As clear as glass: How figurativeness and familiarity impact simile processing in readers with and without dyslexia

Egan, Ciara; Siyanova, Anna; Warren, Paul; Jones, Manon

Quarterly Journal of Experimental Psychology

Accepted/In press: 29/01/2022

Peer reviewed version

As clear as glass: How figurativeness and familiarity impact simile processing in readers with and without dyslexia

Ciara Egan[1, 2, 3], Anna Siyanova-Chanturia[4, 5], Paul Warren[4], and, Manon W. Jones[1]
[1] School of Human and Behavioural Sciences, Bangor University, Wales, UK, LL57 2AS.
[2] Department of Experimental Psychology, University of Oxford, OX2 6GG.
[3] School of Psychology, National University of Ireland Galway, Ireland.
[4] School of Linguistics and Applied Language Studies, Victoria University Wellington, New Zealand.
[5] College of Foreign Languages, Ocean University of China, China.

Corresponding author:
Dr Ciara Egan
Current Address:
School of Psychology
National University of Ireland Galway
Email: ciara.egan@nuigalway.ie

Abstract

For skilled readers, idiomatic language confers faster access to overall meaning compared with non-idiomatic language, with a processing advantage for figurative over literal interpretation. However, currently very little research exists to elucidate whether atypical readers – such as those with developmental dyslexia – show such a processing advantage for figurative interpretations of idioms, or whether their reading impairment implicates subtle differences in semantic access. We wanted to know whether an initial figurative interpretation of similes, for both typical and dyslexic readers, is dependent on familiarity. Here, we tracked typical and dyslexic readers' eye movements as they read sentences containing similes (e.g. *as cold as ice*), orthogonally manipulated for novelty (e.g. familiar: *as cold as ice*, novel: *as cold as snow*) and figurativeness (e.g. literal: *as cold as ice* [low temperature], figurative: *as cold as ice* [emotionally distant]), with figurativeness being defined by the sentence context. Both participant groups exhibited a processing advantage for familiar and figurative similes over novel and literal similes. However, compared to typical readers, participants with dyslexia had greater difficulty processing similes both when they were unfamiliar, and when the context biased the simile meaning toward a literal rather than a figurative interpretation. Our findings suggest a semantic processing anomaly in dyslexic readers, which we discuss in light of recent literature on sentence-level semantic processing.

Keywords: idioms, similes, developmental dyslexia, eye-tracking, semantics, reading.

1. Introduction

Developmental dyslexia (henceforth, dyslexia) refers to a specific reading impairment in individuals with IQ within the normal range (Lyon et al., 2003). While reading improvement is often seen in adulthood, so-called compensated dyslexia is characterised by deficits in phonological processing, written expression, and slower reading than typically developed readers (Hatcher et al., 2002; Jones et al., 2010; Ramus & Szenkovits, 2008). Recent theorizing suggests that semantic processing is unimpaired in readers with dyslexia, and that the relative strength of conceptual-level knowledge can be used to compensate for orthographic / phonological processing difficulties (Hulme & Snowling, 2014; Nation & Snowling, 1998; Snowling & Hulme, 2013). However, other recent evidence shows a subtle semantic deficit, in which readers with dyslexia are less sensitive to, and have more difficulty processing input that does not match their semantic expectations when compared with typical readers (Jednoróg et al., 2010; Rüsseler et al., 2007; Schulz et al., 2008). Here, we examine how readers with dyslexia process idiomatic language, a form of expression which typically requires flexibility in the comprehension system for accurate interpretation.

Formulaic language is an umbrella term that comprises various types of conventional expressions, such as collocations (e.g., *strong tea*), binomials (e.g., *bride and groom*), idioms (e.g., *kick the bucket*), proverbs (e.g., *better late than never*), lexical bundles (e.g., *in line with*), and so on. Idioms, such as *spill the beans*, are of a particular interest in that they convey a figurative, non-literal meaning. In this particular example (as with many other idioms), the literal and idiomatic meanings were originally aligned,[1] but over time, the literal meaning became obsolete, and only the idiomatic interpretation remained. Formulaic language is very common in day-to-day oral and written language (Siyanova-Chanturia et al., 2011; Siyanova-Chanturia & Pellicer-Sanchez, 2019). A growing number of empirical studies have examined the means by which these highly familiar linguistic forms are processed, in comparison with more declarative, propositional linguistic forms. Previous research has shown that typically developed readers exhibit a processing advantage for idioms and other types of formulaic language over novel control phrases (Gibbs, 1980; Gibbs & Gonzales, 1985; Kessler et al., 2020; Siyanova-Chanturia

[1] *Spilling the beans* originally referred to an ancient voting system, wherein voters placed a bean in a jar. If the jars were toppled before the votes were counted, then one might see, prematurely, which candidate would be the winner, therefore lending itself to the idiomatic interpretation: to reveal a secret.

et al., 2011; Swinney & Cutler, 1979; Tabossi et al., 2009; for a review, see Siyanova-Chanturia & van Lancker-Sidtis, 2019). Event-related potential studies moreover indicate that formulaic sequences are accessed directly from semantic memory (Molinaro & Carreiras, 2010; Rommers et al., 2013; Siyanova-Chanturia et al., 2017; Vespignani et al., 2010), largely bypassing the incremental composition of meaning, usually characteristic of novel, propositional language comprehension.

Similes, such as *as cold as ice*, are a particularly interesting form of ditropically ambiguous idioms, meaning they allow for either a literal (low temperature) or a figurative (emotionally distant) interpretation. Researchers have long been interested in how figurative phrases are processed relative to novel controls, as well as how the two meanings – figurative and literal – are accessed and processed in ambiguous figurative expressions (Ashby et al., 2018; Bobrow & Bell, 1973; Cacciari & Tabossi, 1988; Conklin & Schmitt, 2008; Siyanova-Chanturia et al., 2011; Swinney & Cutler, 1979). For example, Ashby et al (2018) presented sentences containing either metaphors, such as *knowledge is a river*, or matched similes such as *knowledge is like a river*, the key difference being that only the simile is plausible when interpreted literally. In both instances, disambiguating context occurred after the vehicle word (*river*). Eye-movements revealed longer reading and rereading times for metaphors, both on the target word and the word immediately to its right, and regressions from these words were also less likely than for similes (Ashby et al., 2018). These results were taken to indicate that similes are more easily processed than metaphors, due to their dual interpretations being active in parallel. Research on idiom processing in typical readers suggests that idioms are processed faster than control phrases (e.g., *tie the knot* vs. *tie a knot*), while the (identical in form) figurative and literal meanings are processed similarly (e.g., Siyanova-Chanturia et al., 2011).

Similes, therefore, provide an interesting test case for the linguistic processor's temporal access to literal and figurative meanings, as both interpretations are possible. Indeed, familiar, or more frequent, similes are more likely to yield an initial figurative, than literal, interpretation (Siyanova-Chanturia et al., 2011, 2017). Literal/figurative interpretation is found to be modulated by the global sentence context and participant characteristics. For example, in contexts that prime a specific interpretation (i.e., figurative or literal), the idiom is initially interpreted as figurative or literal accordingly. Importantly, studies looking at first (L1) and second (L2) language speakers have found key differences in simile interpretation as a function of language proficiency. When L2 speakers read ambiguous figurative sequences (e.g., *at the end of the day*), they show a processing *disadvantage* when figurative interpretation is intended relative to both literal interpretation and a control phrase (e.g., *at the end of the war),* which is in contrast with the advantage typically shown in L1 speakers (Cieślicka, 2006; Siyanova-Chanturia et al., 2011). Thus, second language speakers may initially activate the literal interpretation, which then requires revision in order to accurately comprehend the intended meaning of the figurative expression. This likely occurs because L2 speakers acquire the literal meanings of idiom constituents' first, mastering figurative language only with increased exposure and proficiency, and in more naturalistic language environments.

A natural extension to the current research into figurative idiomatic language processing, then, is how literal and figurative meanings may be accessed as a function of reading ability. This topic has been the subject of some investigation, with researchers proposing that pragmatic knowledge may be compromised in individuals with reading disorders, such as dyslexia (Cappelli et al., 2018; R. Cardillo et al., 2018; Griffiths, 2007; Kasirer & Mashal, 2017). This line of research has largely relied on offline questionnaire-based measures of idiom and metaphor comprehension, with mixed results. For example, while Griffiths (2007) found differences in metaphor comprehension between adults with and without dyslexia, Kasirer and Mashal (2017) found that the two groups were comparable in novel metaphor comprehension and generation.

However, it may be that any differences in idiom comprehension are too subtle to be captured by off-line measures. Indeed, this seems likely, as the evidence for semantic processing abnormalities comes largely from electrophysiological evidence which is a sensitive measure of implicit responses (Jednoróg et al., 2010; Rüsseler et al., 2007; Schulz et al., 2008). Thus, what we may observe is differences in temporal access to literal and figurative language in individuals with dyslexia versus typically developed readers. A distinct yet related question is to what extent a phrase's familiarity influences dyslexic readers' processing of idiomatic sequences. We know that typically developed readers find familiar phrases easier to process than novel ones (Kessler et al., 2020; Siyanova-Chanturia et al., 2011; Siyanova-Chanturia & van Lancker-Sidtis, 2019). This processing advantage has been attributed to the frequency, familiarity, and predictability of conventional language. For example, upon hearing or reading "*on the other …"*, participants automatically predict the most likely continuation – "*hand".* As such, it seems likely that readers with dyslexia will have comparable or, indeed, greater difficulty with novel items, given their relative insensitivity to detecting unexpected linguistic input (Rüsseler et al., 2007).

