



**HAL**  
open science

## Interaction between orthographic and graphomotor constraints in learning to write

Jérémy Danna, Marieke Longcamp, Ladislav Nalborczyk, Jean-Luc Velay,  
Claire Commengé, Marianne Jover

### ► To cite this version:

Jérémy Danna, Marieke Longcamp, Ladislav Nalborczyk, Jean-Luc Velay, Claire Commengé, et al.. Interaction between orthographic and graphomotor constraints in learning to write. *Learning and Instruction*, 2022, 80 (101622), pp.10.1016/j.learninstruc.2022.101622. 10.1016/j.learninstruc.2022.101622 . hal-03620980

**HAL Id: hal-03620980**

**<https://hal.science/hal-03620980>**

Submitted on 27 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Interaction between orthographic and graphomotor constraints in learning to write

Jérémy Danna 1\*, Marieke Longcamp 1, Ladislav Nalborczyk 1,2, Jean-Luc Velay 1, Claire Commengé <sup>2</sup>, Marianne Jover 3

1 Aix Marseille Univ, CNRS, LNC, Marseille, France.

2 Aix Marseille Univ, CNRS, LPC, Marseille, France.

3 Aix Marseille Univ, PSYCLE, Aix-en-Provence, France.

## \* Corresponding author:

Jérémy Danna

ORCID ID: 0000-0002-3596-880X

Laboratoire de Neurosciences Cognitives

Pôle 3C, Case C, 3 Place Victor Hugo,

13331 Marseille Cedex 03, France

+33 4 13 55 09 56

jeremy.danna@univ-amu.fr

**Acknowledgement:** The authors would like to thank David Wood (English at your Service, [www.eays.eu](http://www.eays.eu)) for revising the English of the paper.

**Funding:** This work, carried out within the Labex BLRI (ANR-11-LABX-0036) and the Institut Convergence ILCB (ANR-16-CONV-0002), has benefited from support from the French Government, managed by the French National Agency for Research (ANR), under the project title DYSTACMAP (ANR-13-APPR-0010).

**Author contributions:** Jérémy Danna: Conceptualization, Methodology, Data acquisition, Data analysis, Writing- Original draft preparation. Marieke Longcamp: Writing- Reviewing and Editing. Ladislav Nalborczyk: Data analysis, Reviewing. Jean-Luc Velay: Writing- Reviewing and

Editing. Claire Commengé: Conceptualization, Methodology, Data acquisition. Marianne Jover: Conceptualization, Methodology, Data acquisition, Writing- Reviewing and Editing.

### **Highlights**

- Handwriting of 1st graders is influenced by both orthographic and motor constraints
- Motor and spelling processes interact from the beginning of handwriting acquisition
- This calls into question the relevance of evaluating them separately

### **Abstract**

We investigated the combined effects of orthographic and graphomotor constraints as a function of handwriting proficiency in children. Twenty-four first graders, 20 third graders, and 21 fifth graders wrote single five-letter words in cursive writing on a sheet of paper affixed to a digitizing tablet. The words were chosen according to two orthographic constraints, namely their lexical frequency and the graphemic complexity of the last three letters, and one graphomotor constraint resulting from the motor difficulty of tracing the first letter. In addition to massive improvements of handwriting with grade, the results revealed, in the youngest group only, an interaction between first-letter difficulty and lexical frequency. This finding suggests that, before handwriting movement becomes automated, the cognitive resources needed for retrieving word spelling interferes with motor processing while writing a difficult letter. When students start learning to write, they are particularly sensitive to the combination of orthographic and graphomotor constraints.

**Keywords:** handwriting; acquisition; spelling; motor control; delayed copy

## 1. Introduction

Learning to write words constitutes a double challenge: students have to learn both how to spell the words and how to trace the letters. The aim of this study was to find out whether the effect of orthographic constraints associated with word spelling is combined with the effect of graphomotor constraints associated with letter formation during handwriting acquisition, in primary school students. The term “constraints”, which will be used throughout this article, is considered from a generic perspective: it refers to factors, or situations, which have an adverse effect on performance. Orthographic constraints modulate spelling processes, independently of the motor output (*e.g.*, typing, oral spelling, or handwriting). Here, two orthographic constraints were investigated: a) lexical frequency, and b) letter-sound relationships inside the word, as indexed by the graphemic structure. Graphomotor constraints refer to factors that influence the motor processes involved in controlling handwriting movements, independently of the linguistic content (*e.g.*, at high speed or with eyes closed). Here, we studied the graphomotor constraint arising from motor difficulty in letter formation.

In adults, the relationship between spelling and graphomotor processes has first been considered in van Galen’s model (1991), which proposes parallel processing between spelling (high-level) and graphomotor (low-level) modules in which “higher modules are further ahead to real-time output than lower modules” (p. 182). This vision has been refined in the *parallel and cascading* model of writing (Olive, 2014), where the dynamics of parallel processing is flexible according to the cognitive resources that have to be allocated to either spelling or graphomotor control. This idea has been tested by Lambert and colleagues (2011), who investigated handwriting in adults who were copying series of words that varied in lexical frequency and orthographic regularity (*e.g.*: FEMME (/fam/, woman is irregular in French because the spelling does not respect the most frequent phoneme-grapheme conversion corresponding to /a/ = A). From the combined analysis of eye and pen movements, the authors identified periods during which participants searched for visual information about the template (the word to be copied) while writing the previous word (meaning that they were processing spelling and graphomotor aspects of the task in a parallel and cascaded fashion). They observed that the extent of parallel processing depended on the orthographic constraints of the word being written: planning the next

word ahead while writing was more probable when the current word being written was frequent and regular, thus confirming the parallel and cascaded account.

The relationship between spelling and graphomotor processes has also been investigated through the study of the effects that orthographic constraints exert on handwriting movements (e.g., Delattre, Bonin & Barry, 2006; Kandel, Álvarez & Vallée, 2006; Kandel & Perret, 2015; Lambert, Alamargot, Larocque, & Caporossi, 2011; Roux, McKeef, Grosjacques, Afonso, & Kandel, 2013). The basic assumption of these studies is that factors that impact the retrieval of the word's spelling (e.g., lexical frequency or orthographic regularity) also impact the motor processes involved in the control of handwriting movements. Several indexes calculated from the time-course of the pen trajectory can be used to characterize the efficiency of handwriting behavior (duration, velocity, pauses...). For instance, in a single word copy task, Roux and colleagues (2013) found that letter duration was longer for irregular than for regular words. This convincingly shows that handwriting movement can be impacted by the orthographic constraints imposed by the words being written.

In children, it has been shown that spelling processing changes during the acquisition of handwriting. Humblot, Fayol and Lonchamp (1994) evaluated the performance of Grade 1 and Grade 2 students in a copy task where words varied in frequency and regularity. They concluded that the acquisition of handwriting starts with a letter-by-letter planning, then a syllable-by-syllable planning, and ends with a whole-word planning when the word is frequent. In order to assess whether this evolution affected graphomotor processes, Kandel & Perret (2015) investigated the effects of word frequency and orthographic regularity on word production in students from Grade 3 to Grade 5. They observed that both variables affected handwriting execution in all groups in such a way that the writing of words with high orthographic constraints (irregular or infrequent) took longer than that of words with low orthographic constraints (frequent or regular). To go a step further, Afonso, Coalla, González-Martín and Cuetos (2018) studied the effect of word frequency on handwriting duration in both a spelling-to-dictation task and a copy task in Grade 2, 4, and 6 students. Interestingly, the authors observed that the effect of frequency decreased with the mastery of handwriting: it was significant in grade 2, marginally significant in Grade 4 and

disappeared in Grade 6. This suggests that although students' handwriting is generally sensitive to orthographic constraints, improvement of handwriting with development and practice allows a more independent management of spelling and motor processes.

Another important orthographic constraint seems also to draw on information about letter-sound relationships as indexed by the graphemic structure (e.g., Kandel & Spinelli, 2010; Spinelli, Kandel, Gueressimovitch, & Ferrand, 2012). In French, at least 34 complex graphemes, composed of more than one letter, have been identified (Catach, 1995). Kandel & Spinelli (2010) compared the handwriting of words containing simple graphemes (e.g., *E* in *SET*) to words containing complex graphemes (e.g., *E* in *SEA* for a two-letter grapheme). They observed that both the target letter (e.g., *E*) and the preceding letter (e.g., *S*) took more time to be written in the case of complex graphemes. They concluded that the processing of graphemes starts before the actual production of the grapheme and remains activated in parallel with motor processes during the execution of handwriting. This finding was extended to first graders who had not yet mastered handwriting (Kandel, Soler, Valdois & Gros, 2006).

From a motor point of view, writing by hand requires a sophisticated coordination of the muscles and joints recruited in order to form letters consistently as quickly as possible. Graphomotor skills are acquired through one of the longest and most tedious motor learning processes known to humans. In early stages of learning to write, students program their movements stroke-by-stroke until the letters are learned and stored in long-term memory in motor programs (S raphin Thibon, Gerber & Kandel, 2018). This leads to the question of what motor constraints are, and how they influence performance during the acquisition of handwriting. This question has been explored using tasks in which students had to change their handwriting size, velocity, or script. In 2008, Chartrel & Vinter observed that issuing instructions to write between horizontal lines and to increase speed facilitated the formation of associations between the strokes of a letter and thus benefited students of 5-7 years of age learning to write. Between 5 and 12 years, Bo, Barta, Ferencak, Comstock, Riley and Krueger (2014) showed that printed letters tended to be written with a higher consistency than cursive letters. In both dyslexic and typically developing children, Pagliarini et al. (2015) showed that writing a word with cursive letters turned out to be generally more challenging than writing it with upper-case letters. The authors suggested that the late introduction of

cursive training in the Italian educational system might explain this result. Furthermore, typically developing children, when asked to write big letters or to write quickly, modulated their handwriting movements to maintain the absolute and relative duration of letters (isochrony and homothety principles) whereas children with dyslexia failed to do so (Pagliarini et al., 2015).

Graphomotor constraints also result from the motor difficulty arising from the formation of a letter. Jolly, Huron and Gentaz (2014) analyzed the handwriting of the 26 letters of the Latin alphabet in students followed from preschool, to mid- and end- of the first Grade and second Grade. The results showed a wide variability across letters regarding the students' motor performance and progress. This variability may be explained by the biomechanical constraints of the effector and coordination dynamics (e.g., Kelso, 1995) that determine a preferential direction of rotation (Meulenbroek, Vinter & Mounoud, 1993), preferential orientations (Dounskaia, Van Gemmert and Stelmach, 2000), and preferential curvatures (Athènes et al. 2004) in the formation of the written trace. Finally, the rotation and pointing movements (pen lifts) involved in letter formation are also of interest. Recently, Seraphin-Thibon et al. (2019) studied the developmental trajectory of students' graphomotor skills as a function of the characteristics of upper-case letters, involving either pointing or rotation movements. They observed that letters with rotations were produced with more velocity peaks (an index of movement dysfluency) and a longer duration than letters with straight strokes and pointing movements.

To sum up, the reported literature indicates that 1) the processing of the orthographic content of words has an impact on the production of writing in adults and children, and 2) strong graphomotor constraints related to writing conditions and letter complexity also influence handwriting movements in children. But to our knowledge, the combined effect of orthographic and graphomotor constraints during handwriting acquisition has not yet been investigated. One study explored this question in adults: Sausset, Lambert, Olive and Larocque (2012) manipulated graphomotor constraints through letter case (upper- / lower-case), visual control (with / without), and letter size, and also an orthographic constraint through the number of syllables in the words, considered as processing units involved in spelling processing (Kandel, Peereman, Grosjacques, & Fayol, 2011). They observed that when graphomotor constraints were high, the writers took longer pauses between syllables,

indicating that each syllable was processed directly before its actual production. When they were low, the syllable processing took place before the onset of writing. This demonstrates that the dynamics of spelling and motor processes vary as a function of both orthographic and graphomotor constraints in adults.

The present study aimed to investigate the combined effect of orthographic and graphomotor constraints on handwriting in students who are learning to write. We asked students at different levels of handwriting proficiency (in Grades 1, 3, and 5) to write single words, without the presence of the template (the word to be copied). We manipulated both the graphomotor constraints associated with the difficulty of tracing the first letter, and two orthographic constraints: the word's lexical frequency, and the graphemic complexity associated with the last letters. We tested the interaction between these constraints in the different level groups. Knowing that both orthographic (Afonso et al., 2018) and graphomotor (S raphin Thibon et al., 2018) constraints are larger in students who do not yet master handwriting, we hypothesized that the interaction between the two types of constraints would be larger in younger than in older students. We assume that the younger students will produce a slower and less fluent movement when tracing a difficult letter in a complex or infrequent word than when only one of the two types of constraints is present (tracing an easy letter in a complex or infrequent word or tracing a difficult letter in a simple or frequent word). In order to quantify and compare the graphomotor constraint, we conducted a pre-experiment in which Grade 1 students were asked to copy bigrams in cursive. The aim was to select two letters, respectively representative of high and low graphomotor constraint for students who are at the beginning of cursive acquisition.

## **2. Pre-experiment: Characterizing graphomotor constraints**

### **2.1. Methods**

#### **2.1.1. Participants**

Nineteen students in Grade 1 volunteered for the experiment (6;8 years; 8 girls; 3 left-handers for handwriting). Students who were receiving intervention for reading and/or writing difficulties were not included, neither were students who had resumed or skipped a school grade. The students had normal or corrected-to-normal vision. This study was



conducted between March and April and in accordance with local norms and guidelines for the protection of human subjects. This research was approved by the local educational authorities, the school directors, and the teachers. Furthermore, parents signed an informed consent sheet prior to the experiment.

### **2.1.2. Task**

The experiment was conducted individually at school, in a quiet classroom. After a quick familiarization task consisting of writing their first name, students were asked to write isolated bigrams in cursive handwriting on a white sheet of paper affixed to a digitizing tablet (Wacom Intuos 4L, sampling frequency 100 Hz), using an inking pen. Each bigram was presented on a small white paper (5 x 7 cm label) in lowercase cursive letters (Cursive standard font, font-size: 44) and the template remained present until the student had finished writing the bigram. Students were asked to write at their usual speed. No time limit was applied during the task and the presentation order was randomized between students. For data acquisition, we used JAVA software developed in the laboratory.

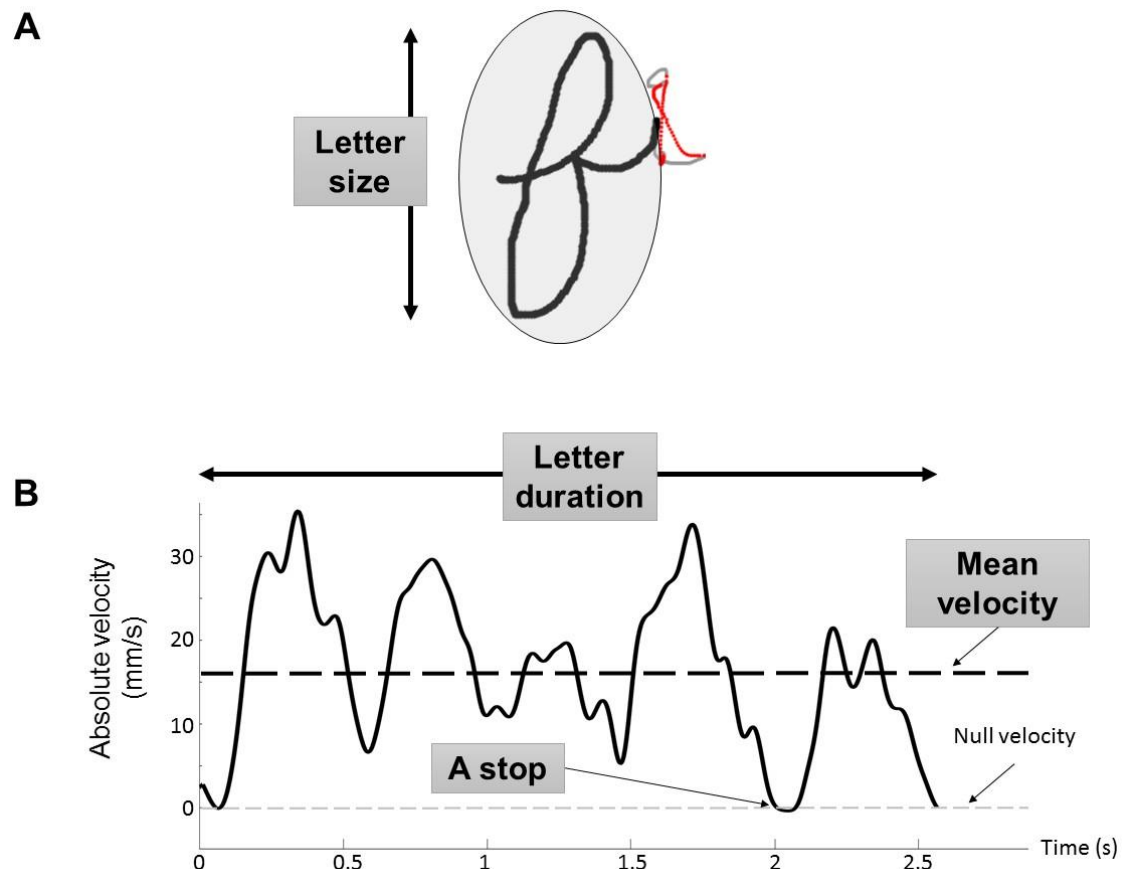
### **2.1.3. Stimuli**

Eleven consonant-vowel bigrams were selected for the task. Only the consonant varied between the bigrams, the vowel was identical for all conditions (letter 'i') to avoid possible motor anticipation on the first letter, a known phenomenon in cursive handwriting (Orliaguet, Kandel & Boë, 1997). The selected consonants were: 'b', 'c', 'f', 'g', 'l', 'm', 'p', 'r', 's', 't', and 'v'. Less frequent letters as 'h', 'j', 'k', 'q', 'w', 'x', and 'z' were not considered because some first graders did not know how to trace them.

### **2.1.4. Data analysis**

The first letter of each bigram was analyzed. In order to identify a common endpoint, the ascendant stroke of the second letter 'i' was included. From the recorded (x, y) coordinates and pen pressure, the first in-air stroke and the strokes produced after the

maxima of the letter *i* were deleted. The segmentation was conducted with Matlab<sup>®</sup>. In order to reduce the statistical risk of Type I errors, four complementary variables were considered: the duration, the mean velocity, the number of stops, and the letter size. These variables respectively describe the temporal (duration), kinematic (mean velocity and number of stops) and spatial (size) content of handwriting performance (see figure 1):



**Figure 1.** Schematic illustration of data processing. (A) Example of the bigram “fi” written by a child. In grey: the whole word recorded. In red: the in-air movements (pen lifts). In black and surrounded: the first letter segmented for further analyses. (B) Absolute velocity profile as a function of time (filtered with a cutoff frequency of 10 Hz) for illustrating the temporal and kinematic variables.

- The duration corresponds to the time elapsed between the beginning and end of letter production, including in-air movement if the writer lifts the pen.
- The mean velocity is the mean of absolute velocity from the time the pen was first in contact with the tablet until the letter was completed. In-air movements were not considered for computing mean velocity.

