

**Long-term written language experience affects grammaticality judgments and  
usage but not priming of spoken sentences**

Saoradh Favier<sup>1</sup> & Falk Huettig<sup>1,2</sup>

<sup>1</sup>Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

<sup>2</sup>Centre for Language Studies, Radboud University, Nijmegen, Netherlands

*Quarterly Journal of Experimental Psychology (in press)*

Correspondence should be addressed to:

Falk Huettig

Max Planck Institute for Psycholinguistics

P.O. Box 310

6500 AH Nijmegen

The Netherlands

E-mail: [falk.huettig@mpi.nl](mailto:falk.huettig@mpi.nl)

phone: +31-24-3521374

fax: +31-24-3521213

## Abstract

‘Book language’ offers a richer linguistic experience than typical conversational speech in terms of its syntactic properties. Here, we investigated the role of long-term syntactic experience on syntactic knowledge and processing. In a pre-registered study with 161 adult native Dutch speakers with varying levels of literacy, we assessed the contribution of individual differences in written language experience to offline and online syntactic processes. Offline syntactic knowledge was assessed as accuracy in an auditory grammaticality judgment task in which we tested violations of four Dutch grammatical norms. Online syntactic processing was indexed by syntactic priming of the Dutch dative alternation, using a comprehension-to-production priming paradigm with auditory presentation. Controlling for the contribution of non-verbal IQ, verbal working memory, and processing speed, we observed a robust effect of literacy experience on the detection of grammatical norm violations in spoken sentences, suggesting that exposure to the syntactic complexity and diversity of written language has specific benefits for general (modality-independent) syntactic knowledge. We replicated previous results by finding robust comprehension-to-production structural priming, both with and without lexical overlap between prime and target. Although literacy experience affected the usage of syntactic alternates in our large sample, it did not modulate their priming. We conclude that amount of experience with written language increases explicit awareness of grammatical norm violations and changes the usage of (PO vs. DO) dative spoken sentences but has no detectable effect on their implicit syntactic priming in proficient language users. These findings constrain theories about the effect of long-term experience on syntactic processing.

**Key words:** literacy, grammaticality judgments, long-term syntactic priming

## Introduction

Syntactic diversity and complexity are key properties of ‘book language’. Sentence structure is often elaborate, with subordination, for instance, occurring 60% more frequently in written narratives than in spoken sentences (Kroll, 1977; cited by Kolinsky & Morais, 2018). Analyses of spoken and written corpora reveal pronounced asymmetry in the distributions of syntactic structures such as passives, object relative clauses, and participial phrases (e.g., Roland, Dick & Elman, 2007). It is important to note that exposure to the richer syntactic environment of ‘book language’ can similarly be gained from listening to audiobooks, or through shared reading for children. The associated benefits for syntactic knowledge can thus be considered a secondary influence of literacy, distinct from primary influences, which arise as a direct consequence of the physical act of reading (Huettig & Pickering, 2019, for further discussion). For example, Crain-Thoreson and Dale (1992) observed a secondary influence of ‘literate activity’ in pre-literate children. Their longitudinal study showed that the frequency of shared story reading with parents at 24 months reliably predicted performance on an auditory standardized test of syntactic comprehension at 30 months.

The current study investigated the contribution of individual differences in *life-long* literacy experience to syntactic processes. Why is that an important question? We suggest that it is a crucial prediction of experience- and usage-based theories of cognitive processing that life-long experience directly affects processing. In the domain of language, for example, usage-based theories assume that all linguistic skills result from accumulated experience with language across the totality of usage events in life (Bybee, 1995; Croft, 1995; Langacker, 1988; Tomasello, 2003). Acquisition is thought to be shaped by the quality and quantity of the input a language user receives (e.g., Abbot-Smith & Tomasello, 2006; Bybee 2006). ‘Book language’ is a source of high-quality input, based on its increased syntactic complexity and diversity relative to conversational speech (Kroll, 1977; Roland et al., 2007). In terms of input quantity, skilled readers encounter a larger volume of language through reading more, in addition to processing information at a faster rate than is possible for

listeners (e.g., skilled readers read English fiction at about 260 words per minute - approximately twice the typical speech rate; Brysbaert, 2019).

Indeed, in support of experienced-based accounts, there are a number of previous findings that suggest that life-long literacy experience affects syntactic processes in spoken language, such as the use of syntactic cues in ambiguous pronoun interpretation (Langlois & Arnold, 2020) and prediction (Mishra et al., 2012; Favier, Meyer, & Huettig, in press). Dąbrowska (2012), moreover, reviewed experimental work investigating the syntactic abilities of adult L1 speakers with varying levels of education and reported converging evidence for considerable individual differences in knowledge of 'core' grammatical constructions (including complementation, quantifiers, and passives, but see Favier & Huettig, in press). Differences were robustly correlated with education: while high educational attainment groups tended to score at or near ceiling, performance among individuals with low educational attainment was often at chance. Regarding the underlying factor driving these effects, it was acknowledged that education could be acting as a proxy for print exposure. The two factors are of course intertwined (print exposure correlates with years of formal schooling, e.g., Dąbrowska, 2018), and later work indeed revealed an independent contribution of print exposure to syntactic proficiency. Street and Dąbrowska (2010) found that print exposure reliably predicted comprehension of passives in a group of adults matched for educational attainment. Reading experience was a weaker predictor of performance on quantifier constructions in the same study, possibly reflecting the more symmetrical distribution of quantified noun phrases across spoken and written modalities (in contrast to full passives which occur seven times more frequently in written texts). In later work, testing comprehension of a range of grammatical constructions frequently heard in everyday conversation, Dąbrowska (2018) observed a significant (albeit small) unique contribution of print exposure.

Comprehension is not the only domain in which life-long written language experience can have consequences for syntactic processing. Montag and MacDonald (2015) examined the effect of

prior reading experience on implicit sentence production choices in children and adults. Individuals who scored highly on the Author Recognition Test (used as an index of print exposure) showed a pattern of production in their spoken language that reflected structural distributions in analysed written language corpora (specifically, increased frequency of passive relative clauses, which are rarely encountered in spoken language). This result leads straightforwardly to the conclusion that long-term exposure to a syntactic structure via reading facilitates its production in speech. The authors posited that becoming a reader entailed a quantitative and qualitative shift in linguistic experience, which continued to shape syntactic behaviour throughout adulthood. In short, there is considerable experimental evidence that life-long literacy experience affects syntactic processes, both in spoken language comprehension and production.

Although experience- and usage-based theories of cognitive processing predict that life-long experience directly affects syntactic processes, it is conceivable that not all syntactic knowledge and processing is affected by literacy experience to the same extent. We chose to explore the effect of long-term literacy experience on grammaticality judgments and syntactic priming because, to the best of our knowledge, there is little to no work that has investigated literacy-related influences on participant performance in these two experimental paradigms. Moreover, the two experimental tasks differ to the extent that they tap offline syntactic knowledge and online syntactic processing. It is important to point out that no psycholinguistic task involves purely explicit or purely implicit processes, but a mixture of both. It is, however, generally agreed that grammaticality judgment and syntactic priming are located at opposite poles of this continuum and it is conceivable that experience influences explicit processes differently than implicit ones.

Individual patterns of long-term syntactic experience related to literacy experience could conceivably play a role in individual differences in syntactic priming behaviour, for example through shaping the base frequencies of structural alternates. Base frequencies have been shown to modulate structural priming in a number of previous studies (e.g., Bernolet & Hartsuiker, 2010; Jaeger &

Snider, 2013; Segaert, Wheeldon, & Hagoort, 2016). In explaining the contribution of literacy to the grammaticality judgment task, some have argued that the decontextualised nature of written language facilitates metalinguistic thinking (e.g., Ravid and Tolchinsky, 2002, see Huettig & Mishra, 2014, for a review). Dąbrowska (2018) posits that inferring meaning from written text requires greater focus on the linguistic form, because of the absence of extra-linguistic cues typically available in speech (e.g., prosody and gesture). As well as being more conducive to learning syntactic structures, this attention to form may also support the ‘meta-syntactic’ processes involved in grammaticality judgment. The idea that literacy brings with it an explicit analytical awareness of language itself is supported by evidence for the causal role of alphabetic literacy acquisition in meta-phonological abilities (e.g., phoneme deletion, Morais et al., 1979). For Kolinsky and Morais (2018), metalinguistic thinking is a key feature of the metaphorical “literate glasses” through which literate people perceive the world.

### *Grammaticality judgment*

Dutch is an interesting case study for grammaticality judgment because of the prevalence of syntactic forms that are prohibited by prescriptive grammar but nevertheless occur frequently in the daily speech of native Dutch speakers. Well-documented examples include the use of the object pronoun *hun* ‘them’ as a subject, and the comparative marker *als* ‘as’ in comparative constructions of inequality, where *dan* ‘than’ is prescribed. Spoken corpus analyses reveal the prevalence of these prescriptive norm violations to be highest among low educated speakers (van Bergen, Stoop, Vogels, & de Hoop, 2011; Hubers & de Hoop, 2013).

### *Syntactic priming*

Syntactic (or structural) priming offers a tool to investigate online syntactic processing. Syntactic priming has more implicit components than metalinguistic tasks such as grammaticality

judgment, though may involve some explicit components as well (e.g., Bernolet et al., 2016). Bock (1986) found that after hearing and repeating a sentence like *The corrupt inspector offered a deal to the bar owner*, participants were more likely to use a prepositional-object (PO) dative to describe an unrelated pictured event (e.g., *The boy is handing a valentine to the girl*), compared with its alternative, the double-object (DO) dative (*The boy is handing the girl a valentine*). Since it was first reported over thirty years ago, the effect of recent syntactic experience on subsequent production has been demonstrated with a variety of tasks, syntactic structures, and languages (see Mahowald et al., 2016, for a meta-analysis). Evidence from pre-literate children (e.g., Branigan & McLean, 2016) shows that syntactic priming can occur without reading experience. However, it is particularly interesting for the purpose of the present study because it has been described both as a short-term (e.g., Pickering & Branigan, 1998) and a long-term phenomenon (e.g., Chang 2002; Chang et al., 2000, 2006).

