, which concerns Incomplete conditions, we expected that repeated exposure to line break and clause boundary clash would render readers likely to flout the line break cue and not assume early closure on line 3, thus promoting the transitive analysis. Consistent with H2, we found that comprehension accuracy was better in the Incomplete-Transitive condition compared to Incomplete-Intransitive one. However, we did not find similar facilitation for the transitive analysis in online measures, as reading times in line 5 (disambiguating region) did not differ between the Incomplete-Transitive condition and the Incomplete-Intransitive one.

Finally, we also observed effects we had not anticipated. The reading time results indicated that Line Completeness affected participants' reading behaviour early on,

in pre-disambiguating regions. Participants were slower while processing line 3 with the ambiguous verb when they had been exposed to structurally incomplete line contexts as opposed to complete ones. In contrast to this, the contents of line 4 were viewed faster in incomplete line contexts compared to complete ones. We consider possible explanations to account for all these findings in the [General Discussion](#).

Overall, these results provided important insights; yet, there are certain limitations to the conclusions that can be drawn. The design of the stimuli meant that the disambiguating region was the final segment readers viewed. Thus, effects of interest cannot be disentangled from wrap-up processes triggered by the end of stimuli (however, wrap-up need not be left till the end of sentences; see Stowe et al., 2018). Additionally, limitations of the line-by-line moving-window paradigm we employed are that (1) only full line reading times are obtained, without specifying which word/section caused the effects and (2) readers cannot regress to previous regions they have already read.

To address these issues in Study 2, we used eye-tracking to gain high-precision data on areas of interest within lines and to have lines of text remain available for re-inspection.

3. Study 2: Eye-tracking during reading

3.1. Participants

Out of the originally recruited 35 participants, 29 native English speakers (22 females, $M_{AGE} = 19.4$, $SD_{AGE} = 1.1$) formed this study’s sample (see [Data Analysis](#) for

Table 3. Example of an Item with the IAs underlined

Condition		Example
Line completeness	Transitivity	
Complete	Intransitive	Alice attended a talk that the speaker named James gave and because Alice <u>heckled</u> James who was sickly and frail <u>was a little</u> mortified
Complete	Transitive	Alice attended a talk that the speaker named James gave and because Alice <u>heckled</u> [I] ^a <u>James</u> who was sickly and frail <u>she was truly</u> mortified
Incomplete	Intransitive	Alice once attended a [I] talk that the speaker named James [I] gave and when Alice <u>heckled</u> James who was sickly and quite <u>frail was a bit</u> mortified
Incomplete	Transitive	Alice once attended a [I] talk that the speaker named James [I] gave and when Alice <u>heckled</u> [I] <u>James</u> who was sickly and quite <u>frail she was so</u> mortified
Question:		Who was it that was mortified?
Options:		a) Alice b) James c) Other(s)

^a[I] marks all lines which preceded or contained the verb region that were structurally incomplete; this was not shown to participants.

exclusions). Participants had normal or corrected-to-normal vision. All were adult students at the University of Cambridge and none participated in Study 1.

3.2. Materials

The same materials as in Study 1 were used. Three interest areas (IA) were defined. On line 3, IA 1 was the locally ambiguous verb ('heckled'). On line 4, IA 2 was the object of the locally ambiguous verb, when the verb was transitive, or the subject of the main clause verb on line 5, when the verb was intransitive ('James'). Finally, IA 3 consisted of the first three words on line 5 so that it would contain the disambiguating word and spillover area. For an example of an item with the IAs underlined, see Table 3.

We decided to combine the first three words on the fifth line into a single IA to allow comparability across conditions. Had these words not been combined, the following complications would have arisen. Firstly, the difference between the disambiguating word in Transitive conditions (a pronoun such as 'he') and that in Intransitive conditions (the main clause verb) was almost double in terms of character count; in the former case, the mean was 4.7 (SD = 1.4), whereas in the latter, it was 2.5 (SD = 0.6). As such, the final IA needed to be length-matched between conditions. Secondly, another consideration was that the disambiguating word in Complete and Incomplete conditions occupied different spatial positions; in the former case, it appeared at the beginning of the fifth line, whereas in the latter case, it appeared in second position. The best solution to the aforementioned problems that could be thought of was to combine the first three words on the fifth line so that (a) in all cases, fixations spanning from the beginning of the line and falling on roughly the same horizontal and vertical area of the screen would be compared across all four conditions, (b) the area examined would not be too small and dissimilar in terms of length across conditions and (c) the resulting IA would include the disambiguating word along with the post-disambiguating word(s), thus taking into account spill-over effects. Table 4 shows the average length of the disambiguating IA, which did not differ between conditions ($p > 0.05$).

Alongside the 32 critical items, participants read 48 five-line texts, using a subset of distractors and unambiguous fillers from the self-paced reading study to avoid making the experiment too long. The question and response options remained the same as in Study 1.

3.3. Procedure

The experiment used an Eyelink 1000 Plus eye-tracker (SR Research) with a desktop-mounted camera at a 52 cm eye distance. Participants' eye movements were recorded

Table 4. Mean length of IA 3 in characters (including spaces)

Condition		Example	Mean	SD
Line completeness	Transitivity			
Complete	Intransitive	was a little	14.3	2.9
Complete	Transitive	she was truly	13.9	2.9
Incomplete	Intransitive	frail was a	13.9	2.2
Incomplete	Transitive	frail she was	13.6	1.5

at 1000 Hz using monocular tracking. A headrest was used to stabilise the participants' head. After instructions, participants completed a 9-point calibration procedure. Following validation, they read four practice items and then proceeded to the main items.

Care was taken to avoid complications related to return sweeps, that is, end-of-line saccadic eye movements that take the reader's eyes to the next line. Return sweeps have been associated with oculomotor error, affecting fixations landing on the following line (see Slattery & Parker, 2019 for review). Thus, we used a gaze-contingent display that presented the text in a cumulative manner, where each line appeared one after the other. To make a line appear, participants had to fixate on a cross positioned to the left of each line. Once the fifth line had appeared, they could look at an arrow sign positioned close to the right border of the screen to move to the comprehension question. This setup was explained to participants at the start of testing sessions using a picture-based demo. Additionally, as in Study 1, participants were instructed to read for comprehension at their normal pace.

