# Effects of individual differences in text exposure on sentence comprehension

**Anastasia Stoops** (agusico2@illinois.edu)
University of Illinois at Urbana-Champaign
**Jessica Montag**
University of Illinois at Urbana-Champaign

**Article**

**Keywords:**

**Posted Date:** May 26th, 2023

**DOI:** https://doi.org/10.21203/rs.3.rs-2846554/v1

## Abstract

Linguistic experience plays a clear role in accounting for variability in sentence comprehension behavior across individuals and across sentence types. We aimed to understand how individual differences in reading experience predict reading behavior. Corpus analyses revealed the frequencies with which our experimental items appeared in written and spoken language. We hypothesized that reading experience should affect sentence comprehension most substantially for sentence types that individuals primarily encounter through written language. Readers with more text exposure were faster and more accurate readers overall, but they read sentence types biased to written language particularly faster than did readers with less text exposure. We see clear effects of text exposure on sentence comprehension in ways that allow explicit links between written and spoken corpus statistics and behavior. We discuss theoretical implications of effects of text exposure for experience-based approaches to sentence processing.

## Introduction

Linguistic experience—experience with words and sentence structures, has implications for the comprehension of those words and sentence structures. The findings that more frequent structures and structure-word combinations are easier to comprehend are central to many theoretical approaches to psycholinguistics including the classic constraint satisfaction that emerged in the 1990s [1; 2] and continues to be a major component of modern psycholinguistic work [3; 4; 5; 6; 7].

The specific aspect of language experience that we investigate here is experience with written language. There are substantial differences between the types of sentences contained in written and spoken language, with written texts containing a greater proportion of rare and complex sentence types, such as passive sentences and sentences containing relative clauses [8; 9]. The effect of written language exposure on sentence comprehension is thus both a critical piece of data in support of experience-based accounts of sentence processing, as well as an important source of individual differences in sentence processing.

Accumulating evidence suggests that reading experience may be an important source of individual differences in various aspects of language processing. Reading experience predicts individual differences in vocabulary size [10; 11], lexical decision times [12], verbal fluency [10], sentence production [13; 14] and various aspects of sentence comprehension [15; 16; 17; 18; 19; 20; 21; 22]. One proposed hypothesis for the observed effects of reading experience on sentence comprehension is greater experience with a subset of sentence types that are more frequent in written language, such as passives [18], relative clauses [15], or constructions containing connectives such as *however* or *since*[23; 22]. Our approach is to link the statistical properties of speech and text to observed patterns of sentence processing. Individuals with more text exposure should have greater exposure to the types of sentences biased to appear in written language and should show facilitation for those sentences.

We contrast our experience-based approach with syntactic complexity approaches, that suggest that the memory demands on comprehension posed by complex syntactic structures underlie differences in processing of different sentence structures as well as individual differences. Under perhaps a straw man version of this approach, reading experience should not affect sentence comprehension because individual differences in sentence processing should be driven by individual differences in memory capacity, which are experience-independent [24; 25]. A more nuanced version of the syntactic complexity approach [4; 26; 27] suggests that both individual differences in experience and memory may affect behavior. This syntactic complexity approach may treat experience and memory as two main effects, with distinct contributions to patterns of behavior. Under our experience-based approach, we predict that sentence structure and experience with sentence structures will interact. For sentence types that are frequent in spoken language, written language exposure should have minimal effects on sentence processing because exposure via spoken language is so frequent the addition of written language exposure provides diminishing returns for behavior. However, for sentence types that are rare in spoken language but more frequent in written language, we should see robust effects of reading experience. While this study was not designed to adjudicate between different approaches to sentence processing, we argue that we can gain significant insight into sentence processing by considering an experience-based approach, and considering the different profiles of experience that individuals might gain from written versus spoken language.

We developed a stimulus set consisting of four types of sentences that varied in comprehension difficulty and in their frequencies in written and spoken language: Simple active sentences, passive sentences, and sentences containing subject and object relative clauses. To hone our predictions for the effects of reading experience on sentence comprehension, we performed a corpus analysis to discover the frequencies of each sentence type in written and spoken language. To assess sentence comprehension, we recorded participant full-sentence reading times and comprehension question accuracy in a web-based sentence reading task. We then related both reading times and comprehension question accuracy to measures of text exposure.

# Corpus Analysis

We predict that text exposure should not predict language comprehension globally, but rather reading experience should lead to better comprehension on sentence types that are more frequent in written language. Individuals with more text exposure should show faster reading times and more accurate sentence comprehension for sentence types that more frequently appear in written language. The goal of this corpus analysis is to determine which sentence types disproportionately appear in written language to understand the aspects of the language environment we expect might change—or not change—with more reading experience.

## Method

Our sentence frequency counts come from a reanalysis of Roland, Dick & Elman (2007)[8], a corpus analysis of the frequencies of a wide range of sentence types in written and spoken corpora. We used the Roland et al. data to calculate frequencies with which our four sentence types, simple transitive sentences, passive sentences, and sentences that contain subject relative clauses (SRCs), and object relative clauses (ORCs), appear in written or spoken language. Our set of simple active sentences do contain some sentence types beyond simple transitive sentences, such as transitive sentences with additional prepositional phrases or conjunctions. Given the available corpus data, and that our sentences indeed all contained transitives, we report the data for simple transitive sentences.

We first calculated frequencies of simple transitive and passive sentences. Roland et al. report passive and simple transitive counts per 100 verb phrases but only overall passive counts. We used the overall corpus size to compute passive counts per million words and use the total passive counts as a reference to convert the simple transitive per 100 verb phrases count to a count of simple transitives per million words.

Computing frequencies for sentences containing relative clauses was slightly more complicated. Roland et al. report separate counts for reduced and full ORCs, so we combine these counts to be more consistent with the counts reported for SRCs, and because we have no reason to believe that the frequencies of both types should not be relevant. Then, due to well-established differences in comprehension patterns between relative clauses with full versus pronominal noun phrases[(28, 29; 30)], we refined our counts to only include SRCs and ORCs with full embedded noun phrases, not pronouns (e.g., ORC: *the teacher that the student met*; SRC: *the teacher that saw the student).* Roland et al. report numbers of full versus embedded phrase type in SRCs and ORCs in the Brown (written) and Switchboard (Spoken) corpora, but only those with *that* as a relative pronoun, but note that other relative clauses follow a similar pattern. We use these counts of full and embedded NPs to extrapolate counts in the entire corpus.

## Results

We observe differences in both the absolute frequencies of different sentence types as well as the ratios of frequencies in written and spoken language. Raw counts, counts per million words, and counts per million words including only SRCs and ORCs with full noun phrase embedded noun phrases are shown in Table 1. Table 1 also shows the ratio of the full noun phrase SRC and ORC and passives in written (Brown corpus) and spoken (Switchboard corpus) language. The counts per million words are also plotted in Figure 1.

| | Brown | | | Switchboard | | | |
|---|---|---|---|---|---|---|---|
| | Raw | Per Million words | Full NP only | Raw | Per Million words | Full NP only | Ratio |
| ctive (ransitive) | 30,641 | 30,641 | NA | 7,075 | 5,054 | NA | 6.1 |
| assive | 10,533 | 10,533 | NA | 566 | 404 | NA | 26.1 |
| RC | 4,622 | 6,433 | 3,897 | 760 | 543 | 394 | 9.9 |
| RC | 2,068 | 2,068 | 1,225 | 870 | 621 | 60 | 20.5 |

**Table 1.** Raw counts, counts per million words, and counts of only subject and object relative clauses with full noun phrase embedded nouns per million words, and ratios of counts in Brown (written) and Switchboard (spoken) corpora.