There are many different theoretical accounts of syntactic priming but a main distinction can be made between activation-based accounts and error-based learning accounts. Traditional activation-based accounts are more compatible with short-term activation. Pickering and Branigan's (1998) account of syntactic priming, for example, posits that verb lemmas and their associated combinatorial nodes (specifying structure) become activated during comprehension, and that residual activation in a given combinatorial node increases the likelihood of reproducing a recently encountered structure (Pickering & Branigan, 1998). Due to the rapid decay of residual activation, syntactic priming according to this account is a relatively short-term phenomenon. Error-based learning accounts, in contrast, assume that syntactic priming is a more long-term phenomenon. Chang and colleagues (Chang 2002; Chang et al., 2000, 2006), for example, propose that during comprehension, the system continuously updates the weighting of mappings between message-level and abstract syntactic representations according to the input it receives. This implicit learning model of syntactic priming thus predicts longer-term effects of experience on syntactic priming. In line with

such an account, it has been observed that syntactic priming can persist over multiple intervening sentences (Bock & Griffin, 2000; Bock et al., 2007) and even a week (Branigan & Messenger, 2016).

An interesting and important finding is that syntactic priming effects are much larger when the verb is repeated across prime and target (the so-called lexical boost effect, e.g., Pickering & Branigan, 1998; Traxler, Tooley, & Pickering, 2014). We believe it is fair to say that researchers (especially those in the error-based learning camp) have struggled to reconcile the lexical boost effect (typically assumed to be a lexical short-term memory phenomenon) with (non-verb repeated) syntactic priming (typically assumed to be based on general structural procedures, i.e., rules). To reconcile long-term persistence with the short-lived boost to syntactic priming that occurs when prime and target sentences share a lexical head (e.g., Hartsuiker et al., 2008), for example, it has been proposed that repeated lexical material may simply cue retrieval of the prime sentence (Bernolet et al., 2016).

An (arguably) more ‘natural’ account of the lexical boost effect in syntactic priming is to abandon the traditional distinction between words and rules. Jackendoff and colleagues (Huettig, Audring & Jackendoff, under revision; Jackendoff, 2002; Jackendoff & Audring, 2020) argue that all rules can be restated in schema form, and as a consequence take on the same format as words, with the only difference to words being that some of a schema’s structure is made up of variables (see Jackendoff & Audring, 2020, for a detailed linguistic discussion of this). In this approach, words and schemas belong to a (single system) extended lexicon (see Bates & Goodman, 1997; Fillmore, 1988; and Langacker, 1987, for similar views). Words and syntactic schemas can prime subsequent occurrences, and accordingly, the lexical boost effect is a natural consequence of activation spreading between words and schemas as both are pieces of stored linguistic structure that are boosted by recent usage. This account also provides a natural explanation of the finding that low frequency items prime more than high frequency items (the so-called inverse priming effect, e.g., Scheepers, 2003; Snider & Jaeger, 2009) because a frequently encountered item has a high resting



state activation, which requires more input activation to raise its resting state activation (by the same amount) than a low frequency item (see Huettig et al, under revision, for further discussion).

It is important to point out here that long-term persistence of syntactic priming in the literature refers to priming over multiple intervening sentences or at maximum a week. Here, we investigate whether *life-long* literacy experience with alternating structures affects syntactic priming, which to our knowledge has not been directly explored. Life-long written language experience, for instance, may influence the usage of alternating structures if a given alternate is more prevalent in print materials. In the present study, we measure the potential bias associated with literacy experience directly by using a baseline measure of Dutch dative usage (PO or DO) in people with varying literacy levels (rather than relying on the small corpora that are available for Dutch, which are prone to biases). If literacy experience changes the usage of Dutch dative alternates, then it is conceivable that this affects their priming. For instance, in the account of language processing proposed by Jackendoff and colleagues, infrequent structures get more of a boost than frequent structures from the same amount of activation, resulting in stronger priming. This is because more activation is required to raise the resting activation of a frequent structure (perhaps reaching ceiling asymptotically). In short, reduced written language experience may make certain structures more infrequent for individuals with lower literacy and thus potentially result in greater priming of those structures.

### *The current study*

The current study investigated the contribution of individual differences in literacy experience to offline and online syntactic processes, as indexed by grammaticality judgment and syntactic priming respectively. Whilst the majority of participants in psycholinguistic research to date have been university students, this group is unrepresentative of the general population in terms of language and literacy skills, which are likely to be skewed towards the upper end of the normal distribution. Given

the theoretical importance of sampling from a broad spectrum of literacy abilities (Tarone & Bigelow, 2005), we focused our efforts on recruiting participants from diverse educational backgrounds. In addition, we tested participants outside of the lab to facilitate community participation.

We integrated correlational and experimental methods, using a correlational design with literacy as a predictor and grammaticality judgment accuracy and syntactic priming magnitude as the predicted variables. We measured a range of literacy-related skills as predictors: word and pseudo-word reading, receptive vocabulary knowledge, misspelling detection, author name recognition, and self-reported reading habits. We performed principal components analysis on these six variables to derive a principal component score, providing an index of literacy for our correlational analyses. We also included tests of working memory capacity, processing speed and non-verbal intelligence in our battery. These served as covariates in the analyses.

We developed an auditory grammaticality judgment task to probe participants' knowledge of four prescriptive grammatical norms in Dutch; specifically, their sensitivity to norm violations that occur in the everyday speech of many native speakers (Hubers, Snijders, & de Hoop, 2016). Whereas previous studies have investigated the grammaticality of these predominantly spoken constructions via the written modality, we used auditory presentation as we aimed to assess the effect of written language experience on spoken language processing. The task required participants to make a binary normative judgment about the syntactic form of each utterance (correct/incorrect). The within-subjects manipulated variable was grammaticality: whether or not the stimulus sentence violated a Dutch grammatical norm. Our outcome measure was the proportion of experimental items correctly judged as grammatical or ungrammatical, according to prescriptive usage. We predicted that native speakers' grammaticality judgments are influenced by their awareness of the syntactic discrepancies between written and spoken Dutch. We assumed that this awareness correlates with reading experience (i.e., exposure to written language) as indexed by our literacy measures. Put another way,

prescriptive grammatical norms are reliably attested in written language, whereas everyday spoken Dutch frequently contains violations of prescribed usage. Therefore, on the basis of differing input, we predicted that participants with less reading experience would have more difficulty recognizing prescriptive norm violations (i.e., their judgments would be more likely to reflect the syntactic patterns of spoken language). A secondary prediction was that grammaticality judgment would correlate positively with vocabulary knowledge, in line with the close association observed between grammar and vocabulary in development (Bates et al., 1995; Hayiou-Thomas et al., 2006).

The syntactic priming experiment focused on the Dutch dative alternation, using a comprehension-to-production paradigm (following Bernolet & Hartsuiker, 2010). Participants alternated between listening to (prime) sentences, performing a picture verification task, and providing spoken responses to target pictures. Rather than generating dative sentences from written verbs, participants in the current study completed dative sentence stems that were presented auditorily. This more constrained elicitation format was intended to minimise the involvement of literacy-related abilities in the task. Primes and their corresponding target pictures were adjacent, as immediate priming effects are expected to be stronger than priming after a lag (Bernolet et al., 2016). We manipulated the structure of the prime (prepositional-object or double-object dative, within items), and the repetition of the head verb between prime and target (verb same or different; within subjects). In line with previous research, we predicted that primed structures would be produced more frequently in the priming conditions. Furthermore, we predicted an increased likelihood of producing the primed structure when the prime verb was repeated in the target sentence (lexical boost). By including a baseline measure (rather than relying on the limited availability of Dutch corpora, Haemers, 2012; Coleman, 2009), we directly measured whether written language experience affects usage (base levels) of Dutch PO/DO dative constructions, and in turn their priming. Finally, based on the account presented by Jackendoff and colleagues, we predicted that reduced written language experience makes some structures more infrequent for individuals with

lower literacy, and thus results in stronger priming of those structures. For the same reason, we predicted that reduced written language experience also results in an increased lexical boost (under the assumption that low literates will get comparatively less exposure to certain verbs than high literates).

*Pre-registered predictions (<https://osf.io/zykp2>)*

Literacy will be positively correlated with accuracy in an auditory grammaticality judgment task (directional).

Vocabulary knowledge will correlate positively with grammaticality judgment accuracy (directional).

Participants will produce more target completions containing the primed structure after hearing a prime sentence versus a structurally unrelated control sentence (directional).

The likelihood of producing the primed structure will be enhanced when the prime verb is repeated in the target sentence (lexical boost) (directional).

We predict a negative correlation between literacy and the magnitude of the syntactic priming effect observed (directional).

The lexical boost will be stronger in participants with lower literacy (directional).

## **Method**

The study was pre-registered with the Open Science Framework, including a sample size appropriate to its correlational, individual differences design. The sampling rationale was based on work by Schönbrodt and Perugini (2013), in which Monte-Carlo simulations of correlational analyses identified  $N=161$  as a point of stability for estimated correlation magnitudes, after which sample estimates do not deviate from a pre-defined ‘corridor of stability’ around the true population value.

## *Participants*

161 Dutch native speakers participated for €10 per hour. We recruited a community-based sample through online and local advertising in Nijmegen, the Netherlands. 20 participants were recruited and tested in their local public library. Email invitations were also sent to eligible 18-35 year-old native Dutch speakers in the Max Planck Institute's participant database. None of the participants had a diagnosed reading disability and all had normal or corrected-to-normal hearing and vision. Thirteen participants were excluded from the analysis: 11 scored less than 2.5 standard deviations below the sample mean on at least one of the individual difference measures, and two had missing data.

## *Materials*

### **1. Individual difference measures**

#### **Literacy-related abilities**

We developed a battery to assess a range of literacy-related abilities, both directly and indirectly. The battery comprised standardised assessments that have been widely used in the psycholinguistics literature, and some measures developed for the current study. Each is briefly described below.

#### Een Minuut Test

We administered a standardised test of word reading ability, consisting of 116 Dutch words that progressively increase in difficulty (Brus & Voeten, 1973). We instructed participants to read the list aloud from top to bottom, as quickly as possible. The score was the number of words read accurately in one minute, precisely as printed on the test sheet. The experimenter timed the test using a stopwatch and scored responses on-line.