In the present investigation, we examine the effects of novelty (familiar vs. novel similes), and figurativeness (literal vs. figurative interpretations) on the sentence processing characteristics of adults with and without dyslexia using an eye-tracking paradigm. Similar to Ashby et al (2018), our similes were embedded in initially non-biasing sentences. In the present study, the correct simile interpretation (figurative or literal) is only available following the simile itself (e.g., *She was as cold as ice with her children*; a familiar simile with a figurative interpretation implied in the final clause). We examined readers' eye movement behaviour in the simile region (*as cold as ice*; initial interpretation of the simile), the target word region (*ice*; precise response to the final word in the simile) and the continuation region (*with her children*; implicating meaning reanalysis). Given previous findings, we expected that typically developed readers would take advantage of linguistic forms that can be chunked and accessed as a single unit from semantic memory (Rommers et al., 2013; Siyanova-Chanturia et al., 2017; Vespignani et al., 2010). Here, the disambiguating context for figurative / familiar meanings should therefore accord with typical readers' online interpretation of the sentence, whereas literal / novel meanings should require re-interpretation. We therefore predicted that for typical readers, familiar meanings would yield shorter initial interpretation time (i.e., first pass times / total times on the simile region and target word), and figurative/familiar meanings would necessitate less investment in reanalysis, i.e., implicating fewer regressive eye-movements back from the continuation region, as well as shorter regression paths, compared with literal / novel meanings.

In relation to dyslexic readers, we also expected slower processing for novel (than familiar) similes and target words. But crucially, we expected disproportionately larger effects for readers with dyslexia, i.e., first pass/total time reading and a greater number of fixations, indicative of processing difficulty (Eden et al., 1994; Rayner, 1998; De Luca et al., 1999; Hutzler & Wimmer, 2004; Jones et al., 2010), compared with typical readers, since context is less salient in the case of novel similes (Hulme & Snowling, 2014; Nation & Snowling, 1998; Snowling & Hulme, 2013). Moreover, if dyslexic readers approach idiomatic language incrementally, indicative of a less efficient comprehension system, we would expect dyslexic readers' eye movements to pattern similarly across figurative and literal sentences. That is, upon reaching the continuation region, dyslexic readers would not have committed to either a literal or figurative interpretation (compared with typical readers' selection of the figurative interpretation), and we therefore expected similar patterns of processing times (first pass and regression path durations) and regressive eye movements in both literal and figurative conditions.

2. Materials and Methods

2.1. Participants

Thirty-nine native English speakers were included in the analysis, comprising 20 typical readers and 19 with developmental dyslexia (a further participant with dyslexia was excluded due to excessive errors on the comprehension task). The latter group self-reported having a diagnosis of developmental dyslexia (n = 19, 12 females, age: M = 24.2, SD = 4.54 years). These participants were recruited via the Miles Dyslexia Centre Specific Learning/Socio-communicative Difficulties Panel at Bangor University. The typical readers reported no history of developmental dyslexia or learning difficulty (n = 20, 14 females, age: M = 23.9, SD = 5.9 years). All participants had normal or corrected-to-normal vision. Ethical approval was granted for this study, and all participants provided written informed consent before taking part.

2.2. Background Cognitive and Literacy Tests

We administered a short battery of cognitive and literacy tests in order to validate the two group differences on key measures. Literacy measures with a focus on latency included word and non-word naming from the Test of Word Reading Efficiency (TOWRE; Torgesen et al., 1999), and rapid automatized naming (RAN) of letters and numbers from the Comprehensive Test of Phonological Processing (CTOPP; Wagner et al., 1999). These are important measures for distinguishing the two groups, as even highly compensated adults with dyslexia perform more poorly on these tasks than typical readers (Berninger et al., 2006; Jones et al., 2010). As readers with dyslexia are thought to have lower print exposure, and to read less than typically developed readers (Stanovich, 2009), we administered the Author Recognition Test, and the accompanying self-report measures about reading habits (ART; Acheson et al., 2008). The ART is a measure of print exposure, and the self-report items measure how many hours per week participants spend reading and writing, and a measure of how they feel their reading habits compare to their peers (Acheson et al., 2008). We also included both verbal and nonverbal IQ measures (expressive vocabulary and matrix reasoning), from the Wechsler Abbreviated Scale of Intelligence (WASI;

Wechsler, 2011). As dyslexia is a reading specific disorder, the groups were expected to differ on the basis of the literacy tests, but not IQ measures (Lyon et al., 2003).

2.3. Experimental Design & Stimuli

We utilized a two-by-two experimental design manipulating novelty (familiar phrases versus novel similes), and figurativeness (literal similes versus figurative similes). Similes were fully rotated across conditions, resulting in a total of 80 phrases, with 20 phrases per condition. Only 76 were included in analyses, however, due to problems with target word position (see statistical analysis section for details). The similes, and the sentences in which they were contained were identical up until the target word (see Table 1 for a depiction of the design). The target word was the same for all familiar phrases (both literal and figurative) as these would be highly familiar to participants (e.g. "as cold as ice"). For novel similes this target was changed to an alternative, which still made sense, but was unfamiliar (e.g. "as cold as snow"). The figurative and literal conditions differ in sentence endings, (e.g. "as cold as ice after her swim", referring to literal temperature, vs. "as cold as ice with her children", using a cold temperature to figuratively refer to an emotion).

Table 1. Experimental design and examples of stimuli. Interest areas for eye-tracking analyses: Simile region (e.g. "as cold as ice") shown in italics below; target word (e.g. "ice") shown in bold below, and continuation region (e.g. "with her children").

Sentence	Figurativeness	Novelty
She was *as cold as* ***ice*** with her children.	Figurative	Familiar
She was *as cold as* ***ice*** after her swim.	Literal	Familiar
She was *as cold as* ***snow*** with her children.	Figurative	Novel
She was *as cold as* ***snow*** after her swim.	Literal	Novel

The 38 target words included in analyses did not differ on word length across the novel (5.36 +/- 1.67) and familiar conditions (5 +/- 1.56; $t(37) = 1.1$, $p = 0.3$). Lexical frequency was calculated as Zipf scores from SUBTLEX-UK (van Heuven et al., 2014), and novel targets had significantly higher frequency (4.34 +/- 0.57) than familiar targets (4.06 +/- 0.68; $t(37) = 2.3$, p =

0.03). Zipf scores were then regressed against length (as these two factors are typically confounded), in order to analyse the residuals of these scores. Both word length and residualised lexical frequency were then included as covariates for the eye-tracking analyses, to ensure that these would not confound the results. Stimuli were normed for plausibility, familiarity, and figurativeness in an online norming study before testing began, with each factor being ranked on a 7-point Likert scale. Each item was rated by six people, and the data were analysed using cumulative link mixed models with the R function *clmm* in the ordinal package (Christensen, 2019). [2] For familiarity (7 being highly familiar, and 1 being highly unfamiliar), novel similes were rated as significantly less familiar (M = 2.94, SE = 0.34) than familiar ones (M = 5.48, SE = 0.31) as expected ($\chi 2[1]=20.8$, $p < .0001$). Similarly for figurativeness (1 being completely literal, and 7 being completely figurative), literal phrases (M = 3.55, SE = 0.30) were rated as being less figurative than those in the figurative condition (M = 5.82, SE = 0.21; $\chi 2[1]=18.2$, $p < .0001$). No other significant effects emerged for those ratings, and both validate our categorisation of items. Plausibility (1 highly implausible, to 7 highly plausible) analyses revealed novel items (M = 3.94 SE = 0.35) were rated as less plausible than familiar items (M = 4.48 SE = 0.34; $\chi 2[1]=19.6$, $p <.0001$), while literal items (M = 4.87, SE = 0.29) were rated as more plausible than figurative ones (M = 3.55, SE = 0.45; $\chi 2[1]=11.2$, $p <.001$). No significant interaction emerged. Whilst we aimed to control for plausibility across conditions as much as possible, it is the case that literal sentences, and more familiar similes both strongly bias raters to judge these items as being more plausible than novel and figurative alternatives (E. R. Cardillo et al., 2010; Lapata et al., 1999; López et al., 2017). Indeed, Cardillo et al (2010) warn that trying to over-match on plausibility in designs such as this one may decrease the metaphorical nature of the figurative conditions.

2.4. Procedure

Participants' eye-movements were recorded from their right eye using an Eyelink 1000 desktop-mounted eye tracker. Stimuli were presented in the centre of a 60 cm wide monitor (with a 60 Hz refresh rate), in black 20-point Courier New on a white background, and participants were seated 60 cm from the screen. The experiment was preceded by a 9-point calibration, which was

[2] Linear mixed effects models of each set of ratings were also carried out using the *lme4* package in *R*, and provided highly similar results to those reported here.

repeated before each block. At the beginning of each trial, participants fixated a small black circle, for drift correction. Following this the experimenter began the trial, at which point the sentence appeared, with the initial word in the position where the drift correction had been. The experiment consisted of four blocks, with each block comprising 78 sentences (20 target sentences, 58 filler sentences), interspersed with a short break. Trial order was pseudorandomized such that each simile only appeared once per block, and that there was full block between first and second presentation of each simile. After a third of all filler trials, participants were presented with a short comprehension question (see Figure 1 for the trial structure). The purpose of this was to ensure that participants were focusing on the sentences and their meaning. After the second block participants were given a longer break, and at this point the cognitive and literacy tests were administered.