- The number of stops corresponds to the total number of periods that the pen was in contact with the paper but did not move for at least 35 milliseconds.
- The letter size corresponds to the distance between the minimal and maximal positions of the letter in the y-axis.

A cluster analysis was applied on the complete data set in order to identify the difficulty level of the letters. All variables were converted to z-scores (i.e. the variables were centered and reduced to a range between 0 and 1) then analyzed by interactive partitioning (K-means), minimizing the within-cluster variability and maximizing the between-cluster variability. This cluster analysis allowed us to separate letters into two categories: easy to write letters (cluster 1) from difficult to write letters (cluster 2).

## 2.2. Results

The cluster analysis on z-scores confirmed the presence of two clusters with an inter-cluster Euclidian distance of 0.74. Cluster 1 included seven letters: *c*, *l*, *m*, *r*, *s*, *t*, and *v*. Cluster 2 included four letters: *b*, *f*, *g*, and *p*. The mean performance for each letter is reported in table 1. The Cluster 1 comprised ‘easy’ letters that, on the whole, were small, written with short duration, low velocity, and few stops. The Cluster 2 comprised ‘difficult’ letters which, on the whole, were bigger, written with a longer duration, a high velocity, and more stops.

	<i>c</i>	<i>l</i>	<i>m</i>	<i>r</i>	<i>s</i>	<i>t</i>	<i>v</i>	<i>b</i>	<i>f</i>	<i>g</i>	<i>p</i>	<i>Mean</i>
<b>Duration</b>	1.7 (<0.1)	1.9 (<0.1)	<b>3.0</b> (<0.1)	2.3 (<0.1)	2.3 (<0.1)	<b>3.0</b> (<0.1)	2.7 (0.1)	<b>3.5</b> (0.1)	<b>3.6</b> (0.1)	<b>3.6</b> (0.1)	<b>3.0</b> (0.1)	<b>2.8</b>
<b>Velocity</b>	<b>12.8</b> (0.3)	17.0 (0.4)	<b>12.1</b> (0.3)	<b>10.4</b> (0.2)	<b>12.5</b> (0.3)	16.5 (0.5)	<b>10.1</b> (0.2)	<b>13.2</b> (0.3)	15.7 (0.4)	14.4 (0.3)	18.1 (0.5)	<b>13.9</b>
<b>Stops</b>	2.0 (0.1)	0.9 (0.1)	<b>3.8</b> (0.1)	<b>4.2</b> (0.1)	<b>3.4</b> (0.1)	3.4 (0.1)	3.1 (0.1)	<b>4.5</b> (0.1)	<b>4.3</b> (0.2)	2.7 (0.1)	<b>3.9</b> (0.2)	<b>3.3</b>
<b>Size</b>	5.7 (0.1)	<b>12.1</b> (0.3)	5.0 (0.1)	5.4 (0.1)	6.0 (0.1)	8.5 (0.2)	5.3 (0.1)	<b>12.7</b> (0.3)	<b>17.6</b> (0.5)	<b>12.95</b> (0.3)	<b>13.6</b> (0.4)	<b>9.5</b>

Table 1: Mean (standard error) performance for each letter. In bold, value above (below for the velocity) the mean value (on the right). Cluster 1: Black letters on white background. Cluster 2: White letters on black background.

### 2.3. Interim discussion

This pre-experiment confirmed that *b*, *f*, *g*, and *p* are more difficult to trace than the other tested consonants for 1<sup>st</sup> grade students. Letters classified as more difficult have the common characteristic of being composed of several strokes that can be large and that demand significant rotational changes between the strokes. Overall, our results extend this finding to lowercase cursive letters and confirms that the graphomotor constraint of a letter depends on how many strokes it consists of, the difficulty of tracing each stroke, how easy transition between them is (*i.e.*, with turning points – Chang & Yu, 2010). Although some letters took longer to be produced, the writing velocity did not appear a relevant index of automation. This finding is consistent with previous studies in which the authors reported that children with dysgraphia did not write more slowly than proficient children because of their tendency to write larger (e.g., Prunty, Barnett, Wilmut, and Plumb, 2013). In the present study, the more difficult letters contained larger strokes than the easy letters, compensating for any tendency to write them more slowly.

For the next experiment, we finally selected the letters *f* and *t* as representative of difficult and simple letter, respectively. This choice was based on the conclusion of a work-meeting with two French teachers, taking account our need to build a strictly controlled material with both high and low motor constraints, comprising 5-letter words with an 'i' in second position, having a low and high frequency, and a simple vs. complex grapheme in the 3 last letters. Out of curiosity, we checked the developmental values published by Jolly, Huron and Gentaz (2014) in their supplementary content for these two letters. These authors measured several temporal, spatial and kinematic variables of letter production in children between preschool, mid-year and end-year of 1st and 2nd Grade. The letter *t* and *f* demonstrated different developmental patterns, especially between 1st grade and 2nd Grade mid-year: at 1st grade mid-year, the number of slow movements and static moments was higher for the *f* than for the *t*, although it was comparable at 2nd grade mid-year. The reduction of slow movement and static moment represented respectively 39% and 53% for the *f* and 28% and 34% for the *t*, confirming more progress (*i.e.* a higher difficulty) for the former than for the later one.

### **3. Main experiment: Interaction between graphomotor and orthographic constraints**

#### **3.1. Methods**

##### **3.1.1. Participants**

Sixty-five French students from two elementary schools volunteered for the experiment, including 24 students in Grade 1 (6;7 years; 13 girls; 3 left-handers for handwriting), 20 in Grade 3 (8;6 years; 11 girls; 2 left-handers for handwriting), and 21 in Grade 5 (10;7 years; 16 girls; 3 left-handers for handwriting). Student who were receiving intervention for reading and/or writing difficulties were not included, neither were students who have repeated or have skipped a school year. The students had normal or corrected-to-normal vision. This study was conducted between March and April and in accordance with local norms and guidelines for the protection of human subjects. This research was approved by the local educational authorities, the school directors, and the teachers. Furthermore, parents signed an informed consent sheet prior to the experiment.

##### **3.1.2. Task**

The experiment was individually conducted at school, in a quiet classroom. After a quick familiarization task consisting of writing their first name, the child was asked to write several single words in cursive on a sheet of paper affixed to a digitizing tablet (Wacom Intuos 4L, sampling frequency 100 Hz), using an inking pen. Each word was written twice to avoid any incorrect production (with errors or unusual performance). The template (word to be copied) was first displayed (in lowercase cursive letters, using the Cursive standard font, font-size: 44) on a label (5 x 7 cm) that the experimenter showed and read aloud to the child. Then, the template was hidden before the child started to write. No time limit was applied during the presentation of the template: the experimenter waited for the child to indicate that the template could be removed. Before the second trial, the previous performance was masked but the child could ask to see and hear the template again. The presentation order was randomized between students. They were asked to write at their usual speed. A black line had been traced on each white sheet to help to write horizontally. For data acquisition, we used JAVA software developed in the laboratory.

### 3.1.3. Stimuli and procedure

A total of eight 5-letter words were selected for the task, including four with the first bigram ‘*fi*’ and four with the first bigram ‘*ti*’. These words were selected on the basis of A) the graphomotor constraint arising from the difficulty of tracing the first letter as characterized in the pre-experiment (easy vs. difficult), B) their lexical frequency from the French database Lexique 3 (high-frequency words, i.e., > 50 occurrences per million words, vs. low-frequency words, i.e., < 2 per million; New 2006; New, Pallier, Brysbaert, & Ferrand, 2004) and from the French children database Manulex (CP-CM2 U > 50 for high-frequency words and CP-CM2 U < 1 for low-frequency words; Lété, Sprenger-Charolles & Colé, 2004), and C) the graphemic complexity of the last three letters (simple vs. complex grapheme) also from the Lexique 3 database. As in the pre-experiment, the second letter was identical for all conditions to avoid possible motor anticipation effect on the first letter. The word length was also identical to avoid possible word length effects on working memory (e.g., Baddeley, Thomson & Buchanan 1975). Finally, the orthographic regularity was also checked, and only regular words were chosen. The stimuli are reported in table 2.

	finir (end)	fille (girl)	firme (firm)	fioul (fuel oil)	titre (title)	tirer (pull)	tibia (tibia)	titan (giant)
<b>First-letter graphomotor constraint</b>	Difficult – “ <i>f</i> ”				Easy – “ <i>t</i> ”			
<b>Word frequency</b>	HF		LF		HF		LF	
<b>Graphemic complexity</b>	simple	complex	simple	complex	simple	complex	simple	complex

Table 2: Stimuli according to the first-letter graphomotor constraint, the word frequency (HF: high-frequency or LF: low-frequency), and the graphemic complexity of the three last letters.

### 3.1.4. Data analysis

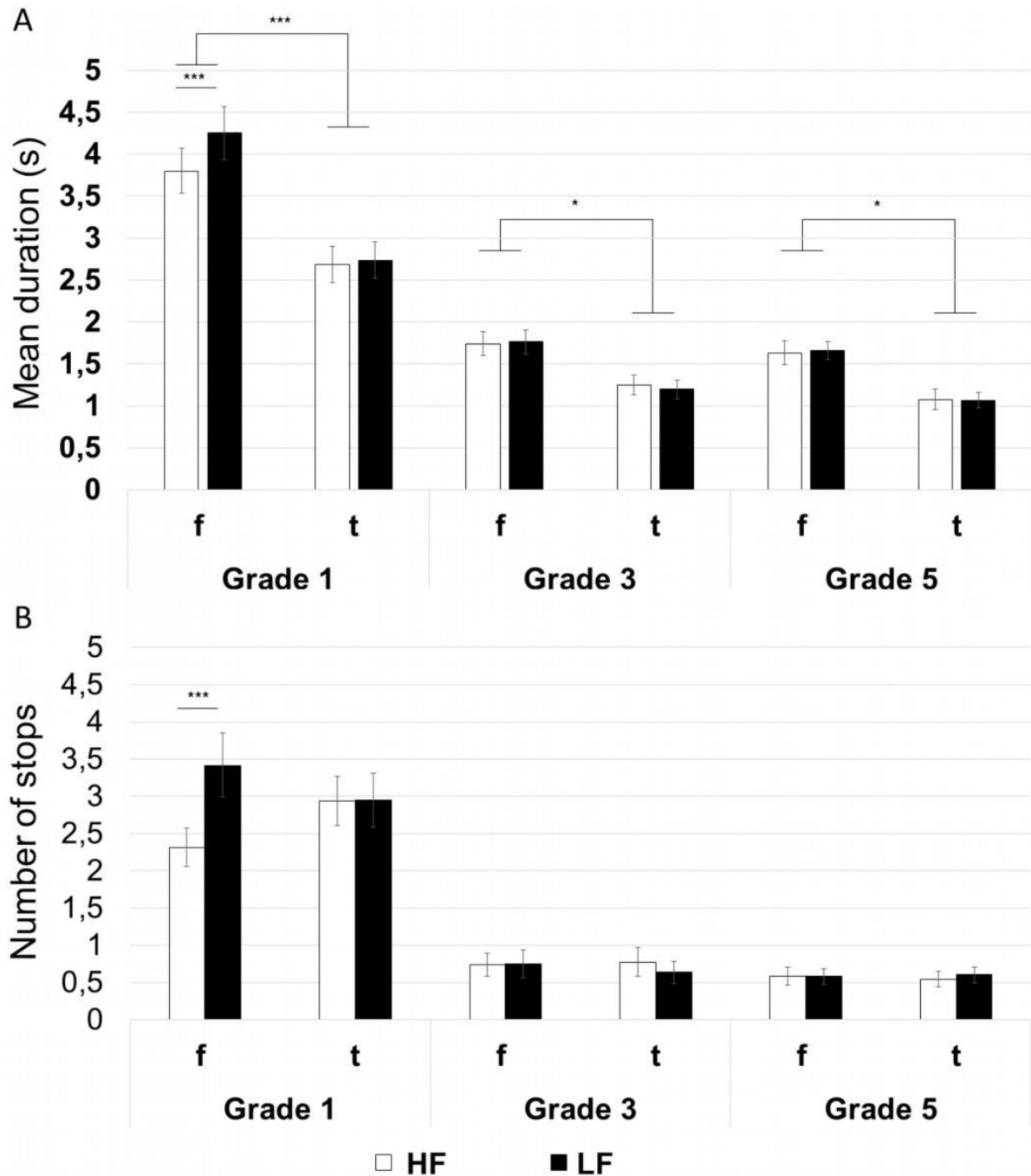
The first letter of each bigram was analyzed in the same way as in the pre-experiment. In order to identify a common endpoint, the ascendant stroke of the second letter ‘*i*’ was also included. Incorrect words were not analyzed. From the 1040 trials collected (65 students \* 8 items \* 2 trials), 13 trials were not considered for analysis. The percentage of

trials with errors was about 1.0 % in 1st grade, 2.5% in 3d grade, 0.3 % in 5th grade. The same temporal (duration), kinematic (mean velocity and number of stops) and spatial (size) variables were considered in this experiment. A table with correlations between the variables, within grades, is presented in Appendix. For these variables, both frequentist and Bayesian statistics were conducted. For the sake of clarity, we report the results of the frequentist statistics (analyses of variance (ANOVAs) complemented with Bonferroni post-hoc tests when necessary) in the following section. The complete description of the Bayesian analyses and their results are reported in supplementary materials. For the frequentist analysis, we averaged the performance of the two trials when they were correct. To make sure that there was no difference related to the inclusion of one or both trials, we inspected the performance of the students who did only one trial with those from the two averaged trials and found no outliers from performance based on a single trial.

## 3.2. Results

### 3.2.1. Duration

The ANOVA revealed a Graphomotor constraint by Grade interaction  $F(2, 62) = 7.84, p < .001, \eta^2_p = 0.20$ , and a Graphomotor constraint by Frequency interaction,  $F(1, 62) = 5.23, p < .05, \eta^2_p = 0.08$ . As can be seen in Figure 2A, the Bonferroni post hoc tests showed that students in Grade 1 spent more time writing the difficult letter *f* when the word was infrequent than when it was frequent ( $p < 0.01$ ). However, this effect of frequency was not observed for the easy letter *t*, nor for the two other groups ( $p = 1$ ). The ANOVA also revealed a main effect of Grade on mean duration,  $F(2, 62) = 63.20, p < .001, \eta^2_p = 0.67$ . The post-hoc analysis revealed that 1<sup>st</sup> graders spent more time writing the first letter than 3<sup>rd</sup> and 5<sup>th</sup> graders ( $p < .001$  for both comparisons) whereas 3<sup>rd</sup> and 5<sup>th</sup> graders did not significantly differ from each other. Finally, the analysis also revealed a main effect of Graphomotor constraint,  $F(1, 62) = 74.24, p < .001, \eta^2_p = 0.54$ , confirming that the letter *f* required more time to be written than the letter *t*. The analysis evidenced neither main effect nor interaction involving graphemic complexity. Interitem correlations between words with simple vs. complex grapheme were between 0.8 and 0.9 ( $R^2$  *finir-fille* = 0.87;  $R^2$  *firme-fioul* = 0.89;  $R^2$  *titre-tirer* = 0.81;  $R^2$  *tibia-titan* = 0.80).



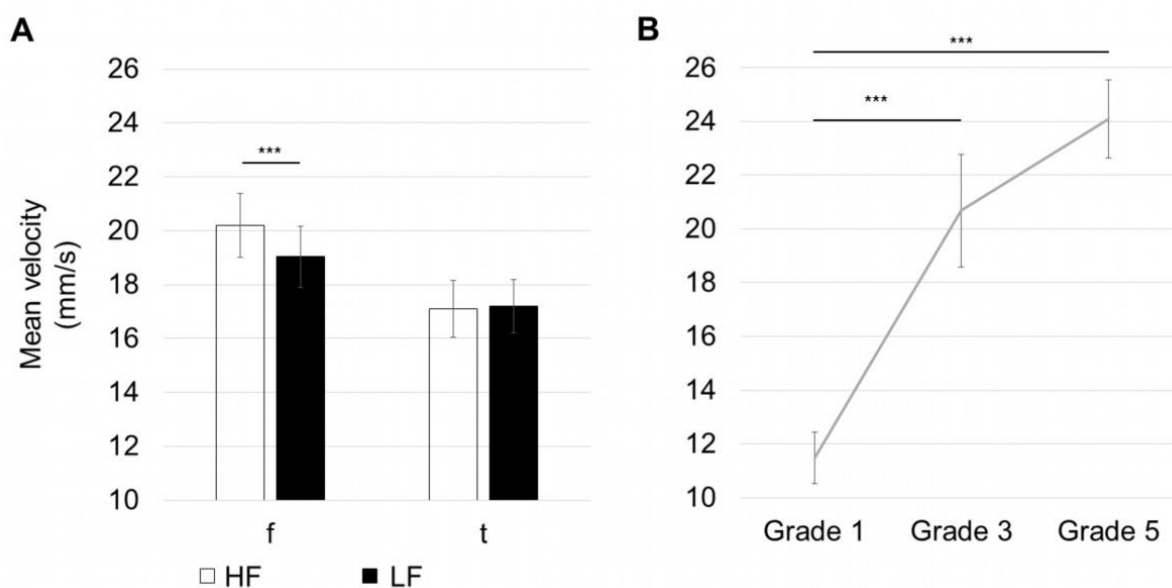
**Figure 2.** (A) Mean duration, and (B) number of stops according to the Grade level (Grade 1, 3, and 5), the word frequency (LF: low-frequency vs. HF: high-frequency), and the first-letter graphomotor constraint (f vs. t). Error bars correspond to standard error of the mean.

### 3.2.2. Mean velocity

The analysis revealed a Graphomotor constraint by Frequency interaction,  $F(1, 62) = 4.98$  ;  $p < .05$ .  $\eta^2_p = 0.07$ , and the post-hoc tests showed that the frequency effect was present in writing the letter *f* ( $p < 0.05$ ) but not in writing the letter *t* ( $p = .81$ , see Figure 3A).



In other words, the first letter of low-frequency words was written more slowly when it imposed stronger graphomotor constraints. The analysis also revealed a main effect of Grade,  $F(2, 62) = 26.16, p < .001, \eta^2_p = 0.46$ , and post-hoc tests confirmed that 1<sup>st</sup> graders wrote more slowly than 3<sup>rd</sup> and 5<sup>th</sup> graders ( $p < .001$  for both comparisons) whereas 3<sup>rd</sup> and 5<sup>th</sup> graders did not significantly differ from each other (see Figure 3B). The ANOVA also showed a main effect of Graphomotor constraint,  $F(1, 62) = 31.21, p < .001, \eta^2_p = 0.33$ : The letter *f* was written faster than the letter *t*. Finally, the analysis evidenced neither main effect of graphemic complexity nor interaction between graphemic complexity and the other factors. Interitem correlations between words with simple vs. complex grapheme were between 0.7 and 0.9 ( $R^2$  *finir-fille* = 0.77;  $R^2$  *firme-fioul* = 0.89;  $R^2$  *titre-tirer* = 0.81;  $R^2$  *tibia-titan* = 0.90).