These ratios are imperfect, and the frequency counts may not perfectly reflect the stimuli in our study. For example, if we had been able to limit our corpus counts to only animate-headed SRC and ORCs (as we use in our experimental items) it is possible that the written to spoken language ratios for the SRCs and ORCs would increase. *Animate* headed relative clauses, especially ORCs, with embedded *full noun-phrases* are especially biased to written language[31; 32] so any error associated with ignoring head noun animacy should make our ratios more conservative. That said, these frequencies help us generate broad predictions for behavior based on the experience an individual encounters from spoken and written language.

As is evident from the ratios, all four constructions appear more frequently in written than spoken language. This effect likely reflects that spoken language consists of large proportion of intransitive utterances, as well as many short utterance and sentence fragments [9]. Despite all utterance types appearing more frequently in written texts, the written to spoken ratios vary: The active transitive sentences have the lowest ratio, appearing only about six times more often in text than speech, while SRC appear 10 times more often, ORCs nearly 21 times and passives over 26 times more frequently in speech than text. Notably, passives are the most text-biased construction, despite not containing any embedded clauses, consistent with many previous investigations of passive use [33; 34; 35].

In addition to ratios, the sentence types also vary in overall frequency. The simple transitive sentences are more frequent than the other sentence types, so despite appearing six time more often in written than spoken language, an individual should accumulate considerable experience with these sentences through speech alone. An important question for linking corpus frequencies with predictions for behavior is the role of both the overall and relative frequencies in written and spoken language. We may observe effects on comprehension based on ratios alone, so for all sentences individuals with more text exposure should show facilitation. However, we also expect that raw frequencies will matter as well. The college-aged participants in our study may have accumulated sufficient experience with simple transitive sentences that the extra experience from written language that an avid reader encounters has little effect on behavior. This scenario is possible if the relationship between experience and behavior is non-linear, reflecting the very old idea that learning rates are often steeper earlier in training [36; 37]. Going from a small amount to a medium amount of experience may have greater consequences for behavior than going from a medium to a large amount of experience. So, we expect smaller effects of reading experience for the more globally frequent simple active sentences than for the other sentence types which are overall rare except in written language.

# Main Study: web-based sentence comprehension

The study was pre-registered prior to data collection (https://osf.io/nwk7x).

### Methods

### Participants

All participants were recruited through the Department of Psychology participant pool at the University of Illinois, Urbana-Champaign. All participants gave their informed consent prior to the inclusion in the study. The work was approved by the Office for Protection of Research Subjects (OPRS) at the University of Illinois. All procedures were performed in accordance with the University of Illinois IRB and the Declaration of Helsinki. 221 native English speakers (mean age: 19; 144 female, 77 male) completed all tasks online.

### Materials

**Experimental sentences:** 120 sentences all 12 words each were split in 2 lists in a Latin-square design and presented in a whole sentence self-paced reading fashion. Sentences included 20 simple active sentences, 20 passive main clauses, and 80 sentences containing relative clauses taken from Traxler et al. [38] (40 subject relative clauses (SRC) and 40 object relative clauses (ORC)). Sentences were followed by comprehension questions (See Appendix for a complete list of sentences and questions). Items were pseudorandomized such that no two items of the same kind followed each other. The number of SRC and ORC was doubled relative to simple and passive sentences because SRC and ORC sentences were constructed in pairs (e.g., *The lawyer that the banker...* and *The banker that the lawyer...*) so participants saw only half of the experimental items. The order of the lexical items for the sentential arguments was counterbalanced. For example, if a participant viewed the SRC with the head noun *lawyer* then they would have viewed the ORC with the head noun *banker*. The question phrasing and the order of answer options was counterbalanced as well such that there was an equal number of “yes no” and “no yes” displays and an equal number of “yes” and “no” responses.

**1. Simple sentence**: I went to the store and bought milk, eggs, and green beans.

Did I go to the library?
Yes No

**2. Passive main clause**: Yesterday morning, the nurse was helped by the cowboy in ripped jeans

Did the cowboy help the nurse last week?
No Yes

**3. Subject relative clause**: The lawyer that irritated the banker retrieved the paperwork from the office.

Did the lawyer irritate the banker?
Yes No

**4. Object relative clause**:The lawyer that the banker irritated retrieved the paperwork form the office

Did the banker irritate the lawyer?
No Yes

**Text exposure surveys**

Assessing reading experience is not straightforward. Adults tend to exaggerate reading habits so indirect measures such as Author Recognition Tasks (ART) circumvent social desirability and yield better estimates [(39)].

**Author Recognition Test:** We used an updated version of the Acheson, et al. (2008)[15] by Moore and Gordon (2015)[40]. See Supplemental materials for the full survey. This task asks participants to choose real authors from a list of names (60 real, 60 foil authors). Participants received 1 point for a real author and 1 point was subtracted if participants chose a foil name.

**Reading Enjoyment Survey:** To build converging measures of reading experience, we adapted a survey that measured reading enjoyment in children [(41)] to survey reading enjoyment in adults. This survey consists of 10 statements that asked participants to either agree or disagree on a 1 through 7 Likert scale about various reading attitudes to assess participants attitudes and intrinsic motivation associated with reading (e.g., *I enjoy reading; I enjoy receiving books as gifts*). See Supplemental Materials for the full survey. A composite score was computed as the average of all 10 responses. For the questions that probed negative attitudes the scores were flipped to keep positive values at the higher end of the scale.

One motivation for this survey was to obtain a convergent measure of text exposure to complement the ART. Positive attitudes and intrinsic motivation are associated with reading frequency [(42)] so we hypothesized assessing attitudes towards reading may allow us to indirectly assess reading experience. A second motivation arises from challenges associated with collecting data online. Even software that locks participants' screens and prevents them from surfing the internet while performing a task cannot prevent participants from using their phones to look up whether the author is real or not. Assessing reading attitudes may minimize opportunities for participants to “cheat” even if social desirability may become a greater concern.

**Vocabulary Test** [43]: Participants were asked to choose a synonym for 40 real English words out of 4 possible variants for each word. Given a suspicious number of perfect or very high scores, it was evident that participants used their phones or other devices to look up correct synonyms for this test. We do not discuss the results further because we believe the results are not reliable.

**Demographics Survey**: In-house developed survey that collected basic information pertaining to participants age, gender, SES, reading ability, disability, or dyslexia diagnoses. See Supplemental materials for full version of the survey)

### Procedure

Participants were given a link after they chose to participate in our experiment through the SONA participant pool administration software. First, they gave consent to participate. Then they were directed to the website that displayed the sentence reading portion of the experiment followed by ART, Vocabulary, Reading Experience survey and basic demographics questionnaire. The experiment was implemented in Ibex farm online software [44]. Ibex farm uses JavaScript and html forms to collect participant responses and response times on the participant’s own computer and uploads participants responses to the server only after participants hit “Finish” button on last page of the experiment. Such approach minimized the response time delays for the reading time measures.