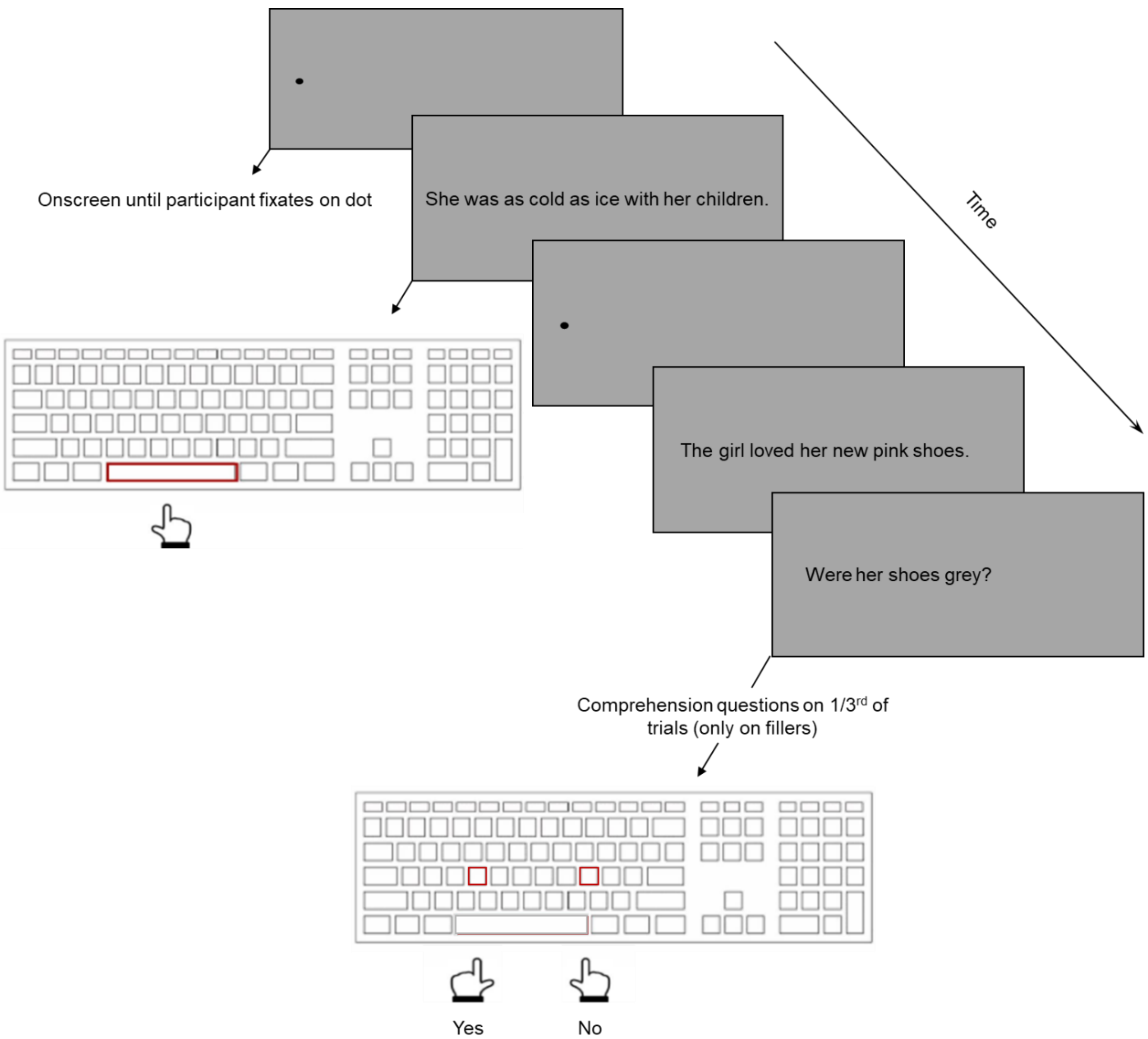

Figure 1. Schematic of the trial procedure with example simile and filler trials.

2.5. Statistical Analyses

Shapiro-Wilks test indicated that the assumption of normality was violated for most of the cognitive and literacy tests, with the exception of weekly writing time, comparative reading and matrix reasoning. Comparison of the two groups for these normally distributed measures were conducted via Welch's t-tests with separate estimates of variance, and the remaining measures were compared via Wilcoxon rank sum test for non-parametric data.

For eye-tracking analyses, three interest areas were defined; the full simile (e.g. "as cold as ice"), the target word (e.g. "ice"), and the continuation region after the simile (e.g. "with her children"). As noted in the experimental design section, two items were excluded from analysis due to their target word not occurring at the end of the item (e.g. "like a *bull* in a china shop"), which rendered them unsuitable for comparison with the other items in analyses. As such, we had 38 items per condition for analyses. Eye-tracking measures for sentence reading are often split into so-called early and late measures (Rayner, 1998, 2009). Our early measures (those relating to lexical and predictive processing facilitated by the earlier sentence context) were: first fixation duration, first pass fixation count, and first pass dwell time (gaze duration). Our later measures (which incorporate regressions to the area preceding the words) were: regression path (the total time/fixations between the participant fixating the interest area, and them exiting it to the right), and selective regression path (as the previous, but only counting first run fixations and refixations directly on the interest area). We also analysed early reading measures (first pass dwell time and fixation counts) in the disambiguating continuation region, as well as regressions back to the simile from this region. The durational measures were log-transformed prior to analysis (either explicitly or as part of the R routine used to model the data).

The analyses were conducted using mixed-effects modelling. The fixed effects for most analyses were Novelty (familiar vs. novel), Figurativeness (figurative vs. literal), and Group (typical readers vs. readers with dyslexia), and their interactions, as well as the covariates target word Length and Frequency (Zipf scores residualised against length). These covariates were included in order to ensure that target word length and frequency were not unduly impacting our results. Continuation region analyses did not include the covariates as these apply to the target word. The random effects structures included intercepts for participants and items, by-participant and by-item slopes for Novelty, Figurativeness (when appropriate), and a scaled and centred value for trial number, as well as by-participant slopes for target word length and (residualised) frequency (when appropriate), and by-item slopes for Group. Analyses of the simile and target word regions did not include Figurativeness as a fixed effect, as the conditions were identical at that point of the sentence.[3]

For most of the analyses we used a zero-inflation approach, using *glmmTMB* (Magnusson et al., 2017) in the *R* programming environment (Team & DC, 2019). This permits a two-part analysis, in which the first part is a logistic model predicting the likelihood of skipping an item, and the second part provides a model, conditional on skip rate, for reading time or fixation counts, using linear (gaussian) or poisson models, respectively. In cases where there were very few skips, straightforward linear or poisson models of non-zero reading times or of fixation counts were carried out, using *lmer* or *glmer* from the *lme4* package (Bates et al., 2018). Our model-building approach was to start from a maximal model and to run model comparisons to find the best model for the data, while retaining both our simple effects and interactions involving our main predictors (Group, Novelty, and Figurativeness). Model selection was automatized using the R package *buildmer* (Voeten, 2020) with settings that first ensured that the model converged and which then ran a backward elimination analysis of non-significant factors. *Buildmer* is able to evaluate all three model types listed above. All fixed effects were retained during the evaluation of the random effects structure, but were then assessed through model comparisons using *Anova* from the *car* package (Fox & Weisberg, 2019). The maximal and final models will be described for each analysis as it is introduced. For data, analysis scripts, and supplementary analysis please see https://osf.io/r37ta/?view_only=ee3155c1c9724e80aac7460750ed145e.

3. Results

3.1. Cognitive and Literacy Tests

The results of the cognitive and literacy tests largely validated the expected differences between groups (see Table 2). Readers with dyslexia had longer rapid naming for letters and numbers, words and nonword reading latencies, and more errors for nonwords than typical readers.

---

[3] See supplementary material for additional analyses in which we used the rating scores we obtained for Familiarity, Figurativeness and Plausibility, in the place of Novelty.

Surprisingly, print exposure and self-reported reading and writing times were comparable between groups. Although in the comparative reading measure participants with dyslexia reported poorer reading than their peers. Notably, both groups had comparable verbal and nonverbal IQ, indicating that any deficits the participants with dyslexia showed were in fact reading specific.

Table 2: Scores on cognitive and literacy tests, significant group differences in bold. Note: [a] Time in seconds; [b] Number of errors; [c] Word from 104 total; [d] Nonword from 63 total; [e] Number of authors (max 30); [f] Time in hours; [g] WASI subtest scaled score; [h] Self-report ability compared to peers (max 30).

	Median (MAD)				
	Typical *n* = 20	Dyslexic *n* = 19	*W*	*p*	*r*
RAN Letters [a]	**12.65 (2.45)**	**19.80 (8.89)**	**54.5**	**<.01**	**-.61**
RAN Numbers [a]	**11.50 (2.22)**	**15.80 (7.71)**	**78.5**	**<.01**	**-.49**
Word Reading (Time) [a]	**53.90 (11.49)**	**91.70 (41.96)**	**34.0**	**<.01**	**-.69**
Nonword Reading (Time) [a]	**46.50 (12.75)**	**95.00 (44.48)**	**37.5**	**<.01**	**-.68**
Word Reading (Errors) [b]	0 (0.00)	1 (1.48)	141.5	.15	-.23
Nonword Reading (Errors) [b]	**1 (1.48)**	**5 (4.45)**	**110.5**	**.02**	**-.36**
Word at 45 seconds [c]	**91 (14.83)**	**74 (17.79)**	**237.0**	**<.01**	**-.52**
Nonword at 45 seconds [d]	**58 (2.97)**	**40.5 (12.60)**	**180.5**	**<.01**	**-.58**
ART [e]	11.00 (5.93)	10.00 (4.45)	231.0	.25	-.18
Average weekly reading [f]	17.00 (5.19)	17.00 (10.38)	192.0	.97	-.01
Verbal IQ [g]	12.00 (1.48)	10.00 (2.97)	245	.12	-.25
	Mean (SD)		*t (df)*	*p*	*Cohen's d*
Average weekly writing [f]	10.90 (4.62)	12.26 (4.68)	-0.92 (36.8)	.37	-0.29
Matrix Reasoning	11.15 (2.56)	11.26 (2.02)	-0.15 (35.8)	.88	-0.05
Comparative reading [h]	**22.25 (3.04)**	**17.21 (4.64)**	**3.99 (30.8)**	**<.01**	**1.29**

3.2. Eye-tracking Measures

3.2.1. Simile Region Analyses

Results from the simile region (cf. Table 1) eye-tracking analyses are reported in Table 3 and Figure 2.