### Klepel Test

We used a standardised test of pseudo-word reading ability, comprising 116 Dutch pseudo-words of progressively increasing complexity (Van den Bos et al., 1994). The administration and scoring procedure were as above, except that participants had two minutes to read aloud as many items as accurately as possible. As some participants completed the list in less than two minutes, we also kept a record of their score after one minute. Digital voice recordings of both reading tests were made, and a native speaker later verified the scores.

### Peabody Picture Vocabulary Test

A large body of research highlights the bi-directional relationship between vocabulary knowledge and reading (e.g., Braze et al., 2007; Lee, 2011; Tannenbaum et al., 2006). In adulthood, most new words are encountered in written texts (Cunningham & Stanovich, 1998; Stanovich et al., 1995), making receptive vocabulary knowledge a useful proxy for literacy experience (not only oral language competence). We used a computerised version of the Dutch Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997; Dutch translation by Schlichting, 2005). Each trial comprised a spoken word and a visual display with four numbered line drawings. Participants selected the picture that illustrated the word's meaning by pressing the corresponding number on the keyboard. The task was self-paced and participants could listen to each word more than once. Trials were presented blocks of twelve, which progressively increased in difficulty. If the number of incorrect responses in a block exceeded seven, the test was discontinued. The raw score was the final item number reached, minus the total number of errors. From this the participant's standardised score and percentile rank were derived, based on Dutch age norms.

### Misspelling Detection Test

We developed a short paper test to assess receptive spelling knowledge, based on norms from Dutch and Flemish university students (Marc Brysbaert, personal correspondence). We selected a subset of 20 high-prevalence words with item scores that correlated the most with total test scores (0.30-0.55 correlation). We chose words from the higher end of the item score distribution (0.87-0.99 correct), to account for the wider range of ability in our community-based sample relative to the norming sample. Each correctly spelled word had a misspelled counterpart featuring a single substitution error, e.g., \*onbemindt (correct spelling: onbemind). Two counterbalanced, pseudo-randomised lists were constructed such that all 20 words appeared in their correct and incorrect versions across the two lists and no more than three of the same condition appeared consecutively. Each word was presented in a plausible sentence context, e.g., Hij stierf onbemind. We instructed participants to indicate whether the underlined word in the sentence was correctly spelled or not by marking a tick or a cross on the test sheet.

### Author Recognition Test

The Author Recognition Test (ART; Stanovich & West, 1989) is widely used as a proxy for engagement in print-related activities. Adapted for the Netherlands and Belgium (Brysbaert et al., 2013), the test comprises 60 author names, known to 66% of the Dutch norming sample, and 30 non-author foils that yielded 13% false alarms. We instructed participants to indicate which authors they knew and advised them against guessing as false alarms would be penalised. The test was completed on paper and untimed. The score was the number of authors correctly identified, minus the number of foils marked.

### Reading Habits Questionnaire

A paper questionnaire was used to evaluate self-reported engagement in print-related activities. This was a Dutch translation of the subtest “Your Reading Activities”, extracted from the

OECD Programme for International Student Assessment (PISA; OECD, 2009). Importantly, the questionnaire also probed time spent reading digital and online media. Participants answered questions on a four or five-point likert scale and the score was the sum of coded responses.

### **Covariates**

We also administered a battery of covariate measures to assess non-verbal intelligence, processing speed, and verbal working memory.

### Raven's Progressive Matrices

To assess participants' non-verbal intelligence, we administered a shortened, computerised version of Raven's advanced progressive matrices test (RPM; Raven, Raven & Court, 1993). The task was to indicate via mouse-click which of eight shapes completed a matrix of geometric patterns. Participants had 20 minutes to complete 36 items. It was possible to skip any item and return to it at the end of the test. The score was the total number of correct responses.

### Letter Comparison

As an index of processing speed, we used the letter comparison task (based on Salthouse & Babcock, 1991). Participants were presented with pairs of capital letter strings containing only consonants, in large black font on a white screen. The task was to indicate whether the strings were the same or different by pressing '1' or '0' respectively on the keyboard. Half of the items consisted of three-letter strings and the other half six-letter strings. Incongruent pairs differed by only one letter. There were six practice trials and 48 test trials, each beginning with a fixation cross, followed by a pair of letter strings that remained on the screen until the participant responded. There was an inter-trial interval of 1000ms. The score was calculated as the mean response time (RT) for all



correct responses that were no more than three standard deviations slower than the participant's grand mean RT.

### Backward Digit Span

We used a computerised version of the backwards-recall digit span task to measure working memory capacity, with auditory presentation of stimuli (adapted from the Wechsler Adult Intelligence Scale, 1997). Participants listened via headphones to sequences of two to eight digits, spoken by a female Dutch native speaker with a consistent rate (1 second pauses) and neutral prosody. The task was to type the sequence heard in reverse order, using the keyboard. There were 14 test trials, comprising seven blocks of two trials. Between blocks, sequences increased in length by an increment of one digit. The test was discontinued if participants responded incorrectly to both items in a block. The score was the number of correctly recalled digit sequences. We also recorded a top recall score for each participant; that is, the number of digits in the longest correctly recalled sequence.

## **2. Grammaticality judgment**

We based our stimuli on previous work by Hubers, Snijder, and de Hoop (2016), which focused on the perception of prescriptive grammatical norm violations in Dutch. For that study, they pre-tested several hundred sentences containing violations of five prescriptive norms. We thus had access to grammaticality ratings from an educationally diverse sample (n= 97; aged 18-35). We excluded one type of violation that was only relevant to written language, and calculated difficulty scores for items in the remaining four categories: *als/dan*; *mij/ik*; *hun/ze*; *die/dat*. As the goal was to develop a task challenging enough to yield a spread of scores, we selected the eight lowest-scoring items from each category, after outliers were excluded. Accuracy scores in the final shortlist ranged from 0.50 to 0.94, with 'die' constructions scoring the highest, and 'hun' constructions the lowest.

For each category of norm violation, we devised eight control sentences featuring prescribed usage of the relevant critical word; *als*, *mij*, *hun*, or *die* (see Table 1). The resulting 64 critical sentences were matched for syllable length, critical word position, as well as the frequency and prevalence of lexical items.

Critical word	Norm violation	Control
<i>hun</i>	Vorige week liepen <b>hun</b> naar de speeltuin. [Last week they walked to the playground]	Gisteren heb ik <b>hun</b> twee boeken gegeven. [Yesterday I gave them two books]
<i>als</i>	De jongen eet minder <b>als</b> zijn grote neef. [The boy eats less as his big cousin]	Zij is net zo groot <b>als</b> Vera op die hoge hakken. [She is just as tall as Vera in those high heels]
<i>mij</i>	Steven heeft eerder dan <b>mij</b> zijn rijbewijs gehaald. [Steven got his drivers licence earlier than me]	Hij vindt Linda aardiger dan <b>mij</b> maar niet grappiger. [He finds Linda kinder than me but not funnier]
<i>die</i>	Is er een bureau <b>die</b> voor mij bedoeld is? [Is there a desk that* is meant for me?]	Kent Kees een supermarkt <b>die</b> nog goedkoper is? [Does Kees know a supermarket that is even cheaper?]

**Table 1.** Example items from each category of norm violation with matched control sentences. \*Norm violation does not translate due to lack of grammatical gender in English.

Given the evident uncertainty amongst many native speakers of Dutch regarding the prescribed usage of these forms, we expected to see inaccuracy, both in the rejection of sentences that adhered to grammatical norms, and acceptance of sentences that violated them.

In addition, we generated 16 filler sentences, half of which were “truly ungrammatical”, featuring syntactic anomalies consistently detected by Dutch native speakers, e.g., errors relating to

subordinate clause word order, or verb tense and number agreement. These unambiguous filler sentences allowed us to ensure that participants were not responding randomly.

Stimuli were recorded by a female native speaker of Dutch, using a Sennheiser ME64 microphone. The speaker was instructed to maintain natural, conversational speech rate and prosody across all items.

All 80 items were presented to all participants in one, pseudo-randomised list, such that no more than three correct or incorrect items appeared consecutively. A maximum of two consecutive sentences could contain the same critical word, but they always contrasted in terms of grammaticality. The task began with two filler trials, one ungrammatical. A full stimulus list is provided in the online supplementary materials<sup>1</sup>

### 3. Syntactic priming

We selected 10 alternating dative verbs that have previously yielded syntactic priming effects in Dutch (Hartsuiker et al., 2008; Bernolet & Hartsuiker, 2010; Bernolet et al., 2016). We used these target verbs to generate 10 sentence stems (e.g., *Hij geeft* [*He gives*], *Ze overhandigt* [*She hands over*]), which could be completed as either a PO or DO dative construction. The gender of the subject pronoun was balanced across items, with 50% of sentence stems using *ze* [*she*].

From the MultiPic database (Duñabeitia et al., 2017), we selected colour pictures of 10 inanimate and 10 animate nouns, matched for Log10 word frequency (SUBTLEX-NL; Keulers et al., 2010), syllable length, picture naming agreement and visual complexity. We used the two sets of pictures to generate 30 different theme-recipient pairs, and assigned each pair to one of the 10 target sentence stems. We conducted a Google Books search to ensure that the transitional probabilities of target verb and animate/inanimate noun combinations were matched within items.

---

<sup>1</sup> Stimuli for both experimental tasks, as well as the data, and other further analyses are available in the supplementary materials on [https://osf.io/eg7pw/?view\\_only=56ebe093366943bfb3e6f3eb995c4fac](https://osf.io/eg7pw/?view_only=56ebe093366943bfb3e6f3eb995c4fac)

For each target item we constructed five prime sentences, corresponding to the following prime conditions: a) PO Verb Same; b) DO Verb Same; c) PO Verb Different; d) DO Verb Different; e) Baseline (see Figure 1).