### Data Exclusion Criteria

Participants who learned English after 5 years of age (N= 40) or reported a history of reading difficulties (N=23) were excluded from the analyses.

A substantial challenge with online data collection is that it tends to be noisier than data collected in the lab. We developed a pipeline to remove trials and participants that did not likely reflect true reading processes (e.g., “button mashing,” careless clicking, or multitasking during study participation). For the response times, a two-step process was used: first, reading times faster than 1500 ms and slower then 138,000 ms (2.3 minutes – highest cut off time used traditionally for one-word-at-a-time self-paced reading studies of 11,500 ms was multiplied by 12 corresponding to 12 words in all our sentences) were excluded (3,513 data points removed out of the total 21,826 points). Second, the individual reading times were trimmed to cut off 2.5 standard deviations above and below the individual conditional mean (additional 853 data points reducing the dataset to 17,460 total data points). Based on these exclusions the total number of participants was reduced from 280 to 241. Additionally, participants were excluded if previous trimming left less than 50% of items for each of the 4 sentence types (additional 20 participants). Finally, based on our prior observations of individuals completing these tests in person in the lab setting it takes about 5 minutes maximum to complete the ART test. As a result, we excluded participants who took longer than 300,000 ms (5 min) to complete the test (N=10 participants). As a result of all the exclusion criteria the final dataset contained 211 participants – 64% of the participants who took part in our experiment (343 total participants). These exclusion rates are consistent with other online studies that find 45% to 53% of participants/trials are removed [45].

## Statistical variables, contrasts, and model fitting considerations

Text exposure surveys (ART, and RE) and sentence type (active, passive, SRC, ORC) were used to predict sentence reading times and comprehension question accuracy. Reading times were analyzed using linear mixed-effects (LME) models, and accuracy results were analyzed with generalized LMM (GLMM) models using the lme4 package (Version1.1-13; [46]) in R (Version 3.2.0; [47]). Three orthogonal contrasts were specified through dummy coding to compare relative clause versus main clause sentences, followed by active versus passive sentences and SRC versus ORC sentences. This coding scheme was preregistered. Three contrasts were defined: relative clauses: SRC "1" and ORC "1" versus main clauses: Active "-1" and Passive "-1" (MC vs RC); ORC versus SRC: SRC "-1" , ORC "1", Active "0" and Passive "0"; passive versus active sentences: SRC "0" , ORC "0", Active "-1" and Passive "1". Additionally, as an exploratory analyses after viewing the results, we used treatment contrast where each of the three complex structures were compared to active sentences that served as a baseline. Measure of text exposure (ART and RE scores) were centered and scaled.

LME models were fitted to untransformed and log-transformed reading times (See Supplemental materials, Tables 1 & 2 for model results). The results of the transformed and untransformed times were remarkably similar, so we report the untransformed models to facilitate interpretation. We note any significant differences in the pattern of results. The random structure was determined following Barr et al (2013)[48] maximal fit approach. LME models were fit by restricted maximal likelihood with the Satterthwaite's method; generalized LME models were fit by maximum likelihood with Laplace approximation. *P*-values were obtained through *summary* function of the *lmerTest* package [49]. The final models for reading times have random slopes for items and participants. The final accuracy models have random slopes for items only due to convergence failure. The exploratory model for accuracy with both ART and RE did not converge with random structure, as a result we fitted this model with regular regression (lm instead of glmer).

Response time and accuracy plots in Figure 2 were inspired by van Langen's open visualizations (2020)[50]

## Results

## Assessments of Text Exposure

The two assessments of text exposure, the Author Recognition Task (Mean: 13.13, SD = 6.21, Range = -1-30) and Reading Enjoyment Survey (Mean: 4.27, SD = 1.55, Range = 1.2-7) were only moderately correlated; readers with more positive attitudes recognized more real authors ($r=.33$, $p<.001$). In subsequent analyses, we probe whether the measures each capture variance in our sentence processing measures.

### Reading Analyses

Whole-sentence reading time analyses were limited to items on which the participant correctly answered the comprehension question. As expected, participants read the simple sentences faster and more accurately (Figure 2 & Table 2) than rarer or more syntactically complex sentences. However, relative rankings across the four sentence types for speed and accuracy were not the same. ORC sentences took the longest time to read, followed by SRC, passive, and simple sentences. However, accuracy was the lowest for passive sentences, followed by ORC, then SRC and simple sentences.

| Sentence Type | N Correct | Reading Times Mean (milliseconds) | SE | N Overall | Accuracy | SE |
|---|---|---|---|---|---|---|
| Active | 3981 | 4009 | 59.71 | 4074 | .98 | .00 |
| Passive | 3153 | 5592 | 98.59 | 4176 | .76 | .01 |
| SRC | 2233 | 6121 | 117.99 | 2520 | .89 | .01 |
| ORC | 2043 | 7099 | 126.89 | 2549 | .80 | .01 |

**Table 2.** Participants means, sentence counts and standard errors (SE) for reading times and accuracy rates by sentence type

### Effects of Text Exposure on Reading Time and Accuracy

### Effects of Author Recognition Test

To test our key hypothesis, we investigated the how reading experience affected both overall reading times and question accuracy, and how reading times interacted with sentence type. For visualization purposes, Figure 3 shows the relationship between ART (top row) and RE survey (bottom row) and reading times (first column) and comprehension question accuracy (second column).

Models predicting reading times revealed main effects and an interaction between sentence types and ART score (Table 3, Model 1). All participants read relative clause sentences slower than main clause sentences (main effect of *MC vs RC*), passives slower than active sentences (main effect of *Active vs Passive*), and ORCs slower than SRCs (*SRC vs ORC*). However, the interactions show that participants with higher ART scores read relative clauses (versus main clauses) and passive sentences (versus the active sentences) faster than participants with lower ART scores (ART interaction with *MC vs RC* and *Active vs Passive*). Participants with more text exposure showed smaller differences in reading times for the easier and harder sentences. Log-transformed data revealed a similar pattern of results, except that the main effect for the ART was not reliable (see Supplemental materials, Exhibit A Table A, Model 1).

When we used our exploratory treatment contrast in the same model (Table 3, Model 2) we get very similar results. All three sentence types were read slower than the active sentences. There was no main effect of the ART but it interacted with all three comparisons. Log-transformed models revealed identical results (Supplemental Material, Exhibit A, Table A, Model 2).