Table 3. Results from simile region (e.g. "as cold as ice") eye-tracking analyses, significant effects are in bold.

	$\chi 2$	*p*
First pass fixation duration		
Model type: lmer (gaussian)		
Length	**4.72**	**<.05**
Frequency	**21.77**	**<.001**
Group	**26.30**	**<.001**
Novelty	**14.99**	**<.001**
Group * Novelty	2.65	.10
Fixation counts		
Model type: glmer (poisson)		
Length	**5.57**	**<.05**
Frequency	**13.66**	**<.001**
Group	**26.14**	**<.001**
Novelty	**12.63**	**<.001**
Group * Novelty	1.38	.24

Relating to our predictions, the primary findings from this region were significant main effects of Novelty and Group. Similes with novel completions had longer total reading times and more fixations (total first pass durations: M = 651 ms, SE = 37.3 vs. M = 607 ms, SE = 31.2, with average fixation counts of 3.10, SE = 0.17 vs. 2.88, SE = 0.16). Readers with dyslexia had longer reading times than typical readers (M = 820 ms, SE = 61.1 vs. M = 483 ms, SE = 35.2), and more fixations (M = 3.63, SE = 0.26 vs. M = 2.46, SE = 0.14), as is shown on Figure 2.

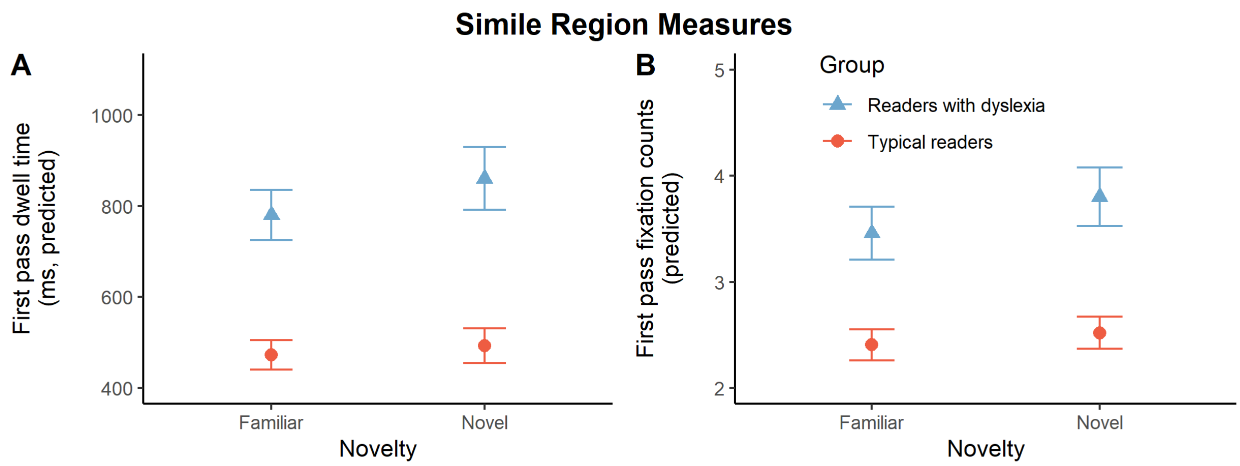

Figure 2. Eye-tracking measures for the simile region (e.g. "*as cold as ice"*), showing the effects of Novelty (familiar vs. novel) and Group (typical readers vs. readers with dyslexia). **A:** First pass dwell time. **B:** Count of first pass fixations.

3.2.2. Target Word Analyses

Results from the target word region (cf. Table 1) analyses are reported in Table 4. Again, the primary factors of interest here were Novelty and Group.[4]

Table 4. Results from target word region (e.g. "ice") eye-tracking analyses, significant effects are in bold.

	χ^2	*p*
First pass fixation duration		
Model type: zero-inflation		
i. zero-inflation component (skip rate; logit)		
Length	**26.43**	**<.001**
Frequency	**8.74**	**<.01**
Group	2.01	.16
Novelty	**5.17**	**<.05**
Group * Novelty	0.41	.52
ii. conditional component (duration; gaussian)		
Length	0.002	.97
Frequency	0.002	.96
Group	**5.05**	**<.05**
Novelty	**5.10**	**<.05**
Group * Novelty	0.014	.91

[4] We found that all readers spent longer reading, made more fixations on, and were more likely to regress to longer and less frequent target words. This is in line with previous research (Hyönä & Olson, 1995), and as this did not differ between groups, or interact with our variables of interest, it is reasonable to conclude that these factors are not influencing our findings. As such, we now turn to a discussion of the main results of interest, beginning with the effect of familiarity.

First pass dwell time		
Model type: zero-inflation		
i. zero-inflation component (skip rate; logit)		
Length	**23.13**	**<.001**
Frequency	**7.54**	**<.001**
Group	1.35	.25
Novelty	3.59	.06
Group * Novelty	0.38	.54
ii. conditional component (duration; gaussian)		
Length	**23.13**	**<.001**
Frequency	**7.54**	**<.01**
Group	**14.45**	**<.001**
Novelty	**4.98**	**<.05**
Group * Novelty	**4.29**	**<.05**
Fixation counts		
Model type: glmer (poisson)		
Length	**33.20**	**<.001**
Frequency	**5.71**	**<.05**
Group	**4.89**	**<.05**
Novelty	**4.78**	**<.05**
Group * Novelty	0.57	.45
Regression path reading time		
Model type: zero-inflation		
i. zero-inflation component (skip rate; logit)		
Length	**23.50**	**<.001**
Frequency	**7.31**	**<.01**
Group	2.17	.14
Novelty	2.51	.11
Group * Novelty	0.45	.50
ii. conditional component (duration; gaussian)		
Length	**5.33**	**<.05**
Frequency	1.80	.18
Group	**30.44**	**<.001**
Novelty	1.72	.19
Group * Novelty	0.98	.32
Regression path fixation counts		
Model type: glmer (poisson)		
Length	**26.91**	**<.001**
Frequency	**5.76**	**<.05**
Group	**20.47**	**<.001**
Novelty	**10.61**	**<.001**
Group * Novelty	**4.09**	**<.05**
Selective regression path reading time		
i. zero-inflation component (skip rate; logit)		
Length	**19.94**	**<.001**
Frequency	**7.21**	**<.01**
Group	1.35	.15
Novelty	3.81	.05
Group * Novelty	0.41	.52

ii. conditional component (duration; gaussian)		
Length	**27.67**	**<.001**
Frequency	2.06	.15
Group	**22.35**	**<.001**
Novelty	**18.52**	**<.001**
Group * Novelty	**11.27**	**<.001**
Selective regression path fixation counts		
Model type: glmer (poisson)		
Length	**48.06**	**<.001**
Frequency	**6.24**	**<.05**
Group	**8.42**	**<.01**
Novelty	**11.07**	**<.001**
Group * Novelty	1.55	.21

For first pass measures (shown in Figure 3) we once again predicted faster reading times / fewer fixations for familiar than unfamiliar similes, and for typically developed readers compared to readers with dyslexia. In line with this, first fixation durations on the target word, conditional on fixation taking place (see Table 4; First fixation duration, ii), were significantly affected by Novelty (longer fixations for novel words), and Group (longer fixations for readers with dyslexia). For total first pass dwell time on the target word (i.e. total fixation time starting from the first fixation until the eyes fixate either to the left or to the right of this word). The conditional component revealed significant effects on total first pass reading time for both Novelty and Group, and a significant interaction of Group and Novelty. Post hoc analysis in *emmeans* (Lenth et al., 2019) indicated that the Novelty effect was significant for the readers with dyslexia (t = -2.96, $p < .01$), but not for typical readers(t = -0.37, p = 0.71; see Figure 3). Analysis of first pass fixation counts also returned significant effects of Novelty and Group.

