



**HAL**  
open science

## The Dynamics of Multiword Sequence Extraction

Leonardo Pinto Arata, Laura Ordonez Magro, Carlos Ramisch, Jonathan Grainger, Arnaud Rey

► **To cite this version:**

Leonardo Pinto Arata, Laura Ordonez Magro, Carlos Ramisch, Jonathan Grainger, Arnaud Rey. The Dynamics of Multiword Sequence Extraction. Quarterly Journal of Experimental Psychology, 2024, 10.1177/17470218241228548 . hal-04409016

**HAL Id: hal-04409016**

**<https://hal.science/hal-04409016>**

Submitted on 22 Jan 2024

**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.

1  
2  
3  
4  
5  
6 The Dynamics of Multiword Sequence Extraction  
7  
8  
9

10 Leonardo Pinto Arata<sup>1,2,3</sup>, Laura Ordonez Magro<sup>1,4</sup>, Carlos Ramisch<sup>2,3</sup>,  
11 Jonathan Grainger<sup>1,2</sup>, and Arnaud Rey<sup>1,2</sup>  
12

13  
14 <sup>1</sup> Aix-Marseille Université, CNRS, LPC, Marseille, France

15 <sup>2</sup> Institute of Language Communication and the Brain, Aix-Marseille Université, France

16 <sup>3</sup> Aix-Marseille Université, Université de Toulon, CNRS, LIS, Marseille, France

17 <sup>4</sup> Psychological Sciences Research Institute, Université catholique de Louvain, Louvain-la-  
18 Neuve, Belgium  
19

20  
21  
22  
23  
24 Running head: MULTIWORD SEQUENCES  
25  
26  
27  
28  
29  
30

31 Corresponding author:

32 Leonardo Pinto Arata

33 Laboratoire de Psychologie Cognitive

34 CNRS – Aix-Marseille Université

35 3, place Victor Hugo – Case D

36 13331 Marseille Cedex 3 - France

37 E-mail: leonardo.pinto-arata@univ-amu.fr  
38

39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57

**Abstract**

Being able to process multiword sequences is central for both language comprehension and production. Numerous studies support this claim, but less is known about the way multiword sequences are acquired, and more specifically how associations between their constituents are established over time. Here we adapted Rey et al.'s (2020) Hebb naming task into a Hebb lexical decision task to study the dynamics of multiword sequence extraction. Participants had to read letter strings presented on a computer screen and were required to classify them as words or pseudowords. Unknown to the participants, a triplet of words or pseudowords systematically appeared in the same order and random words or pseudowords were inserted between two repetitions of the triplet. We found that RTs for the unpredictable first position in the triplet decreased over repetitions (i.e., indicating the presence of a *repetition* effect) but more slowly and with a different dynamic compared to items appearing at the predictable second and third positions in the repeated triplet (i.e., showing a slightly different *predictability* effect). Implicit and explicit learning also varied as a function of the nature of the triplet (i.e., unrelated words, pseudowords, semantically related words, or idioms). Overall, these results provide new empirical evidence about the dynamics of multiword sequence extraction, and more generally about the role of statistical learning in language acquisition.

*Keywords:* language processing, multiword chunking, statistical learning

59  
60

## Introduction

61 Humans are constantly exposed to and produce an unlimited number of novel  
62 utterances and this generative ability has long been considered as a hallmark of human  
63 language. For decades, generative linguists have argued that this phenomenon is explained by  
64 an innate system of abstract grammatical rules known as the “universal grammar hypothesis”  
65 (e.g., Chomsky, 1957). Distinct cognitive abilities supported by different neural systems may  
66 allow people to generate complex utterances (Ullman et al., 2005). For example, a mental  
67 lexicon including simple linguistic forms (e.g., individual words, morphemes) combined with  
68 a mental grammar including combinatorial rules would enable the formation of an infinite  
69 number of sentences (Pinker, 1991; Pinker & Ullman, 2002).

70 More recently, usage-based approaches to language have provided an alternative view  
71 to account for the mechanisms involved in language acquisition (e.g., Croft, 2001; Goldberg,  
72 2006; Tomasello, 2003). According to this view, language gradually emerges through the  
73 interaction between general cognitive mechanisms and the repeated exposure to concrete  
74 items (Ibbotson, 2013). Learners are thought to store incoming utterances and to generate  
75 knowledge about the properties of these utterances (e.g., grammatical categories, semantics)  
76 by generalising over these stored multiword sequences (Abbot-Smith & Tomasello, 2006).

77 Over the last two decades, this approach has received multiple computational  
78 implementations to illustrate this learning and generalisation process. For instance, Solan et  
79 al. (2005) developed an algorithm (ADIOS for automatic distillation of structure) capable of  
80 generalising over different kinds of sentences from a given corpus using the statistical  
81 information present in the same data. In the same vein, Borensztajn et al. (2009) used an  
82 automatic data-oriented parsing procedure to identify the most likely multiword sequences  
83 used in child speech and model the evolution of their abstractness over time. Similarly,  
84 Meylan et al. (2017) developed a Bayesian statistical model to study the contribution of

85 language productivity and abstractness to children’s linguistic knowledge by focusing on their  
86 early capacity to use the determiners “a” and “the” along with a noun.

87         Whilst these computational modelling studies have successfully captured multiword  
88 learning process, the emergence of grammatical knowledge and different developmental  
89 patterns more broadly, their reliance on mathematical algorithms and comprehensive corpus  
90 analysis undermines their psychological plausibility, as they lack realistic learning  
91 mechanisms and memory constraints inherent to the real-time nature of language processing  
92 (e.g., Christiansen & Chater, 2016). Chunk-based models, on the other hand, rely on a simple  
93 but a powerful mechanism (i.e., associative learning) that can account for both memory  
94 constraints and language processing, ranging from single word segmentation (e.g., Perruchet  
95 & Vinter, 1998) to multiword sequence acquisition (e.g., Jones & Rowland, 2017). For  
96 instance, McCauley and Christiansen (2019) developed a computational model of language  
97 perception and production that assumes language acquisition takes place in an incremental  
98 fashion, through local shallow processes based on chunking and statistical learning  
99 mechanisms. Processing occurs on a word-by-word basis by assembling words into chunks  
100 (i.e., sequences of words), rather than via a full syntactic analysis as assumed by generativist  
101 theories. Given that language perception and production are thought to be interwoven  
102 processes in this model, both are assumed to rely on the same chunks and distributional  
103 statistics learnt during language acquisition. Thereby, this model relies on a chunk-by-chunk  
104 process instead of whole-sentence optimization. Note that McCauley and Christiansen’s  
105 model is the first usage-based model having used a large number of natural language corpora  
106 (i.e., 79 single-child corpora for perception and 200 for production evaluation, representing a  
107 total of 29 languages).

108         In line with McCauley and Christiansen’s (2019) model, numerous studies suggest that  
109 language users are sensitive to distributional properties at different levels of the linguistic

110 input, and that statistical learning plays a key role in language acquisition (e.g., Aslin, 2018;  
111 Conway et al., 2010; Saffran et al., 1996). For instance, word frequency is known to affect  
112 word recognition (e.g., Grainger, 1990) and speech production (e.g., Jescheniak & Levelt,  
113 1994). There is also evidence that linguistic processing is not only affected by word frequency  
114 but also by multiword frequency (Ambridge et al., 2015; Carrol & Conklin, 2020). In these  
115 studies, a multiword sequence is often defined as a number of consecutive words stored and  
116 retrieved from memory as a whole (Wray, 2002), acting as a single unit and resulting in a  
117 processing advantage (e.g., “How are you doing?”). It is worth noting, however, that it has  
118 also been suggested that this processing advantage could arise from either the simultaneous  
119 access to the component parts of a sequence, or from the priming of multiple combinations  
120 via the base components, rather than from storing the sequence as a whole (Wray, 2012,  
121 p.234).

122         Many developmental studies have also tested this hypothesis. For instance, Bannard  
123 and Matthews (2008) used a sentence repetition task and found that 2- and 3-year-old children  
124 are more likely to repeat frequent sentences correctly (e.g., *you want to play*) compared to less  
125 frequent ones (e.g., *you want to work*). Arnon and Clark (2011), showed that 4-year-olds are  
126 better at producing irregular plurals when presented in a familiar context (e.g., *On your feet*).  
127 In the same vein, Janssen and Barber (2012) found multiword frequency effects in adults’  
128 production latencies during a task where participants had to name drawings of noun and  
129 adjective pairs. Arnon and Snider (2010) also showed that comprehension is affected by  
130 multiword frequency. In a grammatical judgement task, adults processed frequent four-word  
131 phrases faster than less frequent ones, even when the frequency of the individual final words,  
132 bigrams and trigrams were controlled for. It is worth noting that sensitivity to statistical  
133 properties of multiword sequences seems to be present early on. Indeed, it has been shown  
134 that eleven- and 12-month-olds can already discriminate frequent multiword sequences from

135 infrequent ones (e.g., *take it off* vs. *shake it of*, Skarabela et al., 2021). Moreover, it has been  
136 demonstrated that multiword sequences acquired early in childhood are processed faster in  
137 adulthood (Arnon et al., 2017).

138 Similarly, written language abounds with distributional cues (Arciuli & Simpson,  
139 2012; Snell & Theeuwes, 2020; Treiman et al., 2014). Reading behaviour, for example, has  
140 also been shown to be influenced by the frequency and predictability of multiword phrases.  
141 For instance, frequent three-word binomial phrases (e.g., *black and white*) are read faster than  
142 their reversed forms (i.e., *white and black*) (Siyanova-Chanturia, Conklin, & van Heuven,  
143 2011) and idioms (e.g., *at the end of the day* – ‘ultimately’) are read faster than non-idiomatic  
144 structurally equivalent counterparts (e.g., *at the end of the war*) (Conklin & Schmitt, 2008;  
145 Siyanova-Chanturia, Conklin, & Schmitt, 2011).

146 In the past decades, research has mainly focused on isolated word learning (e.g.,  
147 Pelucchi et al., 2009; Perruchet & Vinter, 1998; Saffran et al., 1996, 1997), leaving aside the  
148 question of how multiword sequences are acquired in real-time. To date, only one study has  
149 addressed this issue in the context of first language acquisition. In an eye-tracking study,  
150 Conklin and Carrol (2020) presented participants with short stories containing existing  
151 English binomials in their canonical form (e.g., boys and girls), which were seen once, and  
152 novel binomials (e.g., goats and pigs), which were seen one to five times during the task.  
153 Participants were then presented with the existing and novel binomials in reverse (e.g., girls  
154 and boys, pigs and goats). They found that participants were sensitive to the co-occurrences of  
155 the novel binomials, which translated into faster reading times for the novel binomials as the  
156 number of co-occurrences increased. In addition, the results showed an advantage for forward  
157 novel binomials over their reverse forms after only four to five exposures, suggesting that  
158 participants very quickly detected and encoded the structure of the repeated pattern (see  
159 Sonbul et al., 2022, for a replication in second language acquisition).

160           Here, we propose to investigate how associations between multiword constituents  
161 other than binomials are established over time by using a visual lexical decision task. Based  
162 on the assumption that vocabulary acquisition and performance on the Hebb repetition  
163 learning paradigm (Hebb, 1961) are subserved by the same processes (Mosse & Jarrold, 2008;  
164 Norris et al., 2018; Page & Norris, 2009; Page et al., 2013; Smalle et al., 2016; Szmalec et al.,  
165 2009), we used an adaptation of Rey et al.'s (2020) Hebb letter naming task to study the  
166 learning dynamics of repeated words triplets.

167           In the original Hebb repetition task, participants had to recall sequences of digits  
168 where one particular sequence was repeated every third trial. Hebb (1961) found that  
169 participants' performance gradually improved for the repeated sequences compared to the  
170 non-repeated ones. In Rey et al. (2020), participants had to read aloud the names of single  
171 letters that were presented one at a time on a computer screen. Unknown to the participants, a  
172 triplet of letters (i.e., the Hebb sequence) was repeated with its constituent letters  
173 systematically presented in the same order. As in the standard Hebb learning paradigm,  
174 random letters (i.e., fillers) were inserted between two repetitions of the critical letter triplets.  
175 The extraction dynamics of the repeated triplet was tracked by looking at the evolution of  
176 response times (RTs) to the second and third letters of the triplet. RTs for these two letters  
177 decreased with repetition as they progressively became predictable when learning occurred.  
178 To study the extraction dynamics of multiword sequences in the present experiment, we  
179 replaced the triplet of letters used in Rey et al. (2020) by a triplet of words and instead of  
180 using a naming task, we used a lexical decision task hence simplifying online data collection  
181 and providing a better proxy for the silent reading that occupies the vast majority of skilled  
182 reading behaviour.

183           The reasons for using the Hebb paradigm to investigate multiword acquisition are two-  
184 fold. First, as the Hebb paradigm is an implicit learning measure, it allowed us to study the



185 extraction dynamics of multiword sequences in conditions where participants were not  
186 necessarily aware of the repetitions. Indeed, as participants are asked to read words without  
187 further instructions, knowledge of patterns of sequences can be attributed to implicit learning  
188 through regularity extraction. Second, it allowed us to study the online learning trajectory of  
189 multiword sequences rather than solely the “offline” end-product of what has been learned.  
190 Indeed, participants’ knowledge can be the same at the end of the task (offline knowledge),  
191 but their learning trajectories may differ (Siegelman et al., 2017). By using an online learning  
192 task, we sought to provide a comprehensive characterization of the process of word-to-word  
193 associative learning.

194         Measuring the evolution of response times for a repeated triplet of items also allowed  
195 us to study separately the *repetition* effect from the *predictability* effect. Indeed, because a  
196 random number of filler items occurred between two repetitions of the triplet, the first item in  
197 the triplet was not predictable and the evolution of RTs for this item can be considered as  
198 providing a good estimate of the repetition effect. In contrast, items occurring at Positions 2  
199 and 3 of the triplet benefit from the immediately preceding item that systematically occurs  
200 before them and that should help participants anticipating and predicting the next item.  
201 Previous studies in sequence learning (e.g., Minier et al., 2016; Rey et al., 2019, 2020, 2022)  
202 even reported a stronger predictability effect on the third item of the triplet (i.e., a greater  
203 decrease in RTs) due to the richer contextual information provided by the two previous items.  
204 This experimental paradigm therefore allowed us to study the differential effect of repetition  
205 and predictability on the memory trace of each item belonging to a repeated triplet and on the  
206 processing gains generated by these effects.

207         Note that the predictability effect is closely linked to chunking mechanisms since it  
208 reflects the emergent association between several words that appear repeatedly in a sequence.  
209 As previously mentioned, chunking mechanisms are also considered central to several models

210 of sequence learning and language acquisition (e.g., French et al., 2011; Jones & Rowland,  
211 2017; McCauley & Christiansen, 2019; Perruchet & Vinter, 1998; Robinet et al., 2011).  
212 However, less is known about the precise dynamics related to the repeated presentation of a  
213 sequence of words and empirical evidence is needed to constrain models that assume a central  
214 role for chunking mechanisms in the development of language processing skills. The present  
215 set of experiments has been designed to provide such empirical evidence about the dynamics  
216 of these fundamental associative learning mechanisms.

217 In the present study, the learning and chunking dynamics of repeated triplets was  
218 studied in four Hebb lexical decision experiments. In Experiment 1, the repeated word triplet  
219 was composed of three unrelated words. In Experiment 2, the repeated triplet was composed  
220 of three pseudowords in order to test if lexicality had an effect on the learning dynamics of  
221 the triplet. In Experiment 3, the repeated triplet was composed of three semantically related  
222 words in order to test if semantic relatedness would facilitate the development of word  
223 associations. In Experiment 4, the repeated triplet corresponded to an existing idiomatic  
224 expression to test if the learning trajectory of the repeated triplet would be facilitated by  
225 activating the pre-existing long-term memory representation of the triplet. These experiments  
226 were conducted remotely by using a platform for online experimentation that has been  
227 frequently used in experimental psychology to conduct experiments during the COVID-19  
228 pandemic (e.g., Fournet et al., 2022; Isbilen et al., 2022; Ordonez Magro et al., 2022). It is  
229 worth noting that recent research has shown that JavaScript-based online experiment  
230 platforms, such as LabVanced and PsychoJS, allow researchers to collect reliable data that  
231 replicate the findings of in-lab studies (e.g., Angele et al., 2022; Mirault et al., 2018).