- |  |                          |
|--|--------------------------|
| a) Ze schenkt een piano aan de priester.<br>[She gives a piano to the priest]    | <i>PO Verb Same</i>      |
| b) Ze schenkt de priester een piano.<br>[She gives the priest a piano]           | <i>DO Verb Same</i>      |
| c) Ze bezorgt een piano aan de priester.<br>[She delivers a piano to the priest] | <i>PO Verb Different</i> |
| d) Ze bezorgt de priester een piano.<br>[She delivers the priest a piano]        | <i>DO Verb Different</i> |
| e) De priester speelt piano.<br>[The priest plays piano]                         | <i>Baseline</i>          |



 Hij schenkt [he gives]...  
 ...een hoed aan de piraat [a hat to the pirate]  
 ...de piraat een hoed [the pirate a hat]

**Figure 1.** Example prime sentences with corresponding target picture, sentence stem, and expected completions. *Schenken* means “to give” (as in a gift).

We selected a separate set of 10 inanimate and 10 animate nouns from the MultiPic database, which provided 30 different combinations of prime themes and recipients. Prime sentences in conditions (a) and (b) repeated the 10 target verbs. For priming conditions (c) and (d), we selected 10 additional alternating dative verbs, on the basis of corpus and experimental data (Colleman, 2009; Colleman & Bernolet, 2012; more detail provided in the online supplementary materials). For the baseline condition (e), we combined monotransitive and intransitive verbs with the same set of nouns, to generate sentences that were syntactically unrelated to the dative. Like the target stems, all

dative prime sentences featured *hij* or *ze* as the agent of the dative action. The gender of the pronoun alternated between prime sentences and their corresponding target stems.

Auditory stimuli were recorded by a female Dutch native speaker, using a Sennheiser ME64 microphone. To create the set of target stems, we recorded both PO and DO versions of the complete dative sentences and cut them down using Praat software (Boersma & Weenink, 2017). This resulted in two versions of the same sentence stem for each of the 10 target verbs. We counterbalanced the PO and DO versions across target items, to reduce any differential influence of prosodic cues on participants' syntactic choices.

We constructed 90 filler items that were syntactically and semantically unrelated to experimental items. To reduce the saliency of the dative alternation, filler sentences varied in structure. 70 were simple transitives and intransitives in the present tense; 20 were complex sentences with a complement clause in the past tense, e.g., *Hij zei dat het hemd werd gestreken* [*He said that the shirt was ironed*]. As auxiliary-participle word order is reversible in Dutch subordinate clauses, we constructed a set of complex sentences using both word orders, counterbalanced across experimental lists. Filler trials followed the same two-part structure as experimental trials: a) sentence comprehension (picture verification) and, b) sentence completion (picture description). For the purposes of the cover task, 60% of items in the verification set featured a semantic mismatch between the sentence stimulus and the content of the visual display. For example, in one incongruent filler trial, participants saw a display with a man and a glass of milk and heard the sentence *De man drinkt koffie* [*The man drinks coffee*]. All dative trials were congruent.

The picture description set comprised non-critical pictures and corresponding sentence stems that would be unlikely to elicit a dative construction, e.g., *De vrouw draagt* [*The woman wears*]. The precise format of elicitation varied, depending on the type of filler. For example, where participants were required to complete a sentence with a verb phrase or noun phrase, they heard part of the target phrase immediately before the sentence stem was presented, e.g., *Rode. Hij slaapt in een...* [*Red. He*

*sleeps in a...*], or *Geslapen. Ze zei dat de koala...* [*Slept. She said that the koala...*]. Participants were familiarised with the different types of fillers through practice trials at the start of the priming experiment.

We constructed five pseudo-randomised lists of critical stimuli, such that across the lists every item appeared once in each of the five experimental conditions. To give a more reliable measure of participants' structural biases when not primed, we included six additional items in the neutral baseline condition, bringing the total number of dative trials to 36. These were interleaved with the 90 filler items, creating five lists of 126 trials, which we presented in a pseudo-random order such that each dative trial was preceded by at least two filler trials.

### *Procedure*

Participants individually attended two sessions within the same week, each lasting approximately one hour. The first session consisted of the syntactic priming experiment, followed by the grammaticality judgment task. In the second session, participants completed the following sequence of tasks in the same order: Een Minuut Test; Klepel Test; Backward Digit Span; Letter Comparison; PPVT; Misspelling Detection Test; RPM; Author Recognition Test; Reading Habits Questionnaire. Computerised tasks were carried out on a PC in a soundproofed experiment booth at the Max Planck Institute, or on a laptop in a reserved quiet room in the public library. Participants completed the remaining tasks at a desk, under the supervision of the experimenter. Alternating between the two types of activities was intended to help sustain attention levels and balance task demands.

The grammaticality judgment task was carried out on a PC or laptop, with auditory stimuli presented via headphones. Participants were instructed to listen to each sentence and respond to the question, “Is dit een correcte Nederlandse zin?” (Is this a correct Dutch sentence?), by pressing ‘1’ or ‘0’ on the keyboard (for yes and no respectively), and to guess if they were unsure. Each

sentence was presented once, along with a visual prompt showing ‘ja = 1’ on the left of the screen and ‘nee = 0’ on the right. There was no time limit on responses and no feedback given. As soon as a button press was recorded, the screen “Volgende zin” appeared and the next trial began. The task took 10-15 minutes to complete.

The syntactic priming experiment began with written instructions, followed by a series of examples to illustrate the verification task and demonstrate how sentence stems were to be completed in the picture description part of the trial. These demonstrations featured pre-recorded responses from a Dutch native speaker. Two of the example dative trials used PO target completions, and two DO, so that participants’ exposure to the alternates was balanced before beginning the priming experiment. The dative example trials were interleaved with six transitive filler examples to reduce the saliency of the dative alternation. Pilot testing indicated that such a demonstration was necessary to ensure that experimental stimuli reliably elicited dative responses. The passive demonstration phase was followed by four active practice trials, which were semantically and syntactically unrelated to the subsequent experimental trials. At the start of each trial, participants saw a fixation cross, followed by a pair of pictures, positioned in the lower left and upper right corner of the visual display (see Figure 1). The position of the animate and inanimate pictures on the screen was counterbalanced across all items and randomised within each experimental list. Participants then heard a pre-recorded prime sentence that referred to the displayed pictures. As a cover task, they were instructed to press ‘1’ or ‘0’ on the keyboard to indicate whether the content of the sentence and the picture were, respectively, congruent or incongruent. Participants received immediate on-screen feedback: “Correct!” or “Helaas, volgende keer beter!” (Better luck next time!) In the case of a correct response, the feedback screen also displayed reaction time in milliseconds (intended to increase motivation and engagement with the task). The second part of the trial comprised a new visual display with two target pictures (semantically unrelated to the prime pictures), and an auditorily presented dative sentence stem. Participants completed the sentence

aloud with either a PO or DO construction by naming the theme and recipient displayed. To reduce any influence of looking bias on their syntactic choices, participants had a 1000ms preview of the visual display before they heard the target sentence stem. Responses were recorded via a microphone attached to the headset.

## Results<sup>2</sup>

### *Scoring*

Correct responses in the grammaticality judgment task were coded as ‘1’ and incorrect responses as ‘0’.

Responses in the syntactic priming experiment were manually coded as prepositional-object datives (PO), double-object datives (DO), or Others. A response was coded as PO if the theme of the action was supplied first, followed by the preposition *aan* (to), and the recipient (e.g., after the target stem *Hij schenkt* [*He gives*] in Figure 1, *een hoed aan de piraat* [*a hat to the pirate*]). A response was coded as DO if the recipient was supplied first with no preposition, followed by the theme of the action (e.g., after the same target stem, *de piraat een hoed* [*the pirate a hat*]). Non-dative responses were coded as Other.

### *Descriptive summary of individual difference measures*

Means, standard deviations and ranges for each measure are reported in Table 2, as well as a descriptive summary of age. Correlations among the individual difference measures can be found in the online supplementary materials (Table S2).

Measure	Mean	SD	Range	Max
---------	------	----	-------	-----

<sup>2</sup> To interpret our results, we adopt a magnitude estimation approach where applicable, in line with recommendations that the field should be shifting focus away from significance testing, towards estimation based on effect sizes and confidence intervals instead (Cumming, 2013; see also Huettig & Janse, 2016; Hintz et al., 2017).



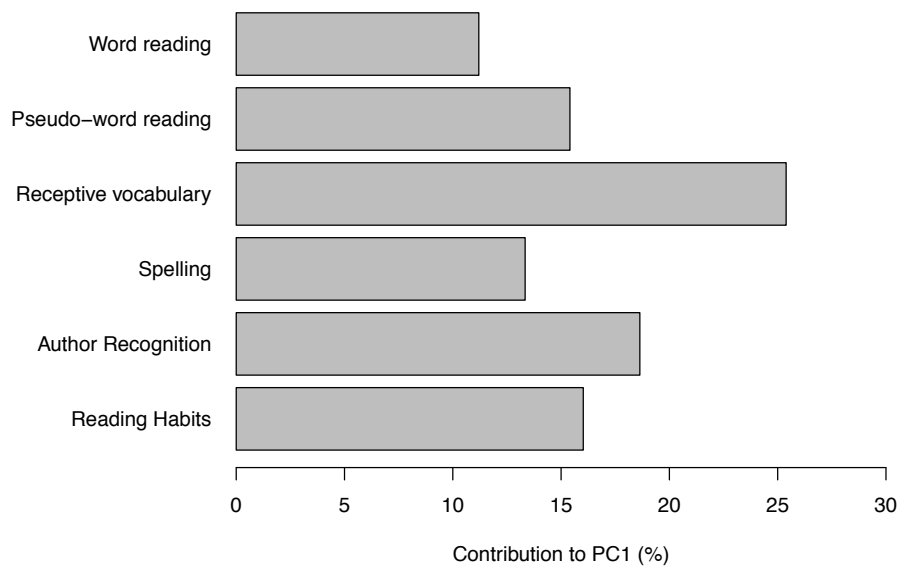
<i>Literacy related</i>				
1. Word reading (Een Minuut)	94.68	14.60	56–116	116
2. Pseudo-word reading (Klepel)	104.90	9.83	77–116	116
Klepel 1 min	65.69	10.76	40–94	–
3. Vocabulary (PPVT)	101.20	10.56	74–128	139
PPVT percentile rank	53.18	23.73	4–97	100
4. Misspelling detection	18.66	1.37	14–20	20
5. Author recognition	8.12	7.74	0–50	60
6. Reading habits questionnaire	80.04	8.88	41–105	114
<i>Covariates</i>				
7. Non-verbal IQ (RPM)	19.72	5.91	5–32	36
8. Processing speed (LC)	1076.00	191.51	673–1644	–
9. Working memory (BDS)	7.90	2.15	2–13	14
BDS top recall	5.45	1.23	2–8	8
10. Age (years)	23.41	3.44	18.00–34.58	–

**Table 2.** Means, standard deviations, ranges, and maximum possible scores for the individual difference measures. N = 148. Max = Maximum possible score; PPVT = Peabody Picture Vocabulary Test; RPM = Raven’s Progressive Matrices; LC = Letter Comparison task; BDS = Backward Digit Span task.