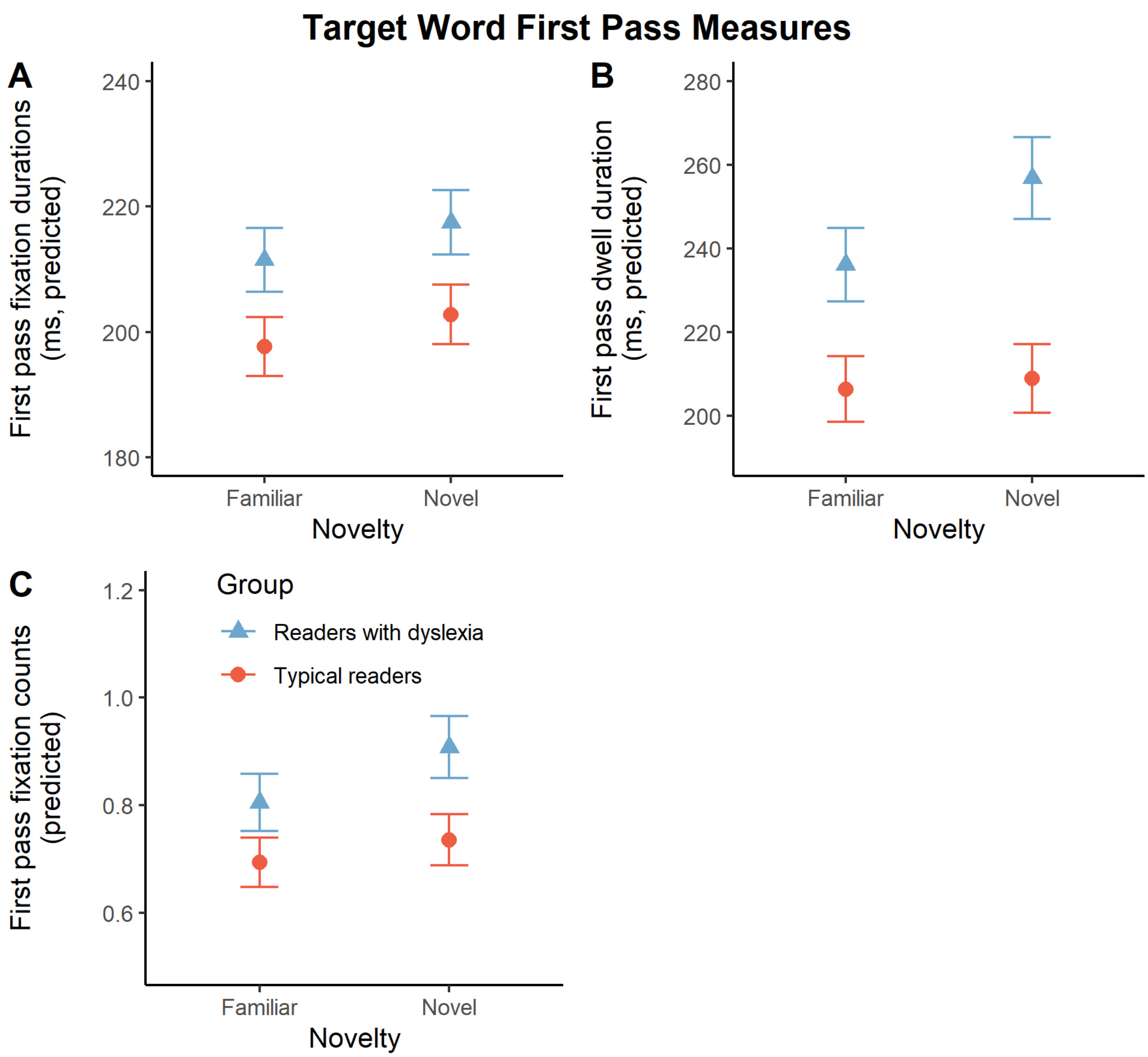

Figure 3. Early (first pass) eye-tracking measures for the target word (e.g. *ice*), showing the effects of Novelty (familiar vs. novel) and Group (typical readers vs. readers with dyslexia). **A:** First pass fixation durations. **B:** First pass dwell time. **C:** Count of first pass fixations.

For later reading measures (shown in Figure 4) we predicted that familiar similes would require less reanalysis than novel similes, as shown by fewer and shorter regressions. For regression path reading time (i.e. total time spent from the first fixation on the target word before moving on to the right of the target word, including regressions to earlier material), analysis of conditional reading time returned a main effect of Group. The analysis of regressive fixation counts also returned a significant effect of Group, as well as a significant interaction between Novelty and Group as shown in Figure 4. Post-hoc analyses using *emmeans* showed a significant effect of Novelty for the readers with dyslexia ($t = -3.83$, $p < .001$), but not for the typical readers ($t = -1.43$, $p = 0.15$). Moving on to the analysis of selective regression path reading time (i.e. the sum of all fixations on the final word before moving onto the following region), the model's conditional component showed a fixed effect of Group, as well as a significant interaction of Group and Novelty (see Figure 4). Post-hoc analysis using *emmeans* revealed that readers with dyslexia spent significantly longer reading novel than conventional simile endings ($t = -5.41$, $p < .001$), but the typical readers did not ($t = -0.64$, $p = 0.523$). Analyses also revealed significantly more selective regression path fixations on target words that were unexpected, as well as by readers with dyslexia.

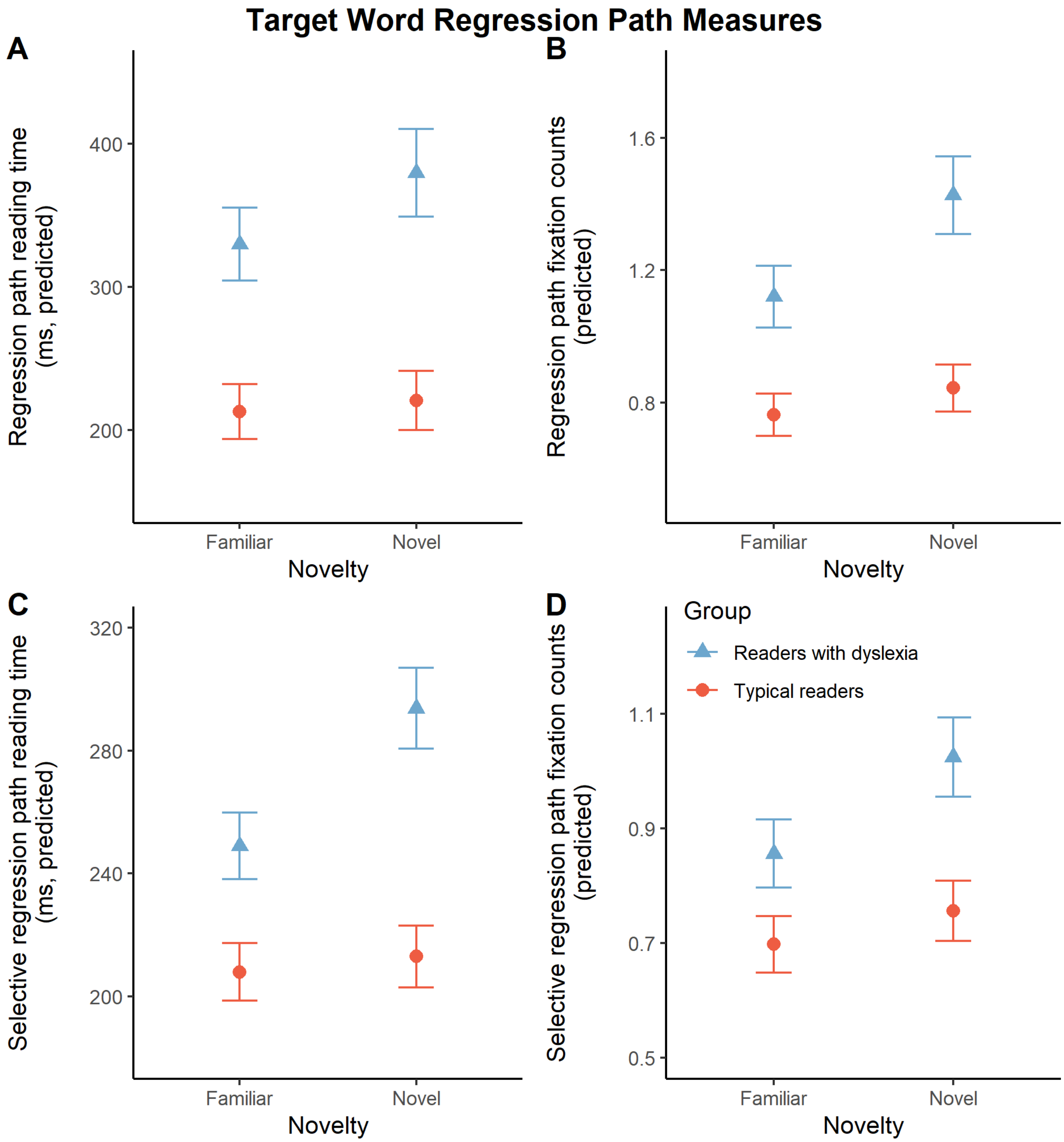

Figure 4. Later eye-tracking measures for the target word region (e.g. "*ice*"), showing the effects of Novelty (familiar vs. novel) and Group (typical readers vs. readers with dyslexia). **A:** Regression path reading time. **B:** Regression path fixation counts. **C:** Selective regression path reading time. **D.** Selective regression path fixation counts.

3.2.3. Continuation Region Analyses

Results from the continuation region (cf. Table 1) eye-tracking analyses are reported in Table 5. Here the factors of interest were Group, Novelty, and Figurativeness (as readers would become aware of the intended interpretation of the simile upon reading the continuation).

Table 5. Results from continuation region (e.g. "with her children") eye-tracking analyses, significant effects are in bold.