This presentation mode offers both methodological benefits and some trade-offs, such as the potential for prolonging fixations from the end of one line to the beginning of the next. Although presenting all lines of text at once would have allowed for more natural reading conditions, this approach was chosen for its ability to yield more reliable data at the beginning of lines. Additionally, the cross on the first line served as a drift check, triggering recalibration if the eye-tracker failed to detect a 250 ms fixation. If a fixation to subsequent crosses (left of lines 2–5) was not detected within 5 seconds, trials would proceed as normal but be excluded from analysis. Participants read 20 texts in each one of the four blocks of this eye-tracking experiment. In between blocks, they could take a short break and did not complete any additional tasks.

3.4. Data analysis

As in Study 1, participants with more than 20% erroneous responses to the questions following unambiguous fillers were excluded from analyses ($N = 2$). Trials in which participants had responded with 'Other(s)' to the question for the critical items or had been timed out (no response recorded within 10 seconds) were excluded; data from one participant were discarded due to high data loss (> 37.5%) in some conditions. Trials in which the cross fixation trigger for lines 2–5 did not work were excluded; data from three participants were discarded due to high data loss (> 37.5%) in some conditions. Thus, these elimination steps led to the exclusion of six individuals' data out of the originally recruited 35 individuals, yielding a sample size of 29. Within this sample, the trial-level data loss due to the aforementioned eliminations steps was at 2.6%.

An automatic cleaning procedure was applied to the eye-tracking data using the Data Viewer software (SR Research). Firstly, fixations shorter than 80 ms were merged into the largest fixation that was in close proximity (threshold of 0.5 degrees of visual angle). Secondly, fixations shorter than 80 ms or longer than 1000 ms were discarded (Warren et al., 2009). The cleaning steps affected 11.1% of the data used in analyses.

Three standard eye-tracking measures were computed. First pass time corresponds to the sum of fixation durations in an IA, beginning with the first fixation until

the IA is exited either to the left or to the right. Rereading time corresponds to the sum of refixation durations in an IA occurring during regression episodes; that is, all cases where the IA is reentered after later material to the right has been viewed. Finally, total time is the sum of all fixation durations in an IA, regardless of their order.

Data from these eye-tracking measures for each IA were log-transformed and residualised for LMEM analyses, as in Study 1 (that is, we regressed the logged data on character count of each IA and analysed the residuals). If missing data were obtained for a measure (e.g. an IA was skipped), these observations were omitted from analyses. The analysis of question responses and all other modeling steps were the same as in Study 1.

Finally, we note that our hypotheses concern the disambiguating IA in which we expected to observe an interaction between Line Completeness and Transitivity. Since we tested for this interaction three times corresponding to our three eye-tracking measures (first pass, rereading and total time), we applied false discovery rate corrections for multiple testing, following guidelines by García-Pérez (2023). Below, we report both corrected and uncorrected *p* values.

3.5. Results

Mean accuracy and eye-tracking results are plotted in Figures 3 and 4, respectively. A summary of the statistical results can be found in Table 5. The full model outputs and descriptive statistics are provided in the Supplementary Materials.

Comprehension accuracy was well above chance in all conditions (> 80%). Only an effect of Transitivity was found, as the odds of providing an accurate response were greater in Transitive conditions.

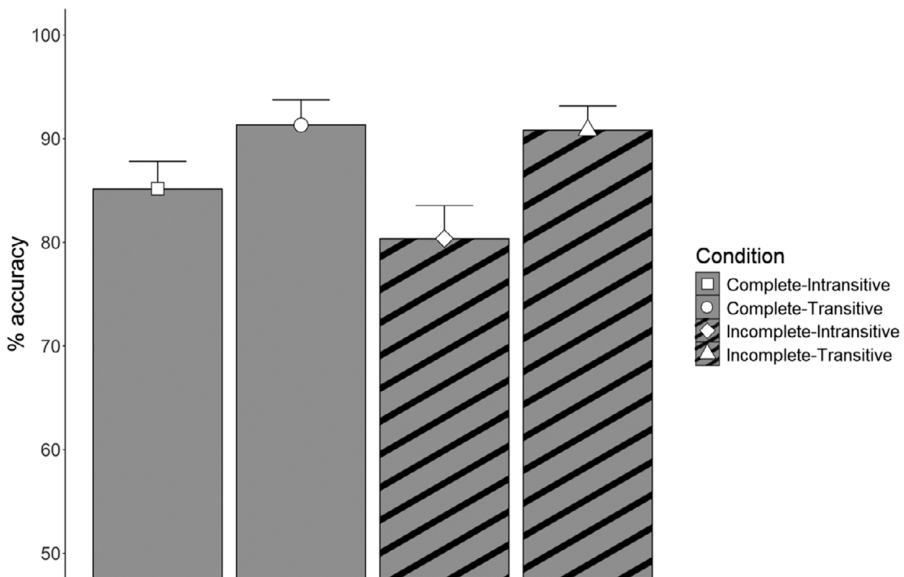


Figure 3. Mean percent correct responses to the comprehension question by condition in Study 2 (SE error bars).