## 232 **Experiment 1**

### 233 Methods

### 234 Participants

235 Forty-two participants (20 females;  $M_{age} = 24$  years,  $SD = 3$ ) were paid for taking part  
236 in the experiment via Prolific (www.prolific.co). All participants reported to be native French  
237 speakers, having no history of neurological or language impairment. Before starting the  
238 experiment, participants accepted an online informed-consent form. Ethics approval was  
239 obtained from the “Comité de Protection des Personnes SUD-EST IV” (17/051).

240 Given that participants were recruited online, their proficiency in French was  
241 measured with the LexTALE language proficiency test (Brysbaert, 2013) before starting the  
242 main task. This test consists of a lexical decision task with no time pressure where  
243 participants are presented with 84 single-item trials (56 real French words, 28 French-looking  
244 pseudowords), and are instructed to decide whether each presented letter sequence is a real  
245 French word or not. Their average LexTALE vocabulary score was 86.53% ( $SD = 5.76$ ). Any  
246 participant whose score was below 2.5 standard deviations from the average LexTALE  
247 vocabulary score was excluded from the analysis. No participant was excluded based on this  
248 criterion. The final dataset consisted of 1890 data points per condition, meeting the 1600  
249 measurements per condition recommendation from Brysbaert and Stevens (2018). A summary  
250 of the participants’ scores and standard deviations on the LexTALE task for each experiment  
251 is provided in Appendix A.

## 252 Materials

253 We adapted Rey et al. (2020)’s naming task into a lexical decision task. The task was  
254 composed of 3 blocks of 120 trials, each trial corresponding to the presentation of a single  
255 item (word or pseudoword) in the middle of the screen. A set of 66 words and 180  
256 pseudowords were used as items in this experiment. All words were monosyllabic or  
257 disyllabic singular nouns. They were composed of four-to-six letters and were selected from  
258 the French database Lexique 3.83 (New & Pallier, 2020). Each word of the triplet had a  
259 freqfilms2 frequency ranging from 2 to 10 occurrences per million. We decided to use low-

260 frequency words to maximise repetitions effects and increase the chances of revealing any  
261 processing differences between positions within the triplet. Indeed, low-frequency words  
262 elicit larger repetition effects compared to high-frequency words in lexical decision tasks  
263 (e.g., Scarborough et al., 1977). Filler words had a frequency ranging from 10 to 100  
264 occurrences per million. Pseudowords were drawn from the French Lexicon Project (Ferrand  
265 et al., 2010). They were monosyllabic or disyllabic and had a length from four to six letters.

266 A Latin-square design was used such that each word of the triplet appeared in every  
267 possible position within the triplet across participants, leading to six possible combinations of  
268 the same triplet of words (ABC, ACB, BAC, BCA, CAB, CBA). Seven triplets of words were  
269 used and were seen in one of the six possible combinations (for a total of  $7*3 = 21$  words).  
270 Each participant saw one triplet in a specific combination, leading to  $7*6 = 42$  participants  
271 (e.g., Participant 1 saw ABC while Participant 2 saw ACB instead throughout the task). Each  
272 triplet appeared 15 times per experimental block (resulting in a total of 45 repetitions across  
273 the 3 blocks) and was separated by three to six filler words or filler pseudowords (75 per  
274 block). Every block was composed of 60 words (the 15 repeated triplets, i.e., 45 words, and  
275 15 filler words) and 60 pseudowords. Therefore, there were an equal number of ‘yes’ and ‘no’  
276 responses in the experiment (i.e., 180 for each type of response). Among the 66 selected  
277 words, 21 served to construct the 7 triplets and 45 served as filler words during the  
278 experiment. The set of word triplets and fillers are listed in Appendix B.

279 In order to obtain more detailed information about participants’ explicit knowledge of  
280 the task, all participants responded to a short questionnaire after the experiment (similarly to  
281 Rey et al., 2020; Tosatto et al., 2022). The first question was: “Did you notice anything  
282 particular in this experiment?”, in case of a “Yes” response, the follow-up question was “Can  
283 you explain what you noticed?” If participants reported noticing the presentation of a repeated  
284 sequence of words, they were asked “Can you recall the words in their correct serial order?”.

285 If the answer to the first question was “No”, the following questions were displayed “Did you  
286 notice that a sequence of words was systematically repeated?” and “Can you recall the words  
287 in their correct serial order?”.

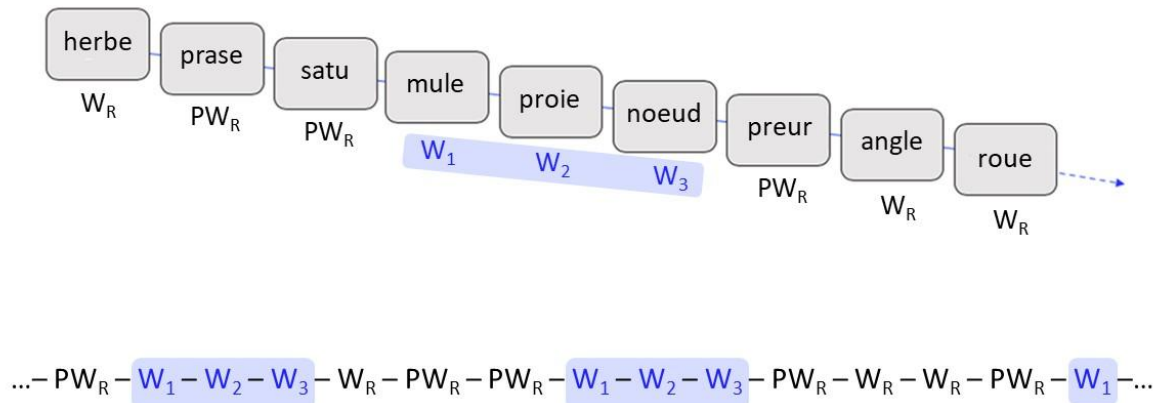
### 288 Apparatus

289 The experiment was implemented in LabVanced, an online experiment builder (Finger  
290 et al., 2017) and participants were recruited via the Prolific platform (www.prolific.co).

291 Participants participated via their personal computer and we made sure that the experiment  
292 would not work on smartphones or tablets in order to keep the testing conditions as similar as  
293 possible across participants. All words and pseudowords were presented in the centre of the  
294 computer screen using a 20-point Lato black font on a white background.

### 295 Procedure

296 Before the experiment, written instructions were displayed on the screen. Participants  
297 were instructed to decide as fast as possible whether the letter sequence displayed on the  
298 screen formed or not a French word. They were required to press “M” (for words) or “Q” (for  
299 pseudowords) on their keyboards (which are at extreme positions on the left and right of  
300 French AZERTY keyboards). RTs and accuracy were recorded for each word and  
301 pseudoword. Each target stayed on the screen until the participant’s response. Subsequently,  
302 the next target appeared immediately after the participant’s response. To encourage the  
303 participants, the number of remaining trials was displayed at the end of each block. The  
304 experiment lasted approximately 10 minutes. Figure 1 provides a schematic description of this  
305 experimental paradigm.



306

307 **Fig. 1.** Experimental procedure for the Hebb lexical decision paradigm. Upper part: items are  
 308 presented one at a time at the centre of the computer screen. Participants had to classify each  
 309 string as a word or a pseudoword. A repeated triplet of three words (e.g.,  $W_1$ : “mule” – mule ;  
 310  $W_2$ : “proie” – prey ;  $W_3$ : “noeud” – knot) always appearing in the same order was intermixed  
 311 with random filler words ( $W_R$ ) or random filler pseudowords ( $PW_R$ ). Words in blue belong to  
 312 the repeated triplet. Lower part: one triplet of words ( $W_1W_2W_3$ ) is repeated several times and  
 313 a variable number of random words or pseudowords ( $W_R$  or  $PW_R$ ) are presented between two  
 314 repetitions of the triplet.

### 315 Results

316 Only correct trials were analysed (97.06 % of the data), and we excluded RTs  
 317 exceeding 1500 ms (0.98 % of data) as well as RTs greater than 2.5 standard deviations above  
 318 a participant’s mean per block and for each of the three possible positions within the triplet  
 319 (2.47 %). The mean RTs and standard deviations computed over the entire sample and for  
 320 each block are presented in Table 1. Data analysis was performed with the R software  
 321 (version 4.2.1) using linear mixed-effects models (LMEs) fitted with the lmerTest (version  
 322 3.1-3; Kuznetsova et al., 2017) and the lme4 packages (version 1.1-29; Bates et al., 2015).

323 The model included the maximum random structure that allowed convergence (Barr, 2013;  
 324 Barr et al., 2013), that is, Position (1 to 3), Repetition (1 to 45) and their two-way interaction  
 325 as fixed effects, participant and Item sets were used as random effects. It is worth noting that  
 326 Position was coded using repeated contrast coding (i.e., Position 1: -0.7 -0.3; Position 2: 0.3 -  
 327 0.3; Position 3: 0.3 0.7) in order to perform pairwise comparisons (Schad et al., 2020), and  
 328 Repetition was mean centred here and in the following analyses. Word length and log-  
 329 transformed word frequency for each word in the triplet were included as covariates to control  
 330 for any word-level differences. Given that the distribution of RTs was close to normal and  
 331 provided good fit (established through visual inspection of QQ plots and histograms), no data  
 332 transformation was performed prior to the analysis. The results of the model are shown in  
 333 Table 2.

334 **Table 1**

335 Mean response times (in milliseconds) and standard deviations (in parentheses) for each block  
 336 and each position in Experiment 1.

	Block 1	Block 2	Block 3
Position 1	617 (100)	572 (95)	581 (120)
Position 2	530 (153)	410 (146)	358 (136)
Position 3	523 (148)	401 (126)	374 (135)

337

338 **Table 2**

339 Fixed effects of the mixed model for Experiment 1.

Predictors	Estimate	SE	95% CI	<i>p</i>
(Intercept)	450.75	74.30	[305.12, 596.37]	<.001
Position 2 - 1	-156.54	3.69	[-163.78, -149.31]	<.001
Position 3 - 2	-1.29	3.67	[-8.48, 5.90]	.725
Repetition	-3.95	0.12	[-4.17, -3.72]	<.001
Position 2 - 1 x Repetition	-4.63	0.28	[-5.19, -4.07]	<.001
Position 3 - 2 x Repetition	0.87	0.28	[0.32, 1.43]	.002
Word length	4.38	11.48	[-18.11, 26.88]	0.707
Word frequency (log)	5.58	15.90	[-25.58, 36.75]	0.730

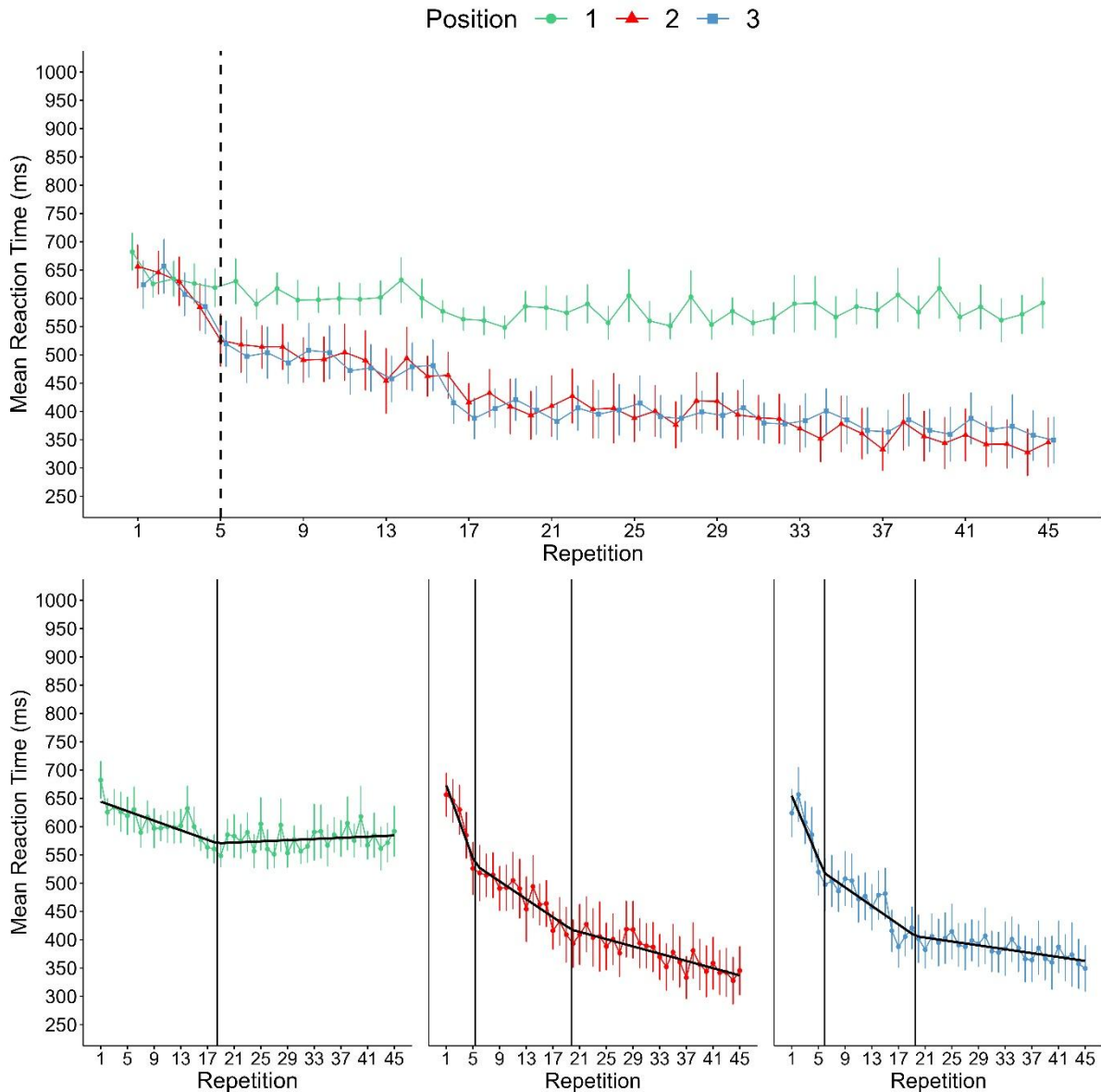
340 *Note.* CI: confidence interval; SE: standard error.

341 We found a significant effect of Repetition with an overall decrease of RTs across the  
342 experiment. As predicted, response times for Position 2 were significantly faster than those  
343 for Position 1, but they did not differ from Position 3. Moreover, there was a significant  
344 negative interaction coefficient for the difference between Position 2 and Position 1, and  
345 Repetition, and a significant positive interaction coefficient for the difference between  
346 Position 3 and Position 2, and Repetition, indicating that response time differences increased  
347 across repetitions. No significant effects were found for word length and word frequency. To  
348 investigate where the significant difference between Position 1 compared to Positions 2 and 3  
349 emerges, we ran a series of paired sample *t*-tests on the RTs for Position 1 and the average  
350 RTs for Positions 2 and 3 on each repetition of the triplet. We found that a significant  
351 difference emerged on the fifth trial,  $t(38) = 5.26$ , Bonferroni-adjusted  $p < .001$ .