	χ^2	*p*
First pass dwell time		
Model type: lmer (gaussian)		
Length	0.15	.70
Frequency	1.42	.23
Group	**16.54**	**<.001**
Novelty	**6.97**	**<.01**
Figurativeness	0.49	.48
Group * Novelty	1.98	.16
Novelty * Figurativeness	0.01	.92
Group * Figurativeness	**4.27**	**<.05**
Group * Novelty * Figurativeness	1.34	.25
First pass fixation counts		
Model type: glmer (poisson)		
Length	0.003	.96
Frequency	1.01	.31
Group	**18.31**	**<.001**
Novelty	**6.31**	**<.05**
Figurativeness	1.88	.17
Group * Novelty	1.67	.20
Novelty * Figurativeness	0.32	.57
Group * Figurativeness	3.03	.08
Group * Novelty * Figurativeness	0.71	.40
Regression duration to simile region		
Model type: zero-inflation		
i. zero-inflation component (skip rate; logit)		
Length	0.15	.70
Frequency	1.15	.28
Group	3.75	.05
Novelty	1.41	.24
Figurativeness	1.02	.31
Group * Novelty	0.53	.46
Novelty * Figurativeness	0.23	.63
Group * Figurativeness	0.02	.91
Group * Novelty * Figurativeness	0.29	.59
ii. conditional component (duration; gaussian)		
Length	0.07	.80
Frequency	3.08	.08
Group	**6.42**	**<.05**
Novelty	1.17	.28
Figurativeness	1.42	.23
Group * Novelty	**33.94**	**<.001**
Novelty * Figurativeness	0.96	.33

Group * Figurativeness	0.19	.67
Group * Novelty * Figurativeness	0.49	.49
Regression counts to simile region		
Model type: zero-inflation		
i. zero-inflation component (skip rate; logit)		
Length	0.46	.50
Frequency	0.37	.54
Group	2.66	.10
Novelty	1.79	.18
Figurativeness	**22.82**	**<.001**
Group * Novelty	0.00	.98
Novelty * Figurativeness	**50.77**	**<.001**
Group * Figurativeness	0.01	.91
Group * Novelty * Figurativeness	1.94	.16
ii. conditional component (counts; poisson)		
Length	**3.97**	**<.05**
Frequency	0.23	.63
Group	**4.38**	**<.05**
Novelty	**7.71**	**<.01**
Figurativeness	**30.76**	**<.001**
Group * Novelty	**25.23**	**<.001**
Novelty * Figurativeness	**34.66**	**<.001**
Group * Figurativeness	0.88	.35
Group * Novelty * Figurativeness	0.75	.39

For first pass measures (shown in Figure 5) we once again predicted faster reading times / fewer fixations for familiar than novel similes, and for typically developed readers compared to those with dyslexia. We predicted also that typically developed readers would have faster reading times / fewer fixations for figurative similes, whereas readers with dyslexia would not differ between figurative and literal continuations on these measures. The analysis of first-pass dwell times returned significant effects for Group and Novelty, but no overall simple effect for Figurativeness. There was however a significant interaction of Group and Figurativeness, with shorter first pass reading times for figurative than literal continuations for the readers with dyslexia, but not for the typical readers, which was contrary to our predictions. We also ran this analysis on a continuation region which excluded the sentence-final word (e.g. “with her” as opposed to “with her children”) in order to control for the possible impact of sentence wrap-up effects - wherein the sentence-final word has longer reading times (Rayner et al., 2000; Warren et al., 2009). The above pattern of results was consistent, except that the interaction between Group and Figurativeness was no longer significant ($p = 0.51$).

For first pass fixation counts significant effects only emerged for Group and Novelty. Readers with dyslexia had more first pass fixations in this region, and there were fewer fixations when similes ended with novel words. A similar pattern of results emerged when figurativeness ratings from the norming study were used in place of the factor Figurativeness, and when the sentence-final word was not included in the analyses (see supplementary material).

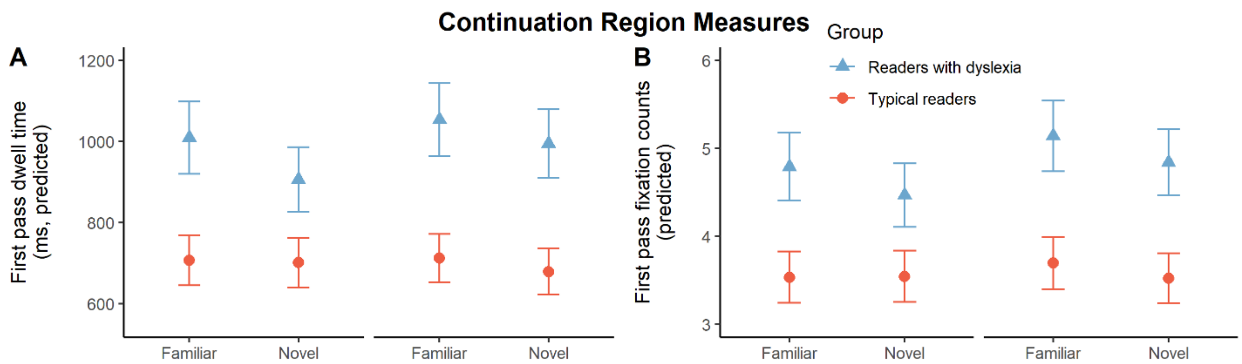

Figure 5. Eye-tracking measures for the continuation region (e.g. *"with her children"*), showing the effects of Novelty (familiar vs. novel), Figurativeness (figurative vs. literal), and Group (typical readers vs. readers with dyslexia). **A:** First pass dwell time. **B:** Count of first pass fixations.

Additional analyses considered the durations and counts of the first set of regressions from the continuation back into the region of the simile. For these later measures (shown in Figure 6) we predicted more (and longer) regressions back to the simile region for novel and literal similes, but again that this would be the case only for typically developed readers. The average duration of regressions to the simile region showed a significant effect of Group, and an interaction of Group and Novelty. As can be seen in Figure 6, while there was no significant difference between the participant groups in the case of the conventional similes (t = -0.31, p = 0.804), there were significantly shorter regression times to the novel similes for the typical readers than for the readers with dyslexia (t = -4.82, p < .001). Within the groups, the readers with dyslexia showed no effect of Novelty (t = -0.94, p = 0.35), while the typical readers had shorter regression reading times for novel than conventional similes (t = 3.97, p < .001).

The zero-inflation component of analysis of regressive fixation counts to the simile region returned a significant interaction of Novelty and Figurativeness as well as a simple effect of Figurativeness (Table 5, Regression counts to simile region, i). Here we see that regression to the simile is most likely (the zero inflation factor is lowest) when the novel simile is followed by a continuation indicating a literal interpretation. Regression to the simile is least likely (the zero inflation factor is highest) when the conventional simile is followed by the figurative interpretation. The regressive fixation counts (conditional on skipping) showed significant effects for Group, Novelty, and Figurativeness, as well as significant interactions of Novelty with both Group and Figurativeness. The three-way interaction between Novelty, Group, and Figurativeness did not reach significance however.

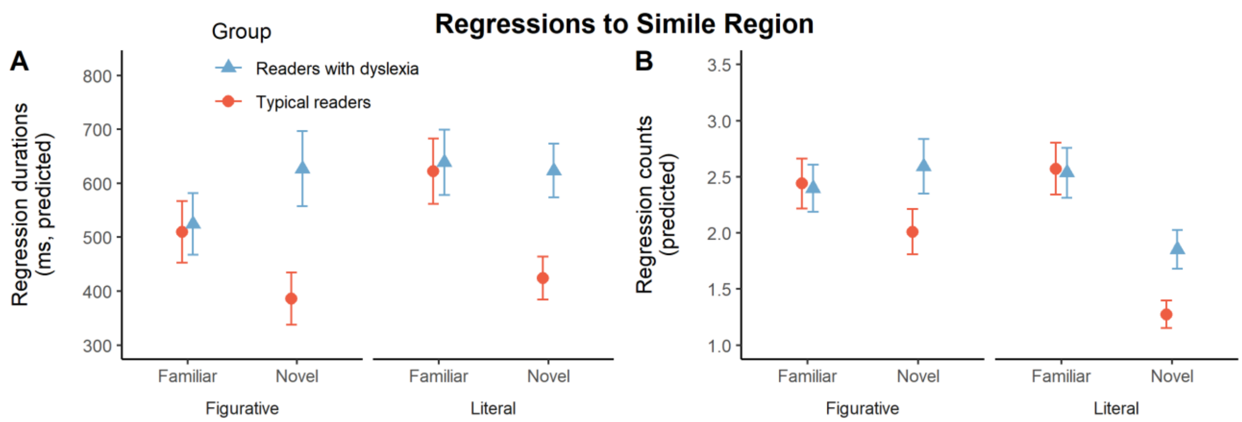

Figure 6. Eye-tracking measures for regressions from the continuation region (e.g. *"with her children"*) back to the simile region (e.g. "as cold as ice"), showing the effects of Novelty (familiar vs. novel), Figurativeness (figurative vs. literal), and Group (typical readers vs. readers with dyslexia). **A:** Duration of fixations back to the simile region. **B.** Count of fixations back to the simile region.

4. Discussion

Idioms and other types of formulaic language consistently show a processing advantage over novel sequences in normal readers (Ashby et al., 2018; Conklin & Schmitt, 2008; Kessler et al., 2020; Siyanova-Chanturia et al., 2011; Siyanova-Chanturia & van Lancker-Sidtis, 2019). However, very little is currently known about idiomatic language processing and interpretation in readers with different levels of ability, particularly readers with difficulties such as developmental dyslexia. Here, we conducted the first eye-tracking experiment comparing idiomatic language processing in adults with and without dyslexia. We predicted that whilst typical readers would show a processing advantage in accessing familiar figurative simile meanings over literal phrases, readers with dyslexia would show similar processing times for both literal and figurative stimuli due to semantic processing delay (Jednoróg et al., 2010; Schulz et al., 2008). This effect was expected to be augmented when similes were novel rather than familiar, since readers with dyslexia are known to experience greater processing difficulty with unexpected sentence continuations (Rüsseler et al., 2007). To summarise our main findings, readers with dyslexia were slower than typical readers to process similes when they were novel, irrespective of whether the phrase was idiomatic or literal.

