BRIEF REPORT

Literacy Can Enhance Syntactic Prediction in Spoken Language Processing

Saoradh Favier¹, Antje S. Meyer¹, and Falk Huettig¹, ²

¹ Department of Psychology of Language, Max Planck Institute for Psycholinguistics
² Centre for Language Studies, Radboud University

Language comprehenders can use syntactic cues to generate predictions online about upcoming language. Previous research with reading-impaired adults and healthy, low-proficiency adult and child learners suggests that reading skills are related to prediction in spoken language comprehension. Here, we investigated whether differences in literacy are also related to predictive spoken language processing in non-reading-impaired proficient adult readers with varying levels of literacy experience. Using the visual world paradigm enabled us to measure prediction based on syntactic cues in the spoken sentence, prior to the (predicted) target word. Literacy experience was found to be the strongest predictor of target anticipation, independent of general cognitive abilities. These findings suggest that (a) experience with written language can enhance syntactic prediction of spoken language in normal adult language users and (b) processing skills can be transferred to related tasks (from reading to listening) if the domains involve similar processes (e.g., predictive dependencies) and representations (e.g., syntactic).

Keywords: literacy, prediction, syntax

Supplemental materials: https://doi.org/10.1037/xge0001042.supp

Prediction has become the dominant theoretical framework for understanding the functioning of the mind and brain. In addition to playing an integral role in perception, action, and learning (Clark, 2013; Friston, 2005), the preactivation of predictable information is argued to reduce processing load and increase efficiency across multiple cognitive systems (e.g., Bar, 2003). Notably, psycholinguistic research increasingly emphasizes the importance of anticipatory mechanisms in language processing (for example, Altmann & Mirković, 2009; Dell & Chang, 2014; Federmeier, 2007; Ferreira & Chantavarin, 2018; Gibson et al., 2013; Hale, 2001; Hickok, 2012; Huettig, 2015; Kuperberg & Jaeger, 2016; Levy, 2008; Norris et al., 2016; Pickering & Gambi, 2018; Pickering & Garrod, 2013; Van Petten & Luka, 2012). However, most studies thus far have ignored the question of individual variation in predictive language processing (but see Federmeier et al., 2010; Hintz et al., 2017; Huettig & Janse, 2016; Kukona et al., 2016; Rommers et al., 2015).

The determinants (and extent) of individual differences in anticipatory language processing may offer important insights about the role and mechanisms of prediction in language and cognition more generally. There is now mounting evidence to suggest that individual variation in reading skills may be an important factor, even in anticipatory spoken language processing. Studies with reading-impaired adults (Huettig & Brouwer, 2015), healthy low-proficiency adults (Mishra et al., 2012), and child learners (Mani & Huettig, 2014) provide converging evidence that reading skills have a bearing on prediction in spoken language. But are these effects the hallmark of an impaired/developing system or are they a general feature of proficiency transfer between two related domains (reading and speech prediction)? This is an important question because it promises to illuminate the relationship between predictive processing and proficiency within a given domain as well as the transfer of training/experience between related domains. Such near transfer effects have previously been demonstrated, for example, between working memory and reading comprehension (Karbach et al., 2015; Novick et al., 2014).

Here, we aimed to address these issues by investigating prediction in language processing (defined as the preactivation of linguistic representations before incoming bottom-up input has had a chance to activate them; Huettig, 2015) in non-reading-impaired healthy adults with varying levels of literacy experience. One previous study provides some tentative evidence. Ng et al. (2018) asked a community sample of adults with varying literacy levels to listen to spoken sentences. Proficient readers showed a reduced
event-related potential (ERP) negativity for strongly predictable target words over anterior channels, in a time window from 170 to 300 ms after target word onset, but the same ERP response was not found in less-skilled readers. Neural correlates of prediction measured on the target word (and not before), however, can be explained by a number of other (nonpredictive) accounts, such as differences in the integration of nonpredicted representations (for further discussion see Baggio & Hagoort, 2011; Huettig, 2015; Mantegna et al., 2019; Nieuwland et al., 2018, 2020).