Across all models, we see a clear effect of text exposure on reading times. We see some evidence that participants who had higher ART score read faster overall, and converging evidence that participants who had higher ART scores were especially faster to read passive, SRC and ORC sentences, the sentence types more frequent in written language, than participants who had lower ART scores.

| Measure | Contrast | b | SE | *t/z* |
|---|---|---|---|---|
| **Model 1: Preregistered dummy coding scheme** | | | | |
| RT ~ MC_RC*ART+A_P*ART+SRC_ORC*ART+(1\|Item)+(1\|Participant) | | | | |
| RT in ms | Intercept | **5765.46** | **187.87** | **30.688*** ** |
| | MC vs RC | **969.14** | **101.29** | **9.57*** ** |
| | Active vs Passive | **821.65** | **149.54** | **5.49*** ** |
| | SRC vs ORC | **587.01** | **134.84** | **4.35*** ** |
| | ART | **-502.95** | **163.77** | **-3.07** ** |
| | MC vs RC: ART | **-136.07** | **43.52** | **-3.13** ** |
| | Active vs Passive: ART | **-164.09** | **52.94** | **-3.10** ** |
| | SRC vs ORC: ART | -36.79 | 68.74 | -0.54 |
| **Model 2: Condition treatment contrasts** | | | | |
| RT~ A_P*ART+A_SRC*ART+A_ORC*ART+ (1\|Item)+(1\|Participant) | | | | |
| RT in ms | Intercept | **3974.67** | **261.88** | **15.18*** ** |
| | Active vs Passive | **1643.29** | **299.09** | **5.49*** ** |
| | Active vs SRC | **2172.91** | **282.69** | **7.69*** ** |
| | Active vs ORC | **3346.93** | **285.07** | **11.74*** ** |
| | ART | -202.78 | 172.75 | -1.17 |
| | Active vs Passive: ART | **-328.18** | **105.88** | **-3.10** ** |
| | Active vs SRC: ART | **-473.02** | **118.21** | **-4.00*** ** |
| | Active vs ORC: ART | **-399.44** | **122.24** | **-3.27** ** |
| **Model 3: Preregistered dummy coding scheme** | | | | |
| RT ~ MC_RC*ART+A_P*ART+SRC_ORC*ART+(1\|Item)+(1\|Participant) | | | | |
| Accuracy | Intercept | **3.20** | **.17** | **18.85*** ** |
| | MC vs RC | **-1.17** | **.16** | **-7.24*** ** |
| | Active vs Passive | **-2.85** | **.39** | **-7.36*** ** |
| | SRC vs ORC | **-.44** | **.15** | **-3.01** ** |
| | ART | **.23** | **.07** | **3.11** ** |
| | MC vs RC: ART | .06 | .06 | 1.12 |

| | | | | |
|---|---|---|---|---|
| | Active vs Passive: ART | .11 | .11 | .95 |
| | SRC vs ORC: ART | .02 | .04 | .44 |
| **Model 4: Condition treatment contrasts** | | | | |
| RT~ A_P*ART+A_SRC*ART+A_ORC*ART+ (1\|Item)+(1\|Participant) | | | | |
| Accuracy | Intercept | **4.36** | **.29** | **14.97*** ** |
| | Active vs Passive | **-2.85** | **.39** | **-7.36*** ** |
| | Active vs SRC | **-1.89** | **.36** | **-5.32*** ** |
| | Active vs ORC | **-2.77** | **.35** | **-7.88*** ** |
| | ART | **.30** | **.12** | **2.55*** |
| | Active vs Passive: ART | .11 | .11 | .95 |
| | Active vs SRC: ART | .11 | .13 | .87 |
| | Active vs ORC: ART | .15 | .12 | 1.24 |

**Table 3.** LME Models predicting the untransformed reading times with condition and ART for pre-registered and exploratory (treatment) contrasts. Note: ^p<.1; *p<.05; **p<.01; ***p<.001

Generalized LME predicting comprehension question accuracy with ART scores and sentence type revealed only main effects of sentence type and text exposure for both pre-registered and exploratory contrasts (Table 3 Models 3 and 4). Participants were overall less accurate on relative clauses and passive sentences than active sentences and participants with more text exposure were overall more accurate on all sentence types.

### Effects of Reading Enjoyment Survey

Reading Enjoyment scores showed an identical pattern of effects on reading times as did ART scores for both our preregistered and exploratory model contrasts (Table 4, Model 1 & 2). Log-transformed data revealed identical pattern of results to the raw data with two exceptions: only passive versus active sentences contrast (not the main versus relative clause) yielded significant interaction with the Reading Enjoyment score and both models with pre-registered and exploratory contrasts revealed the main effect of Reading Enjoyment (Supplemental Materials, Exhibit B, Table B, Models 1 & 2). Effects of Reading Enjoyment scores on comprehension question accuracy were also nearly identical to those of the ART, when using the preregistered contrasts (Table 4 Model 3). However, the same model with the exploratory treatment contrasts converged only with random slopes for items, not participants. Given the potential problems with model fit, we additionally include results from a linear regression model (Table 4, Model 5). We observe no main effect of Reading Enjoyment but see significant interactions between Reading Enjoyment and sentence types such that participants who reported higher degrees of reading enjoyment tend to be more accurate in comprehending all three types of rare or complex sentences relative to the simple sentences than participants who enjoy reading less.

| Measure | Contrast | b | SE | *t/z* |
|---|---|---|---|---|
| **Model 1: Preregistered dummy coding scheme** | | | | |
| RT~MC_RC*RE+A_P*RE+SRC_ORC*RE+(1\|Item)+(1\|Participant) | | | | |
| RT in ms | Intercept | **5759.64** | **188.64** | **30.53*** ** |
| | MC vs. RC | **972.47** | **102.43** | **9.49*** ** |
| | Active vs Passive | **820.13** | **151.40** | **5.42*** ** |
| | SRC vs. ORC | **590.06** | **136.15** | **4.33*** ** |
| | RE | **-486.69** | **165.03** | **-2.95** ** |
| | MC vs. RC: RE | **-119.61** | **43.27** | **-2.76** ** |
| | Active vs Passive: RE | **-124.09** | **53.42** | **-2.32*** |
| | SRC vs. ORC: RE | -57.89 | 67.54 | -0.86 |
| **Model 2: Condition treatment contrasts** | | | | |
| RT~ A_P*RE+A_SRC*RE+A_ORC*RE + (1\|Item)+(1\|Participant) | | | | |
| RT in ms | Intercept | **3967.04** | **264.12** | **15.02*** ** |
| | Active vs Passive | **1640.26** | **302.81** | **5.42*** ** |
| | Active vs SRC | **2175.01** | **285.90** | **7.61*** ** |
| | Active vs ORC | **3355.13** | **288.30** | **11.64*** ** |
| | RE | -242.99 | 174.09 | -1.40 |
| | Active vs Passive: RE | **-248.17** | **106.84** | **-2.32*** |
| | Active vs SRC: RE | **-421.19** | **117.74** | **-3.58*** ** |
| | Active vs ORC: RE | **-305.41** | **120.80** | **-2.53*** |
| **Model 3: Preregistered dummy coding scheme** | | | | |
| Accuracy~MC_RC*RE+A_P*RE+SRC_ORC*RE+(1\|Item)+(1\|Participant) | | | | |
| Accuracy | Intercept | **3.19** | **.17** | **18.84*** ** |
| | MC vs. RC | **-1.15** | **.16** | **-7.13*** ** |
| | Active vs Passive | **-2.82** | **.39** | **-7.29*** ** |
| | SRC vs. ORC | **-.44** | **.15** | **-3.00** ** |
| | RE | **.24** | **.07** | **3.32*** ** |
| | MC vs. RC: RE | .03 | .05 | .56 |