352 To get a clearer picture of the learning dynamics for each position in the triplet of  
353 words, Figure 2 represents the evolution of the mean response times for each position in the  
354 triplet and for the successive 45 repetitions of the triplet. Given that linear regression only  
355 captures the overall change of Position across repetitions, we conducted a broken-stick linear  
356 regression, using the segmented package (version 1.6-0; Muggeo, 2008), in order to account  
357 for the evolution of the learning pattern across the task. In broken-stick regression, multiple  
358 linear regressions are fitted and connected at certain estimated values referred as breakpoints.  
359 At the breakpoint the relationship between the variables changes to model non-linear  
360 relationships between two variables. Thus, each position was regressed onto Repetition  
361 separately. To estimate the number of breakpoints for each position, a broken-stick regression  
362 model was built incrementally (i.e., we added a breakpoint estimate to each successive  
363 model). For each model, an initial guess for the breakpoint was provided, and then the optimal  
364 breakpoints were calculated by the model using an iterative fitting procedure with the default  
365 package parametrization (see Muggeo, 2008, for technical details). We compared each new



366 model with the previous one (based on chi-squared analysis) and selected the most  
367 parsimonious as the final model. For Position 1, the analysis revealed a breakpoint at  
368 repetition 18.46, 95% CI [14.26, 22.67], with RTs decreasing from repetitions 1 to 18.46,  $b =$   
369  $-4.22$ , 95% CI [-5.70, -0.29], followed by a slow increase,  $b = 0.52$ , 95% CI [-0.29, 1.33]. For  
370 Position 2, we estimated two breakpoints at repetitions 5.35, 95% CI [3.54, 7.16] and 19.81,  
371 95% CI [13.63, 25.98], with RTs rapidly decreasing from repetitions 1 to 5.35,  $b = -32.40$ ,  
372 95% CI [-46.24, -18.57], continuing to decrease, but at a slower rate, from repetitions 5.35 to  
373 19.81,  $b = -7.92$ , 95% CI [-10.78, -5.07], followed by a slower decrease until the end of the  
374 task,  $b = -3.20$ , 95% CI [-4.34, -2.07]. For Position 3, we also estimated two breakpoints at  
375 repetitions 5.88, 95% CI [3.52, 8.24] and 19.54, 95% CI [15.36, 23.71], with a fast decrease  
376 in RTs from repetitions 1 to 5.88,  $b = -27.92$ , 95% CI [-41.01, -14.85], continuing to decrease  
377 at a slower rate from repetitions 5.88 to 19.54,  $b = -8.18$ , 95% CI [-10.86, -5.50], and with an  
378 even slower decrease from repetition 19.54 until the end of the experiment,  $b = -1.72$ , 95% CI  
379 [-2.78, -0.65].



380

381 **Fig. 2.** Upper panel: mean response times in Experiment 1 as a function of word position and

382 number of repetitions of the triplet. The vertical dashed line indicates the first repetition at

383 which there was a significant difference between Position 1 vs. Positions 2 and 3. Error bars

384 indicate 95% confidence intervals. Lower panel: results from the broken-stick regressions for

385 each Position in the triplet. Vertical bars indicate the breakpoints.

386 Questionnaire

387 Forty-one of the 42 participants reported noticing a recurrent word sequence; 16 were

388 able to recall the whole triplet, 12 correctly recalled one sub-sequence (words 1 and 2 or

389 words 2 and 3), four could recall non-adjacent words (words 1 and 3), seven only recalled one  
390 word, and the three remaining participants did not recall any word.

### 391 Discussion

392 As expected, the results from Experiment 1 showed faster RTs for predictable words  
393 (i.e., words 2 and 3) within the repeated triplet, and the difference between unpredictable  
394 (word 1) and predictable items increased as the task progressed. Furthermore, this difference  
395 between unpredictable and predictable items emerges early on, around the fifth repetition of  
396 the triplet. The analysis of the mean response times over the 45 repetitions of the triplet  
397 further indicated that learning occurred also for words appearing in Position 1 of the triplet.  
398 Although unpredictable, these words were repeated and their processing was facilitated by  
399 this repetition. The broken-stick regression analysis suggested that learning occurred during  
400 the first 18 repetitions and subsequently reached a plateau performance. While the mean RT  
401 for the first occurrence of these words was 682 ms, the mean RT was 561 ms after 18  
402 repetitions, and 592 ms at the 45<sup>th</sup> repetition, indicating a processing speed up of 90 ms  
403 between the first and last occurrence of the word. These data therefore provide an estimate of  
404 the dynamics of the repetition effect for words that are not predictable.

405 In contrast, RTs for predictable words (i.e., on Positions 2 and 3) followed a totally  
406 different dynamic. According to the broken-stick regression analysis, they indeed decreased  
407 very rapidly during the first 5 repetitions (640 ms at the first repetition, and 523 ms at the 5<sup>th</sup>  
408 repetition - RTs are averaged over Position 2 and 3) and the decrease was slower between  
409 repetition 5 and 18 (419 ms at the 18<sup>th</sup> repetition). After the 18<sup>th</sup> repetition, RTs continued to  
410 decrease but at an even slower rate (347 ms at the 45<sup>th</sup> repetition). Clearly, compared to the  
411 results obtained for words at Position 1 of the repeated triplet, we found that the predictability  
412 effect was much larger than the repetition effect and followed different learning dynamics.  
413 For example, for the 3<sup>rd</sup> position of the triplet, the mean response times were 624 ms for the

414 first occurrence of the word and 349 ms for the 45<sup>th</sup> repetition, resulting in a processing gain  
415 of 275 ms between the first and last occurrence of these words.

416 Interestingly, there was no evidence for an advantage of the third over the second  
417 word in the triplet, contrary to what was observed by previous studies. Indeed, prior findings  
418 indicated faster RTs for the final stimulus in a repeated triplet, as it benefits from the  
419 cumulative information provided by the two preceding stimuli (e.g., Minier et al., 2016; Rey  
420 et al., 2019, 2020, 2022). Regarding our study, although the words clearly benefited from  
421 immediate contextual information (i.e., the preceding word in the triplet that systematically  
422 appeared before them), we did not observe any additional predictability effect regarding the  
423 final word of the triplet when the context was richer (i.e., words in Position 3 of the triplet  
424 benefit from the contextual information provided by words in Position 1 and 2). This  
425 intriguing result likely reflects some limitations of associative and Hebbian learning  
426 mechanisms due to the specific time-scale of the present experimental paradigm. We will  
427 return to this issue in the general discussion.

428 Despite a clear decrease in RTs for the predictable positions in the triplet, indicating  
429 that learning of this repeated sequence occurred, most participants were unable to correctly  
430 recall the whole triplet, even though most of them noticed the presence of a repeated  
431 sequence. This result suggests that part of the triplet learning was explicit but that most of the  
432 learning was probably implicit. Participants did not have to explicitly encode the triplet  
433 repetition to anticipate the occurrence of words appearing on predictable positions.

434 In contrast to Experiment 1, which was conducted with triplets of unrelated words,  
435 Experiment 2 was conducted with triplets of pseudowords. We decided to use pseudowords  
436 because tasks consisting of the repetition and encoding of pseudoword sequences have been  
437 shown to mimic novel word learning (Norris et al., 2018; Schimke et al., 2021). Indeed,  
438 whereas words are likely to have long-term memory representations, pseudowords cannot

439 benefit from such representations as they have not yet been encountered by participants. It is  
440 worth noting that the Hebb paradigm has also been described as a laboratory analogue of  
441 novel word learning (Szmalec et al., 2009, 2012). Therefore, studying triplets of pseudowords  
442 will allow us to compare the learning dynamics of completely novel multiword sequences  
443 with those obtained for already known words in Experiment 1.

## 444 **Experiment 2**

### 445 Methods

#### 446 Participants

447 Forty-six participants (22 females;  $M_{age} = 25$  years,  $SD = 3$ ) were recruited from  
448 Prolific (www.prolific.co) for the experiment. All participants indicated that French was their  
449 native language and declared no neurological or language impairment. Four participants were  
450 excluded from the analyses due to chance-level performance on the main task.

451 As in Experiment 1, participants' French proficiency was measured with the  
452 LexTALE test (Brysbaert, 2013). Participants' average scores were 85.13% ( $SD = 7.08$ ). No  
453 participant was excluded from the analysis. The final number of participants was 42, which  
454 corresponds to a dataset of 1890 data points per condition.

#### 455 Materials

456 In contrast to the previous experiment, here the target triplets were composed of  
457 pseudowords whereas the words served only as fillers items. We selected 180 words from the  
458 French database Lexique 3.83 (New & Pallier, 2020). All words were monosyllabic or  
459 disyllabic singular nouns and had a length from four to six letters. Their freqfilms2 frequency  
460 was between 10 and 100 occurrences per million. A set of 66 pseudowords was selected from  
461 the French Lexicon Project (Ferrand et al., 2010). Twenty-one were drawn therefrom to  
462 construct triplets and the remaining 45 were used as filler pseudowords. All pseudowords  
463 were four-to-six letter long and monosyllabic.

464 Seven triplets were generated and counterbalanced across participants using a Latin-  
 465 squared design. Every triplet repetition of pseudowords (15 per block) was always separated  
 466 by three to six filler words or filler pseudowords (75 per block). As in Experiment 1, each  
 467 block was composed of 60 words and 60 pseudowords. There were an equal number of ‘yes’  
 468 and ‘no’ responses in the experiment (i.e., 180 for each type of response). The sets of  
 469 pseudoword triplets and fillers are listed in Appendix C.

#### 470 Apparatus and procedure

471 The apparatus and procedure were identical to the one used in Experiment 1.

#### 472 Results

473 As the target triplets were made up of pseudowords, only correct “no” responses were  
 474 analysed (95.87% of the data), and RTs exceeding 1500 ms (1.43% of data), as well as RTs  
 475 beyond 2.5 standard deviations from a participant’s mean per block and for each of the three  
 476 possible positions within the triplet (2.01%) were excluded. Means and standard deviations  
 477 per block are shown in Table 3. The linear mixed model we fitted included the maximum  
 478 random effect structure allowing convergence (Barr, 2013; Barr et al., 2013). This model  
 479 included Position, Repetition, and the interaction term as fixed effects. Item and participant  
 480 were used as crossed random effects, with by-participant random slopes for Position. The  
 481 results of the mixed model are summarised in Table 4. Figure 3 provides the evolution of  
 482 mean response times for each position in the triplet and for the 45 repetitions of the triplet.

#### 483 **Table 3**

484 Mean response times (in milliseconds) and standard deviations (in parentheses) for each block  
 485 in Experiment 2.

	Block 1	Block 2	Block 3
Position 1	719 (134)	670 (125)	641 (112)
Position 2	633 (220)	457 (177)	425 (187)
Position 3	595 (190)	439 (161)	401 (169)

486

487 **Table 4**

488 Fixed effects of the mixed model for Experiment 2

Predictors	Estimate	SE	95% CI	<i>p</i>
(Intercept)	554.71	12.50	[530.21, 579.21]	<.001
Position 2 - 1	-171.02	18.51	[-207.30, -134.73]	<.001
Position 3 - 2	-28.43	7.81	[-43.73, -13.13]	<.001
Repetition	-5.39	0.15	[-5.68, -5.10]	<.001
Position 2 - 1 x Repetition	-4.85	0.37	[-5.58, -4.12]	<.001
Position 3 - 2 x Repetition	0.91	0.36	[0.21, 1.62]	.01

489 *Note.* CI: confidence interval; SE: standard error.

490 Results indicated a significant effect of Repetition and faster RTs for pseudowords in  
 491 Position 2 compared to those in Position 1, as well as for Position 3 compared to Position 2.  
 492 Moreover, there was a significant negative interaction coefficient for the difference between  
 493 Position 2 and Position 1, and Repetition, and a significant positive interaction coefficient for  
 494 the difference between Position 3 and Position 2, and Repetition. Similarly to Experiment 1,  
 495 paired sample *t*-tests comparisons showed a significant difference between Position 1  
 496 compared to Positions 2 and 3 on the sixth trial,  $t(36) = 3.12$ , Bonferroni-adjusted  $p = .021$ .

497 As for the first experiment, we conducted a broken-stick regression to study the  
 498 evolution of the learning pattern of Position across the task. The analysis revealed a  
 499 breakpoint at repetition 5.18, 95% CI [3.58, 6.78] for Position 1, with RTs decreasing from  
 500 repetitions 1 to 5.18,  $b = -23.81$ , 95% CI [-36.78, -10.86], followed by a slower decreasing  
 501 rate,  $b = -1.82$ , 95% CI [-2.38, -1.26]. For Position 2, two breakpoints were estimated at  
 502 repetitions 7.44, 95% CI [5.59, 9.29] and 22.00, 95% CI [17.56, 26.44], with RTs rapidly  
 503 decreasing from repetitions 1 to 7.44,  $b = -39.33$ , 95% CI [-50.61, -28.05], continuing to  
 504 decrease, but at a slower rate from repetitions 7.44 to 22.00,  $b = -10.16$ , 95% CI [-14.00, -  
 505 6.32], followed by a slower decrease until the last repetition,  $b = -1.16$ , 95% CI [-2.87, 0.56].  
 506 Regarding Position 3, we estimated two breakpoints at repetitions 6.07, 95% CI [4.40, 7.74]  
 507 and 18.37, 95% CI [14.74, 21.99], with RTs decreasing fast from repetitions 1 to 6.07,  $b = -$

508 41.73, 95% CI [-54.71, -28.74], steadily decreasing at a slower rate from repetitions 6.07 to  
 509 18.37,  $b = -11.66$ , 95% CI [-16.00, -7.33], followed by a slower decrease until the end of the  
 510 task,  $b = -1.86$ , 95% CI [-3.13, -0.59].

511         Given that usage-based theories postulate that novel items become lexicalised when  
 512 they are encountered sufficiently often (e.g., Bybee, 2006; Zang et al., 2023), one might  
 513 expect that after enough repetitions participants would begin to consider the target  
 514 pseudowords to be almost as real words, resulting in more false “yes” judgments as the  
 515 experiment progressed. Therefore, we conducted an additional analysis using a generalised  
 516 (logistic) linear mixed model to compare the mean accuracy between positions across blocks  
 517 (see Figure 4). The model was fitted with Position and Block, and the interaction term as  
 518 fixed effects. The maximal random effects structure that converged was one that included by-  
 519 participant and by-item random intercepts. To explore differences between positions within  
 520 each block, we used the R package emmeans (Lenth, 2023). Helmert contrasts were used to  
 521 compare Position 1 to both Positions 2 and 3, simultaneously, and to compare Position 2 to  
 522 Position 3. The results of the contrasts are summarised in Table 5. The analysis showed that,  
 523 systematically across the three blocks, participants made more false “yes” judgments for  
 524 pseudowords in Position 1 than for those in Positions 2 and 3. In addition, in Block 3,  
 525 participants made more false “yes” judgments for pseudowords in Position 2 compared to  
 526 those in Position 3. Finally, false “yes” judgments for pseudowords in Position 1 increased  
 527 across the blocks, in contrast to those in Positions 2 and 3.