### *Principal components analysis*

Our test battery targeted a range of skills involved in literacy (measures 1–6 in Table 2). Using the *FactoMineR* package in R (Lê et al., 2008; R Core Team, 2012), we performed principal components analysis on this subset of variables to derive an underlying construct that explained the maximal amount of variance in the literacy data. The *FactoMineR* package contains a built-in function to evaluate the intercorrelation of variables with respect to a pre-defined criterion. The analysis extracted six principal components, of which the first explained 37.7% of the variance in the data. The composition of the first principal component is shown in Figure 2. All six literacy-related measures make some contribution, most of all receptive vocabulary (measure 3, Table 2) at 25%. We use the first principal component score as an index of literacy experience (predictor variable) in the

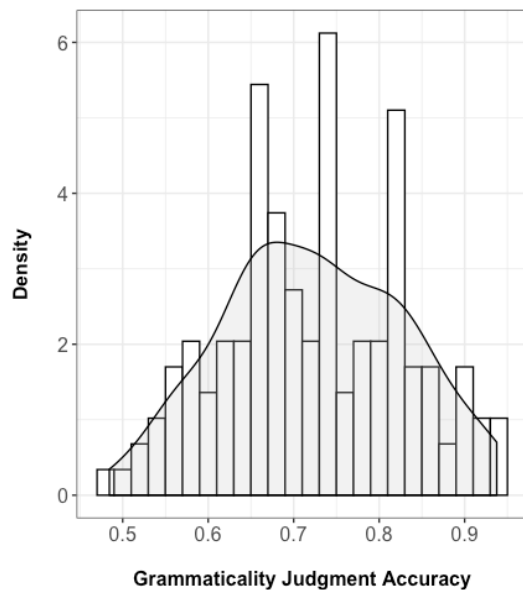
main analyses, since it explains the largest portion of variance in literacy-related skills. Further details about the other five principal components can be found in the online supplementary materials.



**Figure 2.** Bar plot showing the contribution of individual variables to the first principal component for literacy (PC1).

### *Grammaticality judgments*

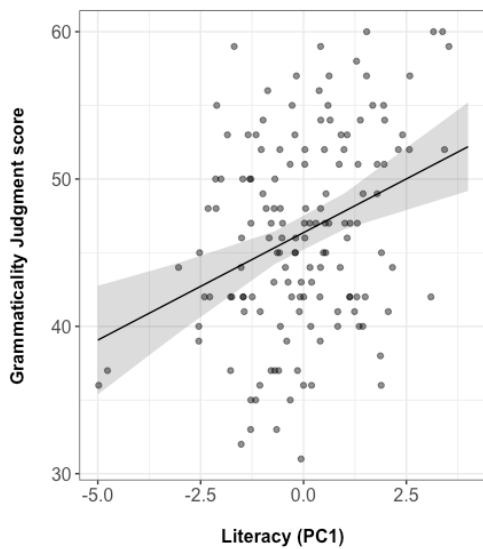
Descriptive statistics on the grammaticality judgment task were 0.72 (mean), 0.11 (standard deviation), 0.72 (median), and 0.48–0.94 (range). Consistent with the equal mean and median values, the histogram in Figure 3 reflects a fairly symmetrical distribution of scores across the sample. Only one participant performed below chance, and while mean accuracy on the task was relatively high, nobody scored at ceiling.



**Figure 3.** Histogram of grammaticality judgment accuracy with a plotted density curve.

#### *Literacy and grammaticality judgments*

We used multiple linear regression analysis to address our research question: Are individual differences in literacy associated with the identification of grammatical norm violations in spoken language? The approach (*lm* in R) enabled us to evaluate literacy as an independent predictor of grammaticality judgment, while controlling for the contribution of non-verbal IQ, verbal working memory, and processing speed. Scores on the RPM, backward digit span, and letter comparison task were entered into the model as covariates, with literacy score as a predictor. The fitted model with an  $R^2$  of .211 revealed an independent contribution of literacy to participants' grammaticality judgment accuracy (unstandardised  $\beta = 1.459$ ,  $SE\beta = 0.358$ , 95% confidence interval [.756, 2.161], standardised  $\beta = .328$ ). The standardised beta represents a measure of effect size, roughly equivalent to Pearson's  $r$ . Figure 4 shows a scatterplot of the relationship between literacy and grammaticality judgment, the line for the model is fitted to the data.



**Figure 4.** Scatterplot showing the relationship between literacy and grammaticality judgment. The line represents the regression fit from the model of judgment accuracy as a function of literacy score, controlling for the contribution of non-verbal IQ, verbal working memory, and processing speed ( $R^2 = .211$ , effect size = .328). Dependent variable is raw score in the grammaticality judgment task (scored out of 64). PC1 = First principal component.

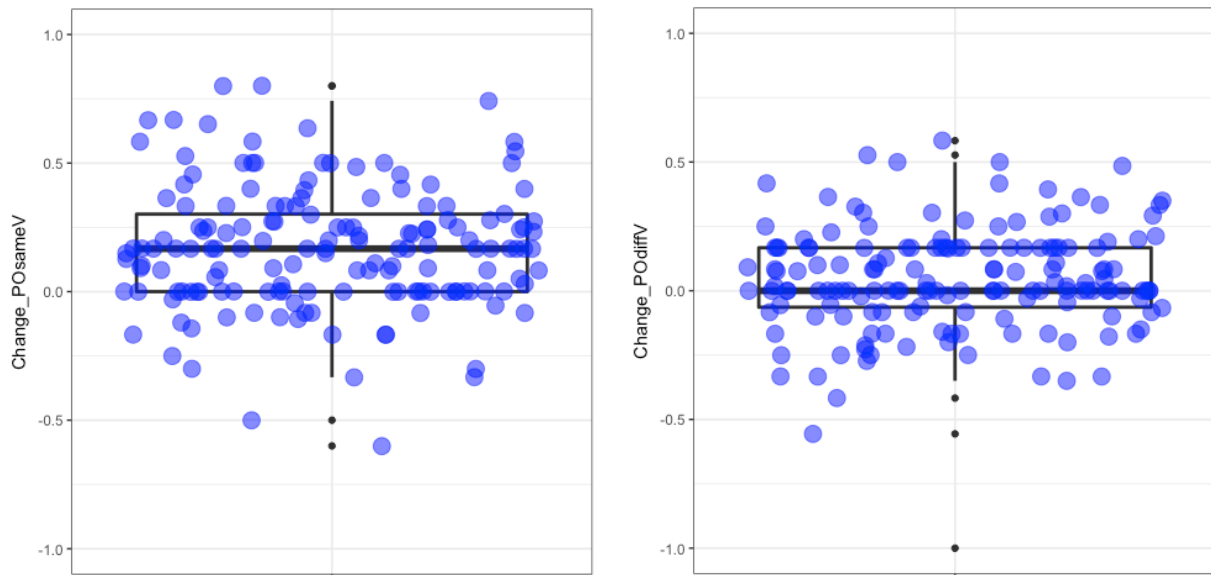
Regarding our secondary prediction that grammaticality judgments would be positively correlated with vocabulary knowledge, simple correlation analysis revealed a magnitude of  $r = 0.39$  (Kendall's  $\tau = 0.25$ ) between judgment accuracy and PPVT.

### *Syntactic priming*

Participants produced 2499 PO responses (43.1%), 2894 DO responses (49.9%), and 403 Other responses (6.9%). To evaluate the consistency of priming behaviour *within* individuals, we conducted a split-half reliability analysis. For subsets of even and odd trials separately, we calculated the proportion of PO responses as a function of prime condition for each participant. The PO proportions for the two subsets were then correlated to provide a measure of within-participant consistency. The split-half correlation magnitude was  $r = 0.66$  (Kendall's  $\tau = 0.56$ ), suggesting that priming behaviour was moderately consistent at the individual level.

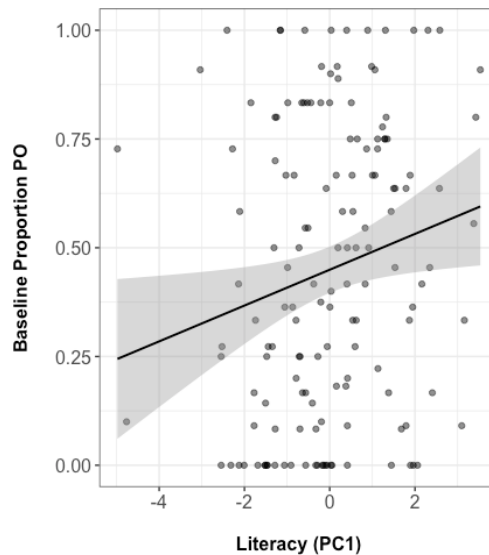
There was considerable between-participant variability in priming behaviour across both structures and verb conditions. Figure 5 illustrates the individual variability in PO priming

magnitude (and direction) for Same and Different verb conditions. The points below the zero line reflect a negative effect, i.e. people who produced fewer POs in the priming conditions compared to baseline.



**Figure 5.** Box plot showing the distribution of difference scores for PO priming (i.e., proportion PO after PO primes, minus proportion PO at baseline). Left panel = Same Verb condition; Right panel = Different Verb condition. Jittered data points correspond to individual participants' difference scores.