In the present study, we chose to investigate syntactic prediction because syntactic proficiency is expected to increase with literacy experience. Readers are exposed to a rich syntactic environment that is considerably more complex and diverse than typical conversational speech. Written narratives contain 60% more instances of subordination than speech (Kroll, 1977), and relatively infrequent structures like passives, object relative clauses, and participial phrases are predominantly attested in written corpora (e.g., Roland et al., 2007). The impact of written language exposure on syntactic proficiency has been demonstrated both for comprehension (Langois & Arnold, 2020; Street & Daibrowska, 2010) and production (Montag & MacDonald, 2015) in children and adults (e.g., Crain-Thoreson & Dale, 1992; Daibrowska, 2018). Several studies have investigated syntactic prediction (e.g., Arai & Keller, 2013; Chen et al., 2005; Kamide et al., 2003; Staub & Clifton, 2006), but (to the best of our knowledge) none assessed the influence of reading experience on syntactic anticipation in healthy literate adults.

Here, we used the visual world paradigm to measure eye gaze as a straightforward marker of prediction. We asked Dutch adults with high- and low-literacy experience to listen to passive sentences such as “Het raam wordt inderdaad gebroken door een stier [The window is indeed broken by a bull]” in conjunction with visual displays such as Figure 1.

The auxiliary wordt is an early but unreliable indicator of passive voice because it could also be parsed as an intransitive main verb (e.g., “Hij wordt rijp [He gets ripe]”). Only at the participle (gebroken) can the grammatical function of wordt be disambiguated and the preverbal argument (Het raam) be assigned the role of patient. In other words, there is sufficient information at the participle to parse the unfolding sentence as passive and to predict that a prepositional complement specifying the agent may follow. The preposition door, by, provides the final and unequivocal cue to expect an agent.

Method

Participants

Thirty-eight native Dutch speakers volunteered to participate in the experiment (M age = 25.2 years; 25 female). We recruited from a pool of 161 participants with varying degrees of literacy experience who had completed a battery of individual difference measures as part of a different study (Favier & Huettig, in press). We performed a preregistered principal components analysis on six literacy measures (receptive vocabulary, author recognition, reading habits, spelling, word, and pseudoword reading) to derive an underlying construct that explained the maximal amount of variance in the literacy data (literacy PC1 in Table 1). For the current study, we tested all participants in the top and bottom quartiles for the literacy PC1 who responded to our invitation for the eye-tracking experiment. Whereas this constrained the ultimate sample size, the potential to observe an effect of literacy experience was increased by our extreme-groups design. In addition, the sample size (N = 38) is similar to previous visual world studies that observed syntactic prediction effects with 30–40 participants (Arai & Keller, 2013; Kamide et al., 2003). One participant was excluded because the eye tracker failed to calibrate to the eyes. There is a pronounced group difference in literacy experience, based on literacy PC1 (t = 8.70, p < .001). The small difference in nonverbal IQ (Raven’s) scores between high- (HLE) and low-literacy experience (LLE) groups (t = 2.10, p = .04) was expected and is in line with previous research (e.g., Hervais-Adelman et al., 2019; Olivers et al., 2014; Skeide et al., 2017). The groups did not differ significantly in age (t = 1.79, p = .09). Because literacy PC1 was measured 18–24 months prior to the current study, we readministered the Author Recognition Test (Brysbaert et al., 2013; Stanovich & West, 1989) to account for change in literacy experience over time that could affect the size of the group difference. Author Recognition Test scores remained significantly higher in the HLE group (t = 4.01, p < .001).

Figure 1

Example Visual Display for the (Passive) Spoken Sentence, “Het raam wordt inderdaad gebroken door een stier [The Window Is Indeed Broken by a Bull],” in which “stier” is the Target Picture

Note. Distractors were always inanimate objects, and none were plausible agents of the events described. There were no statistical differences in mean frequency of picture labels (Table 2) in the four sets of pictures (F = 0.09, p = .97). Picture positions were pseudorandomized such that patient and target entities appeared equally often in all four positions on the visual display. There was no repetition of pictures in the experiment. See the online article for the color version of this figure.