528 **Table 5**

529 Summary of Helmert contrasts between positions across blocks for Experiment 2

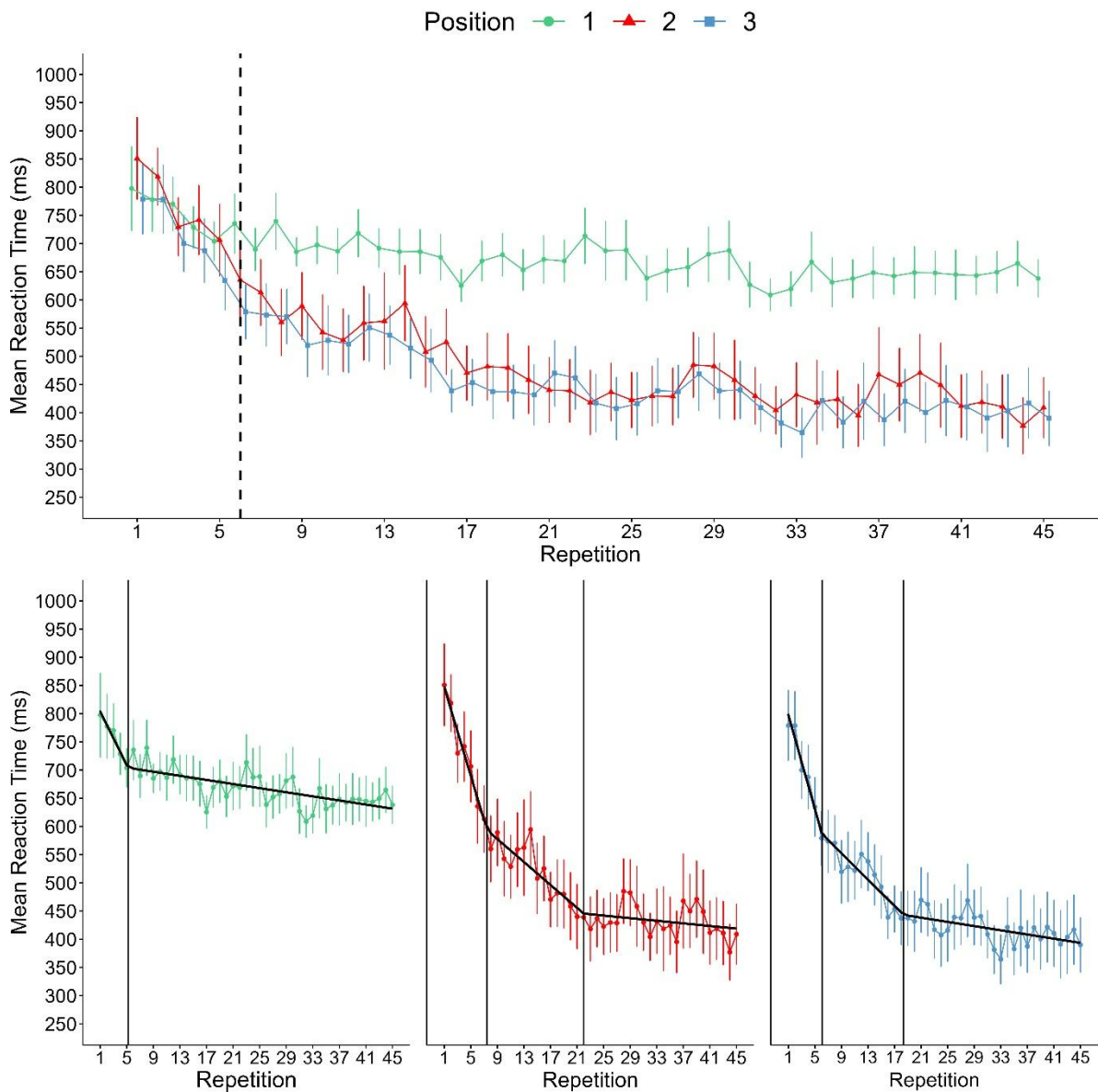
Predictors	Block 1			Block 2			Block 3		
	<i>b</i>	SE	<i>p</i>	<i>b</i>	SE	<i>p</i>	<i>b</i>	SE	<i>p</i>
P1 vs P2-P3	-0.76	0.22	<.001	-2.15	0.25	<.001	-2.48	0.25	<.001
P2 vs P3	-0.24	0.30	.42	-0.22	0.43	.60	-1.40	0.44	.002



530 *Note.* P: Position; SE: standard error.

531 Questionnaire

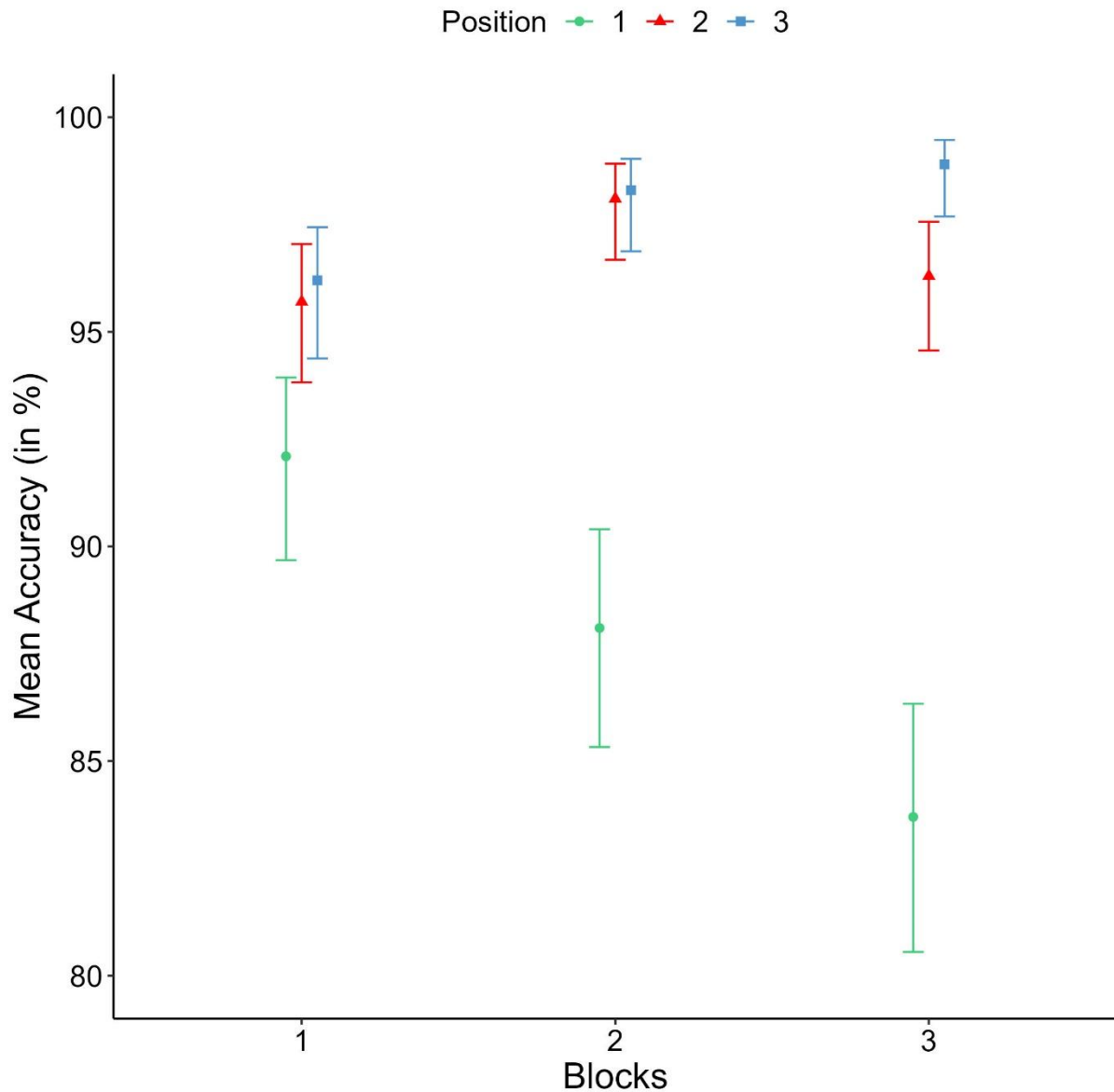
532 Thirty-nine participants reported noticing a recurrent pseudoword sequence; 12 were  
 533 able to recall the whole triplet, one could recall one subsequence (words 2 and 3), eight  
 534 correctly recalled non-adjacent pseudowords (words 1 and 3), eight only recalled one  
 535 pseudoword, and the 13 remaining could not recall any pseudoword.



536

537 **Fig. 3.** Upper panel: Mean response times in Experiment 2 as a function of pseudoword  
 538 position in the repeated triplet and number of repetitions. The vertical dashed line indicates

539 the first repetition at which there was a significant difference between Position 1 vs. Positions  
 540 2 and 3. Error bars indicate 95% confidence intervals. Lower panel: Results from the broken-  
 541 stick regressions for each Position in the triplet. Vertical bars indicate the breakpoints.



542

543 **Fig. 4.** Mean accuracy in Experiment 2 as a function of pseudoword position in the repeated  
 544 triplet and block number. Error bars indicate 95% confidence intervals .

#### 545 Discussion

546 The results of Experiment 2 partly replicated those of Experiment 1. A first main  
 547 difference between the two experiments concerns the overall slower RTs obtained for

548 pseudowords compared to words: when averaging the RTs of all three positions, the mean  
549 RTs on their first occurrence was 654 ms for words and 810 ms for pseudowords; on their last  
550 occurrence (i.e., at the 45<sup>th</sup> repetition), the mean RTs for words was 429 ms and 470 ms for  
551 pseudowords. Apart from these longer RTs, the learning dynamics also produced noticeable  
552 differences compared with the one observed for words.

553         Regarding the repetition effect that is measured by the evolution of RTs for  
554 pseudowords occurring at Position 1 of the triplet, the dynamics was clearly different  
555 compared to words with a fast decrease of response times during the first 5 repetitions (with a  
556 mean RT of 798 ms for the first occurrence and of 704 ms for the 5<sup>th</sup> repetition), followed by  
557 a smoother decrease until the last repetition (with a mean RT of 638 ms for the 45<sup>th</sup>  
558 repetition). While the beta coefficient of the first regression line was -4.22 for words, it was  
559 much larger for pseudowords (-23.81). The processing gain for pseudowords at Position 1  
560 (i.e., the difference between mean RTs for the last repetition and the first occurrence) was 160  
561 ms, which is much larger than the one obtained for words (90 ms). Pseudowords seem  
562 therefore to benefit to a larger extent from the repetition effect indicating that repetitions  
563 produced a fast change in the way these pseudowords were processed and in the way their  
564 trace developed in memory.

565         For predictable pseudowords (i.e., in Position 2 and 3 of the triplet), the broken-stick  
566 regression analysis also identified two break points that were slightly different from those  
567 obtained with words (for pseudowords, 7.44 and 22 at Position 2, and 6.07 and 18.37 at  
568 Position 3; for words, 5.35 and 19.81 at Position 2, and 5.88 and 19.54 at Position 3). Apart  
569 from these differences, the learning dynamics were similar with a fast decrease in RTs during  
570 the initial repetitions followed by an intermediate decrease and a slower one during the last  
571 repetitions. Compared to the repetition effect, the predictability effect was again much larger  
572 and produced a much stronger processing gain (i.e., for the 3<sup>rd</sup> Position, when subtracting the

573 mean RTs for the 45<sup>th</sup> repetition, 390 ms, from the mean RT for the first occurrence, 779 ms,  
574 the processing gain was  $779-390 = 389$  ms).

575         Contrary to Experiment 1, the data revealed a significant difference between Position  
576 2 and 3, with faster RTs on Position 3 of the triplet. This difference seems to emerge around  
577 the same time as in Experiment 1, namely on the sixth repetition of the triplet. Although this  
578 result is consistent with previous finding in sequence learning, here it might be an artifact due  
579 to the fact that participants were slower to classify the pseudowords in Position 2 at the  
580 beginning of the task, resulting in a higher estimation of the regression intercept compared to  
581 the one of Position 3. Due to this unexpected initial difference (that should have been  
582 cancelled by the Latin square design), this difference between Position 2 and 3 is difficult to  
583 interpret.

584         Additionally, we found that as the task progressed, it became more difficult for  
585 participants to classify the first item of the triplet as being a pseudoword. Indeed, they  
586 systematically made more false “yes” judgments for pseudowords in Position 1 than for those  
587 in Positions 2 and 3. Interestingly, false “yes” judgments for pseudowords in Position 1  
588 increased over the course of the task, in contrast to those for pseudowords in Positions 2 and  
589 3. This finding, consistent with usage-based theories, suggests that participants gradually  
590 became familiar with the first pseudoword of the repeated triplet, which presumably became  
591 lexicalised over time. As a result, participants were more likely to respond incorrectly to the  
592 first pseudoword in the triplet. Once they recognised the first pseudoword, they simply had to  
593 respond correctly to the rest of the triplet. It is worth noting that in Block 3, participants were  
594 also more likely to consider the second pseudoword in the triplet to be a word compared to  
595 the third, suggesting that the triplet was becoming progressively lexicalised as well.

596         As for Experiment 1, the number of participants who reported detecting a recurring  
597 sequence was high (93%) but the number of participants who were able to fully recall the

598 triplet was much lower (29% in Experiment 2 compared to 38% in Experiment 1). Here again,  
599 the data suggest that learning occurred both implicitly and explicitly, and the rate of explicit  
600 learning (i.e., with a full recall of the triplet) was lower for pseudowords (29%) than for  
601 words (38%).

602 Overall, Experiment 1 and Experiment 2 yielded similar results regarding the learning  
603 dynamics of the repeated triplet, that is, a slower learning rate on the first unpredictable  
604 position due to a simple repetition effect, and a much larger learning rate for the predictable  
605 positions (i.e., the 2<sup>nd</sup> and the 3<sup>rd</sup>) due to the predictability effect. However, in both  
606 experiments and contrary to natural language, words and pseudowords were totally unrelated  
607 and apart from systematically occurring one after the other, there was no other reason to  
608 associate these items. In Experiment 3, we tested whether the use of a triplet composed of  
609 semantically related words (e.g., belonging to the same word category, like for example, the  
610 fruit category: strawberry, banana, cherry) could have an effect on the learning dynamics of  
611 the triplet. We expected semantic relatedness to facilitate learning both at the implicit level  
612 (i.e., on RTs) and at the explicit level (i.e., on the recall of the triplet).

### 613 **Experiment 3**

#### 614 Methods

#### 615 Participants

616 Forty-two participants (22 females;  $M_{age} = 23$  years,  $SD = 4$ ) were paid and recruited  
617 via Prolific (www.prolific.co). All participants were native French speaker and reported  
618 having no neurological or language disorders. The average LexTALE vocabulary score  
619 (Brysbaert, 2013) was 85.08% ( $SD = 6.37$ ), and no participant was excluded.

#### 620 Materials

621 To construct seven semantically related triplets, we selected 21 low-frequency words  
622 from the database Lexique 3.83 (New & Pallier, 2020). All words were four-to-six letters

623 monosyllabic or disyllabic singular nouns and had a freqfilms2 frequency ranged from 2 to 10  
624 occurrence per million. Forty-five additional words and 180 pseudowords were selected and  
625 used as filler items between two repetitions of the target triplet. All filler words were  
626 monosyllabic or disyllabic singular nouns and were composed of four to six letters. Their  
627 freqfilms2 frequency ranged from 10 to 100 occurrences per million. Pseudowords were  
628 retrieved from the Lexicon Project (Ferrand et al., 2010), were monosyllabic or disyllabic,  
629 and were composed of four to six letters.

630 A Latin-square design was used, leading to the generation of seven triplets for the 42  
631 participants (i.e., 6 participants per triplet). Every triplet repetition (15 per block) was  
632 separated by three to six filler words or filler pseudowords (75 per block). Sixty words and 60  
633 pseudowords were presented in each block. There were an equal number of ‘yes’ and ‘no’  
634 responses in the experiment (i.e., 180 for each type of response). Stimuli are listed in  
635 Appendix D.