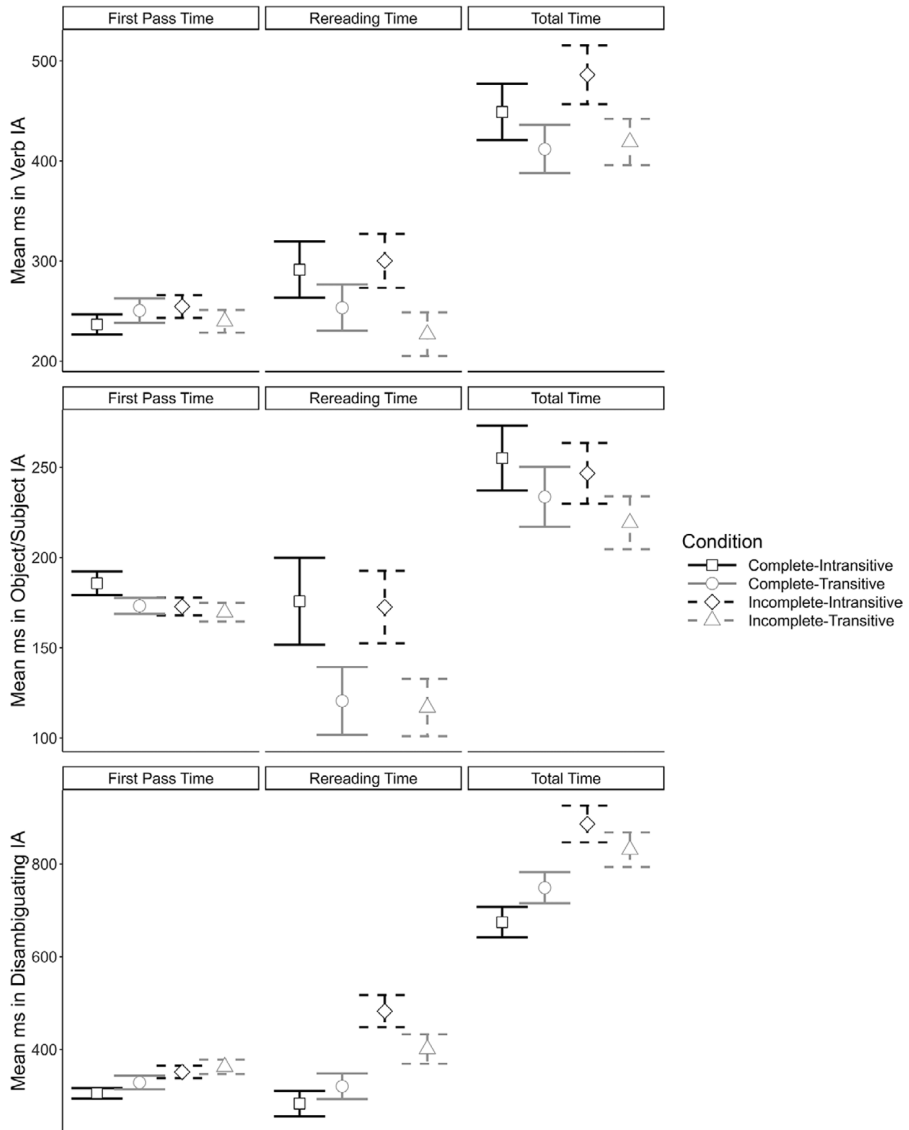


Figure 4. Mean first pass time, rereading time and total time in the three interest areas, namely IA 1 (i.e., verb; top), IA 2 (i.e., object/subject; middle) and IA 3 (i.e., disambiguating Region; bottom), by condition in Study 2 (SE error bars).

Regarding the eye-tracking data in IA 1 (verb), there was an effect of Transitivity in rereading time, as the locally ambiguous verb received shorter refixations in Transitive conditions. The same effect of Transitivity emerged for total time, suggesting again that when the verb was transitive, less total time was spent in the verb IA.

Regarding IA 2 (object/subject), analyses revealed an effect of Transitivity on rereading time and total time. Hence, similarly to what was reported above for IA

Table 5. Summary of the statistical results in Study 2.

Measure	Predictors	Model estimate (CI)	p-value	OR/d effect size (CI)
Accuracy	Line completeness	−0.24 (−0.65, 0.18)	0.264	OR = 0.79 (0.52, 1.19)
	Transitivity	0.83 (0.42, 1.24)	< 0.001***	OR = 2.29 (1.51, 3.47)
	Line completeness×Transitivity	0.26 (−0.57, 1.08)	0.541	
First pass time IA 1 (verb)	Line completeness	−0.005 (−0.07, 0.06)	0.869	d = −0.01 (−0.17, 0.14)
	Transitivity	−0.01 (−0.08, 0.05)	0.725	d = −0.02 (−0.18, 0.12)
	Line completeness×Transitivity	−0.08 (−0.21, 0.05)	0.240	
Rereading time IA 1 (verb)	Line completeness	−0.05 (−0.17, 0.07)	0.443	d = −0.07 (−0.27, 0.12)
	Transitivity	−0.15 (−0.28, −0.03)	0.013*	d = −0.24 (−0.44, −0.05)
	Line completeness×Transitivity	0.007 (−0.24, 0.25)	0.952	
Total time IA 1 (verb)	Line completeness	0.02 (−0.06, 0.11)	0.605	d = 0.03 (−0.10, 0.18)
	Transitivity	−0.13 (−0.21, −0.04)	0.004**	d = −0.21 (−0.35, −0.06)
	Line completeness×Transitivity	−0.09 (−0.26, 0.08)	0.320	
First pass time IA 2 (object/subject)	Line completeness	−0.03 (−0.08, 0.01)	0.162	d = −0.12 (−0.30, 0.05)
	Transitivity	−0.03 (−0.08, 0.02)	0.245	d = −0.10 (−0.28, 0.07)
	Line completeness×Transitivity	0.03 (−0.07, 0.12)	0.573	
Rereading time IA 2 (object/subject)	Line completeness	0.07 (−0.09, 0.23)	0.412	d = 0.11 (−0.15, 0.38)
	Transitivity	−0.17 (−0.32, −0.006)	0.041*	d = −0.27 (−0.54, 0.006)
	Line completeness×Transitivity	0.22 (−0.10, 0.54)	0.179	
Total time IA 2 (object/subject)	Line completeness	−0.07 (−0.16, 0.02)	0.124	d = −0.12 (−0.27, 0.03)
	Transitivity	−0.10 (−0.19, −0.02)	0.022*	d = −0.18 (−0.33, −0.02)
	Line completeness×Transitivity	0.01 (−0.17, 0.19)	0.903	
First pass time IA 3 (disambiguating region)	Line completeness	0.14 (0.06, 0.21)	< 0.001***	d = 0.32 (0.14, 0.50)
	Transitivity	0.04 (−0.03, 0.11)	0.254	d = 0.09 (−0.07, 0.27)
	Line completeness×Transitivity	−0.02 (−0.13, 0.09)	0.756 (uncorrected)	
Rereading time IA 3 (disambiguating region)	Line completeness×Transitivity	−0.02 (−0.13, 0.09)	0.756 (corrected)	
	Line completeness	0.09 (−0.05, 0.22)	0.215	d = 0.10 (−0.06, 0.27)
	Transitivity	−0.02 (−0.16, 0.11)	0.721	d = −0.03 (−0.20, 0.13)
Total time IA 3 (disambiguating region)	Line completeness×Transitivity	−0.14 (−0.41, 0.13)	0.308 (uncorrected)	
	Line completeness×Transitivity	−0.14 (−0.41, 0.13)	0.461 (corrected)	
	Line completeness	0.24 (0.15, 0.33)	< 0.001***	d = 0.41 (0.25, 0.58)
	Transitivity	0.06 (−0.05, 0.17)	0.301	d = 0.09 (−0.03, 0.23)
	Line completeness×Transitivity	−0.16 (−0.31, −0.01)	0.035* (uncorrected)	
	Line completeness×Transitivity	−0.16 (−0.31, −0.01)	0.103 (corrected)	