**Figure 3.** Mean velocity, according to (A) the word frequency (LF: low-frequency vs. HF: high-frequency) and the first-letter graphomotor constraint (f vs. t), and (B) the Grade level (Grade 1, 3, and 5). Error bars correspond to standard error of the mean.

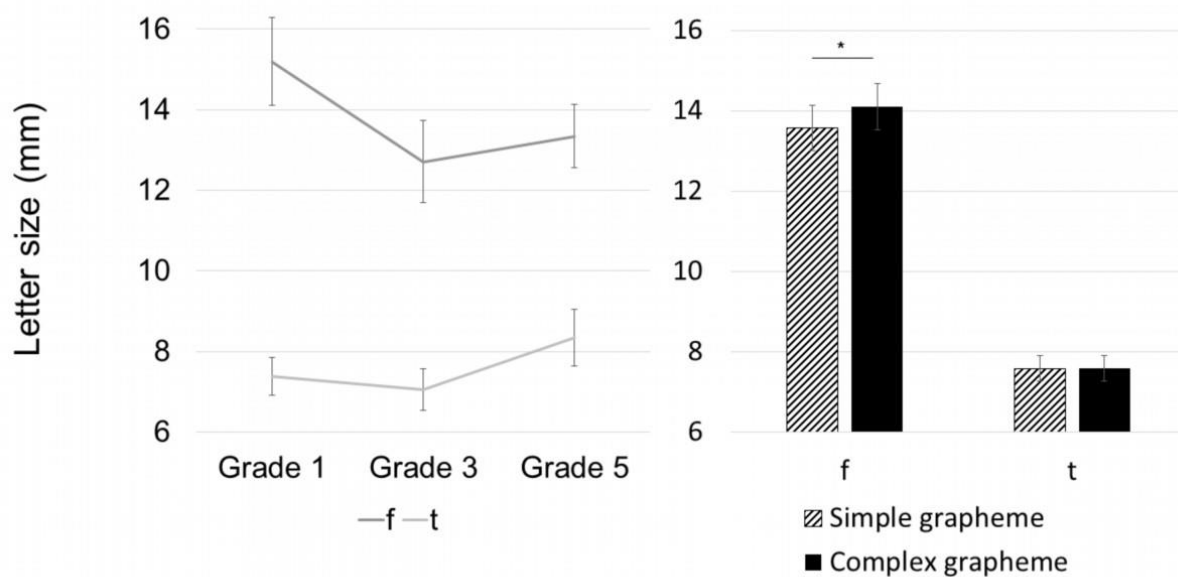
### 3.2.3. Number of stops

The ANOVA revealed a Frequency by Grade interaction,  $F(2, 62) = 3.25, p < .05, \eta^2_p = 0.09$ , a Graphomotor constraint by Frequency interaction,  $F(1, 62) = 4.82, p < .05, \eta^2_p = .07$ , and the Graphomotor constraint by Frequency by Grade double interaction,  $F(2, 62) = 4.09$ ,

$p < .05$ .  $\eta^2_p = 0.12$ . As can be seen in Figure 2B, the Bonferroni's post-hoc tests showed that the Frequency effect was only present in 1<sup>st</sup> graders writing the letter *f* ( $p < .001$ ). The analysis also revealed a main effect of Grade,  $F(2, 62) = 47.70$ ,  $p < .001$ .  $\eta^2_p = 0.61$ : 1<sup>st</sup> graders wrote with more stops than 3<sup>rd</sup> and 5<sup>th</sup> graders ( $p < 0.001$  for both comparisons). Here again, 3<sup>rd</sup> and 5<sup>th</sup> graders did not significantly differ from each other and the analysis evidenced neither main effect of graphemic complexity nor interaction between graphemic complexity and the other factors. Interitem correlations between words with simple vs. complex grapheme were between 0.5 and 0.8 ( $R^2$  *finir-fille* = 0.52;  $R^2$  *firme-fioul* = 0.71;  $R^2$  *titre-tirer* = 0.77;  $R^2$  *tibia-titan* = 0.66).

### 3.2.4. Letter size

The analysis revealed a Graphomotor constraint by Grade interaction,  $F(2, 62) = 6.27$ ,  $p < .01$ .  $\eta^2_p = 0.17$ . As can be seen in figure 4A, the size of letter *f* tended to decrease between the 1<sup>st</sup> grade and the 3<sup>rd</sup> but the post-hoc tests did not confirm these tendencies ( $p > 0.10$ ). As expected, the analysis revealed a main effect of Graphomotor constraint,  $F(1, 62) = 313.07$ ,  $p < .001$ .  $\eta^2_p = 0.83$ : the written letter *f* was larger than the letter *t*. Finally, the analysis revealed a main effect of Graphemic Complexity  $F(1, 62) = 4.36$ ,  $p < .05$ ,  $\eta^2_p = 0.06$ , and the Graphomotor constraint by Graphemic Complexity interaction tended to be significant,  $F(1, 62) = 3.88$ ,  $p = .053$ .  $\eta^2_p = 0.05$ . As can be seen in figure 4B, the Bonferroni post-hoc analysis revealed that the size of *f* increased with graphemic complexity ( $p < .05$ ) whereas the size of *t* did not. Interitem correlations between low-frequency and high-frequency words were between 0.7 and 0.9 ( $R^2$  *finir-firme* = 0.79;  $R^2$  *fille-fioul* = 0.90;  $R^2$  *titre-tibia* = 0.81;  $R^2$  *tirer-titan* = 0.89).



**Figure 4.** Mean size of letters f and t, according to the grade level (left) and according to graphemic complexity (right).

#### 4. Discussion

This study aimed to evaluate the combined effects of orthographic and graphomotor constraints on handwriting in beginning, intermediate and advanced writers. From the pre-experiment, we selected two letters with different graphomotor constraint: an easy letter 't' and a difficult letter 'f'. As expected, in the main experiment, the letter 'f' was more difficult to write than the letter 't' in all groups. The longer duration, as well as the higher mean velocity are obviously related to the size of the 'f' compared to the 't'. At first glance, a longer duration seems to be contradictory with a higher velocity. In fact, the principle of isochrony states a proportional relationship between movement speed and trajectory length, so the total duration of execution remains approximately constant (Viviani & Terzuolo 1982). In the present case, because of a greater difficulty of 'f' compared to 't', the increase in velocity is not proportional to the increase in trajectory length. This explains why the students took more time to write the letter 'f'.

The results showed an effect of grade on all variables. According to the literature (Meulenbroek and Van Galen, 1988; Mojet, 1991; Zesiger, Mounoud & Hauert, 1993), the greatest differences were observed between Grade 1 and Grade 3. The letter size evolved

differently for each letter: it tended to decrease for the 'f' whereas it remained relatively stable for the 't'. It is well established that the tendency to write larger is one of the criteria that characterize non-mastered handwriting (e.g., Hamstra-Bletze & Blöte, 1993). Taken together, our analyses confirm that the letter 'f' was less automated than the letter 't' for Grade 1 students only. Knowing that automation refers to the fact that writing is produced with minimal attentional involvement (Tucha et al., 2008), more attentional resources are allocated to the motor processes when writing the difficult letter than the easy one. If children pay too much attention to handwriting movements, they may have some difficulties in the allocation of cognitive resources to other processes (Berninger & Swanson, 1994; Jones & Christensen, 1999).

In this study, the orthographic constraints were generated by the lexical frequency and the graphemic complexity of the word. Lexical frequency had no direct main effect on any of the variables, but it interacted with letter difficulty and decreased with increasing school grade. This interaction occurred on the number of stops (observed in both frequentist and Bayesian analyses) and, to a lesser extent, on the duration and velocity (observed in the frequentist analysis only). The frequency effect was significant for Grade 1 students but not in the other two groups, and it was observed for letter 'f' but not for letter 't'. In other words, the effect of orthographic constraints was only observed for beginners in the production of the less automated letter. This finding supports the idea of an interaction between motor and spelling processes from the first stage of handwriting acquisition, which later disappears or changes through automation. Previous studies have revealed a frequency effect on handwriting movements in a word copy task in adults (Delattre et al., 2006; Lambert et al., 2011) and in children (Alfonso et al., 2018; Kandel & Perret 2015). Alfonso et al. (2018) studied words written in Grade 2, 4 and 6, and confirmed that the automation of writing skills through learning reduces the effect of word frequency that is significant only in Grade 2. We also observed an effect of frequency for the youngest students, in Grade 1 only. Furthermore, Kandel & Perret (2015) observed, in all groups of children tested (8-10 years), that the effect of frequency was not significant at the beginning but in the middle of the words. In the present study, we did observe a frequency effect when producing the difficult letter. This difference could result, in our study, from the absence of the template, which was removed before the child started to write. We suppose that for frequent words,

students in Grade 1 have already memorized the correct spelling in orthographic long-term memory. Knowing that the acquisition of handwriting starts with a letter-by-letter planning (Humblot et al., 1994), beginners allocate their attention to the motor processes if the letter is non-automated. For low-frequency words, removing the template led the students to retain in orthographic working memory the entire word they had to write (Kandel & Valdois, 2005). According to Olive and Kellogg (2002) who proposed that central and peripheral processes in written production compete for a common cognitive resource (see also and Sausset et al., 2012), our findings suggest that when handwriting movements are not yet automated, the significant cognitive resources required for maintaining the spelling of low-frequency words in working memory, interfere with the motor processing required for writing a non-automated difficult letter.

When focusing on the interaction between graphomotor and orthographic constraints in the younger group of children, the Bayesian analyses (supplementary materials) revealed that the strongest evidence for this interaction effect was in the number of stops. In other words, Grade 1 students discretized their movement. This mode of production was previously reported by Van Mier (2006) in a drawing task. She observed that young children made continuous movements (zigzag) in a discrete manner, producing short strokes separated by stops. She interpreted this finding as a less proficient motor planning in younger children whereas older children program and plan the upcoming segment during the production of the previous segment (see also Badan, Hauert and Mounoud, 2000). This finding is compatible with the results of Sausset et al. (2012) who evidenced similar effects of spelling processes on writing execution in adults when the graphomotor constraints induced by the task were high.

Finally, the effects of graphemic complexity were marginal, with an effect on the letter size in the frequentist analysis only. It has been shown that in a dictation task, the phoneme-grapheme conversion underlying complex graphemes impacts the latency, i.e. the time before starting to write once the word had been heard (Bonin & Delattre 2010). Those authors proposed that the different phoneme-to-grapheme correspondences that are available for a given phoneme are simultaneously activated and compete in spelling to dictation. In the present study, the visual presentation of the words before starting to write

may have limited this competition, and thus the effects of graphemic complexity on writing kinematics.

## **5. Conclusions: Implications for Primary Education, limitations and perspectives**

The combined effects of orthographic and graphomotor constraints in 1st graders indicate that graphomotor and spelling processes build up and interact from the beginning of learning. This finding supports the idea of a parallel activation of spelling (high-level) and graphomotor (low-level) modules that was proposed in the van Galen's model (1991). It calls into question the relevance of evaluating spelling and graphomotor processes separately. The present study has practical implications for helping teachers to both teach and evaluate handwriting quality and spelling proficiency. In terms of evaluation, if the word to be written is not frequent and contains difficult letters, handwriting performance could be impacted, and the teacher could wrongly interpret this as a graphomotor difficulty. This might explain some of the handwriting difficulties sometimes observed in children with reading disorders (Cheng-Lai, Chan, & Lo, 2013; Kandel, Lassus-Sangosse, Grosjacques, & Perret, 2017).

In terms of instruction, teachers should be careful not to put students who do not yet fully master graphomotor skills in a situation where they have to deal both with unusual graphomotor demands while writing words that are too complex orthographically. Otherwise they risk to slow down one or both learning processes. A very good example is the application of instructional methods that involve modifying visual feedback (Bara & Bonneton-Botté, 2021), while children are used to control both pen movements and word spelling visually. Although this type of method has been tested only with single letters, its extension to whole words should take the combination of orthographic and graphomotor demands of the task into account.

A logical continuation of the present study would be to evaluate how this combined effect of constraints evolves longitudinally between Grade 1 and Grade 2, and differs according to the children's level of handwriting proficiency, especially at the very beginning of learning. Specific difficulties could emerge in students whose graphomotor coordination is impaired. Indeed, these students would face a double challenge: not only do they have trouble tracing letters in a fast, fluid, and legible manner, but this cost associated with

moving their pen correctly could disrupt the attention that needs to be given to the spelling of words, especially if they are not well known by the child.

In conclusion, our study reveals that when children start to write a word, orthographic constraints underlying the word frequency and, to a much lesser extent, the grapheme structure at the end of the word, influence the production of the first letter when this letter has not yet been automated. Of course, the present study has from some limitations. The main limit is the reduced number of items used, which is due to the very controlled material (5-letter words with the same vowel at the second position and with a complex and simple grapheme within the 3 last letters). Furthermore, this study has not been designed to evaluate the total temporal course of the writing movement, from initialization (with latency) to the end of the word. Further studies are necessary to confirm this combined effect, and to better understand the emergence of the *cascading* phenomenon between the spelling and motor processes in learning to write (Olive, 2014).

## Appendix

		DURATION	MEAN VELOCITY	NUMBER OF STOPS	LETTER SIZE
<b>GRADE 1</b>	DURATION	1,00	-0,29	0,47	0,44
	MEAN VELOCITY	-0,29	1,00	-0,53	0,47
	NUMBER OF STOPS	0,47	-0,53	1,00	-0,17
	LETTER SIZE	0,44	0,47	-0,17	1,00
<b>GRADE 3</b>	DURATION	1,00	-0,52	0,57	0,17
	MEAN VELOCITY	-0,52	1,00	-0,61	0,62
	NUMBER OF STOPS	0,57	-0,61	1,00	-0,29
	LETTER SIZE	0,17	0,62	-0,29	1,00
<b>GRADE 5</b>	DURATION	1,00	-0,31	0,47	0,49
	MEAN VELOCITY	-0,31	1,00	-0,34	0,51
	NUMBER OF STOPS	0,47	-0,34	1,00	0,06
	LETTER SIZE	0,49	0,51	0,06	1,00

Table 3: Correlation matrix between the four dependent variables in the main experiment. In bold, significant correlation coefficients.

As can be seen, the dependent variables are partially correlated, which is not surprising knowing that when handwriting is not mastered, students take more time to write the velocity is lower, they make more stops, and they write larger. These correlations evolve

in a non-monotonic way with the school level, and they decrease when the handwriting is mastered (in Grade 5).

## References

- Afonso, O., Suárez-Coalla P., González-Martín N., & Cuetos F. (2018). The impact of word frequency on peripheral processes during handwriting: A matter of age. *The Quarterly Journal of Experimental Psychology*, 1-9. doi: 10.1080/17470218.2016.1275713
- Athènes, S., Sallagoity, I., Zanone, P. G., & Albaret, J.-M. (2004). Evaluating the coordination dynamics of handwriting. *Human Movement Science*, 23, 621–641.
- Badan, M., Hauert, C. A., & Mounoud, P. (2000). Sequential pointing in children and adults. *Journal of Experimental Child Psychology*, 75, 43–69.
- Baddeley, A.D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 575-589.
- Berninger, V., & Swanson, H. L. (1994). Children's writing: Toward a process theory of the development of skilled writing. In: E. C. Butterfield, & J. Carlson (Eds.). *Modifying Hayes and flower's model of skilled writing to explain beginning and developing writing* (pp. 57–81). London: JAI Press.
- Bara, F., & Bonneton-Botté, N. (2021). Handwriting isolated cursive letters in young children: Effect of the visual trace deletion. *Learning and Instruction*, 74, 101439.
- Bo, J., Barta, J., Ferencak, H., Comstock, S., Riley, V., & Krueger, J. (2014). Developmental characteristics in cursive and printed letter-writing for school-age children. *Journal of Motor Learning and Development*, 2(1), 1-8.
- Bonin, P., & Delattre, M. (2010). The sublexical conversion procedure in spelling to dictation. *Psychologie Française*, 110, 495-516.
- Catach, N. (1995). *L'orthographe française* [French spelling]. Paris: Nathan.
- Chang, S.H., & Yu, N.-Y. (2010). Characterization of motor control in handwriting difficulties in children with or without developmental coordination disorder. *Developmental Medicine & Child Neurology* 52, 244-250.
- Chartrel, E., & Vinter, A. (2008). The impact of spatio-temporal constraints on cursive letter handwriting in children. *Learning and Instruction*, 18, 537-547.



- Cheng-Lai, A., Li-Tsang, C. W. P., Chan, A. H. L., & Lo, A. G. W. (2013). Writing to dictation and handwriting performance among Chinese children with dyslexia: Relationships with orthographic knowledge and perceptual-motor skills. *Research in Developmental Disabilities, 34*, 3372-3383.
- Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Sound-to-spelling regularity affects both writing latencies and durations. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32*, 1330-1340.
- Dounskaia, N., Van Gemmert, A. W. A., & Stelmach, G. E. (2000). Interjoint coordination during handwriting-like movements. *Experimental Brain Research, 135*, 127–140.
- Hamstra-Bletz, L., & Blöte, A. (1993). A longitudinal study on dysgraphic handwriting in primary school. *Journal of Learning Disabilities, 26*, 689–699.
- Humblot, L., Fayol, M., & Longchamp, K. (1994). La copie de mots en CP et CE1 [Words copy in Grade 1 and Grade 2]. *Repères. Recherches en didactique du français langue maternelle, 9*, 47-59.
- Jolly, C., Huron, C., & Gentaz, E. (2014). A one-year survey of cursive letter handwriting in a French second-grade child with developmental coordination disorder. *L'Année psychologique, 114*(3), 421-445.
- Jones, D., & Christensen, C. A. (1999). Relationship between automaticity in handwriting and students' ability to generate written text. *Journal of Educational Psychology, 91*(1), 44–49.
- Kandel, S., Álvarez, C., & Vallée (2006). Syllables as processing units in handwriting production. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 18–31.
- Kandel, S., & Perret, C. (2015). How does the interaction between spelling and motor processes build up during writing acquisition? *Cognition, 136*, 325-336.
- Kandel, S., & Valdois S. (2005). The effect of orthographic regularity on children's handwriting production. *Current Psychology Letter: Brain Behaviour and Cognition, 17*, 3. <http://cpl.revues.org/463>
- Kandel, S., & Spinelli, E. (2010). Processing complex graphemes in handwriting production. *Memory & Cognition, 38*, 762–770.