#### 636 Apparatus and procedure

637 The apparatus and procedure were identical to that used in Experiments 1 and 2.

#### 638 Results

639 Only correct responses were analysed (96.86% of the data). RTs exceeding 1500 ms  
640 (1.32% of data) and RTs greater than 2.5 standard deviations from a participant’s mean per  
641 block and for each of the three possible positions within the triplet (2.26%) were removed.  
642 Means and standard deviations per block are shown in Table 6. We constructed a linear  
643 mixed-effects model with the maximum random effect structure allowing convergence (Barr,  
644 2013; Barr et al., 2013). This model included Position, Repetition and the interaction term as  
645 fixed effects, participant and Item were used as random intercepts with by-participant random  
646 slopes for Position. We included word length and log-transformed word frequency for each  
647 word in the triplet as covariates. Given that word associations have been shown to influence

648 processing times in multiword sequences (e.g., Carrol and Conklin, 2020), and that the order  
 649 of presentation of the words in the triplets varied across participants (because of the Latin-  
 650 squared design), potentially affecting processing times as some words were more strongly  
 651 associated than others, we also included a measure of association strength between triplet  
 652 words as a covariate. As existing free-association databases in French don't contain all the  
 653 items we used, we decided to calculate the indirect association strength between the words  
 654 using the JeuxDeMots database (Lafourcade & Joubert, 2008). This database is based on a  
 655 collaborative online project where participants see a word and provide an association, which  
 656 is only validated if other peers have suggested the same association. These associations are  
 657 then weighted according to the number of associations given by the participants to obtain the  
 658 association strength. To calculate the indirect association strength between two target words,  
 659 we generated a list of the most frequently associated words with the target word, then selected  
 660 the most frequent common word between two target words and averaged the association  
 661 strengths to obtain the indirect association strength measure. For instance, both banana and  
 662 strawberry were associated with fruit (i.e., 526 and 480, respectively). To obtain the indirect  
 663 association strength, we then averaged the two values, resulting in an indirect association  
 664 strength of 503. The results of the model are summarised in Table 7. Figure 5 provides the  
 665 evolution of mean response times for each position in the triplet and for the 45 repetitions of  
 666 the triplet.

667 **Table 6**

668 Mean response times (in milliseconds) and standard deviations (in parentheses) for each block  
 669 in Experiment 3.

	Block 1	Block 2	Block 3
Position 1	613 (115)	581 (110)	576 (103)
Position 2	496 (162)	369 (143)	344 (147)
Position 3	475 (162)	366 (130)	354 (142)

670

671 **Table 7**

672 Fixed effects of the mixed model for Experiment 3.

Predictors	Estimate	SE	95% CI	<i>p</i>
(Intercept)	377.62	68.40	[243.57, 511.68]	<.001
Position 2 - 1	-184.91	13.72	[-211.79, -158.03]	<.001
Position 3 - 2	-6.73	7.38	[-21.19, 7.73]	.368
Repetition	-3.56	0.11	[-3.77, -3.35]	<.001
Position 2 - 1 x Repetition	-4.02	0.26	[-4.53, -3.50]	<.001
Position 3 - 2 x Repetition	0.87	0.26	[0.36, 1.38]	<.001
Word length	8.79	10.43	[-11.65, 29.22]	0.411
Word frequency (log)	2.83	10.78	[-18.31, 23.96]	0.798
Association strength bigram 1	0.04	0.08	[-0.11, 0.19]	0.612
Association strength bigram 2	0.08	0.08	[-0.07, 0.23]	0.294

673 *Note.* CI: confidence interval; SE: standard error.

674 The results showed a significant negative effect of Repetition reflecting a decrease in  
675 RTs. We also found faster RTs for words in Position 2 compared to those in Position 1, but  
676 not to those in Position 3. Finally, there was a significant negative interaction coefficient for  
677 the difference between Position 2 and Position 1, and Repetition, and a significant positive  
678 interaction coefficient for the difference between Position 3 and Position 2, and Repetition.  
679 No significant effects were found for Word length, Word frequency and Association strength  
680 for both bigrams. Paired sample *t*-tests comparisons showed that a significant difference  
681 between Position 1 compared to Positions 2 and 3 emerged on the third trial,  $t(39) = 3.39$ ,  
682 Bonferroni-adjusted  $p = .005$ .

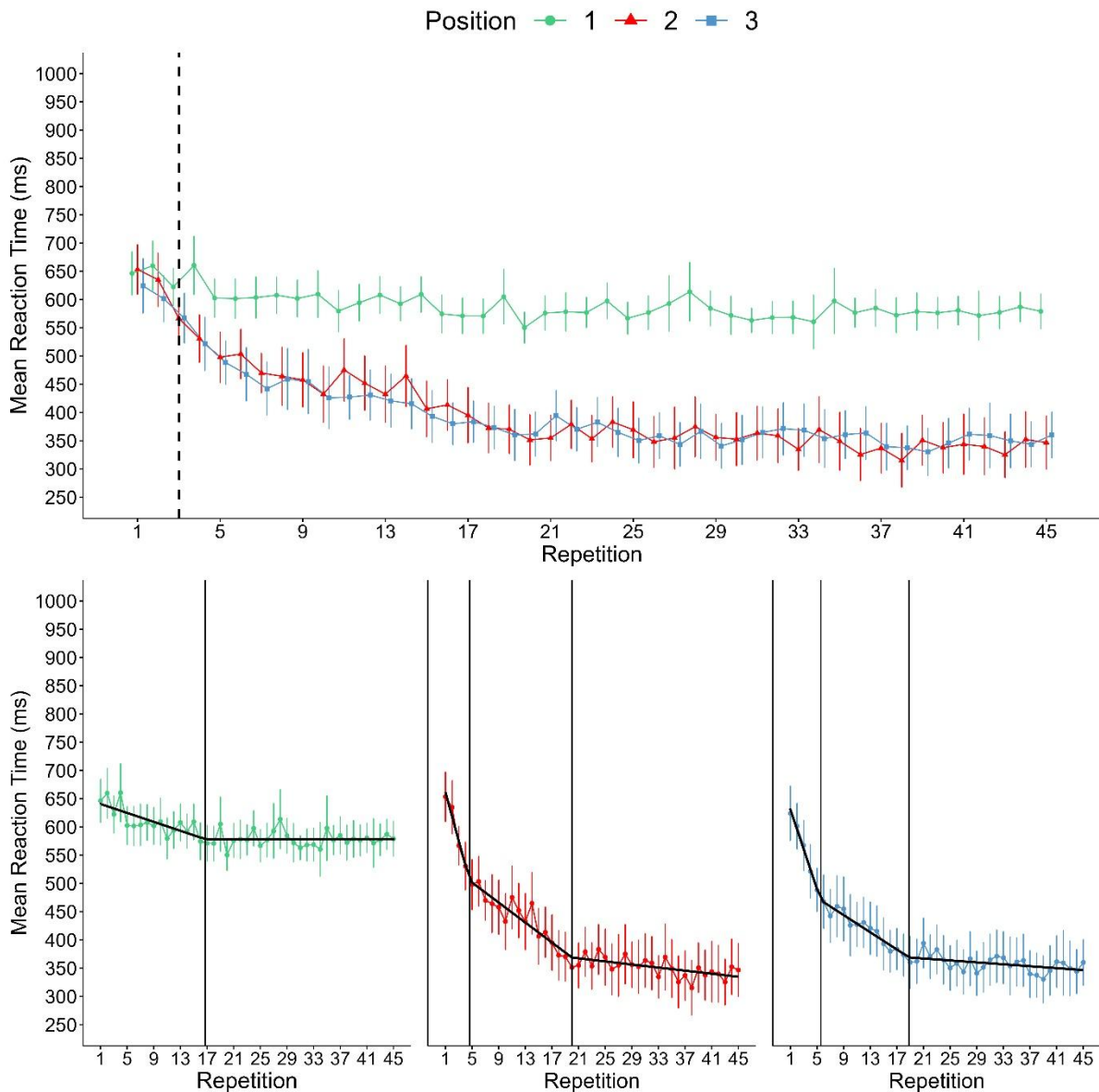
683 Following the same procedure as in Experiments 1 and 2, we performed a broken-stick  
684 regression on each Position of the repeated triplet. For Position 1, the analysis revealed a  
685 breakpoint at repetition 16.72, 95% CI [11.34, 22.10], with RTs decreasing from repetitions 1  
686 to 16.72,  $b = -3.96$ , 95% CI [-5.80, -2.11], followed by an almost flat slope,  $b = 0.01$ , 95% CI  
687 [-0.76, 0.77]. Concerning Position 2, we estimated two breakpoints at repetitions 4.64, 95%  
688 CI [3.08, 6.20] and 20, 95% CI [16.20, 23.80], with RTs rapidly decreasing from repetitions 1  
689 to 4.64,  $b = -43.01$ , 95% CI [-63.16, -22.86], continuing to decrease, but at a slower rate from



690 repetitions 4.64 to 20,  $b = -8.85$ , 95% CI [-11.27, -6.42], followed by a slower decrease until  
691 the end of the task,  $b = -1.37$ , 95% CI [-2.62, -0.11]. For Position 3, we also estimated two  
692 breakpoints at repetitions 5.57, 95% CI [3.89, 7.24] and 18.85, 95% CI [14.49, 23.22], with a  
693 fast decrease in RTs from repetitions 1 to 5.57,  $b = -35.23$ , 95% CI [-49.00, -21.47],  
694 continuing to decrease at a slower rate from repetitions 5.57 to 18.85,  $b = -7.62$ , 95% CI [-  
695 10.77, -4.46], followed by a slower decreasing until the last repetition,  $b = -0.85$ , 95% CI [-  
696 1.92, 0.21].

### 697 Questionnaire

698           Forty-one of the 42 participants reported noticing a recurrent word sequence; 29 were  
699 able to recall the whole triplet, one recalled one subsequence (words 2 and 3), four could  
700 recall non-adjacent words (words 1 and 3), five recalled all the words but in the wrong order,  
701 and the three remaining could not recall any word.



702

703 **Fig. 5.** Upper panel: mean response times in Experiment 3 as a function of word position and  
 704 number of repetitions. The vertical dashed line indicates the first repetition at which there was  
 705 a significant difference between Position 1 vs. Positions 2 and 3. Error bars indicate 95%  
 706 confidence intervals. Lower panel: results from the broken-stick regressions for each Position  
 707 in the triplet. Vertical bars indicate the breakpoints.

708 Discussion

709 Experiment 3 produced similar results as in Experiment 1. Concerning the repetition  
 710 effect, we did not expect any advantage of the semantic relatedness because there is no reason

711 to observe any effect of this variable on the first word of the triplet. And indeed, the dynamics  
712 of the repetition effect was very similar to the one obtained in Experiment 1.

713 For predictable items (in Position 2 and 3 of the triplet), the beta coefficient of the first  
714 regression line (from the broken-stick regression analysis) was larger (-43.01 for related  
715 words compared to -32.4 for unrelated words) and the first breakpoint occurred earlier (4.64  
716 compared to 5.35), suggesting that the initial learning phase was much steeper in the  
717 semantically related condition compared to the unrelated words from Experiment 1. The  
718 semantical relatedness between these words helped producing a larger predictability effect  
719 that certainly took advantage of the pre-existing semantic associations between these words.  
720 This was also confirmed by the fact that a difference between unpredictable and predictable  
721 items emerges earlier than in Experiment 1 (i.e., around the third rather than the fifth  
722 repetition of the triplet). Note that this advantage was only present at the early phase of  
723 learning because the processing gain for words in Experiment 1 is similar to the one obtained  
724 in Experiment 3. Indeed, the difference between the mean response times on Position 3 for the  
725 first and last occurrence of these items was  $624 \text{ ms} - 349 \text{ ms} = 275 \text{ ms}$  in Experiment 1 and  
726  $624 \text{ ms} - 360 \text{ ms} = 264 \text{ ms}$  in Experiment 3. Finally, as for Experiment 1, there was no  
727 additional advantage for items occurring in Position 3 of the triplet compared to those being  
728 in Position 2.

729 Like Experiment 1, the number of participants who reported detecting a recurring  
730 sequence was high (98%) but the number of participants who were able to fully recall the  
731 triplet was much larger (69% compared to 38% in Experiment 1). Clearly, the semantic  
732 relatedness may have helped participants encoding the triplet in an explicit way which  
733 probably also explains the stronger predictability effect observed during the early phase of  
734 learning.

735 As expected, semantic relatedness had a facilitatory effect on the predictability effect  
736 but also on the ability of participants to explicitly memorize the repeated triplet and to recall  
737 it. However, this situation is rather artificial given that words belonging to the same semantic  
738 category rarely appear in a sequence when reading texts, apart from special cases such as  
739 binomials (e.g., *salt and pepper*, *boys and girls*, *knife and fork*), which are often composed of  
740 words belonging to the same semantic category. It has been shown that the association  
741 strength of the component words in binomials influences reading times in a natural reading  
742 task (Carrol & Conklin, 2020). We therefore tested whether the learning dynamics of a triplet  
743 would be improved by using words that often cooccur, like idioms. A recent study has indeed  
744 shown that meaningful three-word sequences (e.g., idioms: *on my mind*; phrase: *is really*  
745 *nice*) are easier to process and lead to faster RTs compared to fragment sequences (e.g.,  
746 *because it lets*) in a phrasal decision task (Jolsvai et al., 2020). Similarly, Northbrook et al.  
747 (2022) presented Japanese English speakers with a series of short stories containing repeated  
748 three-word lexical bundles, each seen three times, followed by a phrasal decision task. They  
749 found that repeated lexical bundles (e.g., *set off home*, *tired and hungry*) were processed faster  
750 than non-repeated bundles in the phrasal decision task, with faster RTs at each subsequent  
751 repetition. This advantage for repeated lexical bundles emerged from the first repetition and  
752 was still present a week later. In Experiment 4, we therefore used three-word idioms as  
753 repeated triplets to study whether the presence of frequently cooccurring words increases the  
754 predictability effect. We expected idioms to facilitate learning as they have already been  
755 encountered and encoded in memory as whole sequences by the participants.

## 756 **Experiment 4**

### 757 Methods

### 758 Participants

759 Forty-two participants (21 females;  $M_{age} = 24$  years,  $SD = 4$ ) were recruited and paid  
760 to take part in the study via Prolific (www.prolific.co). All participants were native French  
761 speakers and reported having no neurological or language impairments. Their average  
762 LexTALE vocabulary score (Brysbaert, 2013) was 86.18% ( $SD = 6.19$ ), no participant was  
763 excluded from the analysis.

#### 764 Materials

765 We constructed the triplets by selecting seven three-word idiomatic expressions from  
766 two databases of French idioms rated by native speakers (Bonin et al., 2013, 2018). Filler  
767 items that were inserted between two repetitions of the triplet were 45 words and 180  
768 pseudowords. Words were monosyllabic or disyllabic singular nouns and were chosen from  
769 the database Lexique 3.83 (New & Pallier, 2020). All words were four to six letters long and  
770 had a freqfilms2 frequency between 10 and 100 occurrences per million. Pseudowords were  
771 selected from the French Lexicon Project (Ferrand et al., 2010). All pseudowords were  
772 monosyllabic or disyllabic and were composed of four to six letters.

773 In contrast to previous experiments in which we used a Latin-square design, here the  
774 triplets were not scrambled, and therefore participants saw the idioms in their canonical form.  
775 Indeed, reversing the word order of existing idiomatic expressions has been shown to result in  
776 a processing penalty (Conklin & Carrol, 2020). Each of the 7 idiomatic expressions was  
777 presented to six participants ( $6 \times 7 = 42$ ). Every triplet repetition (15 per block) was separated by  
778 three to six filler words or pseudowords (75 per block). As in the previous experiments, every  
779 block was composed of 60 words and 60 pseudowords. Therefore, there were an equal  
780 number of ‘yes’ and ‘no’ responses in the experiment (i.e., 180 for each type of response).  
781 Stimuli are listed in Appendix E.