4.1. Idiomatic language processing common to both typical and dyslexic readers

Typical readers and readers with dyslexia experienced more difficulty comprehending similes containing a novel (e.g., "as cold as *snow*") over a conventional familiar ending (e.g., "as cold as *ice*"), as evidenced by longer first pass reading times and more fixations within the simile region in the novel versus familiar condition (Figure 2). These novel target words were also skipped less often, and had more fixations, more selective regression path fixations, and longer first fixations than familiar words. Thus, all readers demonstrated slower comprehension when the final word in the simile was unfamiliar and unpredictable in the context, consistent with our predictions and with previous research on unexpected but plausible phrasal endings in sentence contexts (Abbott & Staub, 2015; Ashby et al., 2005; Rayner et al., 2004). In the area immediately following the simile, the continuation region, these effects were reversed, with shorter first pass dwell times and fewer first pass fixations when the sentence contained a novel simile (Figure 5). This may reflect a preview benefit, as the longer reading times on the final word of the simile would lead to greater parafoveal processing of the words immediately after.

The effect of figurativeness was only assessed in the continuation region (cf. Table 1), at which point the correct interpretation of the simile (literal or figurative) became clear. It is important to consider that results from this region may be influenced by sentence wrap-up effects, wherein readers spend longer reading sentence-final sections of text than earlier parts of a sentence (Warren et al., 2009). The zero-inflation analysis of regressive fixation counts showed that all readers were less likely to make regressions from the continuation region back to the simile when a figurative rather than a literal interpretation was required, suggesting that both groups found figurative continuations easier to process and integrate than literal ones. Further, a figurativeness-by-novelty interaction in the same analysis revealed that regressions back to the simile were least likely when the final word in the simile was conventional and therefore expected, and the continuation region stipulated a figurative interpretation. Interestingly, the condition that was found in the zero-inflation component to be most likely to have a regression from the continuation (i.e. novel similes with literal continuations), was also the condition with the lowest average number of regressive fixations (conditional on the likelihood of skipping). In other words, readers are likely to return to these novel-literal similes after encountering the later disambiguating context, but they make relatively few fixations when they do return to them. This is likely driven by the mismatch between participants' expectations (i.e. that similes are usually figurative) and the actual sentence ending, as well as the additional difficulty of the final word of the simile being unexpected. As such participants return to the simile region in order to reprocess the novel content once they have comprehended the full context (Reichle et al., 2009).

Taken together, these findings suggest that *all* readers expect similes to have a familiar figurative meaning, implicating less effort in integrating the non-literal interpretation with the rest of the sentence (Laurent et al., 2006; Siyanova-Chanturia et al., 2011; Strandburg et al., 1993; Vespignani et al., 2010), and are consistent with our predictions for typical readers and with reports of a processing advantage for idiomatic over declarative propositional language or matched metaphors (Ashby et al., 2018; Conklin & Schmitt, 2008; Kessler et al., 2020; Siyanova-Chanturia et al., 2011). While this finding has been widely reported in studies with typically developed readers (Siyanova-Chanturia & van Lancker-Sidtis, 2019), this is the first investigation to show that adult dyslexic readers also benefit from the conventional, predictable nature of figurative sequences, processing them more efficiently than novel/literal sequences. We note however that these findings relate specifically to the processing of similes, a form of figurative formulaic language, which has a highly predictable format (they are well-known and figurative), and may not generalise to other types of figurative language (cf. Frisson & Pickering, 1999; Pickering & Frisson, 2001).

4.2. Processing of idiomatic language specific to readers with developmental dyslexia

As expected, readers with dyslexia showed slower and more effortful reading on all measures (cf. Jones et al., 2010). Importantly, beyond this global effect, readers with dyslexia also showed greater difficulty in integrating novel similes than typical readers. Specifically, readers with dyslexia had longer reading times (Figure 3), as well as more regression path fixations and longer regression / selective regression path times (Figure 4) to the target word for novel than familiar similes, compared with typical readers. Note, however, that longer first pass reading times for novel similes in the simile / target word regions was traded off in the continuation region, in which shorter first pass durations were observed, compared with familiar similes. This likely indicates that the initially longer processing times for novel similes subsequently enabled faster eventual integration of the simile with the larger semantic context of the sentence.[5] Further, and contrary to our expectations, readers with dyslexia showed an advantage for figurative compared with literal interpretations, manifest in shorter first pass reading times for figurative than literal continuations; such an effect was not similarly manifest in typical readers' reading time. Taken together, our data suggest that dyslexic readers' processing of idiomatic language is more strongly affected – with effects emerging across all continuation region measures – by simile novelty and figurativeness, compared with typical readers.

Our findings suggest that dyslexic readers' access to sentence-level semantic information is delayed when phrasal-level information is *unexpected*; a pattern of results consistent with previous reports (Jednoróg et al., 2010; Rüsseler et al., 2007). Here, our overall pattern of data suggests that readers show an online processing difficulty for unfamiliar language (an erroneous simile), and are less efficient at updating their interpretation of the sentence upon encountering disambiguating information post-simile. We show these effects for the first time in the context of idiomatic language, in which dyslexic readers need to resolve semantic ambiguity.

We now consider our findings in light of recent research on figurative language comprehension in dyslexia (Griffiths, 2007; Kasirer & Mashal, 2017). Whilst these readers with dyslexia exhibited greater difficulty integrating novel, unexpected linguistic information than their typical-reading peers, we found that figurative language comprehension *per se* was not compromised. On the contrary, it appears dyslexic readers' processing was facilitated when familiar figurative similes were presented. Our findings thus appear to contradict earlier research suggesting that figurative language processing may be impaired in dyslexia (Griffiths, 2007; Kasirer & Mashal, 2017). However, methodological differences may help explain these differences. Previous studies on this topic have relied on offline pen-and-paper measures of pragmatic competence and figurative language comprehension, and thus examined explicit

[5] A caveat to this interpretation is that the novel vs. familiar conditions had different target words, which may have impacted on the processing times for these measures. However it seems unlikely that this would be unduly influencing the results due to the words being matched for length, and our target word and simile region analyses showing that target word length and lexical frequency did not interact with any of our factors of interest.

pragmatic knowledge rather than implicit language processing. In the present study, we employed eye-tracking, a highly sensitive online measure, allowing the analysis of multiple areas of interest and of early and late stages of processing, which enabled us to obtain an indication of online processing of figurative versus literal phrasal configurations.

Whilst developmental dyslexia is broadly defined by difficulty in forming adequate surface level word representations, such as orthographic form and phonology, our findings suggest a subtle semantic processing anomaly in processing unexpected or novel idiomatic phrases. We propose two interpretations for this effect, which cannot be adjudicated by the current data, but provide an avenue for further investigation. First, a large body of literature suggests that readers with dyslexia bootstrap poor orthotactic and phonological processing by relying on high-level semantic processing (Hulme & Snowling, 2014; Nation & Snowling, 1998; Snowling & Hulme, 2013; van Rijthoven et al., 2018). Under this interpretation, a target word that was unexpected in the sentence context would be more likely to cause temporary disruption to dyslexic readers' interpretation of the sentence, given their increased reliance on semantic information at the sentence level. This interpretation is likely, given that our data shows rapid recovery in processing times following the novel stimulus. Second, our findings may bely a subtle semantic *impairment*, in which dyslexic readers may be more sensitive to transitional probabilities in sentence processing for rapid semantic access. Under this account, unexpected or unfamiliar word continuations would result in a small processing delay, which is ameliorated by repeated exposures to these same sequences (Rüsseler et al., 2007).

4.3. Conclusion

In typically developed readers, conventional idiomatic language (such as similes) is more easily processed than declarative or novel text. Our study aimed to examine how dyslexic readers deal with idiomatic text using eye movements to analyse reading of sentences which contained similes that were manipulated for novelty and figurativeness. Our data show that both typical *and* dyslexic adult readers show a processing advantage for familiar and figurative similes. We show, for the first time, that readers with dyslexia benefit from the conventional and predictable nature of formulaic language, to an even greater extent than typical readers. Further research is needed to clarify how adult dyslexic readers flexibly interpret text with multi-layered meanings.

Acknowledgements: The authors would like to thank the Miles Dyslexia Centre at Bangor University for support with participant recruitment, and Dr Adam Parker for his helpful and insightful comments on the manuscript.

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 the Faculty Research Grant [grant number 215139], Victoria University of Wellington, awarded to Anna Siyanova-Chanturia.

References

Abbott, M. J., & Staub, A. (2015). The effect of plausibility on eye movements in reading: Testing E-Z Reader's null predictions. *Journal of Memory and Language*, *85*, 76–87. https://doi.org/https://doi.org/10.1016/j.jml.2015.07.002

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, *40*(1), 278–289. https://doi.org/10.3758/BRM.40.1.278

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye Movements of Highly Skilled and Average Readers: Differential Effects of Frequency and Predictability. *The Quarterly Journal of Experimental Psychology Section A*, *58*(6), 1065–1086. https://doi.org/10.1080/02724980443000476

Ashby, J., Roncero, C., de Almeida, R. G., & Agauas, S. J. (2018). The early processing of metaphors and similes: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *71*(1), 161–168.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., & Green, P. (2018). Package 'lme4.' *Version*, *1*, 17.

Berninger, V. W., Abbott, R. D., Thomson, J., Wagner, R. K., Swanson, H. L., Wijsman, E. M., & Raskind, W. (2006). Modeling Phonological Core Deficits Within a Working Memory Architecture in Children and Adults With Developmental Dyslexia. *Scientific Studies of Reading*, *10*(2), 165–198. https://doi.org/10.1207/s1532799xssr1002_3

Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition*, *1*(3), 343–346.