We first explored whether literacy is associated with a bias towards PO or DO dative constructions in line with our prediction that written language experience affects the usage of structural alternates. Figure 6 shows that higher literacy scores are associated with producing a PO dative following a neutral baseline sentence (correlation coefficient, Kendall's  $\tau = .19$ ). This is consistent with the notion that literacy experience affects (baseline) usage of PO or DO dative constructions.



**Figure 6.** Literacy score plotted against proportion of POs produced in the baseline condition. PC1 = First principal component.

Table 3 reports the proportion of POs and DOs out of all datives produced in each priming condition (excluding Other responses). The baseline proportions shown in Table 3 reflect the overall bias towards DO datives observed in this experiment (cf. Bernolet et al., 2016). The likelihood of producing a PO dative following a neutral baseline sentence was 45%. This increased to 51% when a PO prime was presented, resulting in a 6% priming effect in the absence of any lexical overlap between prime and target. When prime and target verbs were the same, there was a 62% chance of a PO response following a PO prime (17% priming effect). The 11% change in priming magnitude as a function of verb overlap demonstrates a lexical boost effect (see also the interaction between prime structure and verb condition shown in Table 4). DO datives showed weaker priming and lexical boost effects. Compared to baseline, the chance of a DO response was 4% higher in the different verb priming condition, and 10% higher in the same verb condition, indicating a 6% lexical boost.

	<b>Proportion PO</b>	<b>Proportion DO</b>
Baseline	0.45	0.55
PO Different Verb	0.51	0.49
PO Same Verb	0.62	0.38
DO Different Verb	0.41	0.59
DO Same Verb	0.35	0.65

**Table 3.** PO and DO responses as a proportion of datives produced in the different priming conditions.

We fit a linear mixed effects logistic regression model to the participants' responses across conditions, in line with the current standard for analysing categorical data (e.g., Barr et al., 2013; Jaeger, 2008). We used the 'lme4' package in R version 1.0.153 (Bates et al., 2015; R Core Team, 2012) to create the model, which predicts the logit-transformed likelihood of a PO response (see Table 4). PO responses were as coded as '1' and DO responses were coded as '0' (other responses were excluded from this analysis). The first model comprised three fixed effects: Prime Type (Baseline/Dative), Prime Structure (PO/DO), and Verb Condition (Same/Different). We used contrast coding to capture the nested design, whereby structure and verb condition were manipulated only within the dative primes (not the baseline primes). In addition, we were interested in the interaction between Prime Structure and Verb Condition (i.e., the lexical boost effect). The model included random intercepts for participants and target verbs, as well as a random effect of Prime Structure by participant and by target verb, and a random effect of Verb Condition by participant. We assumed that the priming effect would be influenced to varying degrees by individual target verbs' PO- or DO bias, hence the inclusion of target verb in the model's random effects structure. All random effects were de-correlated. The model results are summarised in Table 4.

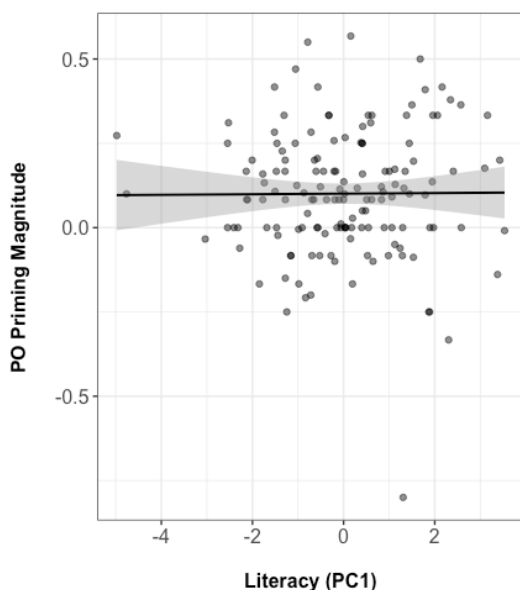
Predictor	Coefficient	SE	z value	CI
Intercept	-0.30	0.24	-1.25	-0.78 – 0.17
Prime Type (Dative)	0.00	0.12	0.01	-0.23 – 0.24
Prime Structure (PO)	1.27	0.10	12.68	1.07 – 1.46
Verb Condition (Same)	0.18	0.09	1.92	-0.00 – 0.36
Interaction = <i>Prime Structure &amp; Verb Condition</i>	1.20	0.19	6.35	0.83 – 1.58

**Table 4.** Summary of fixed effects in the mixed logit model (N = 4984, log-likelihood = -2425.7). The intercept represents the grand mean log-odds of a PO response, averaged across conditions. SE = Standard Error; CI = Confidence Interval.

Table 4 reveals a large syntactic priming effect (Prime Structure,  $z = 12.68$ ). It also reveals a robust lexical boost effect (*Prime Structure & Verb Condition*,  $z = 6.35$ ). These data therefore reflect a successful replication of the syntactic priming phenomenon (including the lexical boost effect) in a large, community-based sample of native Dutch speakers with varying literacy levels.

#### *Literacy and syntactic priming*

Figure 7 plots literacy score against priming magnitude (calculated as raw number of POs in PO prime condition minus POs at baseline). It shows that literacy did not modulate syntactic priming.





**Figure 7.** Literacy score plotted against priming magnitude (calculated as raw number of POs in PO prime condition minus POs at baseline). Correlation coefficient, Kendall's  $\tau = .081$ . PC1 = First principal component.

We fit another mixed logit model to participants' responses, this time incorporating Literacy (first principal component score) as a continuous predictor variable. The model results are summarised in Table 5.

Predictor	Coefficient	SE	z value	CI
Intercept	-0.30	0.24	-1.27	-0.77 – 0.16
Prime Type (Dative)	0.00	0.12	0.01	-0.23 – 0.24
Prime Structure (PO)	1.26	0.10	12.67	1.07 – 1.46
Verb Condition (Same)	0.18	0.09	1.95	0.00 – 0.36
Literacy PC1	0.28	0.12	2.43	0.05 – 0.51
Interaction = <i>Prime Structure &amp; Verb Condition</i>	1.19	0.19	6.32	0.82 – 1.56
Interaction = <i>Prime Structure &amp; Literacy PC1</i>	0.00	0.07	-0.05	-0.13 – 0.12

**Table 5.** Summary of fixed effects in the mixed logit model (N=4984, log-likelihood = -2422.8). The intercept represents the grand mean log-odds of a PO response, averaged across conditions. SE = Standard Error; CI = Confidence Interval.

Table 5 reveals a large syntactic priming effect (Prime Structure,  $z = 12.67$ ), which was not modulated by participants' literacy skills (*Prime Structure & Literacy*,  $z = -0.05$ ). Interestingly, the model shows that higher literacy scores are associated with a greater tendency to produce PO constructions in general (Literacy coefficient,  $z = 2.43$ , i.e., a main effect of literacy on the log-odds of a PO response, averaged across conditions). This adds to the finding that literacy experience increases usage of PO dative constructions following a neutral baseline sentence (Figure 6). In other words, in our large, community-based sample (N=161) of individuals with varying literacy levels, we observed that literacy experience affects the usage of Dutch PO or DO datives, yet literacy experience did not modulate the syntactic priming of these constructions.

Some accounts of syntactic priming predict an inverse effect of structural preference on priming. One way of examining this is via prime verbs' individual subcategorisation biases (e.g., Jaeger & Snider, 2013; Bernolet & Hartsuiker, 2010). Another way is to look at the relationship between participants' baseline structural preferences and their tendency to be primed. To the mixed logit model described in Table 4, we added PO base rate as a continuous predictor (PO base rate = rate of PO production as a proportion of dative responses in the baseline condition). Importantly, we included its interaction with PO prime to explore the question: Does a lower PO base rate predict an increased likelihood of PO production after a PO prime? The model results are summarised in Table 6. The confidence interval for the coefficient suggests only a very marginal interaction effect in the predicted direction; i.e., people with a lower PO base rate show a tendency to prime more for PO.

Predictor	Coefficient	SE	z value	CI
Intercept	-2.82	0.21	-13.74	-3.22 - -2.42
Prime Type (Dative)	-0.02	0.12	-0.13	-0.26 - 0.22
Prime Structure (PO)	1.01	0.23	4.35	0.55 - 1.46
Verb Condition (Same)	0.16	0.10	1.54	-0.04 - 0.36
PO base rate	5.45	0.22	24.38	5.01 - 5.89
Interaction = <i>Prime Structure</i> & <i>Verb Condition</i>	1.21	0.19	6.50	0.84 - 1.58
Interaction = <i>Prime Structure</i> & <i>PO base rate</i>	-0.67	0.37	-1.81	-1.40 - 0.05

**Table 6.** Summary of fixed effects in the mixed logit model (N = 4984, log-likelihood = -2286.9). The intercept represents the grand mean log-odds of a PO response, averaged across conditions. SE = Standard Error; CI = Confidence Interval.

Finally, to test our prediction that the strength of the lexical boost would be negatively associated with literacy, we created a further model, identical to the one described in Table 5, except for the addition of a three-way interaction between Prime Structure, Verb Condition and Literacy

PC1. We did not find evidence that literacy experience modulated the lexical boost, nor did the inclusion of the interaction improve model fit.

## **General discussion**

We investigated the contribution of individual differences in literacy experience to syntactic processes in spoken language. We administered a battery of tests to assess a range of literacy-related skills and their covariates (non-verbal IQ, verbal working memory, and processing speed). We used two experimental tasks, grammaticality judgment and syntactic priming, to target offline and online syntactic processes respectively. Four of our pre-registered predictions were confirmed: Literacy was positively correlated with accuracy in an auditory grammaticality judgment task; vocabulary knowledge correlated positively with grammaticality judgment accuracy; participants produced more target completions containing the primed structure after hearing a prime sentence versus a structurally unrelated control sentence, and the likelihood of producing the primed structure was enhanced when the prime verb was repeated in the target sentence. Two of our pre-registered predictions were not confirmed: There was no negative correlation between literacy and the magnitude of the syntactic priming effect observed, nor was there evidence for a stronger lexical boost in participants with lower literacy. We will now discuss these results in turn.