#### 782 Apparatus and procedure

783 The apparatus and procedure were identical to the one used in Experiments 1, 2 and 3.

784 Results

785 Only correct responses were analysed (96.34% of the data). RTs exceeding 1500 ms  
 786 (1.54% of data), and RTs beyond than 2.5 standard deviations from a participant's mean per  
 787 block and for each of the three possible positions within the triplet (2.19%) were removed.  
 788 Mean response times and standard deviations per Block and Position are shown in Table 8.

789 We constructed a linear mixed-effects model with the maximum random effect  
 790 structure allowing convergence (Barr, 2013; Barr et al., 2013). This model included Position,  
 791 Repetition and the interaction term as fixed effects. Item and participant were used as crossed  
 792 random effects, with by-participant random slopes for Position. In addition to Word length  
 793 and Word frequency, we also included Idiom frequency and Bigram and Trigram mutual  
 794 information<sup>1</sup> (MI) scores as covariates in our analysis. Indeed, previous research on idioms  
 795 has shown that these factors can influence the processing of multiword sequences (e.g., Carrol  
 796 & Conklin, 2020). Idiom frequency and MI scores were calculated based on the French web  
 797 corpus frTenTen20 (Jakubíček et al., 2013), which consists of 20.9 billion words. All  
 798 frequencies were log-transformed prior to analysis. The results of the model are summarised  
 799 in Table 9. Figure 6 provides the evolution of mean response times for each position in the  
 800 triplet and for the 45 repetitions of the triplet.

801 **Table 8**

802 Mean response times (in milliseconds) and standard deviations (in parentheses) for each block  
 803 and position in Experiment 4.

	Block 1	Block 2	Block 3
Position 1	607 (126)	582 (121)	570 (108)
Position 2	478 (179)	337 (156)	309 (136)
Position 3	464 (166)	347 (147)	315 (120)

804

805 **Table 9**


---

<sup>1</sup> MI estimates the predictability of observing a word given the preceding words in the sequence (Ramisch, 2015, p. 66).

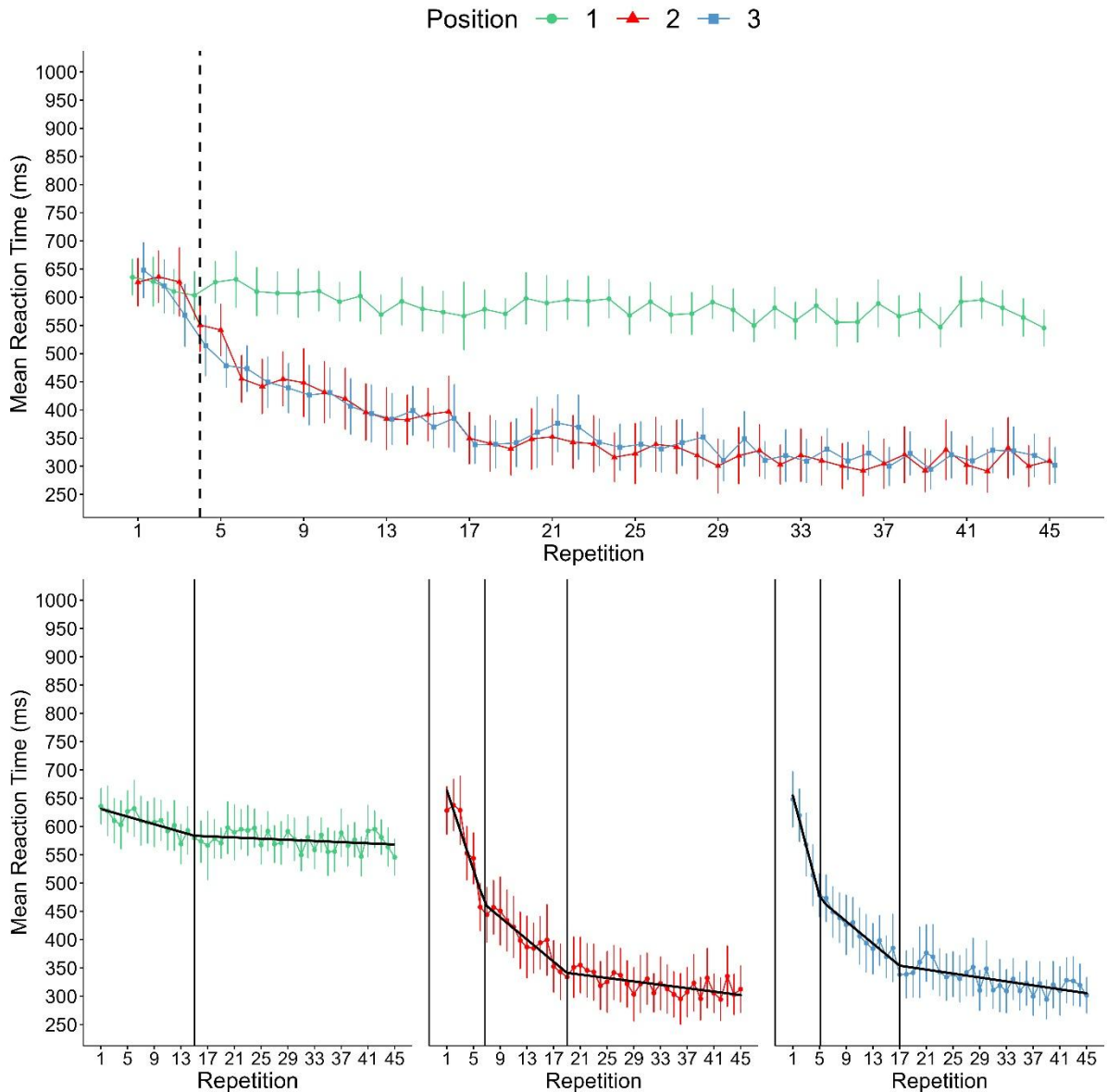
806 Fixed effects of the mixed model for Experiment 4.

Predictors	Estimate	SE	95% CI	<i>p</i>
(Intercept)	509.16	84.82	[342.91, 675.42]	<.001
Position 2 - 1	-219.80	33.53	[-285.52, -154.07]	<.001
Position 3 - 2	3.65	12.23	[-20.32, 27.63]	.774
Repetition	-4.09	0.12	[-4.32, -3.87]	<.001
Position 2 - 1 x Repetition	-4.66	0.28	[-5.21, -4.10]	<.001
Position 3 - 2 x Repetition	0.82	0.28	[0.27, 1.38]	.003
Word length	9.43	6.19	[-2.70, 21.56]	.157
Word frequency (log)	-0.01	6.48	[-12.71, 12.69]	.999
Idiom frequency (log)	20.24	9.19	[2.23, 38.25]	.033
Bigram MI	-16.12	4.68	[-25.28, -6.95]	.002
Trigram MI	0.47	4.41	[-8.17, 9.11]	.916

807 *Note.* CI: confidence interval; SE: standard error.

808 The results showed a significant negative effect of Repetition reflecting a decrease in  
809 response times with repetitions. We also found faster response times for words in Position 2  
810 compared to those in Position 1, but there was no difference between Position 3 and 2. There  
811 were a significant interaction coefficient for the difference between Position 2 and Position 1,  
812 and Repetition, as well as for the difference between Position 3 and Position 2, and  
813 Repetition. Finally, there was a significant effect of Idiom frequency, with less frequent  
814 idioms eliciting faster responses, and of Bigram MI, with faster response times for bigrams  
815 with stronger MI. This unusual pattern is most likely due to the fact that one of our less  
816 frequent idioms in the experiment (i.e., *qui dort dîne*) has a high bigram MI score (i.e., 4.72),  
817 which may have speeded up participants' responses even though the idiom frequency was  
818 low. In fact, any collocation above an MI score of 3 is considered to be strong. When this  
819 idiom is excluded from the analysis, the effect of Idiom frequency is no longer significant,  $b =$   
820 16.65,  $SE = 11.22$ , 95% CI = [-5.33, 38.64],  $p = .147$ . Similarly to Experiment 3, paired  
821 sample *t*-tests comparisons showed a significant difference between Position 1 compared to  
822 Positions 2 and 3 on the fourth trial,  $t(40) = 2.70$ , Bonferroni-adjusted  $p = .04$ .

823



824

825 **Fig. 6.** Upper panel: mean response times in Experiment 4 as a function of word position and  
 826 number of repetitions. The vertical dashed line indicates the first repetition at which there was  
 827 a significant difference between Position 1 vs. Positions 2 and 3. Error bars indicate 95%  
 828 confidence intervals. Lower panel: results from the broken-stick regressions for each Position  
 829 in the triplet. Vertical bars indicate the breakpoints.

830 We then performed a broken-stick regression to better account for the evolution of the  
 831 learning pattern throughout the task. A breakpoint was estimated at repetition 15, 95% CI  
 832 [6.51, 23.49] for Position 1, with RTs decreasing from repetitions 1 to 15,  $b = -3.38$ , 95% CI



833 [-5.81, -0.95], followed by a slower decrease in RTs,  $b = -0.51$ , 95% CI [-1.25, 0.23].  
 834 Regarding Position 2, we estimated two breakpoints at repetitions 6.72, 95% CI [4.88, 8.57]  
 835 and 19, 95% CI [15.37, 22.63], with a fast decrease in RTs from repetitions 1 to 6.72,  $b = -$   
 836 35.45, 95% CI [-46.72, -24.17], continuing to decrease but at a slower rate from repetitions  
 837 6.72 to 19,  $b = -9.87$ , 95% CI [-13.27, -6.47], followed by a slower decrease until the end of  
 838 the task,  $b = -1.51$ , 95% CI [-2.72, -0.31]. For Position 3, two breakpoints were estimated at  
 839 repetitions 5.17, 95% CI [3.92, 6.41] and 17, 95% CI [13.50, 20.50], with a strong decrease in  
 840 RTs from repetitions 1 to 5.17,  $b = -44.36$ , 95% CI [-58.21, -30.51], continuing with a slower  
 841 decrease from repetitions 5.17 to 17,  $b = -9.79$ , 95% CI [-13.31, -6.28], followed by an even  
 842 slower decrease until the end of the task,  $b = -1.74$ , 95% CI [-2.73, -0.75].

#### 843 Questionnaire

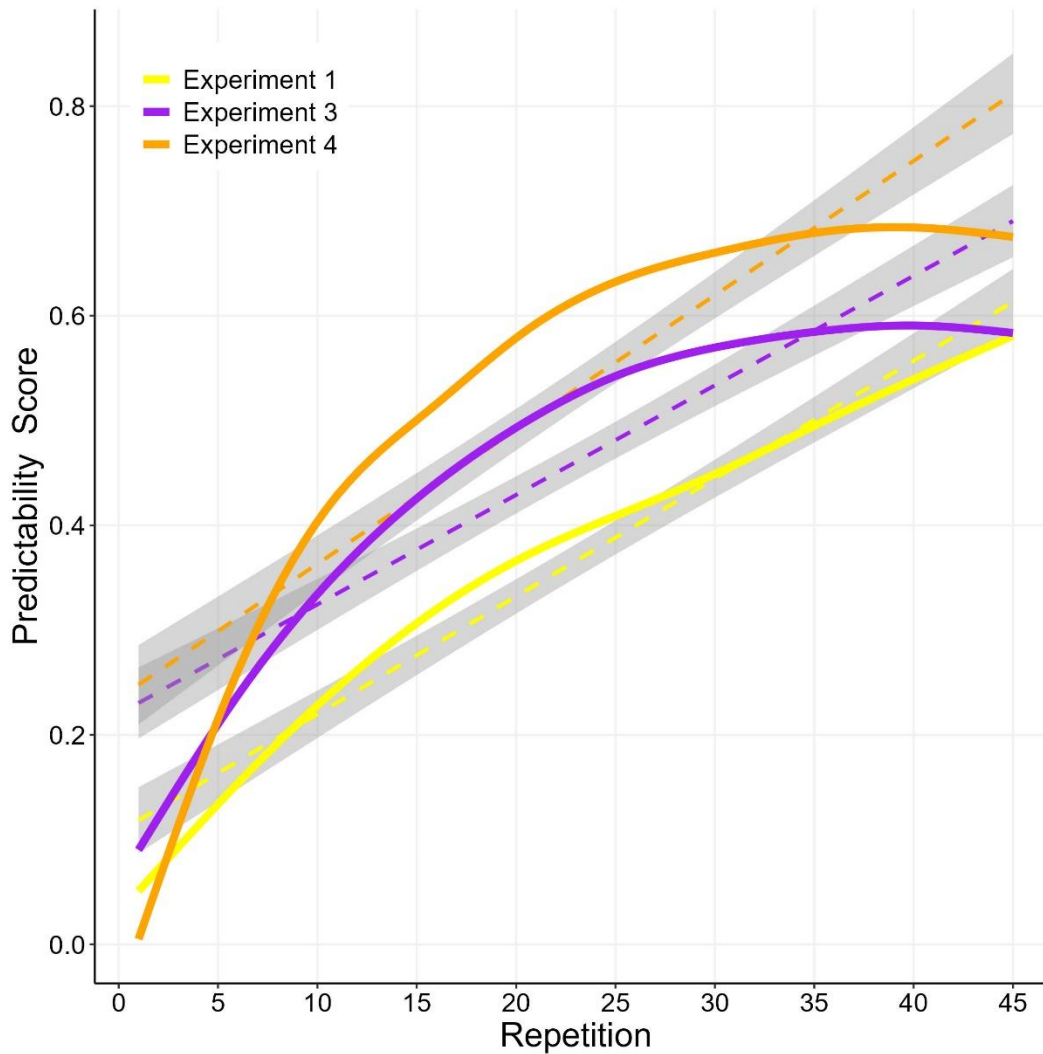
844 All participants reported noticing a recurrent word sequence; 37 were able to recall the  
 845 whole triplet, two recalled one subsequence (words 2 and 3), one could recall non-adjacent  
 846 words (words 1 and 3), one recalled all the words but in the wrong order, and the last one  
 847 could not recall any word.

#### 848 Additional analysis

849 To compare the predictability effects observed in Experiment 1, 3 and 4, we computed  
 850 a predictability score for these experiments by calculating a difference between log-  
 851 transformed RTs for unpredictable words (Position 1) versus the log-transformed mean RT  
 852 for predictable words (Position 2 and 3) for each repetition. Here, a positive score reflects a  
 853 predictability effect. We decided to use log-transformed values to control for baseline  
 854 differences in the participants' responses (see Siegelman, Bogaerts, Kronenfeld, et al., 2018).  
 855 For instance, let us consider two participants with a mean difference of 100 ms between  
 856 predictable and unpredictable words, but with a different baseline RT: P1 unpredictable = 600  
 857 ms, predictable = 500 ms; P2 unpredictable = 400 ms, predictable = 300 ms. Without this

858 transformation, these participants would have the same difference score, even if the relative  
859 acceleration of P2 to predictable words is much higher. After log-transformation, the  
860 difference between predictable and unpredictable words reflects better this acceleration: log  
861 difference of P1 = 0.18, P2 = 0.29.

862 We then ran a linear mixed-effects model on the predictability scores, using  
863 Experiment, Repetition and the interaction term as fixed effects, and participant as random  
864 effect. Experiment was coded using repeated contrast coding (Experiment 1: -0.7 -0.3;  
865 Experiment 3: 0.3 -0.3; Experiment 4: 0.3 0.7). We observed higher predictability scores in  
866 Experiment 4 (idioms) compared to Experiment 3 (semantically related words),  $b = 0.07$ ,  $SE$   
867  $= 0.01$ ,  $p < .001$ , and higher scores in Experiment 3 compared to Experiment 1 (non-related  
868 words),  $b = 0.09$ ,  $SE = 0.01$ ,  $p < .001$ . In addition, there was a main effect of Repetition,  $b =$   
869  $0.01$ ,  $SE = 0.00$ ,  $p < .001$ , and a significant interaction between Experiment 4 - Experiment 3  
870 and Repetition,  $b = 0.002$ ,  $SE = 0.001$ ,  $p = .008$ , indicating an increasing difference of  
871 predictability scores between both experiments (see Figure 7).