Note: To derive effect sizes for significant interactions, we performed post-hoc tests which we present in the text. Significant p values are highlighted in bold (* denotes p < 0.05; ** denotes p < 0.01; *** denotes p < 0.001).

1, when the verb turned out to be transitive, there were shorter refixations and less total time spent in the object/subject region.

As for IA 3 (disambiguating region), there was an effect of Line Completeness on first pass time, indicating that longer fixations were observed in Incomplete conditions. Similarly, an effect of Line Completeness was observed in total time, suggesting that in Incomplete conditions more total time was spent in the disambiguating IA. Finally, a Line Completeness and Transitivity interaction effect was detected ($p = 0.035$); yet, after applying corrections for multiple testing, the effect was rendered non-significant ($p = 0.103$).

In the spirit of comprehensiveness, we performed post-hoc tests to explore which differences drove this effect, though it should be interpreted with caution given its non-significance. Results revealed that less total time was spent in the disambiguating IA in Complete conditions compared to Incomplete ones (all p 's < 0.05). In fact, this is not surprising, given the aforementioned effect of Line Completeness. Thus, the interaction was likely driven by differences between the two Complete conditions; less total time was spent in the disambiguating IA in the Complete-Intransitive condition compared to the Complete-Transitive one, although this difference was marginal after corrections (beta = -0.14 , $p = 0.057$; $d = -0.24$, 95% CI [-0.48 , -0.002]). There was no significant difference between the two Incomplete conditions in the post-hoc tests ($p = 0.748$), and there were no other significant effects in any of the models reported above.

3.6. Discussion

These eye-tracking results are mostly consistent with the findings of the self-paced reading study.

Regarding H1, we expected that in Complete conditions the intransitive analysis will be promoted. In contrast to Study 1 where we found evidence of significant processing facilitation for the intransitive analysis, in this case we found only trends in the same direction. Specifically, when total time was examined, we found a marginal interaction between Line Completeness and Transitivity, which was likely driven by the relatively faster processing of the disambiguating IA in the Complete-Intransitive condition compared to the Complete-Transitive one. Yet, since this effect was rendered non-significant after applying corrections, this result should be interpreted with caution and warrants further investigation to assess its reliability. For now, we note that the eye-tracking result patterns generally align with the finding of processing facilitation for the intransitive analysis that we observed in the self-paced reading study, which provides support for H1. As for comprehension accuracy, we again found no evidence of facilitation for the intransitive analysis, as in Study 1. The main effect of Transitivity suggests that accuracy was lower in the Complete-Intransitive condition than the Complete-Transitive one.

Regarding H2, we expected that in Incomplete conditions the transitive analysis will be promoted. Consistent with H2, comprehension accuracy was better in the Incomplete-Transitive condition compared to Incomplete-Intransitive one. However, we did not find similar facilitation for the transitive analysis in online measures, as there were no differences in the disambiguating IA between the Incomplete-Transitive condition and the Incomplete-Intransitive one in the eye-tracking measures.

Additionally, we detected effects in pre-disambiguating regions. While in Study 1, significant effects of Line Completeness were detected in pre-disambiguating regions, in this study these effects were not detected; instead, we found significant effects of Transitivity in IA 1 (verb) and IA 2 (object/subject) in rereading time and total time. Essentially, when disambiguating material rendered the transitive analysis grammatical, readers did not regress to the verb IA or the object/subject IA to the same extent as they did when the verb turned out to be intransitive. We discuss all these findings in more detail below.

4. General discussion

4.1. *Do coinciding line breaks and clause boundaries encourage early closure?*

The first aim of the present studies was to test whether line breaks can help readers avoid the late closure misanalysis of object–subject garden paths. We hypothesised that, in contexts with syntactically complete lines, a line break after an optionally transitive verb will trigger early syntactic–prosodic closure (Hirotani et al., 2016; Kennedy et al., 1989), thus promoting the intransitive analysis.

The self-paced reading time results provide support for this hypothesis. In the two conditions where lines preceding the verb region were structurally complete, similar reading rates were observed for lines 3 and 4 but not for line 5. The final region was read significantly faster when disambiguating material rendered the intransitive analysis grammatical rather than the transitive one, suggesting that participants experienced facilitation when the verb was intransitive.

In the eye-tracking study, the result patterns were generally in the same direction as the ones of the self-paced reading study. For instance, a trend emerged in total time, suggesting that readers were somewhat faster to process the disambiguating region when the texts were resolved towards intransitivity compared to transitivity. However, this difference was not significant after applying corrections, and thus, this result should be interpreted with caution. For now, we note that the numerical differences were in the predicted direction, and consistent with the findings of Study 1. While not conclusive, this consistency across studies indicates that the effect may be subtle rather than absent, warranting further investigation by future research.