1 A more detailed description of the tests used, the principal components analysis, and other further analyses, as well as the data, are available at https://osf.io/pds4w/?view_only=cc7a95d3b0414b2e8c86a86efa0a5d10.
Table 1

Descriptive Summary of High- and Low-Literacy Experience Groups

<table>
<thead>
<tr>
<th>Literacy experience</th>
<th>n</th>
<th>Age</th>
<th>Literacy PC1 score</th>
<th>Nonverbal IQ</th>
<th>Verbal WM</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>20</td>
<td>26.10 (4.41)</td>
<td>1.50 (.89)</td>
<td>22.30 (6.11)</td>
<td>7.30 (2.20)</td>
</tr>
<tr>
<td>Low</td>
<td>18</td>
<td>24.16 (1.89)</td>
<td>-1.69 (1.30)</td>
<td>18.28 (5.72)</td>
<td>8.39 (2.68)</td>
</tr>
</tbody>
</table>

Note. PC1 = first principal component; WM = working memory. Group means are based on raw scores for individual difference measures (except literacy PC1, a derived score). Standard deviations are shown in parentheses. Nonverbal IQ and verbal working memory were assessed using Raven’s progressive matrices and the backward digit span task, respectively.

Materials

Twenty-four Dutch passive sentences were constructed by combining a set of 12 transitive verbs with 24 inanimate patients (objects) and 24 animate agents (animals) such that each verb was presented twice. Our aim was to assess syntactic prediction; thus, we avoided object-animal pairings with salient semantic associations (e.g., shoe-dog) in favor of less typical combinations (e.g., paintbrush-dog). All experimental sentences used the present-tense passive frame as in “Het raam wordt inderdaad gebroken door een stier [The window is indeed broken by a bull],” that is, patient + present passive auxiliary + adverb + participle + preposition + agent. The adverb inderdaad (indeed) served as padding to ensure sufficient power to detect potential syntactic prediction effects. Each experimental sentence was paired with a visual display, comprising four color pictures (see Table 2) from the MultiPc database (Duñabeitia et al., 2018): an inanimate entity (the patient), an animate entity (the agent, henceforth the target), and two semantically unrelated distractors (see Figure 1).

In addition, 60 filler items using intransitive verbs (e.g., “De caravan staat inderdaad best ver weg van de molen [The caravan is indeed quite far away from the windmill]”) were constructed. A subset of 14 fillers was used as reference items to provide a hypertext participant experimental baseline for lexical processing speed. We constructed the reference items such that the sentence-final entity (the target) could not easily be predicted from either the sentence context or the pictures in the visual display (see Figure 2). Auditory stimuli were recorded in a soundproofed booth by a female Dutch speaker, using a Sennheiser ME64 microphone. The recording was sampled at 44 kHz (mono) with 16-bit sampling resolution.

Procedure

Participants were tested individually in a soundproofed experimental booth. An SR Research EyeLink 1000 tower-mounted eye tracker (Ottawa, Canada), was used to record eye movements. Participants were instructed to avoid moving their eyes away from the screen and to listen carefully to the sentences presented via headphones (a well-established protocol, particularly in experiments on prediction in language processing; see Huettig et al., 2011, for further discussion).

Each trial began with a 1-s central fixation, followed by a visual display. There was a 2-s preview of the pictures before the onset of the cue (wordt) in the speech signal. The pictures remained on the screen until 2,500 ms after the onset of the target word. The mean duration of the critical window between cue and target onset was 2,005 ms in passive trials and 2,162 ms in reference trials (here, the cue was defined as the verb, e.g., lijkt [seems]). There were 84 trials in total, of which 29% were passive and 71% were fillers. The order of trials was automatically pseudorandomized for each participant (with a maximum of two consecutive passive trials). The experiment took approximately 20 min, including calibration.