872

873 **Fig. 7.** Predictability scores across repetitions for word triplets in Experiments 1, 3 and 4.

874 Continuous lines represent loess fit for the predictability scores. Dashed lines represent the

875 best linear fit and grey-shaded areas indicate 95% confidence intervals around linear

876 regression lines.

877 Discussion

878 Experiment 4 produced results similar to Experiment 3. However, two notable

879 differences suggest that idioms have benefited to a larger extent from triplet repetition

880 compared to semantically related words. First, the predictability score represented in Figure 7

881 indeed shows that when the repetition effect is subtracted from the predictability effect on

882 each repetition trial, the remaining predictability score is stronger for idioms compared to

883 semantically related words, which is also stronger than the score obtained for unrelated words  
884 from Experiment 1. Idioms, which are supposedly already coded in the brain as semantically  
885 coherent and frequent sequences of words, appear to derive a greater processing advantage  
886 from repetition. Second, while 69% of participants in Experiment 3 were able to recall the full  
887 triplet of semantically related words, 88% of participants in Experiment 4 managed to recall  
888 the full idiom. This improved performance for explicit correct recall of idioms is probably due  
889 to their pre-existing encoding as relevant linguistic sequences, or at least to a facilitated access  
890 to them in memory, and it suggests more generally that frequent multiword sequences (apart  
891 from idioms) do result in a different learning dynamic in this Hebb lexical decision task  
892 compared to less frequent multiword sequences.

### 893 General discussion

894 The goal of the present set of experiments was to provide empirical evidence about the  
895 dynamics of multiword sequence extraction by studying the evolution of response times (RTs)  
896 for a repeated triplet of items in a task where participants were not informed about the  
897 presence of this regularity. Using a Hebb lexical decision task, where a word (Experiments 1,  
898 3 and 4) or a pseudoword (Experiment 2) triplet was repeated throughout a noisy stream of  
899 random words and pseudowords, we found that RTs for the unpredictable first position in the  
900 triplet decreased over repetitions (i.e., the repetition effect) but more slowly and with a  
901 different dynamic compared to items appearing at the predictable second and third positions  
902 in the repeated triplet (i.e., the predictability effect). The learning dynamic also varied as a  
903 function of triplet type (i.e., unrelated words, pseudowords, semantically related words, or  
904 idioms) and there was no evidence of a difference between items appearing at Position 2 and  
905 3 of the triplets. Finally, these results, supported by implicit associative learning mechanisms,  
906 were accompanied by evidence of an explicit learning of the sequence that also varied as a  
907 function of the triplet's type.

908 Repetition is a key mechanism for the development of memory traces for words and  
 909 sequences of words. There is much recent evidence showing that we acquire not only memory  
 910 traces for words but also for multiword sequences (e.g., Arnon & Snider, 2010; Bannard &  
 911 Matthews, 2008; Conklin & Carrol, 2020; Conklin & Schmitt, 2008; Janssen & Barber, 2012;  
 912 Siyanova-Chanturia, Conklin, & Schmitt, 2011; Siyanova-Chanturia, Conklin, & van Heuven,  
 913 2011). The development of these memory traces may facilitate their processing and this  
 914 phenomenon is now considered by several models of language acquisition (e.g., Abbot-Smith  
 915 & Tomasello, 2006; Ambridge, 2020; Bannard & Lieven, 2012; McCauley & Christiansen,  
 916 2019; Perruchet & Vinter, 1998) as being central for the processing of multiword sequences.

917 The present set of experiments provides new empirical evidence allowing to better  
 918 understand the effect of repetitions on the creation of memory traces in the processing of  
 919 multiword sequences and notably, to differentiate the dynamics of the repetition effect and the  
 920 predictability effect. The different dynamics of these effects were notably revealed by the  
 921 broken-stick regression analyses that we conducted on mean response times overall  
 922 repetitions and for all positions in the repeated triplet. A summary of the main results from  
 923 these analyses is provided in Table 10.

924 **Table 10**

925 Broken-stick regressions results: breakpoints (BP) and beta coefficients ( $b$ ) of the regression  
 926 lines for each position in the repeated triplet and for each experiment.

Experiment	Position	BP <sub>1</sub>	BP <sub>2</sub>	$b_1$	$b_2$	$b_3$
Unrelated words (1)	1	18.46		-4.22	0.52	
	2	5.35	19.81	-32.40	-7.92	-3.20
	3	5.88	19.54	-27.92	-8.18	-1.72
Pseudowords (2)	1	5.18		-23.81	-1.82	
	2	7.44	22.00	-39.33	-10.16	-1.16
	3	6.07	18.37	-41.73	-11.66	-1.86
Semantically related words (3)	1	16.72		-3.96	0.01	
	2	4.64	20.00	-43.01	-8.85	-1.37

	3	5.57	18.85	-35.23	-7.62	-0.85
	1	15.00		-3.38	-0.52	
Idioms (4)	2	6.72	19.00	-35.45	-9.87	-1.51
	3	5.17	17.00	-44.36	-9.79	-1.74

---

927

928           Regarding the repetition effect (indexed by the evolution of RTs on the first  
929 unpredictable position of the triplet) for words in Experiment 1, 3, and 4, it was characterized  
930 by a late breakpoint (i.e., BP1) and a relatively slow decrease in RTs indexed by a small beta  
931 coefficient (i.e.,  $b_1$ ). The dynamic was very different for pseudowords in Experiment 2 since  
932 it produced an earlier breakpoint (5.18) and a much larger beta coefficient (-23.81). Similarly,  
933 the processing gain (indexed by the difference in mean RTs between the 45<sup>th</sup> repetition and  
934 the first occurrence of the item) was smaller for words (i.e., 90 ms, 67 ms, and 90 ms, for  
935 Experiment 1, 3, and 4, respectively) than for pseudowords (160 ms). These results suggest  
936 that repetition will differentially affect the processing of items that are already encoded in  
937 memory (i.e., words) compared to novel items (i.e., pseudowords). Thus, in the present study,  
938 we observe that the processing of novel items benefits very rapidly from repetition and  
939 certainly from the transitory development of a memory trace representing these items.

940           The dynamics of the predictability effect, that is indexed by the evolution of RTs on  
941 the second and third predictable positions of the triplet, was characterized, for all items, by a  
942 fast decrease in RTs with an early breakpoint (around 4-7 repetitions of the triplet) for the first  
943 regression line and a large beta coefficient. The processing gain, which can be computed by  
944 subtracting the mean RTs (averaged over Positions 2 and 3) for the last occurrence of the  
945 triplet (i.e., 45<sup>th</sup> repetition) from the mean RTs obtained for the first occurrence of the same  
946 items (e.g., 640 ms – 347 ms = 293 ms, for Experiment 1), indicates that the predictability  
947 effect was much larger than the repetition effect (i.e., it was 293 ms, 417 ms, 285 ms, and 332  
948 ms, for Experiment 1-4 respectively). The emergence of these early breakpoints for  
949 predictable items, as well as of the difference between unpredictable and predictable items

950 (around 3 to 5 repetitions for words), is consistent with the findings of Conklin and Carrol  
951 (2020). Indeed, they found a rapid change in participants' reading behaviour after only 4 to 5  
952 repetitions of the repeated pattern. These results clearly illustrate that encoding multiword  
953 sequences in memory drastically accelerates the processing of these items and that the  
954 predictability effect goes far beyond the repetition effect.

955         We note that an alternative interpretation to the predictability effect described above  
956 can also be provided by the multiconstituent unit (MCU) hypothesis (Zang et al., 2023),  
957 which is very close to the assumptions made in McCauley and Christiansen's (2019)  
958 computational model. According to this hypothesis, frequently encountered linguistic units  
959 consisting of more than a single word can be lexically represented in memory and identified  
960 as single representations during reading. In McCauley and Christiansen's (2019) model, this  
961 lexicalization process is driven by the central mechanism of chunking (see also, Perruchet &  
962 Vinter, 1998; Jessop et al., 2023). Therefore, multiword and pseudoword sequences that co-  
963 occur repeatedly and frequently, as in our study, may gradually become lexicalised and  
964 represented as single units in the individual's mental lexicon. Note that several studies by  
965 Liang et al. (2015, 2017, 2021, 2023) provide empirical data in favour of this hypothesis in  
966 the field of Chinese word reading.

967         In addition, it is worth noting that the different learning dynamics that we observed for  
968 pseudowords can be explained not only by the development of a new memory trace, but also  
969 by the lexical decision task itself and the cognitive processes underlying it. Indeed, while in  
970 Experiment 2 participants had to give a "no" response to the triplet consisting of  
971 pseudowords, it has been shown that producing a "yes" response involves different processes  
972 than producing a "no" response. For instance, based on McClelland and Rumelhart's (1981)  
973 interactive activation model, Grainger and Jacobs (1996) propose that the generation of a  
974 "yes" response occurs when a word is recognised as a result of surpassing a certain activation

975 threshold. In contrast, a "no" response is generated on the basis of global lexical activation,  
976 which varies as a function of the likelihood that the stimulus is a word (see also Dufau et al.,  
977 2012). Experiment 2 is therefore not comparable to the other experiments in this regard.  
978 Nevertheless, like other studies of novel words and multiword sequences using pseudowords  
979 (e.g., Norris et al., 2018; Pellicer-Sánchez, 2017; Pellicer-Sánchez et al., 2022; Szmalec et al.,  
980 2012), it allows us to study the dynamics of the development of a trace in memory and its  
981 influence, in this case, on lexical decision processes. This data may also have direct  
982 consequences for computational models of language acquisition like, for example, the Parser  
983 model (Perruchet & Vinter, 1998). In this model, each time a unit is processed again (i.e., its  
984 processing is repeated), it receives a linear increase of its memory trace (indexed by a weight  
985 value). The present results suggest that this increase may not be linear but rather non-linear  
986 depending on the weight of the item's memory trace. For new memory traces, the increase  
987 seems to be stronger and more rapid than for memory traces that are more strongly encoded in  
988 lexical memory (called "perceptual shaper" in this model).

989         Likewise, and beyond the repetition effect, the repeated temporal co-occurrence of  
990 items provides a strong and non-linear processing advantage for the predictable items.  
991 Following Hebbian learning principles (e.g., Brunel & Lavigne, 2009; Endress & Johnson,  
992 2021; Tovar et al., 2018), the coactivation of populations of neurons coding for each item may  
993 result in the strengthening of the connection weights between these two populations, leading  
994 to the creation of a chunk. Another possibility is to assume that both populations of neurons  
995 are activating a third population of pair-coding neurons (Miyashita, 2004) that would code for  
996 the pairing of these items. Irrespective of these two possible implementations, the present data  
997 suggest that these learning dynamics are non-linear, with a fast development of the memory  
998 trace of the chunk followed by a slower regime of memory consolidation.



999           Although the broken-stick analyses did not permit differentiation of the processing  
 1000 dynamics of the three types of words used in Experiment 1, 3, and 4 (i.e., unrelated words,  
 1001 semantically related words, and idioms, respectively), the predictability scores reported in  
 1002 Figure 6 indicate that the processing of idioms benefited more from the predictability effect  
 1003 than the processing of semantically related words, which also benefited more from the  
 1004 predictability effect than the unrelated words of Experiment 1. This is in line with previous  
 1005 studies showing that prior linguistic knowledge influences and facilitates regularity extraction  
 1006 (e.g., Elazar et al., 2022; Siegelman, Bogaerts, Elazar, et al., 2018). Pre-existing associations  
 1007 between words would then support the predictability effect and notably for idioms which are  
 1008 sequences that are supposedly already represented and supported by memory traces.

1009           It is difficult however to determine whether the advantage for idioms was mainly  
 1010 supported by implicit associative learning or by the participants' prior knowledge of idioms.  
 1011 Table 11 provides a summary of the participants' responses to the final questionnaire, and it  
 1012 clearly suggests that participants' explicit knowledge resulted in stronger learning for idioms  
 1013 compared to semantically related words, which only benefited from implicit learning.  
 1014 Therefore, participants' explicit knowledge of the sequence may have interacted with implicit  
 1015 associative learning mechanisms and the stronger predictability score obtained for idioms  
 1016 may be a product of both factors.

1017 **Table 11**

1018 Participants' responses to the questionnaire expressed in percentages for each experiment.

Experiment	Participants who noticed a repeated sequence	Participants who correctly recalled the sequence	Participants who did not recall any words of the sequence
Unrelated words (1)	98	38	7
Pseudowords (2)	93	29	31
Semantically related words (3)	98	69	7
Idioms (4)	100	88	2

1019

1020 Finally, the present data did not reveal a processing advantage for the third position  
1021 over the second, contrary to previous findings on regularity extraction in naming (Rey et al.,  
1022 2020) and visuomotor tasks (Minier et al., 2016; Rey et al., 2019, 2022). This is likely due to  
1023 the specific time-scale of the present experimental paradigm that does not allow chunking to  
1024 occur beyond two items. Indeed, for Hebbian learning to occur between the first and third  
1025 items in the repeated triplet, it certainly requires maintaining the activation of the neural  
1026 population coding for the first item long enough to be coactivated with the neural population  
1027 coding for the last item. However, contrary to previous experimental paradigms that have  
1028 reported a learning advantage on the last position of a triplet sequence, lexical decision takes a  
1029 longer processing time and requires greater attentional load. Both of these factors may lead to  
1030 a fast deactivation of items that were processed two steps before, avoiding any possible  
1031 association to occur between item one and three of the repeated triplets. This is consistent  
1032 with recent findings suggesting that long-distance associations are harder to establish and only  
1033 occur under very specific conditions (Tosatto et al., 2022; Wilson et al., 2018).

1034 The absence of effect on the third position of the triplets may also be related to a  
1035 limitation of the present study. Indeed, participants may have learnt two-item associations  
1036 during the task because stimulus presentation was sequential. It has been argued that parallel  
1037 presentation is essential for determining the creation of co-word dependencies (Snell et al.,  
1038 2018). Therefore, sequential presentation might have influenced word extraction and hindered  
1039 the formation of a 3-word chunk.

1040 In addition, a number of factors are likely to have influenced the learning dynamics  
1041 during the task, and thus constitute limitations to our study. First, one third of the words  
1042 forming the triplets in Experiments 1 and 3 can be considered as being part of existing  
1043 multiword sequences in French (e.g., collocations: “*noyau dur*”, “*tronc commun*”; idioms:  
1044 “*ramener sa fraise*”, “*battre le rappel*”). When presented with these words, participants may

1045 have already made predictions about the upcoming words in the sequence on the basis of  
1046 these pre-existing multiword sequences. However, thanks to the counterbalancing on  
1047 participants, the effect of these pre-existing co-occurrences should be minimized. In addition,  
1048 most of these existing multiword sequences contain words from parts of speech other than a  
1049 noun in the first and second positions, making the prediction of the last word in the triplet  
1050 nearly impossible given the preceding words.

1051         Second, the fact that the triplets in Experiment 4 consisted of different parts of speech  
1052 (i.e., noun, verb and adjective) compared to those in Experiments 1 and 3 (i.e., nouns) may  
1053 also have influenced their processing during the task. Indeed, reaction times have been shown  
1054 to differ across parts of speech (e.g., Kauschke & Stenneken, 2008; Kostić & Katz, 1987;  
1055 Monaghan et al., 2003; Sereno, 1999; Tyler et al., 2001). Similarly, since the triplets were not  
1056 matched in terms of MI across the experiments, certain words in some triplets are much less  
1057 predictive of the following words in the sequence. This is particularly the case in Experiment  
1058 4, where the verb *faire* (*to do* in English) is the first word in four triplets. Hence, it may be  
1059 difficult to directly compare the learning dynamics observed in Experiment 4 with those of  
1060 the other experiments. Future studies that control for these confounding factors are therefore  
1061 needed.