| | | | | |
|---|---|---|---|---|
| | Active vs Passive: RE | .02 | .11 | .15 |
| | SRC vs. ORC: RE | .01 | .04 | .03 |
| **Model 4: Condition treatment contrasts** | | | | |
| Accuracy~ A_P*RE+A_SRC*RE+A_ORC*RE + (1\|Item) | | | | |
| Accuracy | Intercept | **4.14** | **.28** | **15.03*** ** |
| | Active vs Passive | **-2.72** | **.37** | **-7.36*** ** |
| | Active vs SRC | **-1.84** | **.34** | **-5.39*** ** |
| | Active vs ORC | **-2.68** | **.34** | **-7.97*** ** |
| | RE | **.22** | **.10** | **2.10*** |
| | Active vs Passive: RE | .02 | .05 | .86 |
| | Active vs SRC: RE | .02 | .05 | .40 |
| | Active vs ORC: RE | .01 | .04 | .14 |
| **Model 5 (lm):** Accuracy~ A_P*RE+A_SRC*RE+A_ORC*RE | | | | |
| Accuracy | Intercept | **.98** | **.01** | **184.95*** ** |
| | Active vs Passive | **-.22** | **.01** | **-29.91*** ** |
| | Active vs SRC | **-.09** | **.01** | **-10.66*** ** |
| | Active vs ORC | **-.18** | **.01** | **-20.67*** ** |
| | RE | .01 | .01 | .93 |
| | Active vs Passive: RE | **.03** | **.01** | **3.50*** ** |
| | Active vs SRC: RE | **.02** | **.01** | **2.29*** |
| | Active vs ORC: RE | **.03** | **.01** | **3.79*** ** |

**Table 4.** LME Models for the untransformed reading times, accuracy rates and RE results for pre-registered and exploratory (treatment) contrasts Note: ^p<.1; *p<.05; **p<.01; ***p<.001

**Variance accounted for by ART and Reading Enjoyment scores**

In exploratory follow-up analyses, ART and Reading Enjoyment scores were put in the same model with the pre-registered contrasts and exploratory treatment contrasts to investigate whether the two measures of text exposure accounted for similar or different sources of variance in reading times and comprehension accuracy. Full models are presented in supplementary materials (Supplemental materials, Exhibit C, Table C, Models 1-4). Despite the relatively low correlation between the two measures of text exposure (r = 0.33) and that each independently predicted reading times, we found no evidence that the inclusion of both ART and Reading Enjoyment in a model predicting reading times improved fit over including only a single predictor. Though we find some evidence that ART and RE may account for non-overlapping variance in comprehension question accuracy, given potential issues with model convergence and data sparsity, we cannot strongly draw this conclusion. We delegate it to future studies to investigate the sources of similarities and differences between ART and Reading Enjoyment further.

# Discussion

In a web-based experiment, we found differences in the speed with which participants read, and accuracy with which participants answered comprehension questions about four sentence types: simple active sentences, passive sentences and sentences containing subject and object relative clauses. Crucially, we found robust individual differences such that individuals with more text exposure read passive sentences and the sentences containing relative clauses more quickly and overall answered comprehension questions more accurately than participants with less text exposure.

Our key hypothesis was that text exposure should interact with sentence type. Text exposure should not uniformly affect sentence comprehension but rather we should see the strongest effects for the sentence types for which reading should most dramatically affect one's linguistic experience. We do see some evidence of main effects of text exposure on reading speed and clear evidence of main effects of text exposure on comprehension question accuracy. However, we also found sentence type by reading experience interactions. Individuals with more text exposure were faster particularly for the passive sentences and sentences containing SRC and ORC that individuals should encounter relatively more frequently from written language. For participants with more text exposure, reading times for the rarer, written-language biased sentences approached those of the simple active sentences. We found weaker evidence for similar interactions in comprehension question accuracy.

Our results have clear implications for experience-based accounts of sentence processing. Experience interacts with sentence type in predictable ways. We see stronger effects of text exposure on items for which we expect that experience should come predominantly from written language. A potential concern is whether the interaction between sentence type and text exposure reflects a true interaction or is an artifact of a floor effect in the simple active sentence. We argue that this “floor effect” may in fact be evidence of experience-based sentence processing. Our college-aged participants are sufficiently experienced, through both speech and text, with simple active sentences such that additional experience through higher rates of text exposure had little effect on behavior. It then follows that for less experienced readers, like children or adolescents, we would not expect a floor effect, but rather see robust effects of text exposure on even the simple active sentences.

We describe our hypothesis in Fig. 4. This hypothesis derives from notions in classic learning theory that early in training learning proceeds more quickly than later in training. At overall low rates of experience, a small amount of extra experience should have large effects on behavior. This point on the curve is indicated by a gray star which refers to our text-biased sentence types (passives, and sentences containing relative clauses). Participants overall have minimal experience with these sentence types, given their overall infrequency and bias to written language. As participants gain experience, they move rightward on the curve, so differences in experience (x-axis) lead to measurable changes in behavior, in this instance reading speed (y-axis). In contrast, the simple active sentences, which are overall more frequent in speech and text, are already very far to the right of the curve. Differences in experience thus have little effect on behavior. A change on the x-axis leads to little change on the y-axis. This asymptotic effect of behavior given experience is why we see little effect of text exposure on these sentences. However, in children, who have both less experience with spoken language and substantially less experience with written language, we expect actives to be higher on the curve, such that individual differences in experience, including text exposure, should be associated with a measurable effect on behavior. Our approach provides both a coherent account of the observed data and makes important predictions for patterns of behavior in less experienced readers (e.g., children and adolescents) as well expected patterns of behavior in other sentence types that may appear with different frequencies in written and spoken language.

In this work, we also attempted to establish the utility of a Reading Enjoyment survey that may corroborate or complement the commonly used Author Recognition Task (ART). We found that the ART and Reading Enjoyment survey generally accounted for overlapping variance despite being only moderately correlated themselves. Putting both measures of text exposure in a single model did not improve model fit. Larger samples may be necessary to more clearly understand the overlapping or non-overlapping aspects of text exposure that ART and Reading Enjoyment may capture. However, we identify a clear disadvantage of the ART in web-based studies: participants seem to use their phones or other devices to look up author names. We saw similar evidence of this device use in our Shipley vocabulary scores, which were unrealistically high and as such, unusable. While Reading Enjoyment surveys may not replace the ART given the ART's long history of successful use, web-based data collection may want to consider other means, like the Reading Enjoyment survey of assessing text exposure to complement the ART. Alternatively, adding a time limit on the ART display or presenting author names one at a time might discourage participants use of other devices during online study participation.

One question that remains is why we found such different rankings across our four sentence types for reading times and accuracies. Passive sentences were the second-fastest read sentence (after active sentences), but were the least accurately comprehended. There are several potential explanations. First, it is possible that online (reading times) versus offline measures (comprehension question accuracy) assess subtly different aspects of sentence processing or individual differences. For example, James et al. (2018)[20] finds effects of individual differences in only offline, not online, measures. In a related vein, because reading times were computed only for trials on which participants correctly answered the comprehension question, there may be different compositions of and sources of variability in the reading time and question accuracy measures.