Results

Eyelink DataViewer was used to code fixations, saccades, and blinks. Data from four of 888 experimental trials were missing because of track loss. Fixation locations were coded automatically with respect to predefined regions of the visual display: patient, target, distractor 1, distractor 2, and background (i.e., none of the pictures). In the HLE group, 2% of fixations on experimental trials were coded as background. One participant in the low literacy experience (henceforth LLE group) was excluded because of 79% background fixations. For the remaining LLE group (n = 17), the rate of background fixations was 2%. Figure 3 shows the averaged fixation proportions to the target, patient, and averaged distractors on passive trials for HLE (Panel A) and LLE groups (Panel B), as well as a difference score, that is, the time course of target preference for each group (Panel C). Visual inspection of the plots suggests that the HLE group anticipated the target earlier than the LLE group (see Table 3 for group means).

The 14 nonpredictive reference items were analyzed to provide a baseline for lexical processing speed. This was indexed as the log odds of fixating the target entity in the first 500 ms after target word onset (with 200-ms adjustment), averaged across reference items. The mean rate of target fixations during this window was 35% (SD = 11%) for the HLE group and 30% (SD = 12%) for the LLE group, indicating a relatively small group difference. The empirical logit function (Barr, 2008) was used to transform individuals’ average target gaze durations to log odds, providing a participant-level index of lexical processing speed.

2 The semantic relatedness of patients and agents (indexd as cosine distance) was similar to that of patients and distractors (F = 1.15, p = .32). Cosine distances between verbs and targets versus verbs and distractors also did not differ significantly (F = 1.16, p = .32). The full set of visual stimuli for passive trials is reported in the online supplemental materials.
The present results corroborate previous findings that language comprehenders can use syntactic cues to generate predictions online about upcoming language (e.g., Arai & Keller, 2013; Chen et al., 2005; Kamide et al., 2003; Staub & Clifton, 2006). More importantly, the current study presents the first clear experimental evidence from non-reading-impaired adults that syntactic prediction in spoken language comprehension is related to adults’ literacy experience. The eye-tracking method used here enabled us to measure syntactic prediction in speech processing unequivocally (i.e., before participants heard the anticipated target). Literacy experience emerged as the strongest predictor of target (i.e., agent) preference, independent of general cognitive abilities. The main effect of literacy experience on anticipatory eye movements in our study echoes previous observations of reading-related differences in spoken language prediction based on other types of information (e.g., semantic representations, Mani & Huettaig, 2014; grammatical gender, Huettaig & Brouwer, 2015).

There are a number of alternative explanations of the present literacy-related syntactic prediction effect that can be rejected. First, one may argue that the group difference in anticipatory eye movements simply reflects slower word-object mapping in less-experienced literates because language experience has previously been linked to the efficiency of language-mediated looking (e.g., James, 2014; Mishra et al., 2012). We designed a baseline measure of lexical processing speed to address this potential confound and observed evidence neither for an interaction between literacy experience and lexical processing speed nor for a main effect of lexical processing speed on target preference. Second, one may suggest that differences in prediction simply reflect general ability (g-factor) differences that are measured by intelligence tests rather than differences in literacy experience. We can also reject this alternative explanation. Previous research has found that increased literacy and education results in small increases in Raven’s scores (e.g., Hervais-Adelman et al., 2019; Olivers et al., 2014; Skeide et al., 2017). Moreover, the minimal contribution of nonverbal IQ in our results is consistent with previous findings that Raven’s performance explains very little unique variance in language-mediated prediction (Hintz et al., 2017; Huettaig & Janse, 2016; Rommers et al., 2015).