- Kandel, S., Lassus-Sangosse, D., Grosjacques, G., & Perret, C. (2017). The impact of developmental dyslexia and dysgraphia on movement production during word writing. *Cognitive Neuropsychology*, *34*, 219-251.
- Kandel, S., Peereman, R., Grosjacques, G., & Fayol, M. (2011). For a psycholinguistic model of handwriting production: Testing the syllable-bigram controversy. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1310–1322.
- Kandel, S., Soler, O., Valdois, S., & Gros, C. (2006). Graphemes as motor units in the acquisition of writing skills. *Reading and Writing*, *19*, 313-337.
- Kelso, J.A.S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.
- Lambert, E., Alamargot, D., Larocque, D., & Caporossi, G. (2011). Dynamics of the spelling process during a copying task: Effects of regularity and frequency. *Canadian Journal of Experimental Psychology*, *65*(3), 141–150.
- Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). Manulex: A grade-level lexical database from French elementary-school readers. *Behavior Research Methods, Instruments, & Computers*, *36*, 156-166.
- Meulenbroek, R.G.J., & Van Galen, G.P. (1988). The acquisition of skilled handwriting: Discontinuous trends in kinematic variables. In: Colley A.M., & Beech, J.R. (Eds.). *Advances in Psychology: Cognition and action in skilled behaviour*. Elsevier Science Publishers B.V., North Holland (pp. 273-281).
- Meulenbroek, R.G.J., Vinter, A., & Mounoud, P. (1993). Development of the start-rotation principle in circle production. *British Journal of Developmental Psychology*, *11*, 307-320.
- Mojet, J. (1991). Development of graphic skills. In J. Wann, A. M. Wing, & N. Sovik( Eds.). *Characteristics of the developing handwriting skill in elementary education* (pp. 53–75). London: Academic Press.
- New, B. (2006). Lexique 3 : Une nouvelle base de données lexicales. *Actes de la Conférence Traitement Automatique des Langues Naturelles (TALN 2006)*, Louvain, Belgique.
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*, 516-524.
- Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research*, *6*, 173-194.

- Olive, T., & Kellogg, R. T. (2002). Concurrent activation of high- and low-level production processes in written composition. *Memory & Cognition*, *30*, 594–600.
- Orliaguet, J.-P., Kandel, S., & Boë, L.-J. (1997). Visual perception of motor anticipation in cursive handwriting: influence of spatial and movement information on the prediction of forthcoming letters. *Perception*, *26*, 905-912.
- Pagliarini, E., Guasti, M. T., Toneatto, C., Granocchio, E., Riva, F., Sarti, D. ... & Stucchi, N. (2015). Dyslexic children fail to comply with the rhythmic constraints of handwriting. *Human Movement Science*, *42*, 161-182.
- Prunty, M.M., Barnett, A.L., Wilmut, K., & Plumb, M.S. (2013). Handwriting speed in children with Developmental Coordination Disorder: Are they really slower? *Research in Developmental Disabilities*, *34*(9), 2927-2936.
- Roux, S., McKeeff, T.J., Grosjacques, G., Afonso, O., & Kandel, S. (2013). The interaction between central and peripheral processes in handwriting production. *Cognition*, *127*, 235-241.
- Sausset, S., Lambert, E., Olive, T., & Larocque, D. (2012). Processing of syllables during handwriting: Effects of graphomotor constraints. *The Quarterly Journal of Experimental Psychology*, *65*, 1872-1879. DOI:10.1080/17470218.2012.715654.
- Séraphin Thibon, L., Gerber, S., & Kandel, S. (2018). The elaboration of motor programs for the automation of letter production. *Acta Psychologica*, *182*, 200-211.
- Séraphin Thibon, L., Barbier, G., Vilain, C., Sawallis, T.R., Gerber, S., & Kandel, S. (2019). Investigating how children produce rotation and pointing movements when they learn to write letters. *Human Movement Science*, *65*, 15-29.
- Spinelli, E., Kandel, S., Guerassimovitch, H., & Ferrand, L. (2012). Graphemic cohesion effect in reading and writing complex graphemes. *Language and Cognitive Processes*, *27*, 770-791.
- Tucha, O., Tucha, L., & Lange, K. W. (2008). Graphonomics, automaticity and handwriting assessment: Graphonomics, automaticity and handwriting assessment. *Literacy*, *42*(3), 145–155.
- Van Galen, G.P. (1991). Handwriting: Issues for a psychomotor theory. *Human Movement Science*, *10*, 165-191.

- Van Mier, H. (2006). Developmental differences in drawing performance of the dominant and non-dominant hand in right-handed boys and girls. *Human Movement Science, 25*, 657-677.
- Viviani, P. & Terzuolo, C. (1982). Trajectory determines movement dynamics. *Neuroscience, 7*, 431-437.
- Zesiger, P., Mounoud, P., & Hauert, C.-A. (1993). Effects of lexicality and trigram frequency on handwriting production in children and adults. *Acta Psychologica, 82*, 353-365.

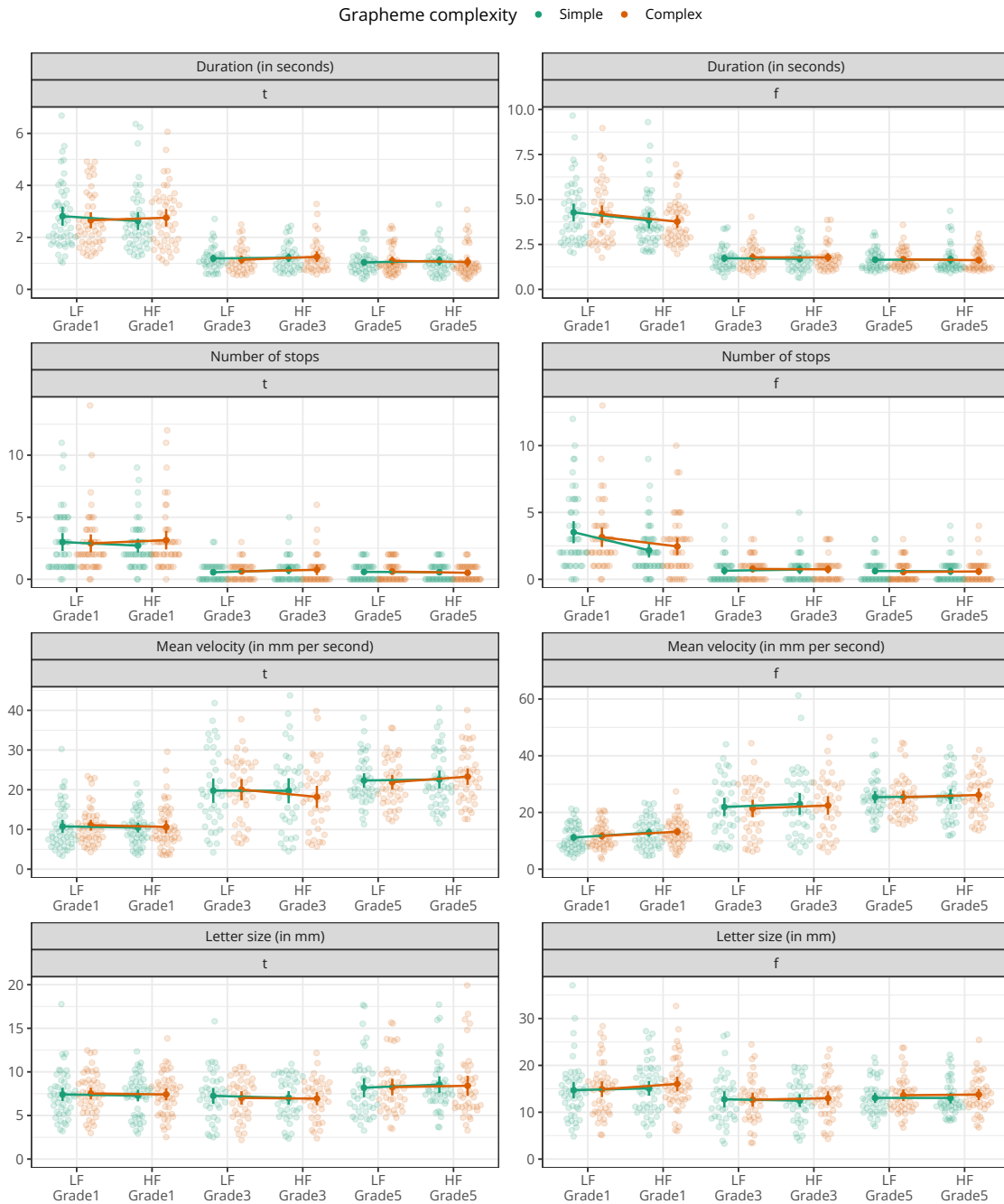
# Supplementary materials to ‘Interaction between orthographic and graphomotor constraints in learning to write’

## Contents

<b>1</b>	<b>Visual data exploration</b>	<b>2</b>
1.1	Univariate raw data exploration . . . . .	2
1.2	Bivariate correlations by grade . . . . .	3
<b>2</b>	<b>Bayesian multilevel modelling</b>	<b>4</b>
2.1	Modelling positive-only values . . . . .	4
2.2	Shifted-lognormal regression model . . . . .	4
2.3	Poisson regression model . . . . .	6
2.4	Fitting the final model . . . . .	6
2.5	Evaluating the model . . . . .	9
2.6	Hypothesis testing . . . . .	10
<b>3</b>	<b>Interpretation of the results for each variable</b>	<b>13</b>
3.1	Letter duration . . . . .	14
3.2	Number of stops . . . . .	18
3.3	Mean velocity . . . . .	22
3.4	Letter size . . . . .	26
<b>4</b>	<b>Acknowledgements</b>	<b>30</b>
<b>5</b>	<b>Session information</b>	<b>31</b>
	<b>References</b>	<b>33</b>

## 1 Visual data exploration

### 1.1 Univariate raw data exploration



*Figure 1.* Effect of grade, word frequency, and grapheme complexity on the total duration, the number of stops, the mean velocity, and the letter size. The error bars represent the 95% confidence intervals of the mean (assuming a Gaussian distribution).

Figure 1 shows the effect of grade, word frequency, and grapheme complexity on the letter duration, the number of stops, the mean velocity, and the letter size. This figure suggests that the average duration (in seconds) seems to decrease monotonically with grade. The number of stops also seems to decrease with grade, with most trials for children from grade 2 being associated with no stop.

### 1.2 Bivariate correlations by grade

Figure 2 shows the overall and by-grade Spearman correlation between each pair of variables. This figure reveals medium to strong positive and negative correlations between each pair of variable. These relations are sometimes non-linear (e.g., between duration and mean velocity), hence the use of Spearman (rather than Pearson) correlation coefficients.

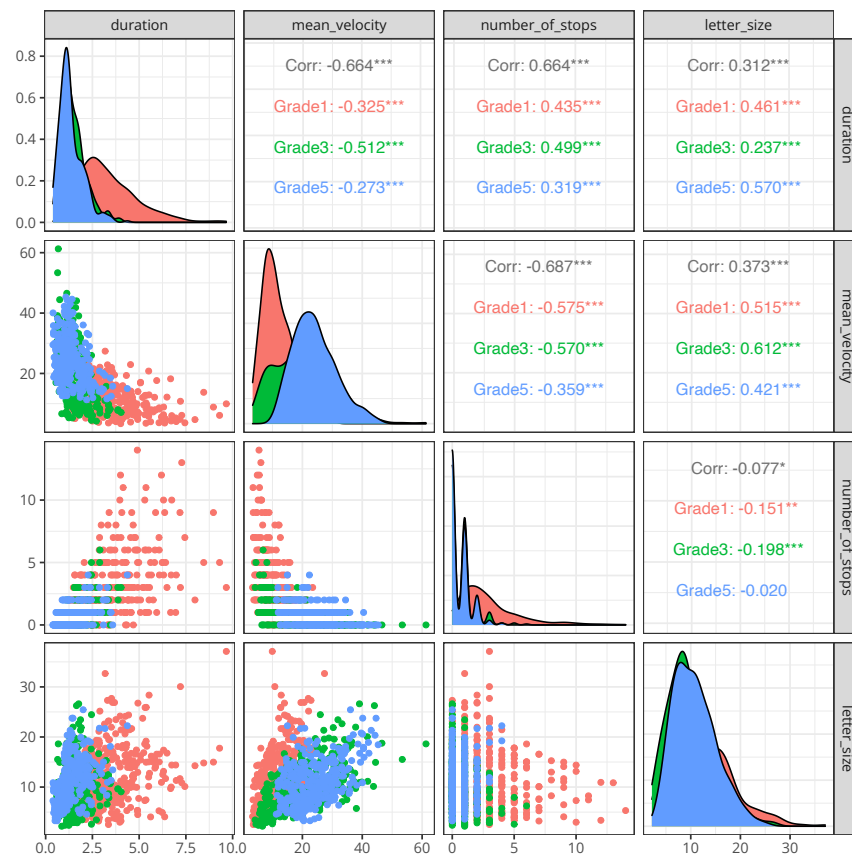


Figure 2. Overall and by-grade Spearman correlation between each pair of measured variables.

## 2 Bayesian multilevel modelling

### 2.1 Modelling positive-only values

A dominant feature of durations (or response times) is that their distribution is generally positively skewed, with the spread and/or the skewness increasing with task difficulty (for review, see for instance Forstmann, Ratcliff, & Wagenmakers, 2016). Therefore, several models have been proposed to account for the peculiarities of the data coming from such tasks as well as to relate it to the underlying cognitive processes. We discuss below why using Gaussian models for this kind of data is generally not a sensible idea and describe our approach in more details. We follow a general “Bayesian workflow” by building our model in an iterative manner and by motivating and validating each modelling choice (for more details, see for instance Gelman et al., 2020).

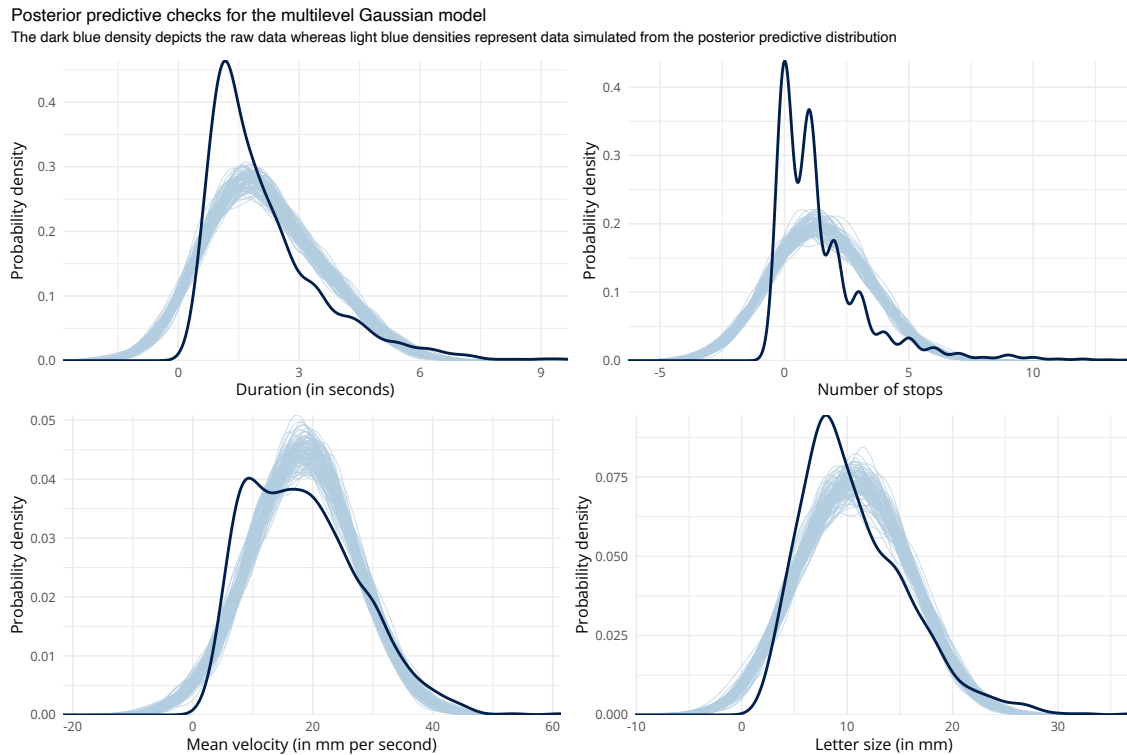
We first fitted a Bayesian multilevel (also known as “mixed-effects”) Gaussian multivariate (i.e., with multiple outcomes) model. One way of evaluating this model is to evaluate its predictions. If this model is a good description of the process that generated the observed data, then it should be able to generate data that looks like the observed data. The process of generating data from the estimated posterior distribution is called *posterior predictive checking* and can be used in many different ways using the `pp_check()` method (Gabry, Simpson, Vehtari, Betancourt, & Gelman, 2019). In Figure 3, we depict the distribution of the raw data along with the distribution of 100 simulated datasets.

This figure reveals that the Gaussian model fails to account for the peculiarities of the data at hand. For instance, it systematically fails to predict the right-skew of all four variables, and more dramatically, sometimes predicts negatives values for these variables, although they are strictly positive. Moreover, using a Gaussian distribution to model the number of stops also leads to nonsensical predictions as the number of stops is necessarily a positive integer (whereas the Gaussian distribution can produce any real number), as illustrated in the upper right panel of Figure 3.

### 2.2 Shifted-lognormal regression model

A useful description of RTs or durations should be able to account for the effects of the difficulty of the task, as well as changes in shift and spread of the distribution. The Log-normal, Ex-Gaussian, or Weibull distributions often provide a good fit to these data, but their parameters are difficult to interpret in terms of difficulty, shift, or spread (i.e., these distributions do not have straightforward interpretable parameters). In contrast, the shifted log-normal distribution has parameters that can easily be interpreted in terms of





*Figure 3.* Posterior predictive checks for the multilevel Gaussian model. The dark blue density depicts the raw data whereas light blue densities represent data simulated from the posterior predictive distribution.

difficulty, shift, and spread.

The log-normal distribution is called “log-normal” because the parameters are the mean and standard deviation of the log-transformed response, which is assumed to be a normal (Gaussian) distribution. The shifted log-normal distribution is then described by three parameters:

- $\mu$  (mu, difficulty): the mean of the log-normal distribution. The median duration is given by  $\text{shift} + \exp(\mu)$ .
- $\sigma$  (sigma, scale): the standard deviation of the log-normal distribution. Increases the mean but not the median of  $\mu$ .
- shift (ndt) indicates the time of the earliest possible response. When  $\text{shift} = 0$ , the shifted log-normal distribution correspond to the conventional log-normal distribution with two parameters.

Regression models using this family usually aim to predict  $\mu$ , the mean of the log-

normal distribution. We employ this strategy as well, while noticing that both  $\sigma$  and shift could be allowed to vary across conditions as well.

### 2.3 Poisson regression model

Whereas the previous models are able to take into account the right-skew of the four variables we are interested in, they are still not able to make proper predictions with regards to the number of stops (because the log-normal distribution is continuous). Yet a better model could be fitted by picking up a discrete probability distribution defined on the positive integer real. The Poisson regression model is appropriate for modelling discrete counts of events (e.g., the number of stops) that happen in a fixed interval of space or time with no upper bound. The Poisson model is simpler than the Gaussian or the lognormal one because it has only one parameter  $\lambda$  that describes its shape. The parameter  $\lambda$  is the expected value of the outcome  $y$  (and also its expected variance). However, we need a link function to relate the predictors with the parameter  $\lambda$  and to ensure that  $\lambda$  is always positive. We use the conventional logarithmic link function, resulting in the following linear model:

$$y_i \sim \text{Poisson}(\lambda_i)$$

$$\log(\lambda_i) = \alpha + \beta_g \cdot \text{grade}_i + \beta_f \cdot \text{frequency}_i +$$

$$\beta_{\text{graphe}} \cdot \text{grapheme}_i + \beta_{\text{grapho}} \cdot \text{graphomotor}_i$$

This kind of model is now able to predict valid number of stops (i.e., positive integers). Note that for simplicity, we omit the varying effects and the priors from the above model (for more details on Poisson regression, see Winter & Bürkner, 2021).