Second, rather unintuitively, given that passives do not contain embedded clauses and prescriptive advice to avoid passives in writing, passives are remarkably biased to appear in written language. In our corpus analysis, passives were more text-biased than the SRC and ORC containing sentences. So perhaps the question ought to be not why passives were so poorly comprehended, but why were they read so quickly. Previous work also finds low rates of comprehension accuracy for passive sentences [51; 21; 53] but no difficulty or even facilitation on online processing measures [54; 55; 56; 57]. These results could be interpreted as a replication of the "good-enough" processing account [57; 50], that suggests that passive sentences are read quickly perhaps because they are interpreting them as actives.

Passive sentences may be particularly prone to misanalysis because of morphological features of the English passive that provides imperfect cues to a passive constriction. Relative clauses all contained the complementizer "that" with full noun phrases: both are strong, unambiguous cues for a subordinate clause so participants can be more sure of the type of sentence they are reading. In English, passive sentences have much weaker cues to their sentence type, at the verb and participle up to the "by"-phrase. In English morphology "was" and "ed" are not exclusive to passive sentences and passives can be interpreted as other sentence types as the sentence is unfolding, or even as a copula construction and an adjective up until the by-phrase as in the sentence "*The nurse was surprised.*" Passive utterances may be read quickly because they are particularly prone to misanalysis. Evidence for misanalyses in other sentence types is primarily reported in off-line accuracy measures [58; 59; 60] just as we see with our passive sentences. Future work can also clarify how the ability to use the imperfect morphological cues to the passive may change with experience, to allow us to understand more precisely what individuals with more or less text exposure may be doing during online and offline sentence processing.

This work provides evidence of effects of text exposure on sentence processing. Moreover, this work suggests pathways by which corpus statistics of spoken and written language could be used to further explore individual differences in language comprehension. The hypothesized pathways introduce clear experimental hypotheses as well as avenues of formal modeling (e.g., formalizing the model in Fig. 4) to better understand the links between input and language behavior. Future work may also benefit from finer-grained measures of sentence processing, including word-by-word reading times which allow experimenters to understand the locus of comprehension difficulty, as well as eye tracking measures that can distinguish between earlier and later measures of processing (e.g., first fixation vs. regressions) that can help us better understand the time course of sentence comprehension processes.

# Declarations

Data availability statement: stimuli, results, and analytical scripts are available on OSF repository https://osf.io/vct7s/

### Author contributions statement

A.S. and J.M conceived the experiment, A.S. conducted the experiment, A.S. and J.M. analyzed the results and wrote the manuscript together. All authors reviewed the manuscript.

### Additional information

We report having no competing interests.

## References


1. MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological review*, *101*(4), 676.
2. Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution.
3. MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in psychology*, *4*, 226.
4. Hahn, M., Futrell, R., Levy, R., & Gibson, E. (2022). A resource-rational model of human processing of recursive linguistic structure. *Proceedings of the National Academy of Sciences*, *119*(43), e2122602119.
5. Huettig, F. (2015). Four central questions about prediction in language processing. *Brain research*, *1626*, 118-135.
6. Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in cognitive sciences*, *22*(2), 154-169.
7. Frank, S. L., Trompenaars, T., & Vasishth, S. (2016). Cross-linguistic differences in processing double-embedded relative clauses: Working-memory constraints or language statistics?. *Cognitive science*, *40*(3), 554-578.
8. Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of memory and language*, *57*(3), 348-379.
9. Biber, D. (1991). *Variation across speech and writing*. Cambridge University Press.
10. Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition*, *20*, 51-68.
11. Stanovich, K. E., West, R. F., & Harrison, M. R. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Developmental Psychology*, *31*(5), 811.
12. Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & cognition*, *28*(1), 143-153.
13. Montag, J. L., & MacDonald, M. C. (2015). Text exposure predicts spoken production of complex sentences in 8-and 12-year-old children and adults. *Journal of Experimental Psychology: General*, *144*(2), 447.
14. Wu, S. H., Henderson, L. M., & Gennari, S. P. (2022). Animacy interactions with individual variability in sentence production and comprehension reveal similar lexically driven competitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(11), 1660.
15. 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.

16. Arnold, J. E., Strangmann, I. M., Hwang, H., Zerkle, S., & Nappa, R. (2018). Linguistic experience affects pronoun interpretation. *Journal of Memory and Language*, *102*, 41-54.
17. Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, *2*(3), 219-253.
18. Dąbrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences*, *28*(6), 604-615.
19. Huettig, F., & Pickering, M. J. (2019). Literacy advantages beyond reading: Prediction of spoken language. *Trends in cognitive sciences*, *23*(6), 464-475.
20. James, A. N., Fraundorf, S. H., Lee, E. K., & Watson, D. G. (2018). Individual differences in syntactic processing: Is there evidence for reader-text interactions?. *Journal of memory and language*, *102*, 155-181.
21. Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research*, *5*(1), 1-10.
22. Tskhovrebova, E., Zufferey, S., & Tribushinina, E. (2022). French-speaking teenagers' mastery of connectives: the role of vocabulary size and exposure to print. *Applied Psycholinguistics*, *43*(5), 1141-1163.
23. McClure, E., & Geva, E. (1983). The development of the cohesive use of adversative conjunctions in discourse. *Discourse processes*, *6*(4), 411-432.
24. Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1-76.
25. Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological review*, *99*(1), 122.
26. Levy, R., Fedorenko, E., & Gibson, E. (2013). The syntactic complexity of Russian relative clauses. *Journal of memory and language*, *69*(4), 461-495.
27. Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, *116*(1), 71-86.
28. Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of experimental psychology: learning, memory, and cognition*, *27*(6), 1411.
29. Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of memory and Language*, *51*(1), 97-114.
30. Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85*(1), 79-112.
31. Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of memory and language*, *57*(1), 1-23.
32. Montag, J. L. (2019). Differences in sentence complexity in the text of children's picture books and child-directed speech. *First language*, *39*(5), 527-546.
33. Biber, D., & Finegan, E. (1988). Adverbial stance types in English. *Discourse processes*, *11*(1), 1-34.

34. Gordon, P., & Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition*, *36*(3), 227-254.
35. Montag, J. L. (2019). Differences in sentence complexity in the text of children's picture books and child-directed speech. *First language*, *39*(5), 527-546.
36. Ebbinghaus, H. (1885). *Über das gedächtnis: untersuchungen zur experimentellen psychologie*. Duncker & Humblot.
37. Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements*, *2*(4), i.
38. Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of memory and language*, *47*(1), 69-90.
39. Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading research quarterly*, 402-433.
40. Moore, M., & Gordon, P. C. (2015). Reading ability and print exposure: Item response theory analysis of the author recognition test. *Behavior research methods*, *47*, 1095-1109.
41. Adelson, J. L., Cash, K. M., Pittard, C. M., Sherretz, C. E., Pössel, P., & Blackburn, A. D. (2019). Measuring reading self-perceptions and enjoyment: Development and psychometric properties of the Reading and Me Survey. *Journal of Advanced Academics*, *30*(3), 355-380.
42. Schiefele, U., Schaffner, E., Möller, J., & Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behavior and competence. *Reading research quarterly*, *47*(4), 427-463.
43. Shipley, W. C. (1986). Shipley institute of living scale. *The Journal of Psychology: Interdisciplinary and Applied*.
44. Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). https://doi.org/10.17605/OSF.IO/MD832
45. Thomas, K. A., & Clifford, S. (2017). Validity and Mechanical Turk: An assessment of exclusion methods and interactive experiments. *Computers in Human Behavior*, *77*, 184-197.
46. Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software*, 67(1), 1–48. doi:10.18637/jss.v067.i01.
47. R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
48. Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278.
49. Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of statistical software*, *82*, 1-26.
50. van Langen, J. (2020). Open-visualizations in R and Python. https://github.com/jorvlan/open-visualizations
51. Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive psychology*, *47*(2), 164-203.