As for comprehension accuracy, we did not find evidence of facilitation for the intransitive analysis in either study. Thus, even though the online reading time results suggest facilitation for the intransitive analysis, this did not lead to better offline comprehension. We believe this is due to the nature of the disambiguating information. That is, the presence of a gendered pronoun on the final line when the verb was transitive probably helped identify the referent that participants were asked about (e.g. the pronoun *he/she* helps clarify whether the referent is James or Alice). This was not the case when the verb was intransitive since there was no such pronoun. Hence, this imbalanced cue may have boosted comprehension accuracy when the verb was transitive. Another possibility is that transitive structures are generally comprehended more easily, as is typically observed in the object–subject garden path literature, or that participants have a preference for the transitive interpretation, as was also observed in our norming study.

Based on our combined results across studies, we tentatively conclude that line break-induced boundaries are relevant during real-time processing and can encourage early closure, affecting online parsing even if offline comprehension responses are dissociable.

4.2. *Do clashing line breaks and clause boundaries discourage early closure?*

The second aim of our studies was to test how strong this line break cue to closure is and whether it can be overridden via repeated exposure to line break and clause boundary clash. We hypothesised that in contexts where participants have been exposed to syntactically incomplete lines, the line break after the optionally transitive verb will not be interpreted as a cue to syntactic–prosodic closure. Thus, this repeated exposure to syntactic incompleteness was expected to make readers likely to flout the line break cue and refrain from early closure, essentially promoting the transitive analysis, rather than the intransitive one.

We found mixed evidence for our second hypothesis. On the one hand, the offline comprehension results suggest facilitation for the transitive analysis. In contexts where line breaks routinely clashed with syntactic boundaries, readers exhibited better comprehension when the ambiguous texts were resolved towards transitivity rather than intransitivity. This was reliably observed in both studies.

On the other hand, the online self-paced reading and eye-tracking results indicate that syntactically incomplete line contexts led to a similar reading behaviour in the disambiguating region, regardless of whether the verb turned out to be transitive or intransitive. Thus, no facilitation for either analysis was observed in reading times. To better appreciate the lack of facilitation, we could also do another comparison between syntactically incomplete and complete contexts. The outcome of this comparison indicates disruption in the former case, as significantly slower reading rates were observed when participants encountered disambiguating information in incomplete line contexts compared to complete ones; this result pattern was reliably observed both in the reading times of the final, disambiguating region in the self-paced reading study as well as in first pass time and total time spent in the disambiguating IA in the eye-tracking study.

Beyond this disruption, there was also some evidence in the self-paced reading study that having successive lines be structurally incomplete affected parsing early on. Participants took longer to process line 3 with the verb when they had been exposed to structurally incomplete lines compared to complete ones. This finding could suggest that, prior to disambiguation, parsers became sensitive to the two competing analyses upon encountering the ambiguous verb. The inflated reading times may point towards high entropy (Hale, 2003), that is, readers experiencing uncertainty about a probabilistic outcome, namely whether they should assume syntactic–prosodic discontinuity (line break cue) or continuity (given previous exposure to incomplete lines). Subsequently, line 4 was read faster in structurally incomplete line contexts compared to complete ones. This could be due to a ‘fast forward’ strategy (Koops van’t Jagt et al., 2014) in search for disambiguating input.

These results were not replicated in the eye-tracking study, probably due to methodological differences between studies. Relatedly, methodological differences are key to consider as we now turn to the rereading effects that were detected in the verb IA and the object/subject IA in the eye-tracking study. Since in the eye-tracking study textual lines remained available for re-inspection, participants regressed to view pre-disambiguating IAs, and did so for longer when the verb proved to be intransitive as opposed to transitive. The finding that the verb IA and the object/subject IA were revisited to a lesser extent when the transitive analysis was imposed points towards reliance on compensatory strategies. Specifically, participants could make use of the gender cue to clarify which one of the gender-differentiated antecedents was coreferent with the pronoun. This also meant that there was no

need to regress to pre-disambiguating IAs as they could rely on the gender information to respond accurately to the comprehension question.

Overall, these findings suggest that presenting participants with repeated line break and clause boundary clash promoted neither early nor late closure during online processing. Evidence of processing disruption was reliably observed across studies, especially when compared with textual versions that did not contain incomplete syntax on successive lines.

4.3. *An alternative proposal*

Another way to address our second hypothesis would be to compare the two conditions where the intransitive analysis was correct (Complete-Intransitive and Incomplete-Intransitive). This allows us to assess what effects our Line Completeness manipulation had in the absence of any aiding cues (that is, no gendered pronoun).

The differences between conditions in which the intransitive analysis was grammatical were much greater than the differences between conditions where the transitive analysis was imposed, as evidenced by the larger effect sizes across studies in online and offline measures. Reading times in the disambiguating region revealed greater processing costs when participants had been presented with structurally incomplete lines as opposed to complete ones, even though in both of the conditions in question the verb turned out to be intransitive. Offline comprehension results mirror this pattern. When there was no evidence in the preceding context to discourage early closure, since all lines routinely coincided with structural boundaries, comprehension was better. By contrast, comprehension was significantly impacted in the condition in which it turned out that early closure should have been assumed, but in the preceding context the discontinuity of the line regularly conflicted with the continuity of the syntax; in fact, accuracy was numerically the lowest in this condition across studies, although the differences with other conditions proved significant only in Study 1.

Since these two conditions were virtually identical lexically and differed only with respect to the structural completeness textual lines, we propose that the results discussed above reflect *priming-induced garden path effects*. Specifically, repeated exposure to fragmented syntax may have generated a bottom-up expectation for structural incompleteness, resulting in parsers holding off on closure at the ending of the line with the critical verb (that is, line 3) and contemplating that the candidate found at the beginning of the subsequent line (that is, line 4) could be the object argument of the verb. If the transitive analysis received any activation, even if briefly, then the lingering memory trace of that interpretation (e.g. Slattery et al., 2013) could account for the comprehension failure that was observed in the self-paced reading study. If we assume that participants had momentarily entertained the transitive misanalysis, then we could explain the processing costs that were reliably observed across studies. These effects were not observed when all lines coincided with clause boundaries since there was no (con)textual feature therein to prompt readers to entertain transitivity. Hence, this assessment of results provides supportive evidence for our second hypothesis, namely that exposing participants to recurrent line break and clause boundary mismatch can discourage early closure, potentially due to a priming-induced expectation for structural incompleteness.