### 2.4 Fitting the final model

To set up the model, we need to invoke the `brms::brmsformula()` function and construct one formula for each of the four dependant variables. We fitted all models using the `brms` package (Bürkner, 2017). We used sum contrasts (i.e., recoding conditions as -0.5 vs. 0.5) for binary predictors (i.e., frequency, grapheme complexity, and graphomotor difficulty) and used the default factor coding scheme (i.e., dummy coding) for grade.

```
# defining the model formula for the generalised multilevel model
formula_generalised <-
  bf(
    duration ~ 1 + group * frequency * grapheme_complexity *
```

```

        graphomotor_difficulty + (1 | subject),
    family = shifted_lognormal()
  ) +
  bf(
    mean_velocity ~ 1 + group * frequency * grapheme_complexity *
      graphomotor_difficulty + (1 | subject),
    family = shifted_lognormal()
  ) +
  bf(
    number_of_stops ~ 1 + group * frequency * grapheme_complexity *
      graphomotor_difficulty + (1 | subject),
    family = poisson()
  ) +
  bf(
    letter_size ~ 1 + group * frequency * grapheme_complexity *
      graphomotor_difficulty + (1 | subject),
    family = shifted_lognormal()
  )

# defining the priors for the multilevel generalised model
priors_generalised <- c(
  prior(normal(1, 0.5), class = Intercept, resp = "duration"),
  prior(normal(0, 0.5), class = b, resp = "duration"),
  prior(exponential(0.1), class = sd, resp = "duration"),
  prior(exponential(0.1), class = sigma, resp = "duration"),
  prior(normal(2, 0.5), class = Intercept, resp = "meanvelocity"),
  prior(normal(0, 0.5), class = b, resp = "meanvelocity"),
  prior(exponential(0.1), class = sd, resp = "meanvelocity"),
  prior(exponential(0.1), class = sigma, resp = "meanvelocity"),
  prior(normal(1, 0.5), class = Intercept, resp = "numberofstops"),
  prior(normal(0, 0.5), class = b, resp = "numberofstops"),
  prior(exponential(0.1), class = sd, resp = "numberofstops"),
  prior(normal(2, 0.5), class = Intercept, resp = "lettersize"),
  prior(normal(0, 0.5), class = b, resp = "lettersize"),
  prior(exponential(0.1), class = sd, resp = "lettersize"),
  prior(exponential(0.1), class = sigma, resp = "lettersize")
)

```

```
)

# centering and reordering predictors
df2 <- df %>%
  mutate(
    group = factor(
      x = group,
      levels = c("CP", "CE", "CM"),
      labels = c("Grade1", "Grade3", "Grade5")
    ),
    frequency = factor(
      x = frequency,
      levels = c("LF", "HF"),
      labels = c("LF", "HF")
    ),
    grapheme_complexity = factor(
      x = grapheme_complexity,
      levels = c("Simple", "Complex"),
      labels = c("Simple", "Complex")
    ),
    graphomotor_difficulty = factor(
      x = graphomotor_difficulty,
      levels = c("EL", "HL"),
      labels = c("t", "f")
    )
  ) %>%
  # removes rows where duration is equal to 0
  filter(duration != 0)

# defining contrasts
contrasts(df2$frequency) <- c(-0.5, +0.5)
contrasts(df2$grapheme_complexity) <- c(-0.5, +0.5)
contrasts(df2$graphomotor_difficulty) <- c(-0.5, +0.5)

# fitting the model
mod2 <- brm(
  formula = formula_generalised + set_rescor(rescor = FALSE),
```

```
prior = priors_generalised,  
chains = 4, cores = 4,  
warmup = 2000, iter = 1e4,  
control = list(adapt_delta = 0.95),  
data = df2,  
sample_prior = TRUE,  
file = "models/multilevel_generalised_model"  
)
```

We then fit this model below using the `brms::brm()` function. We run four chains, each for 10000 iterations and using the first 2000 iterations used as warmup (i.e., the first 2000 samples of each chain are discarded from the final analysis). This results in a total of  $4 \times (10000 - 2000) = 32000$  samples from the (joint) posterior distribution that will be used for inference.

## 2.5 Evaluating the model

One way of evaluating the model is to evaluate its predictions. In Figure 4, we depict the distribution of the raw data along with the distribution of 100 simulated datasets (a posterior predictive check, as introduced previously).

As can be seen from Figure 4, the model seems pretty good at simulating data that looks like the observed data. From this predictive/sampling distribution (i.e., the distribution of simulated data sets), so-called “Bayesian  $p$ -values” can be computed to quantify the compatibility between the observed data and the proposed model.

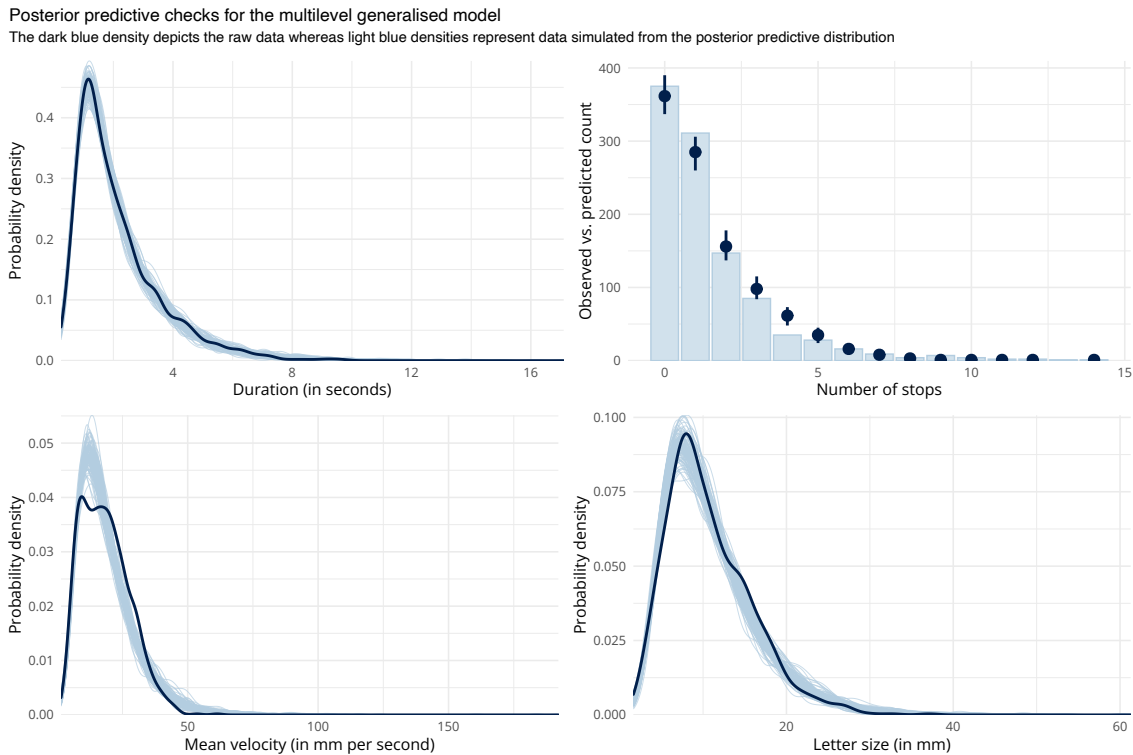


Figure 4. Posterior predictive checking. The dark blue line represents the distribution of raw data whereas light blue lines represent data simulated from the posterior distribution.

## 2.6 Hypothesis testing

We can test any arbitrary hypothesis using the `brms::hypothesis()` method, which is computing a Bayes factor via the Savage-Dickey method (Wagenmakers, Lodewyckx, Kuriyal, & Grasman, 2010). This method consists in comparing the posterior probability density to the prior probability density for some hypothesised value for the parameter of interest (e.g.,  $\theta = 0$ ). For instance, we test below the hypothesis according to which the effect of graphemic complexity in Grade 1 would be null.

```
# testing whether the effect of grapheme complexity on duration equal to 0
hyp <- hypothesis(x = mod2, hypothesis = "duration_grapheme_complexity1 = 0")

# prints the output
print(hyp)

## Hypothesis Tests for class b:
##           Hypothesis Estimate Est.Error CI.Lower CI.Upper Evid.Ratio
## 1 (duration_graphem... = 0           0           0.05   -0.11     0.1         9.87
```

```
## Post.Prob Star
## 1      0.91
## ---
## 'CI': 90%-CI for one-sided and 95%-CI for two-sided hypotheses.
## '*': For one-sided hypotheses, the posterior probability exceeds 95%;
## for two-sided hypotheses, the value tested against lies outside the 95%-CI.
## Posterior probabilities of point hypotheses assume equal prior probabilities.

# plotting it
data.frame(posterior = hyp$samples$H1, prior = hyp$prior_samples$H1) %>%
  gather(type, value) %>%
  ggplot(aes(x = value, fill = type) ) +
  geom_vline(xintercept = 0, linetype = 2, alpha = 1) +
  geom_area(stat = "density", alpha = 0.8, position = "identity") +
  theme_bw(base_size = 12, base_family = "Open Sans") +
  labs(x = expression(beta[grapheme_complexity]), y = "Probability density") +
  scale_fill_brewer(palette = "Dark2") +
  theme(legend.title = element_blank() ) +
  coord_cartesian(xlim = c(-2, 2) )
```

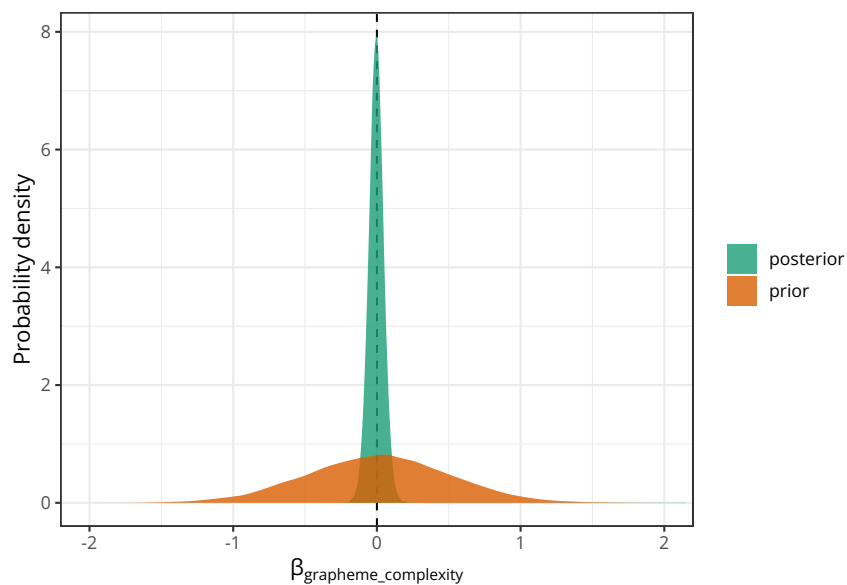


Figure 5. Hypothesis testing via the Savage-Dickey method. The resulting Bayes factor (BF) is the ratio of the height (i.e., the density probability) of the posterior versus prior distribution at some value of interest for the parameter (here it is 0).

The resulting Bayes factor (BF, called “Evid. Ratio” in the output) may be inter-

preted as follows: the observed data are 9.87 more likely under the hypothesis of null effect than under the hypothesis of a non-null effect. From the BF in favour of the null hypothesis (relative to the alternative hypothesis), we can compute the BF in favour of the alternative hypothesis (relative to the null hypothesis), using  $\text{BF}_{10} = 1/\text{BF}_{01}$  (we report the  $\text{BF}_{10}$  in the following). Alternatively, the BF can be interpreted as an *updating factor*, indicating by “how much” we should update our *prior odds* (the ratio of the a priori probability of  $\mathcal{H}_0$  versus  $\mathcal{H}_1$ ) to convert them into *posterior odds* (the ratio of the a posteriori probability of  $\mathcal{H}_0$  versus  $\mathcal{H}_1$ ).



### 3 Interpretation of the results for each variable

Now that we have fitted the model, we are left with the task of interpreting the output from the model. The output of the model is a (joint) posterior distribution over all parameters of the model. We can marginalise this joint distribution to obtain the (marginal) posterior distribution on each parameter. To summarise this distribution, we can retrieve samples from the joint posterior distribution.

```
# retrieves posterior samples (for all parameters)
posterior_samples <- as_draws_df(mod2)

# displays a summary
posterior_summary <- summarise_draws(posterior_samples)

# displays the first six rows
head(posterior_summary)

## # A tibble: 6 x 10
##   variable      mean median    sd   mad    q5    q95  rhat ess_bulk ess_tail
##   <chr>         <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>    <dbl>    <dbl>
## 1 b_duration_~  1.08  1.08  0.0310 0.0304  1.03  1.13  1.00  29023.  25012.
## 2 b_meanveloc~  2.34  2.34  0.0271 0.0263  2.29  2.38  1.00  30342.  20075.
## 3 b_numberofs~  1.04  1.04  0.0367 0.0354  0.977  1.10  1.00  41575.  26004.
## 4 b_lettersiz~  2.29  2.29  0.0234 0.0229  2.25  2.33  1.00  32779.  23505.
## 5 b_duration_~ -0.866 -0.865 0.0386 0.0384 -0.931 -0.804 1.00  23574.  24590.
## 6 b_duration_~ -0.967 -0.966 0.0414 0.0418 -1.04  -0.900 1.00  20887.  23486.
```

The above command outputs a matrix with parameters of the model in columns and posterior samples in rows. Let's examine these results for each parameter in more details. For instance, Figure 6 represents the posterior distribution of the average letter duration in Grade-1 children.

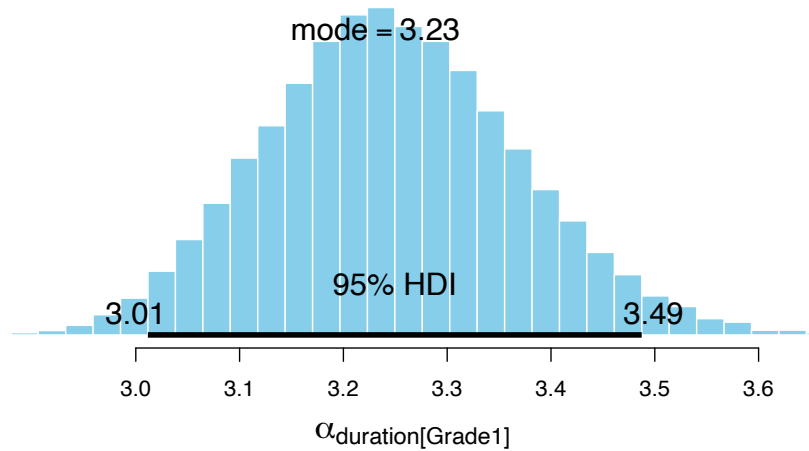
```
# retrieves the posterior samples for the average letter duration in Grade 1
average_duration_grade1 <- posterior_samples$b_duration_Intercept +
  posterior_samples$ndt_duration

# plotting it
plotPost(
  paramSampleVec = exp(average_duration_grade1), showMode = TRUE,
```

```

xlab = expression(paste(alpha[duration][paste("[", Grade1, "]")]) )
)

```



*Figure 6.* Posterior distribution of the intercept (i.e., the average letter duration in Grade 1). The mode (i.e., the most probable value) and the 95% credible (highest density) interval are also displayed.

Recall that we used a logarithmic link function, therefore the median letter duration is given by  $\exp(\alpha + \text{shift})$ .

### 3.1 Letter duration

Table 1 reports the estimates (median of the posterior distribution) and associated 95% credible intervals and BFs for all parameters regarding the letter duration variable.

Table 1  
*Estimates and BF<sub>10</sub> for the slopes for letter duration.*

Term	Estimate	MAD	Lower	Upper	Rhat	BF <sub>10</sub>
Intercept	1.083	0.030	1.020	1.142	1.000	NA
groupGrade3	-0.865	0.038	-0.943	-0.793	1.000	$7.281 \times 10^{17}$
groupGrade5	-0.966	0.042	-1.050	-0.889	1.000	$3.586 \times 10^{15}$
frequency	-0.055	0.050	-0.157	0.049	1.000	0.192
grapheme_complexity	-0.005	0.050	-0.106	0.097	1.000	0.101
graphomotor_difficulty	0.423	0.051	0.320	0.526	1.000	$9.762 \times 10^{15}$
groupGrade3:frequency	0.074	0.062	-0.049	0.196	1.000	0.258
groupGrade5:frequency	0.027	0.062	-0.092	0.148	1.000	0.136
groupGrade3:grapheme_complexity	0.011	0.062	-0.112	0.132	1.000	0.13
groupGrade5:grapheme_complexity	0.004	0.061	-0.115	0.124	1.000	0.124
frequency:grapheme_complexity	0.046	0.097	-0.151	0.244	1.000	0.225
groupGrade3:graphomotor_difficulty	0.001	0.063	-0.121	0.127	1.000	0.132
groupGrade5:graphomotor_difficulty	0.099	0.064	-0.024	0.225	1.000	0.436
frequency:graphomotor_difficulty	-0.071	0.097	-0.268	0.128	1.000	0.261
grapheme_complexity:graphomotor_difficulty	-0.008	0.098	-0.203	0.188	1.000	0.197
groupGrade3:frequency:grapheme_complexity	0.010	0.120	-0.229	0.248	1.000	0.249
groupGrade5:frequency:grapheme_complexity	-0.101	0.119	-0.337	0.131	1.000	0.352
groupGrade3:frequency:graphomotor_difficulty	0.007	0.122	-0.233	0.243	1.000	0.246
groupGrade5:frequency:graphomotor_difficulty	0.068	0.120	-0.167	0.308	1.000	0.285
groupGrade3:grapheme_complexity:graphomotor_difficulty	0.072	0.120	-0.162	0.309	1.000	0.291
groupGrade5:grapheme_complexity:graphomotor_difficulty	0.035	0.119	-0.201	0.270	1.000	0.252
frequency:grapheme_complexity:graphomotor_difficulty	-0.040	0.176	-0.401	0.321	1.000	0.368
groupGrade3:frequency:grapheme_complexity:graphomotor_difficulty	-0.006	0.217	-0.434	0.424	1.000	0.442
groupGrade5:frequency:grapheme_complexity:graphomotor_difficulty	0.134	0.213	-0.283	0.548	1.000	0.529

*Note.* For each slope (for each line), the first two columns represent the estimated most probable value and its standard error (SE). The 'Lower' and 'Upper' columns contain the lower and upper bounds of the 95% CrI, whereas the 'Rhat' column reports the Gelman-Rubin statistic. The last column reports the Bayes factor in favour of the alternative hypothesis, relative to the null hypothesis (BF<sub>10</sub>).