52. Meng, M., & Bader, M. (2021). Does comprehension (sometimes) go wrong for noncanonical sentences?. *Quarterly Journal of Experimental Psychology*, *74*(1), 1-28.
53. Street, J. A., & Dąbrowska, E. (2010). More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers?. *Lingua*, *120*(8), 2080-2094.
54. Carrithers, C. (1989). Syntactic complexity does not necessarily make sentences harder to understand. *Journal of Psycholinguistic Research*, *18*, 75-88.
55. Paolazzi, C. L., Grillo, N., Alexiadou, A., & Santi, A. (2019). Passives are not hard to interpret but hard to remember: evidence from online and offline studies. *Language, Cognition and Neuroscience*, *34*(8), 991-1015.
56. Paolazzi, C. L., Grillo, N., Cera, C., Karageorgou, F., Bullman, E., Chow, W. Y., & Santi, A. (2022). Eyetracking while reading passives: an event structure account of difficulty. *Language, Cognition and Neuroscience*, *37*(2), 135-153.
57. Traxler, M. J., Corina, D. P., Morford, J. P., Hafer, S., & Hoversten, L. J. (2014). Deaf readers' response to syntactic complexity: Evidence from self-paced reading. Memory & Cognition, 42(1), 97–111. https://doi.org/10.3758/s13421- 013-0346-1
58. Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive psychology*, *42*(4), 368-407.
59. Sturt, P. (2007). Semantic re-interpretation and garden path recovery. *Cognition*, *105*(2), 477-488.
60. Wonnacott, E., Joseph, H. S., Adelman, J. S., & Nation, K. (2016). Is children's reading "good enough"? Links between online processing and comprehension as children read syntactically ambiguous sentences. *Quarterly Journal of Experimental Psychology*, *69*(5), 855-879.

# Figures

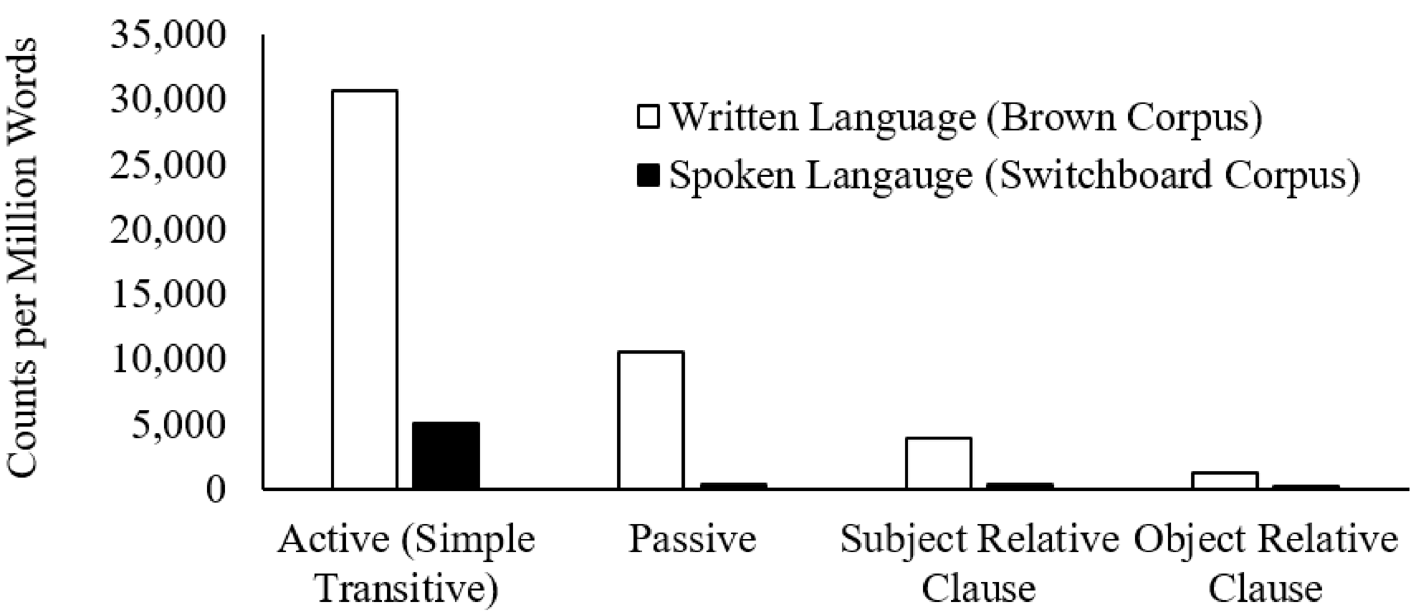

**Figure 1**

Counts per million words of the four experimental sentence types.

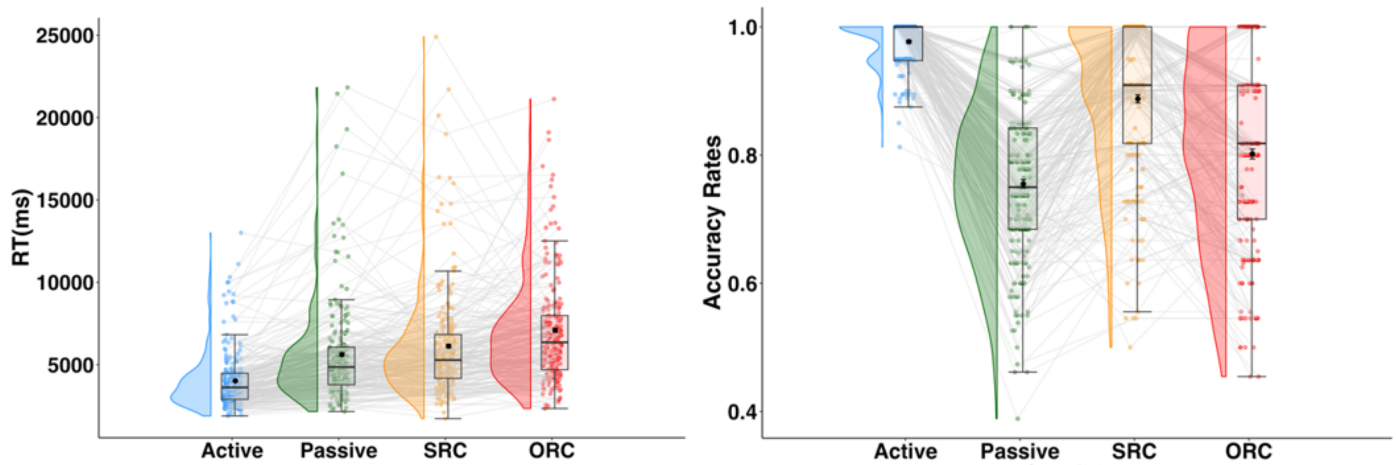


**Figure 2**

Sentence reading times (left) and comprehension question accuracies (right). Colored dots with grey lines = individual means; black dots with point ranges = conditional means with standard errors