Discussion

The present results corroborate previous findings that language comprehenders can use syntactic cues to generate predictions online...
The Time Window of Interest Extended From the Acoustic Onset of “Wordt” to the Onset of the Target Word (Time Zero), Both Adjusted by 200 ms to Account for the Time Taken to Program and Launch a Language-Mediated Saccadic Eye Movement (Saslow, 1967)

**A**

*Panel A plots fixation proportions to patient, target, and averaged distractor entities for the HLE group (n = 19). Panel B plots fixation proportions to the same entities for the low-literacy experience group (n = 18). Panel C plots the difference between target and distractor fixations (i.e., target preference) by group calculated by subtracting the proportion of averaged distractor fixations from the proportion of target fixations at each time step. Patient fixations were not included in this calculation. A target preference of zero means that the target and averaged distractors were fixated equally often, whereas values greater than zero reflect relatively more fixations to the target. The gray shaded areas represent by-participant 95% confidence intervals, computed at each 1-ms sampling step (Masson & Loftus, 2005). Vertical dashed lines indicate the acoustic onsets of *wordt* (M = –2,005 ms) and the target word (time zero). Vertical dotted lines represent a +200-ms adjustment to onset times, reflecting the typical latency of language-mediated eye movements. In Panel C, the approximate onset of the participle in the speech signal (M = –1,003 ms) is indicated by *gebroken*. See the online article for the color version of this figure.*
processing of multiple letters/words in proficient readers; see Huettig & Pickering, 2019. Exposure to the extreme form invariance of printed word forms; parallel informal speech (based on comparative corpus data for English; Roland et al., 2007) is most consistent with a secondary effect of literacy experience on predictive processing. It is also noteworthy that verbal working memory contributed very little to the observed anticipatory eye movements (cf. Huettig & Janse, 2016). To further assess causality and the individual contributions of primary and secondary influences of reading on (syntactic) prediction, a large-scale study with a longitudinal design would be useful.

To sum up, the present study strongly suggests that proficiency is important for predictive processing: Literacy experience enhances anticipation of upcoming language. Strikingly, experience with written language enhances syntactic prediction of spoken language in normal adult language users. Theories of prediction in language processing, and in cognitive science more generally, must be adapted to reflect more clearly that prediction is contingent on experience not only at the task at hand (e.g., spoken language processing) but also at related ones (e.g., reading). Processing skills transfer to related tasks if the domains involve similar processes (e.g., predictive dependencies) and representations (e.g., syntactic).

**Context of the Research**

Prediction has become a very influential theoretical construct of how the human mind works. The idea for this study originated in the research program of Falk Huettig, which explores how predictive processing is related to proficiency with the task at hand. This research had previously shown that reading-impaired adults, healthy low-proficiency adult readers in India, and child learners predict upcoming spoken language less than more literate children and adults. Here, we investigated whether these effects are a hallmark of an impaired or developing system or a general characteristic of language processing. We observed that literacy experience in healthy Dutch adults was the strongest predictor of (syntactic) target anticipation, independent of general cognitive abilities. This shows that proficiency is important for predictive processing and suggests that training/experience in one domain (reading in this case) transfers to a related domain (speech prediction). We plan to extend this research by teasing apart primary (directly linked to the physical act of reading, e.g., written language decoding, extreme form invariance of printed words) from secondary (also attained by listening to book language such as audio books, e.g., enhanced vocabulary knowledge and working memory) influences of reading on the prediction of spoken language. Ultimately, we want to know whether predictions are ‘just’ a natural byproduct of efficient access and processing of mental representations (rather than the deep goal of processing, in line with theories that consider the human mind to be a predictive engine).

### Table 3

<table>
<thead>
<tr>
<th>Group</th>
<th>Target</th>
<th>Average distractors</th>
<th>Mean difference</th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLE</td>
<td>.23 (.22)</td>
<td>.08 (.09)</td>
<td>.15 (.24)</td>
<td>.02 (.08)</td>
</tr>
<tr>
<td>LLE</td>
<td>.19 (.21)</td>
<td>.11 (.10)</td>
<td>.08 (.24)</td>
<td>.02 (.07)</td>
</tr>
</tbody>
</table>