### *Grammaticality judgment*

Violations of four Dutch grammatical norms were tested. We observed systematic variation across individuals in their accuracy on the grammaticality judgment task. Above and beyond the contribution of non-verbal IQ, verbal working memory, and processing speed, literacy uniquely predicted participants' ability to correctly accept and reject spoken sentences according to the prescriptive grammatical norms of their language. Controlling for the contribution of non-verbal IQ, verbal working memory, and processing speed, we observed a robust effect (standardised  $\beta = .33$ ) of

literacy experience on the detection of grammatical norm violations in spoken sentences, suggesting that exposure to the syntactic complexity and diversity of written language has specific benefits for general (modality-independent) syntactic knowledge. This result converges with and extends previous findings concerning the relationship between print exposure and syntactic abilities (Dąbrowska, 2012; Street & Dąbrowska, 2010). The current study used a multi-faceted measure of literacy experience and found a moderate positive correlation with metalinguistic syntactic abilities, adding evidence to support the link between language experience and aptitude in adult native speakers. Our separate finding that grammaticality judgment was positively correlated with vocabulary knowledge in adult native speakers is consistent with longitudinal evidence for the intertwined development of grammar and vocabulary (Bates et al., 1995; Hayiou-Thomas et al., 2006).

In line with the notion that greater vocabulary knowledge is associated with greater written language experience, we had pre-registered that vocabulary knowledge will be positively correlated with accuracy in grammaticality judgments. We had also pre-registered to investigate the influence of verbal working memory as a covariate (<https://osf.io/zykp2>). One may however regard verbal working memory more accurately as a secondary effect of reading (Huettig & Pickering, 2019, cf. Demoulin & Kolinsky, 2016; Smalle et al., 2019). Secondary influences are those influences that can also be attained by listening to ‘book-like’ auditory materials. Listening to audiobooks, for example, supports the acquisition of ‘book language’ because it contains syntactically more elaborate language (with higher demands on verbal memory) and more extensive and sophisticated vocabulary than conversational speech. Primary influences are those that are more directly linked to the physical act of reading (e.g., 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 literacy effect on grammaticality judgment in our study may well be more secondary in nature, likely originating from

exposure to “book language” as opposed to physical reading practice. Future research thus could usefully delineate primary and secondary influences of literacy directly by conducting a confirmatory pre-registered study.

From an experience-based perspective, we had a straightforward prediction about the effect of literacy experience on grammaticality judgment accuracy. The task was to judge the “correctness” of spoken sentences with reference to prescriptive norms that are attested in written texts far more consistently than in spoken language. Therefore, on the basis of quantitative and qualitative differences in the input, prolific readers should have more relevant data to support their judgments. This is consistent with the experience-based notion that variation in grammaticality judgments partly reflects individual patterns of linguistic experience (Kidd et al., 2018).

When considered as a measure of explicit syntactic awareness, grammaticality judgment requires the caveat that the contribution of some implicit syntactic knowledge to task performance cannot be ruled out. Given that no psycholinguistic task is ‘purely’ explicit or implicit, and that participants in grammaticality judgments are explicitly asked to make a meta-linguistic judgment, we can however be reasonably confident that grammaticality judgments involve more explicit processing than the syntactic priming task which only requires participants to complete sentences (and does not explicitly draw attention to the purpose of the task).

### *Syntactic priming*

We successfully replicated the most well-documented effects in the syntactic priming literature. Using comprehension-to-production priming of the dative alternation in Dutch, we observed a statistically large syntactic priming effect ( $z = 12.67$ ) and a robust lexical boost effect ( $z = 6.35$ ) in a large, community-based sample of native Dutch speakers with varying literacy levels. Within our sample ( $N=161$ ) there was considerable individual variability in syntactic priming behaviour, with many participants showing no priming at all, and others showing a negative effect.

Individual differences have been given little attention in the syntactic priming literature to date (cf. Kidd, 2012). There is a tendency to consider effects only at the group level, and to dismiss the absence of priming as experimental noise. Gathered from a large and diverse sample (with respect to ability), our data suggest that between- and to some extent within-participant variability is the norm rather than the exception for syntactic priming.

The large syntactic priming effect we observed was not modulated by participants' literacy experience (*Prime Structure & Literacy*,  $z = -0.05$ ). Importantly, this absence of a modulation of the priming effect was not due to an absence of differences in structure *usage*. Our model revealed that higher literacy scores were in fact associated with a greater tendency to produce PO constructions in general (Literacy coefficient,  $z = 2.43$ , i.e., a main effect of literacy on the log-odds of a PO response, averaged across conditions). Moreover, literacy experience was associated with an increased usage of PO dative constructions (and therefore a decreased usage of DO datives) following a neutral baseline sentence (Figure 6). We therefore conclude that literacy experience affects the usage of Dutch PO/DO dative alternates but does not modulate the syntactic priming of these constructions.

We had predicted a negative correlation between literacy experience and priming magnitude, motivated by the notion that literacy-related differences in usage of the dative alternation would affect syntactic priming of the structures (cf. Huettig et al., under revision). This hypothesis was not supported. One possibility is that literacy-related usage differences only play a role in syntactic priming during the stage of language *acquisition* but (more or less) 'level off' in proficient language users such as the adults who took part in the present study. Future (large N) studies could usefully further explore this possibility. Given that our participants may have reached a plateau in L1 syntactic acquisition, regardless of literacy level, a large-scale developmental study may be particularly fruitful for exploring the influences of literacy on syntactic priming in childhood.

Our pre-registered prediction that participants with less literacy experience would show a stronger lexical boost effect was also not borne out in the data. Despite the robust lexical boost observed at the group level, there was no indication that it was modulated by individual differences in literacy experience. Based on the account presented by Huettig et al. (under revision), inexperienced literates might be expected to show a larger lexical boost because items (in this case, verbs) that are encountered less often (e.g., due to limited written input) are argued to have a lower resting state and thus receive a larger boost from the same amount of activation. One possible reason why we did not observe such a relationship in the current study is that relatively high frequency verbs were used. If the resting state of the prime verbs was generally high due to frequent exposure, it follows that the activation boost caused by a given verb's repetition (i.e., the lexical boost) would not be sensitive to the effects of differential written language input. Future studies could explore this hypothesis by testing low frequency verbs, which would be expected to have a relatively low resting state in people who read very little.

In our experiment, PO datives were produced less frequently than DO datives overall, yet yielded a larger priming effect. Previous studies have shown that infrequent structures tend to prime more reliably, in line with notions that the unexpectedness of the prime structure has a significant effect on priming (i.e., more frequent/predictable structures are assumed to prime less than relatively infrequent/unpredictable structures, the so-called inverse preference effect). Our measure of participants' PO/DO preference following a non-dative prime (baseline condition) allowed us to test this notion. Our exploratory analysis revealed a marginal negative effect of PO base rate on PO priming, suggesting a small tendency in the predicted direction. However, given the big sample size, our data do not provide robust evidence for the inverse preference effect, as has been reported elsewhere in the literature (e.g., Jaeger & Snider 2013). Given that many other (often unpublished) studies have failed to observe the inverse preference effect (e.g., Kaschak & Borreggine, 2008),

future research could be directed at exploring which factors modulate the presence or absence of this effect in priming experiments.

We can reject some alternative explanations of the present data with considerable confidence. Is the absence of a modulation of priming magnitude by literacy ‘just a null effect’ in one study that failed to detect a ‘real’ effect? Given the large (pre-registered) sample size, heterogeneity of participants in terms of literacy level, and carefully selected literacy measures, such an alternative interpretation is very unlikely to be correct. Another suggestion might be that we should have relied on corpus measures to estimate the distribution of PO/DO datives in spoken and written Dutch corpora. We conjecture that measuring PO/DO usage directly with a baseline as in the present study is the more reliable way of assessing influences of written language experience. Note that Dutch corpora (Haemers, 2012; Coleman, 2009) and experimental baseline measures (e.g., Bernolet et al., 2016) are often inconsistent with respect to PO/DO distributions (Hartsuiker, personal communication). This is likely because both Dutch corpora and experiment samples tend to be small in size and prone to bias. It is noteworthy, for instance, that the overall distribution of PO and DO datives in our data was quite different from previous Dutch priming studies (e.g., Bernolet et al., 2016). We found DO to be marginally the more frequent construction, in contrast to the strong PO bias previously reported by Bernolet and colleagues. This divergence likely reflects differences in sample size but also in the populations sampled: In our study, highly experienced literates, like Bernolet’s undergraduate participants, did demonstrate an overall PO bias, while less experienced literates tended to produce more DO datives. Moreover, unlike the ‘averages’ in corpus studies, we found that for the *same* people, literacy experience modulated usage (in line with Montag and MacDonald, 2015, who observed that the spoken production of experienced readers echoes the structural distributions of written language), but not priming of the PO/DO dative alternation in Dutch.



Finally, in the present study, two different measures were used to tap into explicit and implicit syntactic processes. We acknowledge that grammaticality judgment is more of an offline measure than syntactic priming (although both have implicit and explicit components). Thus, the two types of processes may have been indexed by their respective measures to different extents. One could argue that the observed contribution of literacy experience to explicit but not implicit syntactic processes was simply an artefact of the different measures used. Note, however, that this is not what we conclude here. Our conclusion is more specific, namely, that long-term written language experience affects syntactic awareness (as indexed by grammaticality judgment) and usage but not the syntactic priming of spoken sentences (as measured by the syntactic priming paradigm). It is noteworthy that we (Favier, Meyer, & Huettig, in press) have recently found that literacy affects (largely) implicit syntactic prediction in spoken sentence processing. It will be interesting for further research to explore why (largely) implicit syntactic prediction but not (largely) implicit syntactic priming is affected in spoken sentence processing.

### *Conclusion*

We conducted a large-scale correlational study with 161 adult native speakers of Dutch, to examine literacy experience as a predictor of syntactic processing in spoken language, while controlling for the contribution of non-verbal IQ, verbal working memory, and processing speed. As predicted, we found an effect of literacy on explicit syntactic awareness, specifically the detection of grammatical norm violations in spoken sentences. Literacy was also related to the usage of Dutch (PO vs. DO) datives but, contrary to our prediction, had no detectable effect on their implicit syntactic priming in our adult sample. Further research is needed to investigate why usage but not priming of syntactic structures is modulated by life-long syntactic experience in proficient language users.