In summary, the combined findings of the present studies are consistent with the idea that line breaks can act as ‘good’ (that is, helpful) and ‘bad’ (that is, misleading)

cues to sentence structure, similar to the function of disfluencies in speech (Bailey & Ferreira, 2003). Specifically, line breaks in text that align with clause boundaries can help demarcate linguistic structure and lead to processing facilitation. However, when line breaks clash with structural boundaries, disruption rather than facilitation is more likely to be observed, possibly reflecting reading disfluency effects. Additionally, the present findings suggest that readers do not interpret line breaks in isolation, but instead consider the wider textual environment in which they are found. Repeated exposure to line-syntax misalignment appeared to weaken the reliability of line breaks as cues to closure, resulting in distinct processing strategies, possibly underpinned by re-calibrated expectations. Together, these findings underscore the non-trivial role of text segmentation in guiding parsing, and demonstrate that readers flexibly integrate various sources of information in a context-sensitive manner.

5. Limitations

One limitation of our studies concerns the design of stimuli and comprehension questions. Instead of ‘yes-no’ questions (e.g. ‘Did Alice heckle James?’), we opted for ‘Who did what’ questions to avoid favoring particular interpretations (e.g. agreement bias; see Ceháková & Chromý, 2023; Van Gompel et al., 2006). To answer correctly, participants needed to use information from the final line (the absence/presence of a gendered pronoun). However, this introduced an imbalanced cue between conditions, which may have affected findings. To overcome this issue, we presented an alternative approach to addressing our second hypothesis in Section 4.3, comparing conditions without imbalanced cues. For our first hypothesis, despite comparing conditions with imbalanced cues, we observed processing facilitation in the absence of aiding cues, which we view as compelling evidence in support of our hypothesis.

Additionally, we observed certain inconsistencies in results across the two studies. Some effects that were significant in Study 1 became marginal in Study 2. We think this was partly caused by methodological differences, such as the fact that we included fewer filler items in the eye-tracking study as well as changed the presentation mode and the regions examined (for justification, see Materials and Procedure). Another cause is probably the smaller sample size of Study 2 compared to Study 1. Although we had 100% power to detect the effect of Line Completeness, we had 52% power to detect the interaction of interest in the disambiguating IA, as suggested by a post-hoc power analysis using mixedpower in R (Kumle et al., 2021). Given the resources of time and funding available to us, we tried to recruit as many participants as possible for Study 2, ensuring the maximisation of our sample size within these constraints. Despite our best efforts, we had smaller power than desired because of these resource constraints.

Overall, while we acknowledge these limitations, we still consider that this work provides important insight into unexplored effects of text segmentation and hopefully lays the foundation for future research to shed light into outstanding questions.

6. Implications

We now turn to the broader implications of this research for everyday reading, focusing on the visuospatial arrangement of text in real-world settings, such as in books, educational materials and digital content. If we take a popular book as an example, such as *Alice’s Adventures in Wonderland* by Lewis Carroll, we find

- | | |
|-----|--|
| (4) | Alice was beginning to get very tired of sitting by her sister on the bank... |
| (5) | At the show, Izzy was nervous. From the stage, she spotted Ben, who waved. Taking a deep breath, Izzy told the audience her first joke.... |

scissored syntactic units from the beginning of the book, with the first two lines shown in (4). We find similar occurrences in (5), which is an excerpt from the text ‘Izzy’s Talent’. This was used in the 2024 Key Stage 1 Standard Assessment Tests for Reading, which are the UK’s national curriculum tests for second graders who are 6 or 7 years old.

These examples illustrate the arbitrariness of line break placement, where syntactic units are scissored without consideration for linguistic structure. This seems to be a common occurrence, as noted by Levasseur et al. (2006). In their study, they used texts meant for second graders in the USA and found that only 31% of textual lines ended in intact syntactic units.

This raises the question: Does this prevalent scissoring of syntactic units at line endings disrupt reading? Our findings suggest it does, consistent with prior research. Studies with children, including struggling readers and second language learners, suggest that when line breaks are made to coincide with syntactic boundaries, this leads to several benefits, such as better reading fluency, comprehension and retention of information (e.g. Levasseur et al., 2006; Park et al., 2019; Warschauer et al., 2011). Conversely, cases of mismatch can impact fluency and higher-level linguistic processing (syntactic parsing and comprehension). Similar observations have been made in subtitling research. It has been suggested that syntactically complete lines can improve subtitle readability and reduce cognitive effort when processing complex visual scenes (Perego, 2008). Additionally, adults have been shown to prefer syntactically complete lines in subtitles over incomplete ones (Gerber-Morón & Szarkowska, 2018).

All this evidence has practical implications for the way in which widely consumed texts are formatted, ranging from educational materials to digital content and subtitles. By segmenting text in a way that preserves syntactic structure, it may be possible to improve learning and reading experiences in various contexts and populations.

7. Conclusions

The present studies highlight that text segmentation means (line breaks) are far from irrelevant for the parser but rather consequential for online syntactic–prosodic analysis decisions. Line endings that routinely coincide with legitimate structural boundaries seem to lead to processing facilitation and help readers correctly parse clause boundary ambiguities. By contrast, when syntactic constituents are repeatedly ‘scissored’ by line breaks, processing disruption is more likely to occur. Moreover, recurring mismatch between line breaks and clausal boundaries can give rise to an anticipation of structural incompleteness. In turn, if parsers operate based on this anticipation and they end up being proven false, their comprehension may suffer. Overall, all the evidence presented calls attention to the importance of the visuospatial arrangement of text. The way in which text is formatted can affect how linguistic material is analysed and potentially how it is interpreted too. Readers take into account such textual properties to guide linguistic analysis on a context-by-context basis.

Data availability statement. Data, scripts and supplementary materials can be found at: <https://osf.io/v9e6g/>.

Funding statement. This study was funded by the Economic and Social Research Council (Project Reference: 2275541).

Competing interests. The authors declare none.