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, *27*(6), 668–683.

Cappelli, G., Noccetti, S., Arcara, G., & Bambini, V. (2018). Pragmatic competence and its relationship with the linguistic and cognitive profile of young adults with dyslexia. *Dyslexia*, *24*(3), 294–306. https://doi.org/https://doi.org/10.1002/dys.1588

Cardillo, E. R., Schmidt, G. L., Kranjec, A., & Chatterjee, A. (2010). Stimulus design is an obstacle course: 560 matched literal and metaphorical sentences for testing neural hypotheses about metaphor. *Behavior Research Methods*, *42*(3), 651–664.

Cardillo, R., Garcia, R. B., Mammarella, I. C., & Cornoldi, C. (2018). Pragmatics of language and theory of mind in children with dyslexia with associated language difficulties or nonverbal learning disabilities. *Applied Neuropsychology: Child*, *7*(3), 245–256. https://doi.org/10.1080/21622965.2017.1297946

Christensen, R. (2019). *A Tutorial on fitting Cumulative Link Mixed Models with clmm2 from the ordinal Package*. https://cran.r-project.org/web/packages/ordinal/vignettes/clmm2_tutorial.pdf

Cieślicka, A. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research*, *22*(2), 115–144. https://doi.org/10.1191/0267658306sr263oa

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*(1), 72–89.

Fox, J., & Weisberg, S. (2019). An R companion to applied regression (Third). *Thousand Oaks CA: Sage*.

Frisson, S., & Pickering, M. J. (1999). The processing of metonymy: evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1366.

Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, *8*(2), 149–156. https://doi.org/10.3758/BF03213418

Gibbs, R. W., & Gonzales, G. P. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition*, *20*(3), 243–259. https://doi.org/https://doi.org/10.1016/0010-0277(85)90010-1

Griffiths, C. C. B. (2007). Pragmatic abilities in adults with and without dyslexia: a pilot study. *Dyslexia*, *13*(4), 276–296. https://doi.org/https://doi.org/10.1002/dys.333

Hatcher, J., Snowling, M., & Griffiths, Y. M. (2002). Cognitive assessment of dyslexic students in higher education. *British Journal of Educational Psychology*, *72*(1), 119–133. https://doi.org/10.1348/000709902158801

Hulme, C., & Snowling, M. (2014). The interface between spoken and written language: developmental disorders. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1634), 20120395. https://doi.org/10.1098/rstb.2012.0395

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. In *Journal of Experimental Psychology: Learning, Memory, and Cognition* (Vol. 21, Issue 6, pp. 1430–1440). American Psychological Association. https://doi.org/10.1037/0278-7393.21.6.1430

Jednoróg, K., Marchewka, A., Tacikowski, P., & Grabowska, A. (2010). Implicit phonological and semantic processing in children with developmental dyslexia: Evidence from event-related potentials. *Neuropsychologia*, *48*(9), 2447–2457. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2010.04.017

Jones, M. W., Branigan, H. P., Hatzidaki, A., & Obregón, M. (2010). Is the 'naming' deficit in dyslexia a misnomer? *Cognition*, *116*(1), 56–70. https://doi.org/10.1016/J.COGNITION.2010.03.015

Kasirer, A., & Mashal, N. (2017). Comprehension and Generation of Metaphoric Language in Children, Adolescents, and Adults with Dyslexia. *Dyslexia*, *23*(2), 99–118. https://doi.org/10.1002/dys.1550

Kessler, R., Weber, A., & Friedrich, C. K. (2020). Activation of Literal Word Meanings in Idioms: Evidence from Eye-tracking and ERP Experiments. *Language and Speech*, 0023830920943625.

Lapata, M., McDonald, S., & Keller, F. (1999). Determinants of adjective-noun plausibility. *Ninth Conference of the European Chapter of the Association for Computational Linguistics*.

Laurent, J.-P., Denhières, G., Passerieux, C., Iakimova, G., & Hardy-Baylé, M.-C. (2006). On understanding idiomatic language: The salience hypothesis assessed by ERPs. *Brain Research*, *1068*(1), 151–160. https://doi.org/https://doi.org/10.1016/j.brainres.2005.10.076

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). *emmeans: Estimated Marginal Means, aka Least-Squares Means (Version 1.3. 4)*.

López, B. G., Vaid, J., Tosun, S., & Rao, C. (2017). Bilinguals' Plausibility Judgments for Phrases with a Literal vs. Non-literal Meaning: The Influence of Language Brokering Experience . In *Frontiers in Psychology* (Vol. 8, p. 1661). https://www.frontiersin.org/article/10.3389/fpsyg.2017.01661

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, *53*(1), 1–14. https://doi.org/10.1007/s11881-003-0001-9

Magnusson, A., Skaug, H., Nielsen, A., Berg, C., Kristensen, K., Maechler, M., & Brooks, M. (2017). Generalized linear mixed models using template model builder. *GlmmTMB. R Packag, Version 0.1*, *3*.

Molinaro, N., & Carreiras, M. (2010). Electrophysiological evidence of interaction between contextual expectation and semantic integration during the processing of collocations. *Biological Psychology*, *83*(3), 176–190. https://doi.org/https://doi.org/10.1016/j.biopsycho.2009.12.006

Nation, K., & Snowling, M. (1998). Individual Differences in Contextual Facilitation: Evidence from Dyslexia and Poor Reading Comprehension. *Child Development*, *69*(4), 996–1011. https://doi.org/10.1111/j.1467-8624.1998.tb06157.x

Pickering, M. J., & Frisson, S. (2001). Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(2), 556.

Ramus, F., & Szenkovits, G. (2008). What Phonological Deficit? *Quarterly Journal of Experimental Psychology*, *61*(1), 129–141. https://doi.org/10.1080/17470210701508822

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *53*(4), 1061–1080. https://doi.org/10.1080/713755934

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*(6), 1290–1301. https://doi.org/10.1037/0278-7393.30.6.1290

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

Rommers, J., Dijkstra, T., & Bastiaansen, M. (2013). Context-dependent semantic processing in the human brain: evidence from idiom comprehension. *Journal of Cognitive Neuroscience*, *25*(5), 762–776. https://doi.org/10.1162/jocn_a_00337

Rüsseler, J., Becker, P., Johannes, S., & Münte, T. F. (2007). Semantic, syntactic, and phonological processing of written words in adult developmental dyslexic readers: an event-related brain potential study. *BMC Neuroscience*, *8*(1), 52. https://doi.org/10.1186/1471-2202-8-52

Schulz, E., Maurer, U., van der Mark, S., Bucher, K., Brem, S., Martin, E., & Brandeis, D. (2008). Impaired semantic processing during sentence reading in children with dyslexia: Combined fMRI and ERP evidence. *NeuroImage*, *41*(1), 153–168. https://doi.org/10.1016/J.NEUROIMAGE.2008.02.012

Siyanova-Chanturia, A., Conklin, K., Caffarra, S., Kaan, E., & van Heuven, W. J. B. (2017). Representation and processing of multi-word expressions in the brain. *Brain and Language*, *175*, 111–122. https://doi.org/https://doi.org/10.1016/j.bandl.2017.10.004

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, *27*(2), 251–272. https://doi.org/10.1177/0267658310382068

Siyanova-Chanturia, A., & Pellicer-Sanchez, A. (2018). *Understanding formulaic language: A second language acquisition perspective*. Routledge.

Siyanova-Chanturia, A., & van Lancker-Sidtis, D. (2018). What online processing tells us about formulaic language. *Understanding Formulaic Language: A Second Language Acquisition Perspective. New York, NY: Routledge. Http://Dx. Doi. Org/10.4324/9781315206615-3*.

Snowling, M., & Hulme, C. (2013). Children's reading impairments: From theory to practice. *Japanese Psychological Research*, *55*(2), 186–202. https://doi.org/10.1111/j.1468-5884.2012.00541.x

Stanovich, K. E. (2009). Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Journal of Education*, *189*(1–2), 23–55. https://doi.org/10.1177/0022057409189001-204

Strandburg, R. J., Marsh, J. T., Brown, W. S., Asarnow, R. F., Guthrie, D., & Higa, J. (1993). Event-related potentials in high-functioning adult autistics: Linguistic and nonlinguistic visual information processing tasks. *Neuropsychologia*, *31*(5), 413–434. https://doi.org/https://doi.org/10.1016/0028-3932(93)90058-8

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning & Verbal Behavior*, *18*(5), 523–534. https://doi.org/10.1016/S0022-5371(79)90284-6

Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory &*

Cognition, *37*(4), 529–540.

Team, R. C., & DC, R. (2019). A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2012. *URL Https://Www. R-Project. Org*.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency.* PRO-ED.

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. In *The Quarterly Journal of Experimental Psychology* (Vol. 67, Issue 6, pp. 1176–1190). Taylor & Francis. https://doi.org/10.1080/17470218.2013.850521

van Rijthoven, R., Kleemans, T., Segers, E., & Verhoeven, L. (2018). Beyond the phonological deficit: Semantics contributes indirectly to decoding efficiency in children with dyslexia. *Dyslexia*, *24*(4), 309–321. https://doi.org/10.1002/dys.1597

Vespignani, F., Canal, P., Molinaro, N., Fonda, S., & Cacciari, C. (2010). Predictive mechanisms in idiom comprehension. *Journal of Cognitive Neuroscience*, *22*(8), 1682–1700. https://doi.org/10.1162/jocn.2009.21293

Voeten, C. C. (2020). *buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression. R package* (1.7.1).

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processing.* PRO-ED.

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

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence*. TX: Pearson.