These estimations are better understood visually. Thus, we plot the predictions of this model against raw data in Figure 7.

```
# retrieving the model's predictions
duration_predictions <- df2 %>%
  data_grid(graphomotor_difficulty, grapheme_complexity, frequency, group) %>%
  cbind(., fitted(
    object = mod2, newdata = ., resp = "duration",
    scale = "response", probs = c(0.025, 0.975),
    re_formula = NA, robust = TRUE
  ) ) %>%
  ungroup %>%
  dplyr::rename(estimate = Estimate, mad = Est.Error, lower = Q2.5, upper = Q97.5)
```

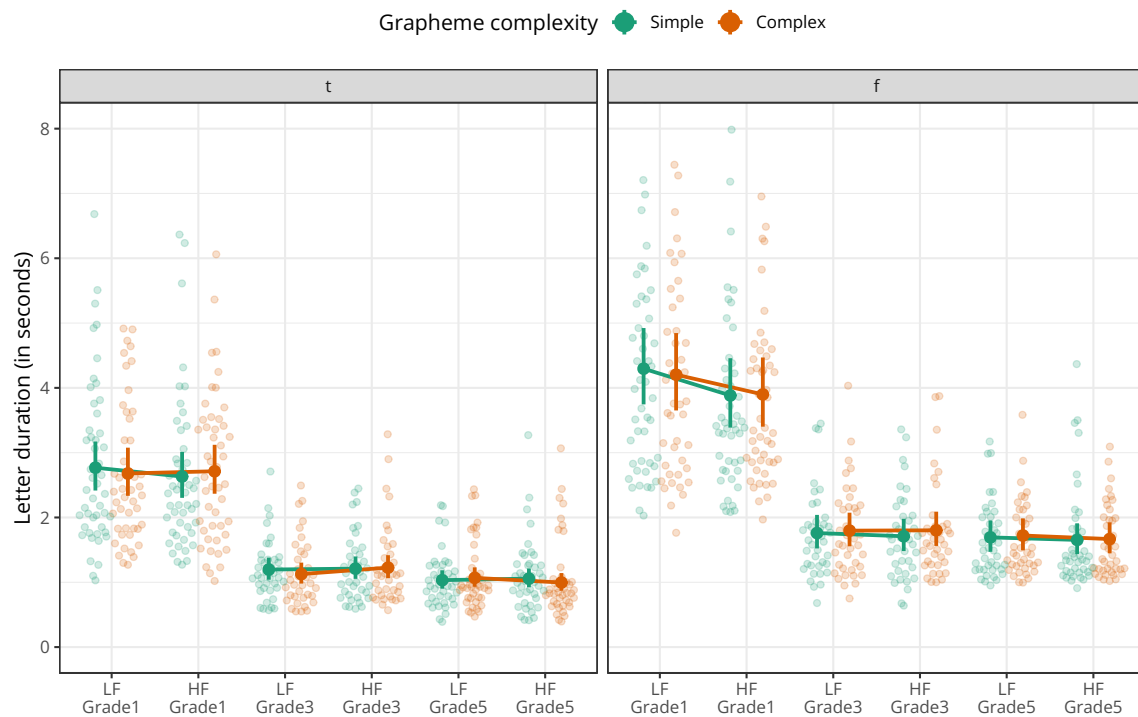


Figure 7. Letter duration by grade and word frequency (x-axis), grapheme complexity (in colour), and graphomotor difficulty (in panels). Transparent points represent individual data per participant. The surimposed dots and intervals represent the model’s predictions (median and 95% credible interval of the posterior distribution).

As can be seen in Figure 7, the model predicts larger letter duration for the difficult letter f as compared to the easy letter t for each grade. As can be seen from Table 1, the

only BFs favouring the alternative hypothesis (relative to the null hypothesis) are the BFs for the difference between Grade 1 and Grade 3 in average letter duration ( $\beta = -0.865$ , 95% CrI [-0.943, -0.793],  $\text{BF}_{10} = 7.281 \times 10^{17}$ ), as well as the difference between Grade 1 and Grade 5 ( $\beta = -0.966$ , 95% CrI [-1.05, -0.889],  $\text{BF}_{10} = 3.586 \times 10^{15}$ ), and the effect of graphomotor difficulty in Grade 1 ( $\beta = 0.423$ , 95% CrI [0.32, 0.526],  $\text{BF}_{10} = 9.762 \times 10^{15}$ ). Predictions from this model for each condition are also summarised in Table 2.

Table 2

*Estimated letter duration in each condition.*

Group	Frequency	Grapheme complexity	Graphomotor difficulty	Estimate	MAD	Lower	Upper
Grade1	LF	Simple	t	2.768	0.186	2.416	3.172
Grade1	LF	Simple	f	4.296	0.287	3.745	4.923
Grade1	LF	Complex	t	2.680	0.179	2.334	3.077
Grade1	LF	Complex	f	4.203	0.291	3.651	4.846
Grade1	HF	Simple	t	2.632	0.176	2.303	3.011
Grade1	HF	Simple	f	3.883	0.263	3.387	4.458
Grade1	HF	Complex	t	2.713	0.182	2.367	3.122
Grade1	HF	Complex	f	3.899	0.261	3.401	4.469
Grade3	LF	Simple	t	1.196	0.084	1.040	1.381
Grade3	LF	Simple	f	1.760	0.127	1.522	2.040
Grade3	LF	Complex	t	1.127	0.080	0.980	1.304
Grade3	LF	Complex	f	1.797	0.127	1.556	2.074
Grade3	HF	Simple	t	1.211	0.085	1.053	1.399
Grade3	HF	Simple	f	1.710	0.121	1.482	1.980
Grade3	HF	Complex	t	1.226	0.087	1.063	1.418
Grade3	HF	Complex	f	1.802	0.129	1.561	2.090
Grade5	LF	Simple	t	1.033	0.070	0.902	1.185
Grade5	LF	Simple	f	1.693	0.120	1.470	1.953
Grade5	LF	Complex	t	1.068	0.072	0.931	1.230
Grade5	LF	Complex	f	1.722	0.121	1.493	1.985
Grade5	HF	Simple	t	1.057	0.072	0.923	1.213
Grade5	HF	Simple	f	1.654	0.115	1.435	1.909
Grade5	HF	Complex	t	0.996	0.066	0.869	1.142
Grade5	HF	Complex	f	1.667	0.117	1.447	1.924

*Note.* For each condition, the 'Estimate' and 'MAD' columns contain the median and the median absolute deviation (MAD) of the posterior distribution, respectively. The 'Lower' and 'Upper' columns contain the lower and upper bounds of the 95% credible interval.

The output of a Bayesian model is a (joint) posterior distribution over all parameters of the model. We can marginalise this joint distribution to obtain the (marginal) posterior distribution on each parameter. To summarise this distribution, we can retrieve samples from the joint posterior distribution. Interestingly, this means we can look at the posterior distribution of any parameter of interest. For instance, and for exploratory purposes, we depict below the posterior distribution of the difference between high-frequency and low-frequency words (i.e., the effect of frequency) separately for each letter (graphomotor difficulty) and each grade. We averaged the predictions across both conditions of graphemic complexity, as this effect appeared to be null.

As can be seen in Figure 8, the posterior distribution for the effect of frequency is

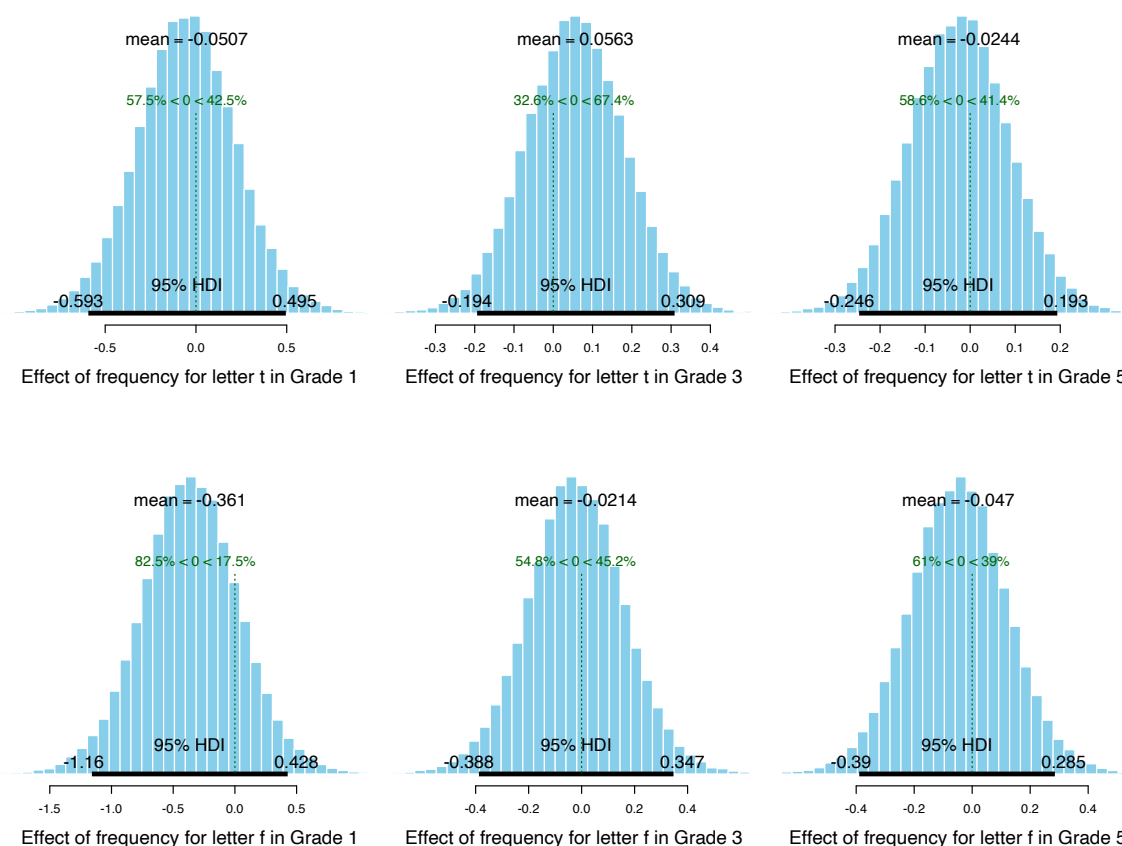


Figure 8. Effect of word frequency on letter duration (in seconds) for each grade (in column) and letter (in row). The histogram contains posterior samples for each effect, where the posterior distribution is summarised by its mean and 95% highest density interval (HDI). The green text indicates the probability that the parameter values is either inferior or superior to 0.

almost perfectly centred on zero in all conditions, except for letter f in Grade 1. Although the 95% credible interval largely encompasses 0 in this condition as well, there is still a 0.82 probability that the effect of frequency on letter duration is negative (given the data and the priors).

### 3.2 Number of stops

Table 3 reports the estimates (median of the posterior distribution) and associated 95% credible intervals and BFs for all parameters regarding the number of stops.

Table 3  
*Estimates and BF<sub>10</sub> for the slopes for the number of stops.*

Term	Estimate	MAD	Lower	Upper	Rhat	BF10
Intercept	1.037	0.035	0.965	1.109	1.000	NA
groupGrade3	-1.384	0.073	-1.529	-1.243	1.000	$4.244 \times 10^{16}$
groupGrade5	-1.556	0.077	-1.709	-1.409	1.000	$6.510 \times 10^{15}$
frequency	-0.180	0.070	-0.324	-0.034	1.000	2.746
grapheme_complexity	0.031	0.070	-0.111	0.175	1.000	0.155
graphomotor_difficulty	-0.051	0.070	-0.192	0.094	1.000	0.184
groupGrade3:frequency	0.292	0.140	0.019	0.568	1.000	2.537
groupGrade5:frequency	0.128	0.147	-0.161	0.409	1.000	0.417
groupGrade3:grapheme_complexity	0.051	0.142	-0.227	0.329	1.000	0.298
groupGrade5:grapheme_complexity	-0.069	0.149	-0.356	0.220	1.000	0.334
frequency:grapheme_complexity	0.177	0.133	-0.096	0.446	1.000	0.653
groupGrade3:graphomotor_difficulty	0.112	0.140	-0.164	0.390	1.000	0.378
groupGrade5:graphomotor_difficulty	0.061	0.144	-0.226	0.346	1.000	0.313
frequency:graphomotor_difficulty	-0.312	0.134	-0.581	-0.032	1.000	3.307
grapheme_complexity:graphomotor_difficulty	-0.047	0.135	-0.315	0.224	1.000	0.285
groupGrade3:frequency:grapheme_complexity	-0.234	0.254	-0.731	0.258	1.000	0.791
groupGrade5:frequency:grapheme_complexity	-0.147	0.260	-0.653	0.358	1.000	0.602
groupGrade3:frequency:graphomotor_difficulty	0.102	0.257	-0.396	0.605	1.000	0.55
groupGrade5:frequency:graphomotor_difficulty	0.326	0.255	-0.176	0.834	1.000	1.153
groupGrade3:grapheme_complexity:graphomotor_difficulty	0.068	0.254	-0.427	0.568	1.000	0.529
groupGrade5:grapheme_complexity:graphomotor_difficulty	0.004	0.264	-0.511	0.509	1.000	0.52
frequency:grapheme_complexity:graphomotor_difficulty	0.042	0.231	-0.430	0.518	1.000	0.464
groupGrade3:frequency:grapheme_complexity:graphomotor_difficulty	-0.057	0.379	-0.801	0.678	1.000	0.766
groupGrade5:frequency:grapheme_complexity:graphomotor_difficulty	0.073	0.384	-0.679	0.836	1.000	0.782

*Note.* For each slope (for each line), the first two columns represent the estimated most probable value and its standard error (SE). The 'Lower' and 'Upper' columns contain the lower and upper bounds of the 95% CrI, whereas the 'Rhat' column reports the Gelman-Rubin statistic. The last column reports the Bayes factor in favour of the alternative hypothesis, relative to the null hypothesis (BF<sub>10</sub>).

These estimations are better understood visually. Thus, we plot the predictions of this model against raw data in Figure 9.

```
# retrieving the model's predictions
stops_predictions <- df2 %>%
  data_grid(graphomotor_difficulty, grapheme_complexity, frequency, group) %>%
  cbind(., fitted(
    object = mod2, newdata = ., resp = "numberofstops",
    scale = "response", probs = c(0.025, 0.975),
    re_formula = NA, robust = TRUE
  ) ) %>%
  ungroup %>%
  dplyr::rename(estimate = Estimate, mad = Est.Error, lower = Q2.5, upper = Q97.5)
```

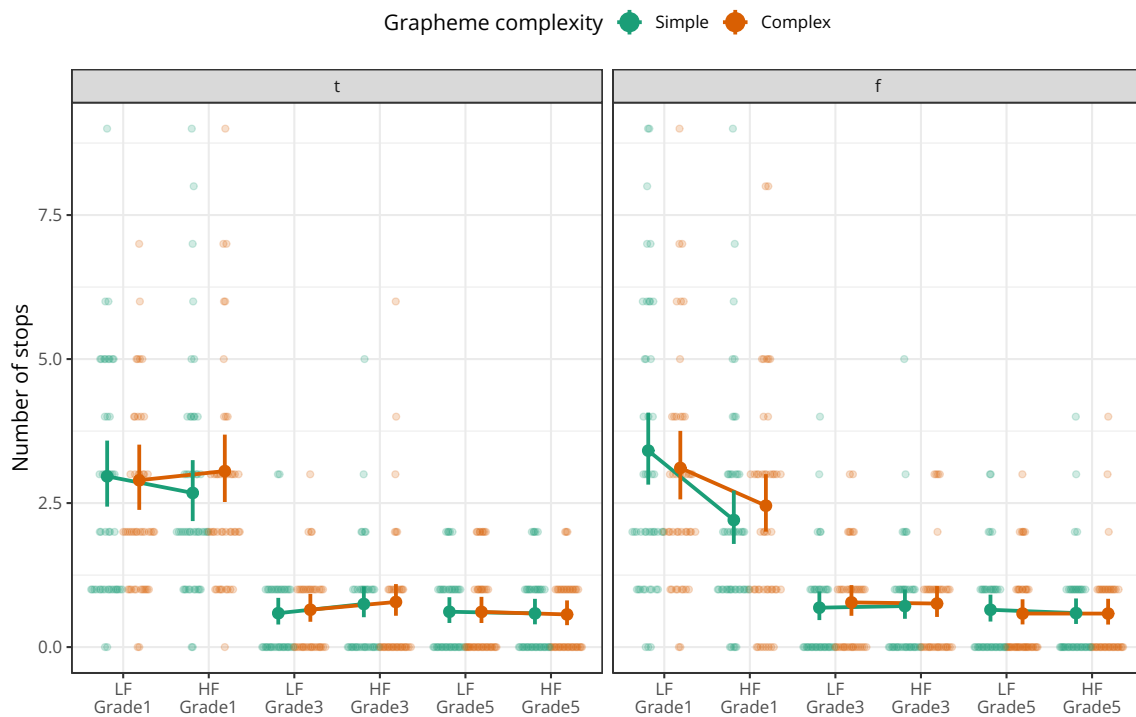


Figure 9. Number of stops by grade and word frequency (x-axis), grapheme complexity (in colour), and graphomotor difficulty (in panels). Transparent points represent individual data per participant. The surimposed dots and intervals represent the model’s predictions (median and 95% credible interval of the posterior distribution).

As can be seen in Figure 9, the model most predicts an interaction between the effect of the word frequency and the effect of first-letter graphomotor difficulty in Grade 1, with



infrequent words leading to a greater number of stops than frequent words for f (difficult letter) more than for t (easy letter) ( $\beta = -0.312$ , 95% CrI [-0.581, -0.032],  $BF_{10} = 3.307$ ). As can be seen from Table 3, others BFs favouring the alternative hypothesis (relative to the null hypothesis) are BFs for the difference between Grade 1 and Grade 3 ( $\beta = -1.384$ , 95% CrI [-1.529, -1.243],  $BF_{10} = 4.244 \times 10^{16}$ ), as well as between Grade 1 and Grade 5 ( $\beta = -1.556$ , 95% CrI [-1.709, -1.409],  $BF_{10} = 6.510 \times 10^{15}$ ), the effect of word frequency in Grade 1 ( $\beta = -0.18$ , 95% CrI [-0.324, -0.034],  $BF_{10} = 2.746$ ), and Grade 3 ( $\beta = 0.292$ , 95% CrI [0.019, 0.568],  $BF_{10} = 2.537$ ). Predictions from this model for each condition are also summarised in Table 4.