References

- Adams, B. C., Clifton, C., & Mitchell, D. C. (1998). Lexical guidance in sentence processing? *Psychonomic Bulletin & Review*, 5(2), 265–270. <https://doi.org/10.3758/BF03212949>.
- Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing: Studies in theoretical psycholinguistics* (Vol. 21, pp. 1–46). Springer. https://doi.org/10.1007/978-94-015-9070-9_1.
- Bailey, K. G., & Ferreira, F. (2003). Disfluencies affect the parsing of garden-path sentences. *Journal of Memory and Language*, 49(2), 183–200. [https://doi.org/10.1016/S0749-596X\(03\)00027-5](https://doi.org/10.1016/S0749-596X(03)00027-5).
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Breen, M. (2014). Empirical investigations of the role of implicit prosody in sentence processing. *Language and Linguistics Compass*, 8(2), 37–50. <https://doi.org/10.1111/lnc3.12061>.
- Ceháková, M., & Chromý, J. (2023). Garden-path sentences and the diversity of their (Mis)representations. *PLoS One*, 18(7), e0288817. <https://doi.org/10.1371/journal.pone.0288817>.
- Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, 42(4), 368–407. <https://doi.org/10.1006/cogp.2001.0752>.
- Chromý, J., & Tomaschek, F. (2024). Learning or boredom? Task adaptation effects in sentence processing experiments. *Open Mind*, 8, 1447–1468. https://doi.org/10.1162/opmi_a_00173.
- Cole, J. (2015). Prosody in context: A review. *Language, Cognition and Neuroscience*, 30(1–2), 1–31. <https://doi.org/10.1080/23273798.2014.963130>.
- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating Behavioral experiments in a web browser. *Behavior Research Methods*, 47(1), 1–12. <https://doi.org/10.3758/s13428-014-0458-y>.
- Drury, J. E., Baum, S. R., Valeriote, H., & Steinhauer, K. (2016). Punctuation and implicit prosody in silent Reading: An ERP study investigating English garden-path sentences. *Frontiers in Psychology*, 7, 1375. <https://doi.org/10.3389/fpsyg.2016.01375>.
- Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11(1), 11–15. <https://doi.org/10.1111/1467-8721.00158>.
- Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research*, 30(1), 3–20. <https://doi.org/10.1023/A:1005290706460>.
- Ferreira, F., & Clifton, C. (1986). The Independence of syntactic processing. *Journal of Memory and Language*, 25(3), 348–368. [https://doi.org/10.1016/0749-596X\(86\)90006-9](https://doi.org/10.1016/0749-596X(86)90006-9).
- Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS One*, 8(10), e77661. <https://doi.org/10.1371/journal.pone.0077661>.
- Fodor, J. D. (1989). Empty categories in sentence processing. *Language and Cognitive Processes*, 4(3–4), SI155–SI209. <https://doi.org/10.1080/01690968908406367>.
- Fodor, J. D. (2002). Prosodic disambiguation in silent Reading. In *Proceedings of the North East Linguistics Society*, 32(1), 113–132. <https://scholarworks.umass.edu/nels/vol32/iss1/8>
- Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. Doctoral Dissertation, University of Connecticut. <https://opencommons.uconn.edu/dissertations/AAI7914150>.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14(2), 178–210. [https://doi.org/10.1016/0010-0285\(82\)90008-1](https://doi.org/10.1016/0010-0285(82)90008-1).
- García-Pérez, M. A. (2023). Use and misuse of corrections for multiple testing. *Methods in Psychology*, 8, 100120. <https://doi.org/10.1016/j.metip.2023.100120>.
- Gerber-Morón, O., & Szarkowska, A. (2018). Line breaks in subtitling: An eye tracking study on viewer preferences. *Journal of Eye Movement Research*, 11(3). <https://doi.org/10.16910/jemr.11.3.2>.

- Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research*, 32, 101–123. <https://doi.org/10.1023/A:1022492123056>.
- 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(3), 425–443. <https://doi.org/10.1016/j.jml.2005.12.001>.
- Hirotani, M., Terry, J. M., & Sadato, N. (2016). Processing load imposed by line breaks in English temporal Wh-questions. *Frontiers in Psychology*, 7, 1465. <https://doi.org/10.3389/fpsyg.2016.01465>.
- Jones, L. L., & Estes, Z. (2012). Lexical priming: Associative, semantic, and thematic influences on word recognition. In J. S. Adelman (Ed.), *Visual word recognition, volume 2: Meaning and context, individuals and development* (pp. 44–72). [Psychology Press. <https://doi.org/10.4324/9780203106976>.
- Jun, S.-A., & Bishop, J. (2015). Priming implicit prosody: Prosodic boundaries and individual differences. *Language and Speech*, 58(4), 459–473. <https://doi.org/10.1177/0023830914563368>.
- Kennedy, A., Murray, W. S., Jennings, F., & Reid, C. (1989). Parsing complements: Comments on the generality of the principle of Minimal Attachment. *Language and Cognitive Processes*, 4(3–4), S151–S176. <https://doi.org/10.1080/01690968908406363>.
- Koops van't Jagt, R., Hoeks, J. C. J., Dorleijn, G. J., & Hendriks, P. (2014). Look before you leap: How enjambment affects the processing of poetry. *Scientific Study of Literature*, 4(1), 3–24. <https://doi.org/10.1075/ssol.4.1.01jag>.
- Kumle, L., Vö, M. L.-H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods*, 53(6), 2528–2543. <https://doi.org/10.3758/s13428-021-01546-0>.
- Lee, E. K., & Watson, D. G. (2012). Sentence processing. In V. Ramachandran (Ed.), *Encyclopedia of human behavior* (2nd ed., pp. 387–395). Academic Press. <https://doi.org/10.1016/B978-0-12-375000-6.00321-9>.
- Levasseur, V. M., Macaruso, P., Palumbo, L. C., & Shankweiler, D. (2006). Syntactically cued text facilitates Oral Reading fluency in developing readers. *Applied PsychoLinguistics*, 27(3), 423–445. <https://doi.org/10.1017/S0142716406060346>.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676–703. <https://doi.org/10.1037/0033-295X.101.4.676>.
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear MixedModels. *Journal of Memory and Language*, 94, 305–315. <https://doi.org/10.1016/j.jml.2017.01.001>.
- Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 601–618). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9781315630427>.
- Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. L. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the selective reanalysis hypothesis. *Journal of Memory and Language*, 59(3), 266–293. <https://doi.org/10.1016/j.jml.2008.06.002>.
- Park, Y., Xu, Y., Collins, P., Farkas, G., & Warschauer, M. (2019). Scaffolding learning of language structures with visual-syntactic text formatting. *British Journal of Educational Technology*, 50(4), 1896–1912. <https://doi.org/10.1111/bjet.12689>.
- Perego, E. (2008). WhatWould we read best? Hypotheses and suggestions for the location of line breaks in film subtitles. *The Sign Language Translator and Interpreter*, 2(1), 35–63.
- Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134(3), 427–459. <https://doi.org/10.1037/0033-2909.134.3.427>.
- Pickering, M. J., McLean, J. F., & Branigan, H. P. (2013). Persistent structural priming and frequency effects during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 890–897. <https://doi.org/10.1037/a0029181>.
- Pickering, M. J., & Van Gompel, R. P. (2006). Syntactic parsing. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 455–503). Academic Press. <https://doi.org/10.1016/B978-012369374-7/50013-4>.
- Prasad, G., & Linzen, T. (2021). Rapid syntactic adaptation in self-paced Reading: Detectable, but only with many participants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(7), 1156. <https://doi.org/10.1037/xlm0001046>.
- Rayner, K., & Frazier, L. (1987). Parsing temporarily ambiguous complements. *The Quarterly Journal of Experimental Psychology*, 39(4), 657–673. <https://doi.org/10.1080/14640748708401808>.

- Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). EyeMovements and on-line language comprehension processes. *Language and Cognitive Processes*, 4(3–4), SI21–SI49. <https://doi.org/10.1080/01690968908406362>.
- Schauffler, N., Schubö, F., Bernhart, T., Eschenbach, G., Koch, J., Richter, S., ... Kuhn, J. (2022). Prosodic realisation of enjambment in recitations of german poetry. In *Proc. Speech Prosody 2022* (pp. 530–534). <https://doi.org/10.21437/speechprosody.2022-108>
- Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, 23(4), 569–588. https://doi.org/10.1207/s15516709cog2304_8.
- Slattery, T. J., & Parker, A. J. (2019). Return sweeps in Reading: Processing implications of Undersweep-fixations. *Psychonomic Bulletin & Review*, 26(6), 1948–1957. <https://doi.org/10.3758/s13423-019-01636-3>.
- Slattery, T. J., Sturt, P., Christianson, K., Yoshida, M., & Ferreira, F. (2013). Lingering misinterpretations of garden path sentences Arise from competing syntactic representations. *Journal of Memory and Language*, 69(2), 104–120. <https://doi.org/10.1016/j.jml.2013.04.001>.
- Staub, A. (2007). The parser doesn't ignore intransitivity, after all. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 550–569. <https://doi.org/10.1037/0278-7393.33.3.550>.
- Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, 86(1), 142–164. [https://doi.org/10.1016/S0093-934X\(02\)00542-4](https://doi.org/10.1016/S0093-934X(02)00542-4).
- Steinhauer, K., & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a Universal Marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, 30(3), 267–295. <https://doi.org/10.1023/A:1010443001646>.
- Stowe, L. A., Kaan, E., Sabourin, L., & Taylor, R. C. (2018). The sentence wrap-up dogma. *Cognition*, 176, 232–247. <https://doi.org/10.1016/j.cognition.2018.03.011>.
- Sturt, P., Scheepers, C., & Pickering, M. (2002). Syntactic ambiguity resolution after initial misanalysis: The role of Recency. *Journal of Memory and Language*, 46(2), 371–390. <https://doi.org/10.1006/jmla.2001.2807>.
- Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, 136(1), 64–81. <https://doi.org/10.1037/0096-3445.136.1.64>.
- Tooley, K. M., & Bock, K. (2014). On the parity of structural persistence in language production and comprehension. *Cognition*, 132(2), 101–136. <https://doi.org/10.1016/j.cognition.2014.04.002>.
- Tooley, K. M., & Traxler, M. J. (2010). Syntactic priming effects in comprehension: A critical review. *Language and Linguistics Compass*, 4(10), 925–937. <https://doi.org/10.1111/j.1749-818X.2010.00249.x>.
- Traxler, M. J. (2008). Lexically independent priming in online sentence comprehension. *Psychonomic Bulletin & Review*, 15(1), 149–155. <https://doi.org/10.3758/PBR.15.1.149>.
- Traxler, M. J. (2009). A hierarchical linear modeling analysis of working memory and implicit prosody in the resolution of adjunct attachment ambiguity. *Journal of Psycholinguistic Research*, 38(5), 491–509. <https://doi.org/10.1007/s10936-009-9102-x>.
- Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33(3), 285–318. <https://doi.org/10.1006/jmla.1994.1014>.
- Van Gompel, R. P., Pickering, M. J., Pearson, J., & Jacob, G. (2006). The activation of inappropriate analyses in garden-path sentences: Evidence from structural priming. *Journal of Memory and Language*, 55(3), 335–362. <https://doi.org/10.1016/j.jml.2006.06.004>.
- Warren, T., White, S. J., & Reichle, E. D. (2009). Investigating the causes of Wrap-up effects: Evidence from eye movements and E-Z reader. *Cognition*, 111(1), 132–137. <https://doi.org/10.1016/j.cognition.2008.12.011>
- Warschauer, M., Park, Y., & Walker, R. (2011). Transforming digital Reading with visual-syntactic text formatting. *The JALT CALL Journal*, 7(3), 255–270. <https://doi.org/10.29140/jaltcall.v7n3.121>.

Cite this article: Tsoukala, A., Vogelzang, M., & Tsimpli, I. M. (2025). The influence of text segmentation on garden path processing: evidence from self-paced reading and eye-tracking, *Language and Cognition*, 17, e58, 1–26. <https://doi.org/10.1017/langcog.2025.10009>