**Note.** Standard deviations are provided in parentheses.

### Literacy and Prediction

How does literacy experience influence syntactic prediction in spoken language processing? It is useful to distinguish between primary and secondary influences of reading experience, both of which affect the core processes and representations that are common to written and spoken language (Huettig & Pickering, 2019). Secondary influences arise from exposure to book language, which is syntactically more elaborate (with higher demands on verbal memory) and lexically more extensive and sophisticated than conversational speech. Secondary influences can be attained not only through reading but also through listening to book-like auditory materials (e.g., audiobooks). It has been shown that the amount of shared book reading with parents at 24 months predicts children’s auditory comprehension of syntactically complex sentences at 30 months (Crain-Thoreson et al., 2001). Moreover, for children and adults alike, literacy results in both increased vocabulary knowledge (Cain & Oakhill, 2011; Cunningham & Stanovich, 1991) and verbal working memory (Démoulin & Kolinsky, 2016; Smalle et al., 2019).

Primary influences are more directly linked to the physical act of reading (for example, efficient decoding of written language; increased exposure to the extreme form invariance of printed word forms; parallel processing of multiple letters/words in proficient readers; see Huettig & Pickering, 2019, for further discussion). The present results cannot conclusively distinguish between primary and secondary influences of reading on spoken language prediction because our (statistically determined) literacy PC1 contained both primary (word and pseudoword reading, spelling, abilities) and secondary (receptive vocabulary, author recognition) characteristics of reading. That being said, the likely higher frequency of the Dutch passive construction in book language than informal speech (based on comparative corpus data for English; Roland et al., 2007) is most consistent with a secondary effect of literacy experience on predictive processing. It is also noteworthy that verbal working memory contributed very little to the observed anticipatory eye movements (cf. Huettig & Janse, 2016). To further assess causality and the individual contributions of primary and secondary influences of reading on (syntactic) prediction, a large-scale study with a longitudinal design would be useful.

To sum up, the present study strongly suggests that proficiency is important for predictive processing: Literacy experience enhances anticipation of upcoming language. Strikingly, experience with written language enhances syntactic prediction of spoken language in normal adult language users. Theories of prediction in language processing, and in cognitive science more generally, must be adapted to reflect more clearly that prediction is contingent on experience not only at the task at hand (e.g., spoken language processing) but also at related ones (e.g., reading). Processing skills transfer to related tasks if the domains involve similar processes (e.g., predictive dependencies) and representations (e.g., syntactic).

### Table 4

**Summary of Fixed Effects in Linear Mixed-Effects Model**

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coefficient</th>
<th>SE</th>
<th>t value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.24</td>
<td>0.38</td>
<td>3.25</td>
<td>[0.49, 1.99]</td>
</tr>
<tr>
<td>Literacy experience, high</td>
<td>1.16</td>
<td>0.56</td>
<td>2.06</td>
<td>[0.06, 2.26]</td>
</tr>
<tr>
<td>Lexical processing speed</td>
<td>0.13</td>
<td>0.52</td>
<td>0.24</td>
<td>[−0.89, 1.14]</td>
</tr>
<tr>
<td>Nonverbal IQ</td>
<td>0.02</td>
<td>0.05</td>
<td>0.37</td>
<td>[−0.07, 0.11]</td>
</tr>
<tr>
<td>Verbal working memory</td>
<td>0.07</td>
<td>0.10</td>
<td>0.69</td>
<td>[−0.13, 0.27]</td>
</tr>
<tr>
<td>Literacy Experience, High × Lexical Processing Speed</td>
<td>0.53</td>
<td>0.90</td>
<td>0.59</td>
<td>[−1.22, 2.29]</td>
</tr>
</tbody>
</table>

**Note.** N = 837. SE = standard error; CI = confidence interval. The intercept represents target preference (log odds of fixating target minus log odds of fixating distractors) for a participant in the LLE group with average lexical processing speed, nonverbal IQ, and verbal working memory. The positive intercept reflects the LLE group’s relatively higher odds of fixating the target versus the distractors (i.e., target preference) during the critical time window.


Received June 10, 2020
Revision received December 17, 2020
Accepted December 28, 2020