Table 4  
*Estimated number of stops in each condition.*

Group	Frequency	Grapheme complexity	Graphomotor difficulty	Estimate	MAD	Lower	Upper
Grade1	LF	Simple	t	2.964	0.275	2.438	3.585
Grade1	LF	Simple	f	3.411	0.302	2.820	4.070
Grade1	LF	Complex	t	2.897	0.272	2.382	3.515
Grade1	LF	Complex	f	3.110	0.289	2.565	3.753
Grade1	HF	Simple	t	2.676	0.258	2.187	3.246
Grade1	HF	Simple	f	2.205	0.227	1.791	2.716
Grade1	HF	Complex	t	3.055	0.283	2.519	3.689
Grade1	HF	Complex	f	2.454	0.240	2.004	3.002
Grade3	LF	Simple	t	0.589	0.114	0.392	0.855
Grade3	LF	Simple	f	0.684	0.126	0.469	0.967
Grade3	LF	Complex	t	0.648	0.120	0.440	0.923
Grade3	LF	Complex	f	0.775	0.135	0.543	1.075
Grade3	HF	Simple	t	0.748	0.133	0.518	1.053
Grade3	HF	Simple	f	0.712	0.127	0.491	1.001
Grade3	HF	Complex	t	0.783	0.138	0.544	1.093
Grade3	HF	Complex	f	0.756	0.134	0.524	1.059
Grade5	LF	Simple	t	0.613	0.113	0.419	0.866
Grade5	LF	Simple	f	0.648	0.116	0.443	0.910
Grade5	LF	Complex	t	0.611	0.115	0.418	0.872
Grade5	LF	Complex	f	0.582	0.110	0.394	0.830
Grade5	HF	Simple	t	0.585	0.112	0.395	0.838
Grade5	HF	Simple	f	0.590	0.111	0.402	0.844
Grade5	HF	Complex	t	0.567	0.109	0.380	0.811
Grade5	HF	Complex	f	0.582	0.112	0.391	0.838

*Note.* For each condition, the 'Estimate' and 'MAD' columns contain the median and the median absolute deviation (MAD) of the posterior distribution, respectively. The 'Lower' and 'Upper' columns contain the lower and upper bounds of the 95% credible interval.

For exploratory purposes, we depict below the posterior distribution of the difference between high-frequency and low-frequency words (i.e., the effect of frequency) separately for each letter (graphomotor difficulty) and each grade. We averaged the predictions across both conditions of graphemic complexity, as this effect appeared to be null.

As can be seen in Figure 10, the posterior distribution for the effect of frequency is almost perfectly centred on zero in all conditions, except for letter f in Grade 1. In this condition, the 95% credible interval excludes 0 and there is a 0.98 probability that the effect

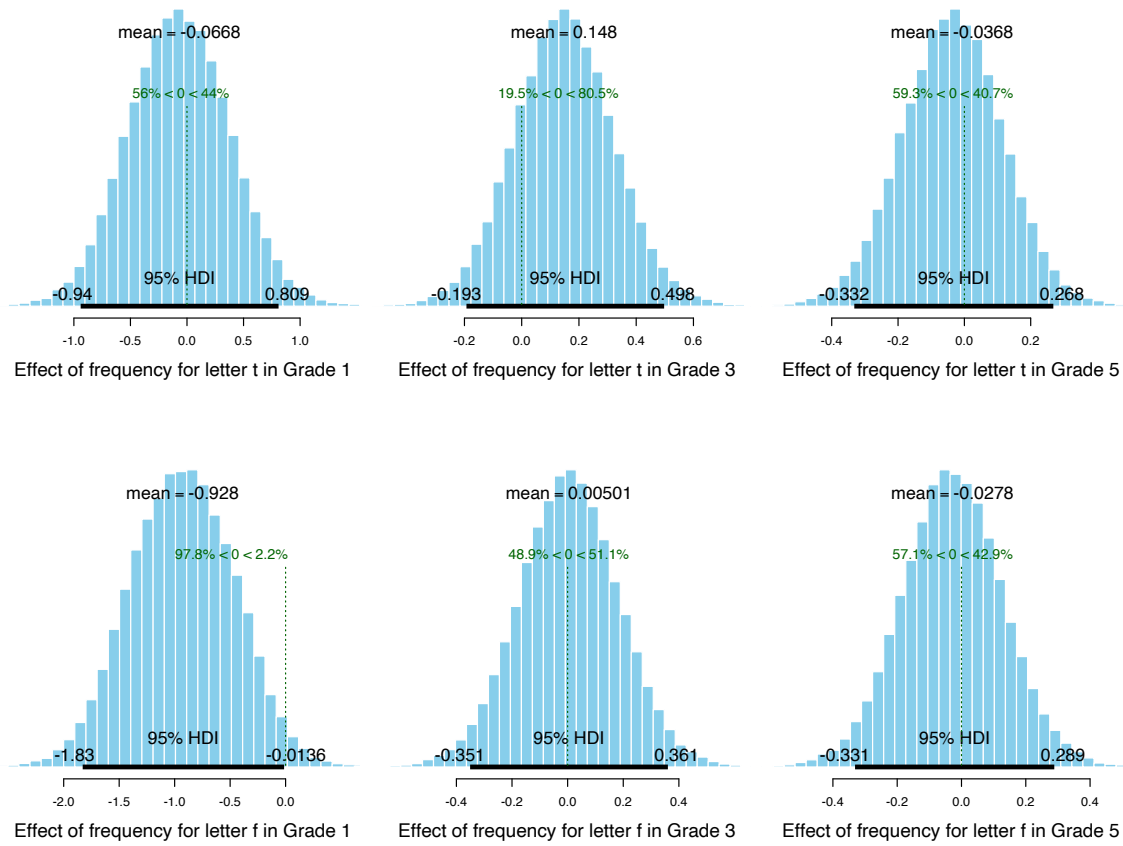


Figure 10. Effect of word frequency on the number of stops for each grade (in column) and letter (in row). The histogram contains posterior samples for each effect, where the posterior distribution is summarised by its mean and 95% highest density interval (HDI). The green text indicates the probability that the parameter values is either inferior or superior to 0.

of frequency on the number of stops is negative (given the data and the priors).

### 3.3 Mean velocity

Table 5 reports the estimates (median of the posterior distribution) and associated 95% credible intervals and BFs for all parameters regarding the mean velocity.

Table 5  
*Estimates and BF<sub>10</sub> for the slopes for the mean velocity.*

Term	Estimate	MAD	Lower	Upper	Rhat	BF10
Intercept	2.338	0.026	2.281	2.387	1.000	NA
groupGrade3	0.562	0.033	0.496	0.628	1.000	1.213 x 10 <sup>15</sup>
groupGrade5	0.793	0.033	0.729	0.859	1.000	1.859 x 10 <sup>16</sup>
frequency	0.049	0.046	-0.043	0.141	1.000	0.162
grapheme_complexity	0.036	0.047	-0.057	0.130	1.000	0.123
graphomotor_difficulty	0.159	0.046	0.066	0.252	1.000	16.037
groupGrade3:frequency	-0.063	0.065	-0.188	0.064	1.000	0.199
groupGrade5:frequency	-0.031	0.064	-0.156	0.095	1.000	0.142
groupGrade3:grapheme_complexity	-0.049	0.066	-0.177	0.080	1.000	0.169
groupGrade5:grapheme_complexity	-0.025	0.065	-0.151	0.100	1.000	0.137
frequency:grapheme_complexity	-0.051	0.089	-0.232	0.129	1.000	0.208
groupGrade3:graphomotor_difficulty	-0.025	0.066	-0.154	0.103	1.000	0.137
groupGrade5:graphomotor_difficulty	-0.031	0.064	-0.156	0.095	1.000	0.14
frequency:graphomotor_difficulty	0.158	0.091	-0.024	0.337	1.000	0.78
grapheme_complexity:graphomotor_difficulty	0.027	0.091	-0.154	0.206	1.000	0.189
groupGrade3:frequency:grapheme_complexity	0.015	0.125	-0.229	0.261	1.000	0.248
groupGrade5:frequency:grapheme_complexity	0.097	0.123	-0.144	0.341	1.000	0.342
groupGrade3:frequency:graphomotor_difficulty	-0.049	0.127	-0.295	0.198	1.000	0.268
groupGrade5:frequency:graphomotor_difficulty	-0.174	0.124	-0.412	0.069	1.000	0.635
groupGrade3:grapheme_complexity:graphomotor_difficulty	-0.036	0.125	-0.284	0.213	1.000	0.255
groupGrade5:grapheme_complexity:graphomotor_difficulty	-0.020	0.124	-0.261	0.224	1.000	0.25
frequency:grapheme_complexity:graphomotor_difficulty	0.070	0.166	-0.249	0.397	1.000	0.357
groupGrade3:frequency:grapheme_complexity:graphomotor_difficulty	0.044	0.227	-0.395	0.487	1.000	0.457
groupGrade5:frequency:grapheme_complexity:graphomotor_difficulty	-0.086	0.222	-0.523	0.351	1.000	0.465

*Note.* For each slope (for each line), the first two columns represent the estimated most probable value and its standard error (SE). The 'Lower' and 'Upper' columns contain the lower and upper bounds of the 95% CrI, whereas the 'Rhat' column reports the Gelman-Rubin statistic. The last column reports the Bayes factor in favour of the alternative hypothesis, relative to the null hypothesis (BF<sub>10</sub>).

These estimations are better understood visually. Thus, we plot the predictions of this model against raw data in Figure 11.

```
# retrieving the model's predictions
velocity_predictions <- df2 %>%
  data_grid(graphomotor_difficulty, grapheme_complexity, frequency, group) %>%
  cbind(., fitted(
    object = mod2, newdata = ., resp = "meanvelocity",
    scale = "response", probs = c(0.025, 0.975),
    re_formula = NA, robust = TRUE
  ) ) %>%
  ungroup %>%
  dplyr::rename(estimate = Estimate, mad = Est.Error, lower = Q2.5, upper = Q97.5)
```

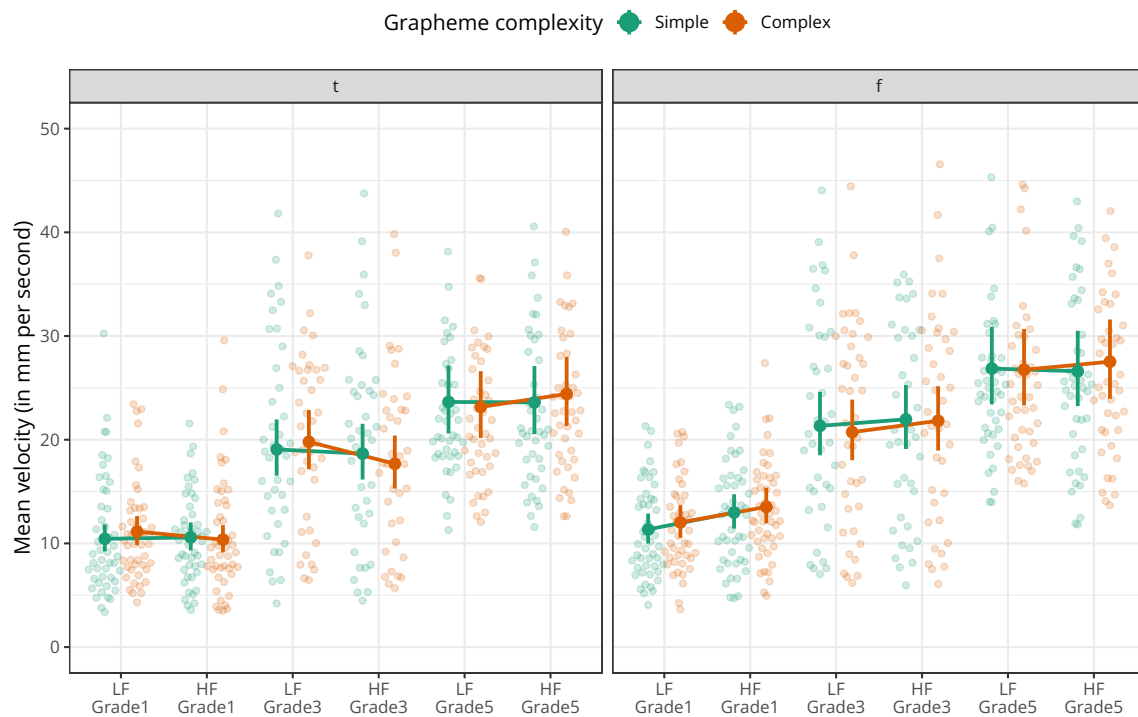


Figure 11. Mean velocity by grade and word frequency (x-axis), grapheme complexity (in colour), and graphomotor difficulty (in panels). Transparent points represent individual data per participant. The surimposed dots and intervals represent the model’s predictions (median and 95% credible interval of the posterior distribution).

As can be seen in Figure 11, the model most notably predicts higher velocity for the difficult letter f as compared to the easy letter t, excepted for low frequency words in Grade

1. First graders seem to have lower velocity than third graders, who themselves seem to have lower velocity than fifth graders on average. As can be seen from Table 5, BFs favouring the alternative hypothesis (relative to the null hypothesis) are BFs for the difference between Grade 1 and Grade 3 ( $\beta = 0.562$ , 95% CrI [0.496, 0.628],  $BF_{10} = 1.213 \times 10^{15}$ ), as well as between Grade 1 and Grade 5 ( $\beta = 0.793$ , 95% CrI [0.729, 0.859],  $BF_{10} = 1.859 \times 10^{16}$ ), and the effect of graphomotor difficulty in Grade 1 ( $\beta = 0.159$ , 95% CrI [0.066, 0.252],  $BF_{10} = 16.037$ ). Predictions from this model for each condition are also summarised in Table 6.

Table 6

*Estimated mean velocity in each condition.*

Group	Frequency	Grapheme complexity	Graphomotor difficulty	Estimate	MAD	Lower	Upper
Grade1	LF	Simple	t	10.439	0.649	9.228	11.835
Grade1	LF	Simple	f	11.350	0.730	9.992	12.881
Grade1	LF	Complex	t	11.140	0.708	9.836	12.636
Grade1	LF	Complex	f	12.020	0.793	10.538	13.677
Grade1	HF	Simple	t	10.576	0.676	9.321	12.025
Grade1	HF	Simple	f	12.975	0.831	11.427	14.742
Grade1	HF	Complex	t	10.362	0.661	9.131	11.763
Grade1	HF	Complex	f	13.521	0.855	11.928	15.358
Grade3	LF	Simple	t	19.061	1.367	16.534	21.957
Grade3	LF	Simple	f	21.336	1.525	18.496	24.615
Grade3	LF	Complex	t	19.793	1.427	17.149	22.855
Grade3	LF	Complex	f	20.730	1.487	18.022	23.867
Grade3	HF	Simple	t	18.659	1.332	16.159	21.537
Grade3	HF	Simple	f	21.965	1.535	19.107	25.269
Grade3	HF	Complex	t	17.680	1.288	15.295	20.402
Grade3	HF	Complex	f	21.822	1.544	18.957	25.160
Grade5	LF	Simple	t	23.643	1.627	20.601	27.133
Grade5	LF	Simple	f	26.864	1.876	23.411	30.888
Grade5	LF	Complex	t	23.158	1.610	20.167	26.606
Grade5	LF	Complex	f	26.743	1.868	23.308	30.681
Grade5	HF	Simple	t	23.599	1.649	20.547	27.100
Grade5	HF	Simple	f	26.604	1.844	23.236	30.499
Grade5	HF	Complex	t	24.416	1.695	21.327	27.982
Grade5	HF	Complex	f	27.526	1.937	23.944	31.594

*Note.* For each condition, the 'Estimate' and 'MAD' columns contain the median and the median absolute deviation (MAD) of the posterior distribution, respectively. The 'Lower' and 'Upper' columns contain the lower and upper bounds of the 95% credible interval.

For exploratory purposes, we depict below the posterior distribution of the difference between high-frequency and low-frequency words (i.e., the effect of frequency) separately for each letter (graphomotor difficulty) and each grade. We averaged the predictions across both conditions of graphemic complexity, as this effect appeared to be null.

As can be seen in Figure 12, the posterior distribution for the effect of frequency is almost perfectly centred on zero in all conditions, except for letter f in Grade 1. In this condition, there is a 0.92 probability that the effect of frequency on the mean velocity (in mm per second) is positive (given the data and the priors).

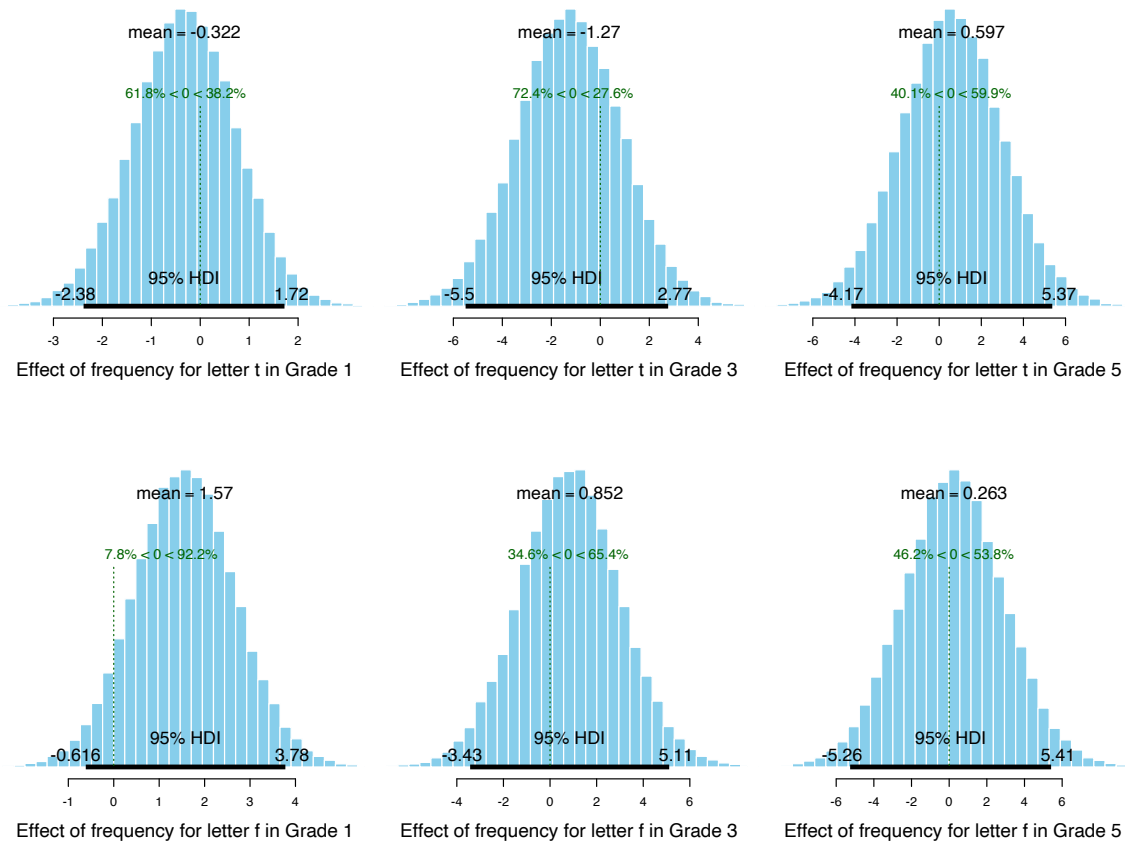


Figure 12. Effect of word frequency on the mean velocity (in mm per second) for each grade (in column) and letter (in row). The histogram contains posterior samples for each effect, where the posterior distribution is summarised by its mean and 95% highest density interval (HDI). The green text indicates the probability that the parameter values is either inferior or superior to 0.