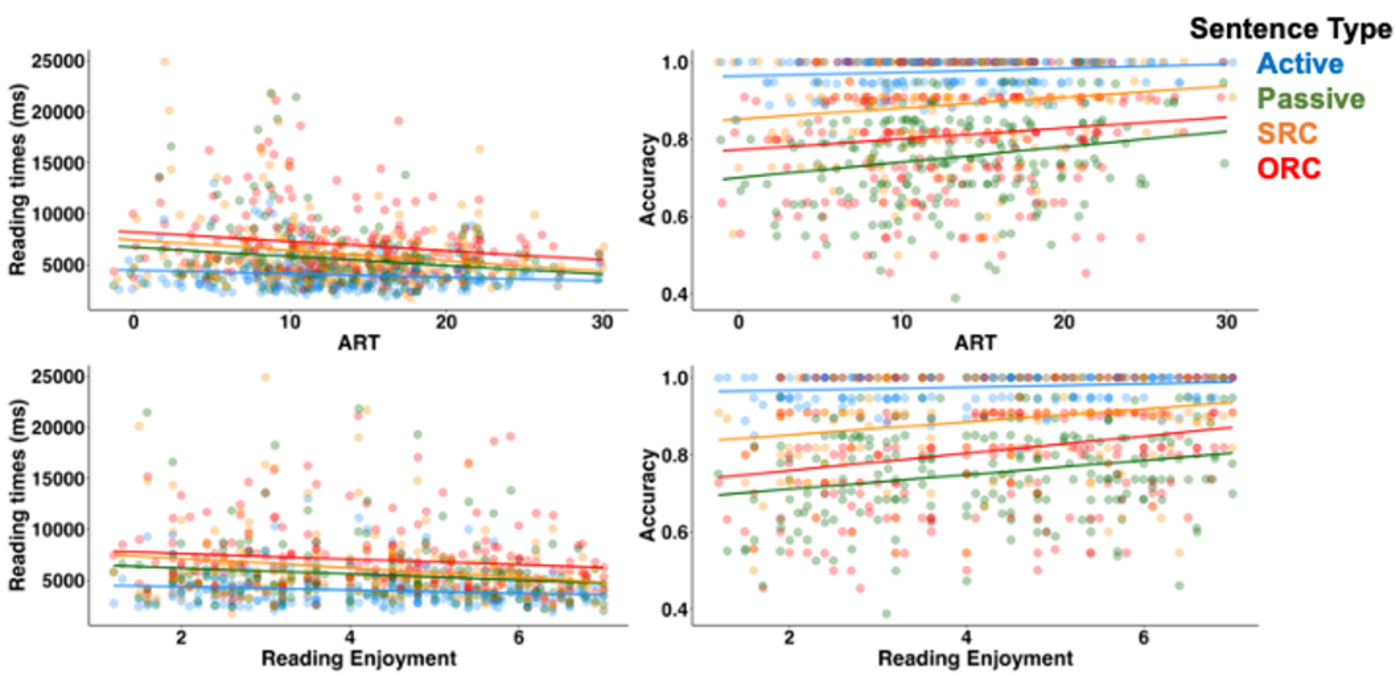


**Figure 3**

Reading times in milliseconds and comprehension question accuracy rates by ART scores (top row) and RE scores (bottom row) by sentence types. Colored dots = individual means

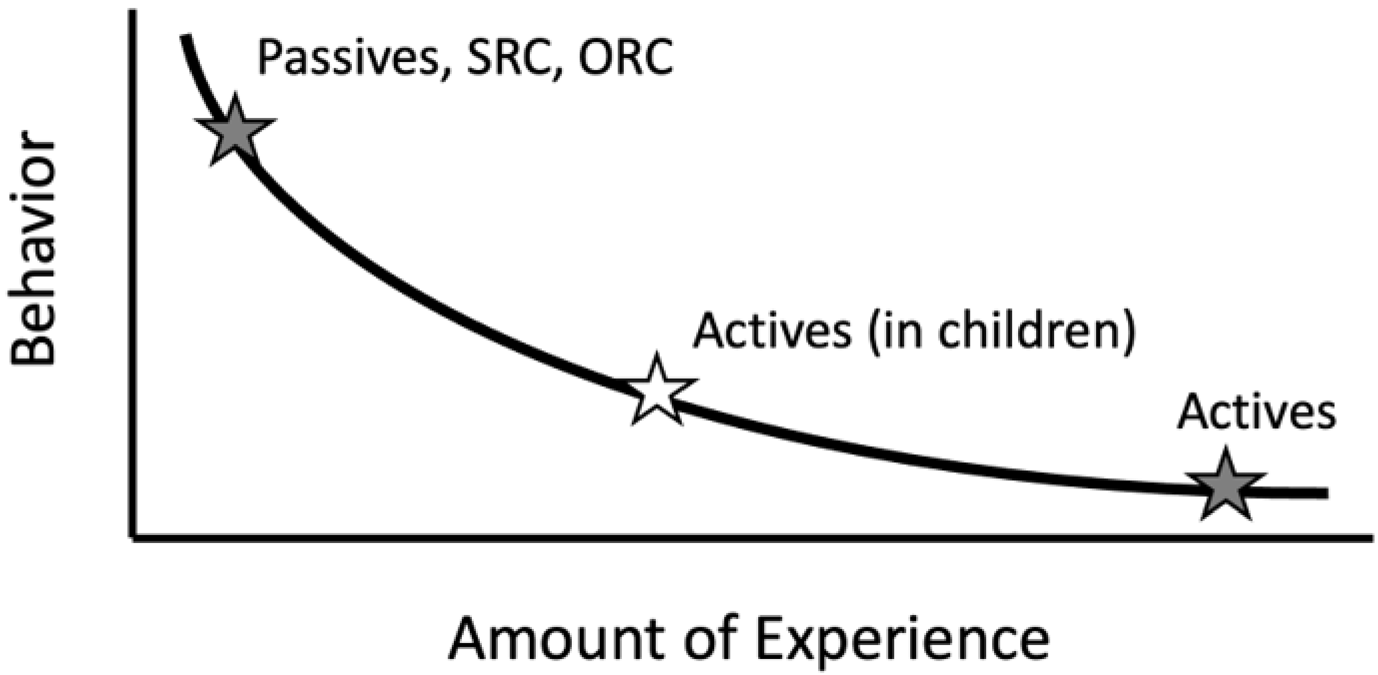


**Figure 4**

A visualization of our hypothesized relationship between language experience and behavior.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ExhibitA.docx
- ExhibitB.docx
- ExhibitC.docx