## References

- Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review*, 23(3), 275–290.
- Barr, D.J., Levy, R., Scheepers, C., & Tilly, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278
- Bates, E., and Goodman, J. C. (1997). On the Inseparability of Grammar and the Lexicon: Evidence from Acquisition, Aphasia and Real-time Processing. *Language and Cognitive Processes*, 12:5-6, 507-584, DOI: 10.1080/016909697386628
- Bates, E., Dale, P., & Thal, D. (1995). Individual differences and their implications for theories of language development. In *Handbook of Child Language*, ed. P Fletcher, B MacWhinney, pp. 96–151. Oxford, UK: Basil Blackwell
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... & Bolker, M. B. (2015). Package ‘lme4’. *Convergence*, 12(1), 2.
- Bernolet, S., Colina, S., & Hartsuiker, R.J. (2016). The Persistence of Syntactic Priming Revisited. *Journal of Memory and Language*, 91, 99–116.
- Bernolet, S., & Hartsuiker, R.J., (2010) Does Verb Bias Modulate Syntactic Priming? *Cognition*, 114(3), 455–61.
- Bock, K. (1986). Syntactic persistence in language production. *Cognitive psychology*, 18(3), 355-387.
- Bock, K., Dell, G., Chang, F., & Onishi, K. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, 104(3), 437–458.
- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129(2), 177.
- Boersma, P., & Weenink, David (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.36, retrieved 11 November 2017 from <http://www.praat.org/>
- Branigan, H. P., & McLean, J. F. (2016). What children learn from adults’ utterances: An ephemeral lexical boost and persistent syntactic priming in adult–child dialogue. *Journal of Memory and Language*, 91, 141-157.
- Branigan, H. P., & Messenger, K. (2016). Consistent and cumulative effects of syntactic experience in children’s sentence production: Evidence for error-based implicit learning. *Cognition*, 157, 250–256.
- Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking Up for Vocabulary: Reading Skill Differences in Young Adults. *Journal of Learning Disabilities*, 40(3), 226–243.
- Brus, B. T., & Voeten, M. J. M. (1973). Een-minuut test. Vorm A en B. *Verantwoording en handleiding*.

- Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language*, 109, 104047.
- Brysbaert, M., Mandera, P., & Keuleers, E. (2013). Naambekendheid van fictieschrijvers in Vlaanderen: Resultaten van de auteursstest 2013.
- Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10(5), 425-455.
- Bybee, J. L. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language*, 82(4), 711-733.
- Chang, F. (2002). Symbolically speaking: A connectionist model of sentence production. *Cognitive science*, 26(5), 609-651.
- Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological review*, 113(2), 234.
- Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, 29(2), 217-230.
- Colleman, T. (2009). Verb Disposition in Argument Structure Alternations: A Corpus Study of the Dative Alternation in Dutch. *Language Sciences*, 31(5), 593-611.
- Colleman, T., & Bernolet, S. (2012). Alternation Biases in Corpora vs. Picture Description Experiments: DO-Biased and PD-Biased Verbs in the Dutch Dative Alternation. *Frequency Effects in Language Representation*, 2, 87-125.
- Crain-Thoreson, C., & Dale, P. S. (1992). Do early talkers become early readers? Linguistic precocity, preschool language, and emergent literacy. *Developmental Psychology*, 28(3), 421.
- Croft, W. (2000). *Explaining language change: An evolutionary approach*. London: Longmans.
- Cumming, G. (2013). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge.
- Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *American educator*, 22, 8-17.
- Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, 2(3), 219-253
- Dąbrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. *Cognition*, 178, 222-235.
- Demoulin, C., & Kolinsky, R. (2016). Does learning to read shape verbal working memory?. *Psychonomic Bulletin & Review*, 23(3), 703-722.

- Duñabeitia, J.A., Crepaldi, D., Meyer, A.S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2017). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*.
- Dunn, L. M., Dunn, L. M., & Schlichting, J. E. P. T. (2005). *Peabody picture vocabulary test-III-NL*. Harcourt Test Publishers.
- Favier, S. & Huettig, F. (in press). Are there core and peripheral syntactic structures? Experimental evidence from Dutch native speakers with varying literacy levels. *Lingua*.  
<https://doi.org/10.1016/j.lingua.2020.102991>
- Favier, S., Meyer, A. S., & Huettig, F. (in press). Literacy enhances syntactic prediction in spoken language processing. *Journal of Experimental Psychology: General*.
- Fillmore, C. (1988). The mechanisms of “Construction Grammar.” *In Proceedings of the fourteenth annual meeting of the Berkeley Linguistics Society*. Edited by Shelley Axmaker, Annie Jaisser, and Helen Singmaster, pp. 35–55. Berkeley, CA: Berkeley Linguistics Society.
- Haemers, E. (2012) *De Datiefalternantie in de Nederlandse Spreektaal: Een Constructioneel Corpus Onderzoek* (Unpublished Master’s thesis) Universiteit Gent, Belgium.
- Hartsuiker, R.J., Bernolet, S., Schoonbaert, S., Speybroeck, S., and Vanderelst, D. (2008) Syntactic Priming Persists While the Lexical Boost Decays: Evidence from Written and Spoken Dialogue. *Journal of Memory and Language*, 58(2), 214–38.\
- Hayiou-Thomas, M. E., Kovas, Y., Harlaar, N., Plomin, R., Bishop, D. V., & Dale, P. S. (2006). Common aetiology for diverse language skills in 4 1/2-year-old twins. *Journal of Child Language*, 33(2).
- Hintz, F., Meyer, A. S., & Huettig, F. (2017). Predictors of verb-mediated anticipatory eye movements in the visual world. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(9), 1352.
- Hubers, F., & de Hoop, H. (2013). The effect of prescriptivism on comparative markers in spoken Dutch. *Linguistics in the Netherlands*, 30(1), 89-101.
- Hubers, F., Snijders, T.M., & de Hoop, H. (2016). How the Brain Processes Violations of the Grammatical Norm: An fMRI Study. *Brain and Language*, 163, 22–31.
- Huettig, F., Audring, J., & Jackendoff, R. (under revision). Prediction as pre-activation: A linguistically and psychologically plausible theory of language processing.
- Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31(1), 80-93
- Huettig, F., & Mishra, R. K. (2014). How literacy acquisition affects the illiterate mind—a critical examination of theories and evidence. *Language and Linguistics Compass*, 8(10), 401-427.

- Huettig, F., & Pickering, M. J. (2019). Literacy advantages beyond reading: Prediction of spoken language. *Trends in Cognitive Sciences*, 23(6), 464-475.
- Jackendoff, R., (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press, USA.
- Jackendoff, R., & Audring, J. (2020). *The texture of the lexicon: relational morphology and the parallel architecture*. Oxford University Press, USA.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language*, 59(4), 434-446.
- Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, 127(1), 57-83.
- Kaschak, M. P., & Borreggine, K. L. (2008). Is long-term structural priming affected by patterns of experience with individual verbs? *Journal of Memory and Language*, 58(3), 862-878.
- Keuleers, E., Brysbaert, M. & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods*, 42(3), 643-650
- Kidd, E. (2012). Individual differences in syntactic priming in language acquisition. *Applied Psycholinguistics*, 33(2), 393-418.
- Kolinsky, R., & Morais, J. (2018). The worries of wearing literate glasses. *L'Annee psychologique*, 118(4), 321-347.
- Kroll, B. (1977). Combining ideas in written and spoken English: Discourse across time and space. In E. O. Keenan & T. L. Bennett (Eds.), *Discourse across Time and Space* (pp. 69-108). Los Angeles, CA: University of Southern California.
- Langacker, R. W. (1987) *Foundations of Cognitive Grammar* (vol. 1). Stanford: Stanford University Press.
- Langacker, R. W. (1988). A Usage-Based Model. *Topics in Cognitive Linguistics*, 127-161.
- Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *Journal of statistical software*, 25(1), 1-18.
- Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, 32(1), 69-92.
- Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, 91, 5-27.
- 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.

- Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 7(4), 323-331.
- OECD (2009). PISA 2009: Assessment framework – Key competencies in reading, mathematics and science. Paris: OECD.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and language*, 39(4), 633-651.
- R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Raven, J., & Court, J.H. (1993). Manual for Raven's progressive matrices and vocabulary scales.
- Ravid, D., & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of child language*, 29(2), 417-447.
- 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.
- Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental psychology*, 27(5), 763.
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize?. *Journal of Research in Personality*, 47(5), 609-612.
- Smalle, E. H. M., Szmalec, A., Bogaerts, L., Page, M. P. A., Narang, V., Misra, D., Araújo, S., Lohagan, N., Khan, O., Singh, A., Mishra, R. K., & Huettig, F. (2019). Literacy improves short-term serial recall of spoken verbal but not visuospatial items – Evidence from illiterate and literate adults. *Cognition*, 185, 144–150.
- Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 402-433.
- 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.
- 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.
- Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships Between Word Knowledge and Reading Comprehension in Third-Grade Children. *Scientific Studies of Reading*, 10(4), 381–398.
- Tarone, E., & Bigelow, M. (2005). Impact of literacy on oral language processing: Implications for second language acquisition research. *Annual Review of Applied Linguistics*, 25, 77-97.

- Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, 11(1/2), 61-82.
- Traxler, M. J., Tooley, K. M., & Pickering, M. J. (2014). Syntactic priming during sentence comprehension: Evidence for the lexical boost. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(4), 905-918.
- van Bergen, G., Stoop, W., Vogels, J., & De Hoop, H. (2011). Leve hun! Waarom hun nog steeds hun zeggen. *Nederlandse taalkunde*, 16(1), 2-29.
- van den Bos, K. P., Lutje Spelberg, H. C., Scheepstra, A. J. M., & De Vries, J. R. (1994). De Klepel. Vorm A en B. *Een test voor de leesvaardigheid van pseudowoorden [The Klepel, Form A and B. A test for readability of pseudo words]*. Nijmegen, The Netherlands: Berkhout.
- Wechsler, D. (1997). Wechsler Adult Intelligence Scale – Third Edition. San Antonio, TX. The Psychological Corporation.