### 3.4 Letter size

Table 7 reports the estimates (median of the posterior distribution) and associated 95% credible intervals and BFs for all parameters regarding the letter size.

Table 7  
*Estimates and BF's for the slopes for the letter size.*

Term	Estimate	MAD	Lower	Upper	Rhat	BF10
Intercept	2.289	0.023	2.241	2.333	1.000	NA
groupGrade3	-0.125	0.029	-0.181	-0.070	1.000	$7.090 \times 10^3$
groupGrade5	0.006	0.028	-0.049	0.061	1.000	0.059
frequency	0.023	0.042	-0.059	0.107	1.000	0.097
grapheme_complexity	0.032	0.042	-0.051	0.116	1.000	0.115
graphomotor_difficulty	0.703	0.042	0.618	0.788	1.000	$-9.629 \times 10^{16}$
groupGrade3:frequency	-0.021	0.057	-0.133	0.090	1.000	0.124
groupGrade5:frequency	-0.004	0.055	-0.112	0.105	1.000	0.112
groupGrade3:grapheme_complexity	-0.027	0.056	-0.138	0.085	1.000	0.126
groupGrade5:grapheme_complexity	-0.012	0.055	-0.121	0.097	1.000	0.111
frequency:grapheme_complexity	0.023	0.081	-0.139	0.186	1.000	0.17
groupGrade3:graphomotor_difficulty	-0.115	0.057	-0.225	-0.003	1.000	0.843
groupGrade5:graphomotor_difficulty	-0.193	0.055	-0.303	-0.084	1.000	42.547
frequency:graphomotor_difficulty	0.074	0.081	-0.090	0.235	1.000	0.251
grapheme_complexity:graphomotor_difficulty	0.022	0.081	-0.142	0.186	1.000	0.169
groupGrade3:frequency:grapheme_complexity	0.004	0.109	-0.208	0.220	1.000	0.222
groupGrade5:frequency:grapheme_complexity	-0.048	0.108	-0.260	0.164	1.000	0.237
groupGrade3:frequency:graphomotor_difficulty	-0.041	0.110	-0.257	0.177	1.000	0.237
groupGrade5:frequency:graphomotor_difficulty	-0.102	0.108	-0.316	0.108	1.000	0.334
groupGrade3:grapheme_complexity:graphomotor_difficulty	0.012	0.111	-0.207	0.228	1.000	0.228
groupGrade5:grapheme_complexity:graphomotor_difficulty	0.030	0.109	-0.185	0.244	1.000	0.229
frequency:grapheme_complexity:graphomotor_difficulty	0.053	0.150	-0.248	0.352	1.000	0.324
groupGrade3:frequency:grapheme_complexity:graphomotor_difficulty	-0.025	0.204	-0.423	0.370	1.000	0.408
groupGrade5:frequency:grapheme_complexity:graphomotor_difficulty	0.023	0.201	-0.369	0.414	1.000	0.409

*Note.* For each slope (for each line), the first two columns represent the estimated most probable value and its standard error (SE). The 'Lower' and 'Upper' columns contain the lower and upper bounds of the 95% CrI, whereas the 'Rhat' column reports the Gelman-Rubin statistic. The last column reports the Bayes factor in favour of the alternative hypothesis, relative to the null hypothesis (BF10).

These estimations are better understood visually. Thus, we plot the predictions of this model against raw data in Figure 13.

```
# retrieving the model's predictions
size_predictions <- df2 %>%
  data_grid(graphomotor_difficulty, grapheme_complexity, frequency, group) %>%
  cbind(., fitted(
    object = mod2, newdata = ., resp = "lettersize",
    scale = "response", probs = c(0.025, 0.975),
    re_formula = NA, robust = TRUE
  ) ) %>%
  ungroup %>%
  dplyr::rename(estimate = Estimate, mad = Est.Error, lower = Q2.5, upper = Q97.5)
```

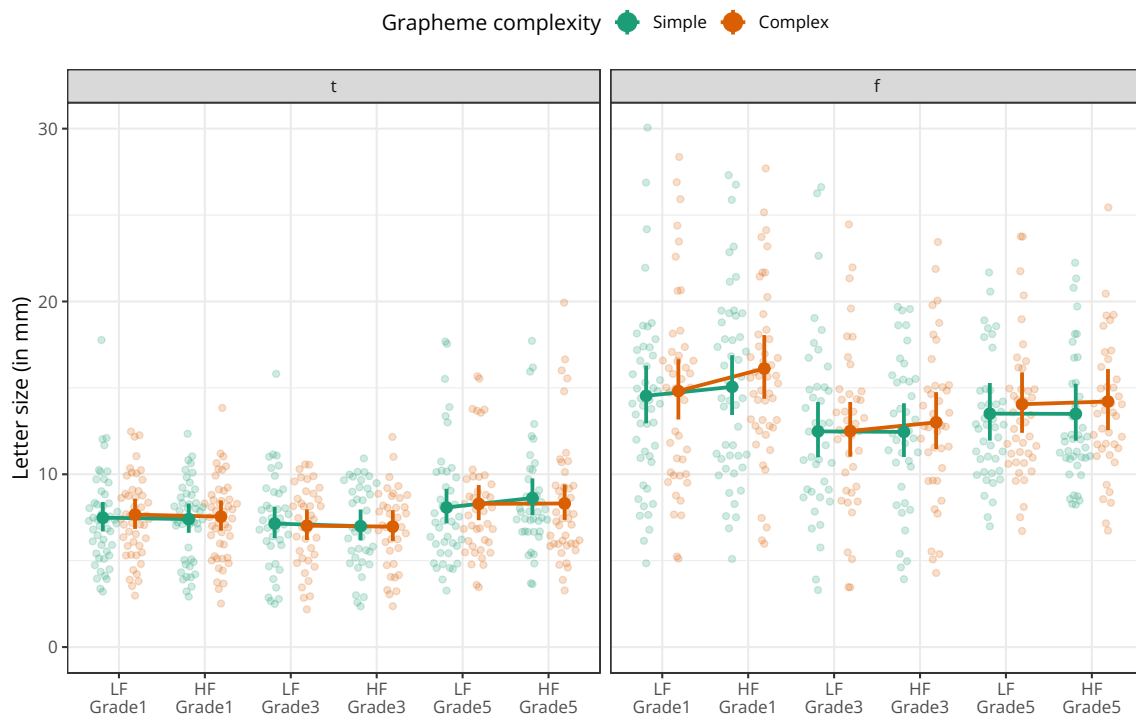


Figure 13. Letter size by grade and word frequency (x-axis), grapheme complexity (in colour), and graphomotor difficulty (in panels). Transparent points represent individual data per participant. The surimposed dots and intervals represent the model’s predictions (median and 95% credible interval of the posterior distribution).

As can be seen in Figure 13, the production of difficult letters was associated with greater letter size than the production of easy letters for all grades. As can be seen from



Table 7, BFs favouring the alternative hypothesis (relative to the null hypothesis) are BFs for the difference between Grade 1 and Grade 3 ( $\beta = -0.125$ , 95% CrI [-0.181, -0.07],  $\text{BF}_{10} = 7.090 \times 10^3$ ), the effect of graphomotor difficulty in Grade 1 ( $\beta = 0.703$ , 95% CrI [0.618, 0.788],  $\text{BF}_{10} = -9.629 \times 10^{16}$ ), and the effect of graphomotor difficulty in Grade 5 ( $\beta = -0.193$ , 95% CrI [-0.303, -0.084],  $\text{BF}_{10} = 42.547$ ). Predictions from this model for each condition are also summarised in Table 8.

Table 8

*Estimated letter size in each condition.*

Group	Frequency	Grapheme complexity	Graphomotor difficulty	Estimate	MAD	Lower	Upper
Grade1	LF	Simple	t	7.489	0.427	6.691	8.391
Grade1	LF	Simple	f	14.538	0.837	12.946	16.292
Grade1	LF	Complex	t	7.665	0.436	6.838	8.586
Grade1	LF	Complex	f	14.815	0.872	13.169	16.659
Grade1	HF	Simple	t	7.405	0.429	6.614	8.310
Grade1	HF	Simple	f	15.058	0.852	13.431	16.889
Grade1	HF	Complex	t	7.550	0.430	6.735	8.483
Grade1	HF	Complex	f	16.120	0.922	14.372	18.056
Grade3	LF	Simple	t	7.147	0.456	6.305	8.115
Grade3	LF	Simple	f	12.482	0.793	10.983	14.186
Grade3	LF	Complex	t	7.022	0.449	6.197	7.969
Grade3	LF	Complex	f	12.503	0.787	11.021	14.182
Grade3	HF	Simple	t	6.998	0.434	6.180	7.962
Grade3	HF	Simple	f	12.454	0.786	11.002	14.111
Grade3	HF	Complex	t	6.972	0.446	6.139	7.925
Grade3	HF	Complex	f	13.006	0.832	11.461	14.752
Grade5	LF	Simple	t	8.081	0.498	7.157	9.157
Grade5	LF	Simple	f	13.509	0.824	11.949	15.281
Grade5	LF	Complex	t	8.288	0.506	7.343	9.378
Grade5	LF	Complex	f	14.054	0.873	12.399	15.896
Grade5	HF	Simple	t	8.625	0.525	7.638	9.761
Grade5	HF	Simple	f	13.495	0.827	11.942	15.235
Grade5	HF	Complex	t	8.310	0.515	7.359	9.412
Grade5	HF	Complex	f	14.209	0.884	12.551	16.088

*Note.* For each condition, the 'Estimate' and 'MAD' columns contain the median and the median absolute deviation (MAD) of the posterior distribution, respectively. The 'Lower' and 'Upper' columns contain the lower and upper bounds of the 95% credible interval.

For exploratory purposes, we depict below the posterior distribution of the difference between high-frequency and low-frequency words (i.e., the effect of frequency) separately for each letter (graphomotor difficulty) and each grade. We averaged the predictions across both conditions of graphemic complexity, as this effect appeared to be null.

As can be seen in Figure 14, the posterior distribution for the effect of frequency is almost perfectly centred on zero in all conditions, except for letter f in Grade 1. In this condition, there is a 0.76 probability that the effect of frequency on the letter size (in mm) is positive (given the data and the priors).

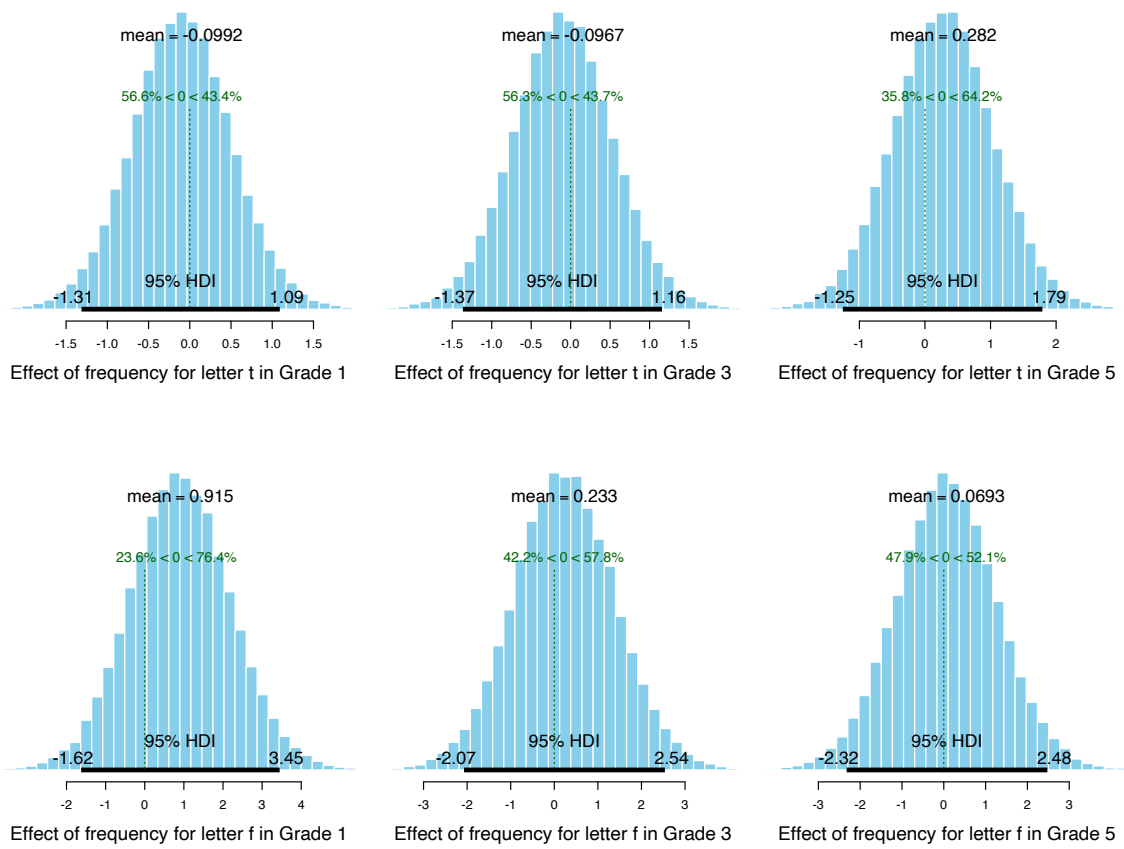


Figure 14. Effect of word frequency on the letter size (in mm) for each grade (in column) and letter (in row). The histogram contains posterior samples for each effect, where the posterior distribution is summarised by its mean and 95% highest density interval (HDI). The green text indicates the probability that the parameter values is either inferior or superior to 0.

#### 4 Acknowledgements

Acknowledgements will be included in the final version of the supplementary materials.

## 5 Session information

```
sessionInfo()
```

```
## R version 4.1.1 (2021-08-10)
## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS Big Sur 10.16
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/4.1/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.1/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats graphics grDevices utils datasets methods base
##
## other attached packages:
## [1] brms_2.16.2 Rcpp_1.0.7 BEST_0.5.3 HDInterval_0.2.2
## [5] glue_1.4.2 knitr_1.36 papaja_0.1.0.9997 readxl_1.3.1
## [9] GGally_2.1.2 modelr_0.1.8 tidybayes_3.0.1 posterior_1.1.0
## [13] patchwork_1.1.1 forcats_0.5.1 stringr_1.4.0 dplyr_1.0.7
## [17] purrr_0.3.4 readr_2.0.2 tidyr_1.1.4 tibble_3.1.5
## [21] tidyverse_1.3.1 ggbeeswarm_0.6.0 ggplot2_3.3.5 extraDistr_1.9.1
##
## loaded via a namespace (and not attached):
## [1] backports_1.3.0 plyr_1.8.6 igraph_1.2.6
## [4] splines_4.1.1 svUnit_1.0.6 crosstalk_1.1.1
## [7] rstantools_2.1.1 inline_0.3.19 digest_0.6.28
## [10] htmltools_0.5.2 rsconnect_0.8.24 fansi_0.5.0
## [13] magrittr_2.0.1 checkmate_2.0.0 tzdb_0.1.2
## [16] RcppParallel_5.1.4 matrixStats_0.61.0 xts_0.12.1
## [19] prettyunits_1.1.1 colorspace_2.0-2 rvest_1.0.1
## [22] ggdist_3.0.0 haven_2.4.3 xfun_0.26
## [25] callr_3.7.0 crayon_1.4.2 jsonlite_1.7.2
## [28] lme4_1.1-27.1 zoo_1.8-9 gtable_0.3.0
```

```
## [31] emmeans_1.7.0      V8_3.4.2      distributional_0.2.2
## [34] pkgbuild_1.2.0      rstan_2.26.3  abind_1.4-5
## [37] scales_1.1.1        mvtnorm_1.1-3 DBI_1.1.1
## [40] miniUI_0.1.1.1      xtable_1.8-4  stats4_4.1.1
## [43] StanHeaders_2.26.3  DT_0.19       htmlwidgets_1.5.4
## [46] httr_1.4.2          threejs_0.3.3 arrayhelpers_1.1-0
## [49] RColorBrewer_1.1-2  ellipsis_0.3.2 pkgconfig_2.0.3
## [52] reshape_0.8.8       loo_2.4.1     farver_2.1.0
## [55] dbplyr_2.1.1        utf8_1.2.2    tidyselect_1.1.1
## [58] rlang_0.4.12        reshape2_1.4.4 later_1.3.0
## [61] munsell_0.5.0       cellranger_1.1.0 tools_4.1.1
## [64] cli_3.1.0           generics_0.1.0 broom_0.7.9
## [67] ggirdges_0.5.3      evaluate_0.14 fastmap_1.1.0
## [70] yaml_2.2.1          processx_3.5.2 fs_1.5.0
## [73] nlme_3.1-152        projpred_2.0.2 mime_0.12
## [76] xml2_1.3.2          compiler_4.1.1 bayesplot_1.8.1
## [79] shinythemes_1.2.0   rstudioapi_0.13 gamm4_0.2-6
## [82] beeswarm_0.4.0      curl_4.3.2    reprex_2.0.1
## [85] stringi_1.7.5       ps_1.6.0      Brodbingnag_1.2-6
## [88] lattice_0.20-44     Matrix_1.3-4  nloptr_1.2.2.2
## [91] markdown_1.1        shinyjs_2.0.0 tensorA_0.36.2
## [94] vctrs_0.3.8         pillar_1.6.4  lifecycle_1.0.1
## [97] bridgesampling_1.1-2 estimability_1.3 httpuv_1.6.3
## [100] R6_2.5.1            bookdown_0.24 promises_1.2.0.1
## [103] gridExtra_2.3       vipor_0.4.5   rjags_4-11
## [106] codetools_0.2-18    boot_1.3-28   MASS_7.3-54
## [109] colourpicker_1.1.1  gtools_3.9.2  assertthat_0.2.1
## [112] withr_2.4.2         shinystan_2.5.0 mgcv_1.8-36
## [115] parallel_4.1.1      hms_1.1.1     grid_4.1.1
## [118] minqa_1.2.4         coda_0.19-4   rmarkdown_2.11
## [121] shiny_1.7.1         lubridate_1.8.0 base64enc_0.1-3
## [124] dygraphs_1.1.1.6
```

### References

- Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown*. Retrieved from <https://github.com/crsh/papaja>
- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology*, *67*, 641–666. <https://doi.org/10.1146/annurev-psych-122414-033645>
- Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization in bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *182*(2), 389–402. <https://doi.org/10.1111/rssa.12378>
- Gelman, A., Vehtari, A., Simpson, D., Margossian, C. C., Carpenter, B., Yao, Y., ... Modrák, M. (2020). Bayesian workflow. *arXiv:2011.01808 [Stat]*. Retrieved from <http://arxiv.org/abs/2011.01808>
- R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the savage–dickey method. *Cognitive Psychology*, *60*(3), 158–189. <https://doi.org/10.1016/j.cogpsych.2009.12.001>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. <https://doi.org/10.21105/joss.01686>
- Winter, B., & Bürkner, P.-C. (2021). *Poisson regression for linguists: A tutorial introduction to modeling count data with brms*. preprint, preprint. <https://doi.org/10.31219/osf.io/93kaf>
- Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Boca Raton, Florida: Chapman; Hall/CRC. Retrieved from <https://yihui.org/knitr/>