1062         Third, given the large number of triplet repetitions (i.e., 45), this task is far from  
1063 mimicking a real reading situation in which multiword sequences are widely spaced from one  
1064 another and occur much less frequently. Nevertheless, the use of a well-controlled  
1065 environment allowed us to characterise the acquisition of multiword sequences in real-time  
1066 and to investigate in depth the process of word-to-word associative learning in different  
1067 linguistic settings (i.e., unrelated words, novel words using pseudowords, semantically related  
1068 words and idioms). To gain a fuller picture of how multiword sequences are acquired, studies  
1069 employing more ecological presentation conditions, such as those of Conklin and Carrol

1070 (2020) and Sonbul et al. (2022), and using different types and larger multiword sequences are  
1071 needed.

1072 **Conclusion**

1073 The current study provides novel information about the learning dynamic of  
1074 multiword sequences when presented in a noisy environment, as is the case in natural  
1075 language. Our data suggests that multiword learning is carried out through chunking of local  
1076 information and shows how repetition affects the development of memory traces and  
1077 improves processing. To further explore and understand the dynamic of multiword sequences  
1078 extraction, future research could manipulate different parameters from the present  
1079 experimental Hebb lexical decision task, like for example, the spacing between two  
1080 repetitions of the repeated sequence or the size of the sequence, to determine the limits of the  
1081 conditions under which associative learning can occur between a sequence of words.

1082

1083

1084

1085

1086

1087

1088

1089

1090

1091

1092

1093

1094 Ethical approval and Informed Consent

1095 Before starting the experiment, participants accepted an online informed-consent form. Ethics  
1096 approval was obtained from the “Comité de Protection des Personnes SUD-EST IV”  
1097 (17/051).

1098 Open Practices Statements

1099 All materials for all experiments, data, and analysis code are available via the Open Science  
1100 Framework at <https://osf.io/h2yyn/>.

1101 Acknowledgments

1102 We would like to thank two anonymous reviewers for their constructive feedback on previous  
1103 versions of this paper.

1104 Declaration of conflicting interests

1105 The author(s) declared no potential conflicts of interest with respect to the research,  
1106 authorship, and/or publication of this article.

1107 Funding

1108 LPA was supported by a doctoral fellowship of the French Ministry of Higher Education,  
1109 Research, and Innovation. This research was supported by ERC grant 742141, the  
1110 Convergence Institute ILCB (ANR-16-CONV-0002), the centre for research in education  
1111 Ampiric, the CHUNKED ANR project (#ANR-17-CE28-0013-02), the COMPO ANR project  
1112 (#ANR-23-CE23-0031) and the HEBBIAN ANR project (#ANR-23-CE28-0008). For the  
1113 purpose of Open Access, a CC-BY 4.0<sup>2</sup> public copyright licence has been applied by the  
1114 authors to the present document and will be applied to all subsequent versions up to the  
1115 Author Accepted Manuscript arising from this submission.

1116

---

<sup>2</sup> <https://creativecommons.org/licenses/by/4.0/>

1117 References

- 1118 Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-  
 1119 based account of syntactic acquisition. *Linguistic Review*, 23(3), 275–290.  
 1120 <https://doi.org/10.1515/TLR.2006.011>
- 1121 Ambridge, B. (2020). Against stored abstractions: A radical exemplar model of language  
 1122 acquisition. *First Language*, 40(5–6), 509–559.  
 1123 <https://doi.org/10.1177/0142723719869731>
- 1124 Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of  
 1125 frequency effects in first language acquisition. *Journal of Child Language*, 42(2), 239–  
 1126 273. <https://doi.org/10.1017/S030500091400049X>
- 1127 Angele, B., Baciero, A., Gómez, P., & Perea, M. (2022). Does online masked priming pass  
 1128 the test? The effects of prime exposure duration on masked identity priming. *Behavior*  
 1129 *Research Methods*, 17. <https://doi.org/10.3758/s13428-021-01742-y>
- 1130 Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children  
 1131 and adults. *Cognitive Science*, 36(2), 286–304. [https://doi.org/10.1111/j.1551-](https://doi.org/10.1111/j.1551-6709.2011.01200.x)  
 1132 [6709.2011.01200.x](https://doi.org/10.1111/j.1551-6709.2011.01200.x)
- 1133 Arnon, I., & Christiansen, M. H. (2017). The role of multiword building blocks in explaining  
 1134 L1–L2 differences. *Topics in Cognitive Science*, 9(3), 621–636.  
 1135 <https://doi.org/10.1111/tops.12271>
- 1136 Arnon, I., & Clark, E. V. (2011). Why brush your teeth is better than teeth—Children’s word  
 1137 production is facilitated in familiar sentence-frames. *Language Learning and*  
 1138 *Development*, 7(2), 107–129. <https://doi.org/10.1080/15475441.2010.505489>
- 1139 Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases.  
 1140 *Journal of Memory and Language*, 62(1), 67–82.  
 1141 <https://doi.org/10.1016/j.jml.2009.09.005>

- 1142 Arnon, I., McCauley, S. M., & Christiansen, M. H. (2017). Digging up the building blocks of  
 1143 language: Age-of-acquisition effects for multiword phrases. *Journal of Memory and*  
 1144 *Language*, 92, 265–280. <https://doi.org/10.1016/j.jml.2016.07.004>
- 1145 Aslin, R. N. (2018). Statistical learning: A powerful mechanism that operates by mere  
 1146 exposure. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8, 1–12.  
 1147 <https://doi.org/10.1002/wcs.1373>.Statistical
- 1148 Bannard, C., & Lieven, E. (2012). Formulaic Language in L1 Acquisition. *Annual Review of*  
 1149 *Applied Linguistics*, 32, 3–16. <https://doi.org/10.1017/S0267190512000062>
- 1150 Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning of four-  
 1151 word combinations. *Psychological Science*, 19(3), 241–248.  
 1152 <https://doi.org/https://doi.org/10.1111/j.1467-9280.2008.02075.x>
- 1153 Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects  
 1154 models. *Frontiers in Psychology*, 4, 3–4. <https://doi.org/10.3389/fpsyg.2013.00328>
- 1155 Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for  
 1156 confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*,  
 1157 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- 1158 Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects  
 1159 models using lme4. *Journal of Statistical Software*, 67(1).  
 1160 <https://doi.org/10.18637/jss.v067.i01>
- 1161 Bonin, P., Méot, A., & Bugajska, A. (2013). Norms and comprehension times for 305 French  
 1162 idiomatic expressions. *Behavior Research Methods*, 45(4), 1259–1271.  
 1163 <https://doi.org/10.3758/s13428-013-0331-4>
- 1164 Bonin, P., Méot, A., Boucheix, J. M., & Bugajska, A. (2018). Psycholinguistic norms for 320  
 1165 fixed expressions (idioms and proverbs) in French. *Quarterly Journal of Experimental*  
 1166 *Psychology*, 71(5), 1057–1069. <https://doi.org/10.1080/17470218.2017.1310269>

- 1167 Borensztajn, G., Zuidema, W., & Bod, R. (2009). Children's grammars grow more abstract  
1168 with age-evidence from an automatic procedure for identifying the productive units of  
1169 language. *Topics in Cognitive Science*, *1*(1), 175–188. <https://doi.org/10.1111/j.1756->  
1170 [8765.2008.01009.x](https://doi.org/10.1111/j.1756-8765.2008.01009.x)
- 1171 Brunel, N., & Lavigne, F. (2009). Semantic priming in a cortical network model. *Journal of*  
1172 *Cognitive Neuroscience*, *21*(12), 2300-2319.
- 1173 Brysbaert, M. (2013). LEXTALE-FR a fast, free, and efficient test to measure language  
1174 proficiency in French. *Psychologica Belgica*, *53*(1), 23–37. <https://doi.org/10.5334/pb->  
1175 [53-1-23](https://doi.org/10.5334/pb-53-1-23)
- 1176 Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models:  
1177 A Tutorial. *Journal of Cognition*, *1*(1), 1–20. <https://doi.org/10.5334/joc.10>
- 1178 Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*,  
1179 *82*(4), 711–733.
- 1180 Bybee, J. (2013). Usage-based theory and exemplar representation. In T. Hoffman & G.  
1181 Trousdale (Eds.), *The oxford handbook of construction grammar* (pp. 49–69). Oxford  
1182 University Press.
- 1183 Carrol, G., & Conklin, K. (2020). Is all formulaic language created equal? Unpacking the  
1184 processing advantage for different types of formulaic sequences. *Language and Speech*,  
1185 *63*(1), 95–122. <https://doi.org/10.1177/0023830918823230>
- 1186 Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- 1187 Christiansen, M. H., & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental  
1188 constraint on language. *Behavioral and Brain Sciences*, *39*.  
1189 <https://doi.org/10.1017/S0140525X1500031X>
- 1190 Conklin, K., & Carrol, G. (2020). Words Go Together Like 'Bread and Butter': The Rapid,  
1191 Automatic Acquisition of Lexical Patterns. *Applied Linguistics*, *42*(3), 1–23.



- 1192 <https://doi.org/10.1093/applin/amaa034>
- 1193 Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly  
 1194 than nonformulaic language by native and nonnative speakers? *Applied Linguistics*,  
 1195 29(1), 72–89. <https://doi.org/10.1093/applin/amm022>
- 1196 Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical  
 1197 learning in language processing: Word predictability is the key. *Cognition*, 114(3), 356–  
 1198 371. <https://doi.org/10.1016/j.cognition.2009.10.009>
- 1199 Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*.  
 1200 Oxford University Press.
- 1201 Dufau, S., Grainger, J., & Ziegler, J. C. (2012). How to say “no” to a nonword: A leaky  
 1202 competing accumulator model of lexical decision. *Journal of Experimental Psychology:*  
 1203 *Learning, Memory, and Cognition*, 38(4), 1117–1128. <https://doi.org/10.1037/a0026948>
- 1204 Dumitru, M. (2016). Gestalt-like representations hijack Chunk-and-Pass processing.  
 1205 *Behavioral and Brain Sciences*, 39.
- 1206 Elazar, A., Alhama, R. G., Bogaerts, L., Siegelman, N., Baus, C., & Frost, R. (2022). When  
 1207 the “tabula” is anything but “rasa:” What determines performance in the auditory  
 1208 statistical learning task? *Cognitive Science*, 46(2). <https://doi.org/10.1111/cogs.13102>
- 1209 Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179–211.  
 1210 [https://doi.org/10.1016/0364-0213\(90\)90002-E](https://doi.org/10.1016/0364-0213(90)90002-E)
- 1211 Endress, A. D., & Johnson, S. P. (2021). When forgetting fosters learning: A neural network  
 1212 model for statistical learning. *Cognition*, 213, 104621.
- 1213 Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Augustinova, M., &  
 1214 Pallier, C. (2010). The French lexicon project: Lexical decision data for 38,840 French  
 1215 words and 38,840 pseudowords. *Behavior Research Methods*, 42(2), 488–496.  
 1216 <https://doi.org/10.3758/BRM.42.2.488>

- 1217 Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: A unified  
 1218 javascript framework for online studies. *2017 International Conference on*  
 1219 *Computational Social Science IC2S2*.
- 1220 Fournet, C., Mirault, J., Perea, M., & Grainger, J. (2022). Effects of letter case on processing  
 1221 sequences of written words. *Journal of Experimental Psychology: Learning, Memory,*  
 1222 *and Cognition, 48*(12), 1–9. <https://doi.org/10.1037/xlm0001179>
- 1223 French, R. M., Addyman, C., & Mareschal, D. (2011). TRACX: A recognition-based  
 1224 connectionist framework for sequence segmentation and chunk extraction. *Psychological*  
 1225 *Review, 118*(4), 614–636. <https://doi.org/10.1037/a0025255>
- 1226 Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*.  
 1227 Oxford University Press.
- 1228 Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision  
 1229 and naming. *Journal of Memory and Language, 29*(2), 228–244.  
 1230 [https://doi.org/10.1016/0749-596X\(90\)90074-A](https://doi.org/10.1016/0749-596X(90)90074-A)
- 1231 Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A  
 1232 multiple read-out model. *Psychological Review, 103*(3), 518–565.  
 1233 <https://doi.org/10.1037/0033-295X.103.3.518>
- 1234 Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye  
 1235 (Ed.), *Brain mechanisms and learning: A Symposium* (pp. 37–46).
- 1236 Ibbotson, P. (2013). The scope of usage-based theory. *Frontiers in Psychology, 4*, 1–15.  
 1237 <https://doi.org/10.3389/fpsyg.2013.00255>
- 1238 Isbilen, E. S., McCauley, S. M., & Christiansen, M. H. (2022). Individual differences in  
 1239 artificial and natural language statistical learning. *Cognition, 225*, 105123.  
 1240 <https://doi.org/10.1016/j.cognition.2022.105123>
- 1241 Jakubiček, M., Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013). The TenTen

- 1242 corpus family. *7th International Corpus Linguistics Conference CL*, 125–127.
- 1243 Janssen, N., & Barber, H. A. (2012). Phrase frequency effects in language production. *PLoS*  
 1244 *ONE*, 7(3). <https://doi.org/10.1371/journal.pone.0033202>
- 1245 Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production:  
 1246 Retrieval of syntactic information and of phonological form. *Journal of Experimental*  
 1247 *Psychology: Learning, Memory, and Cognition*, 20(4), 824–843.
- 1248 Jessop, A., Pine, J., & Gobet, F. (2023). *Chunk-based Incremental Processing and Learning:*  
 1249 *An integrated theory of word discovery, vocabulary growth, and speed of lexical*  
 1250 *processing*. PsyArXiv. <https://doi.org/10.31234/osf.io/dukpt>
- 1251 Jolsvai, H., McCauley, S. M., & Christiansen, M. H. (2020). Meaningfulness beats frequency  
 1252 in multiword chunk processing. *Cognitive Science*, 44(10), e12885.  
 1253 <https://doi.org/10.1111/cogs.12885>
- 1254 Jones, G., & Rowland, C. F. (2017). Diversity not quantity in caregiver speech: Using  
 1255 computational modeling to isolate the effects of the quantity and the diversity of the  
 1256 input on vocabulary growth. *Cognitive Psychology*, 98, 1–21.  
 1257 <https://doi.org/10.1016/j.cogpsych.2017.07.002>
- 1258 Kauschke, C., & Stenneken, P. (2008). Differences in noun and verb processing in lexical  
 1259 decision cannot be attributed to word form and morphological complexity alone. *Journal*  
 1260 *of Psycholinguistic Research*, 37(6), 443–452. [https://doi.org/10.1007/s10936-008-9073-](https://doi.org/10.1007/s10936-008-9073-3)  
 1261 3
- 1262 Kostić, A., & Katz, L. (1987). Processing differences between nouns, adjectives, and verbs.  
 1263 *Psychological Research*, 49(4), 229–236. <https://doi.org/10.1007/BF00309031>
- 1264 Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in  
 1265 linear mixed effects models. *Journal of Statistical Software*, 82(13).  
 1266 <https://doi.org/10.18637/jss.v082.i13>

- 1267 Lenth, R. V. (2023). emmeans: Estimated Marginal Means, aka Least-Squares Means. R  
 1268 package version 1.8.9. <https://CRAN.R-project.org/package=emmeans>.
- 1269 Liang, F., Blythe, H. I., Bai, X., Yan, G., Li, X., Zang, C., & Liversedge, S. P. (2017). The  
 1270 role of character positional frequency on Chinese word learning during natural reading.  
 1271 *PLOS ONE*, 12(11), e0187656. <https://doi.org/10.1371/journal.pone.0187656>
- 1272 Liang, F., Blythe, H. I., Zang, C., Bai, X., Yan, G., & Liversedge, S. P. (2015). Positional  
 1273 character frequency and word spacing facilitate the acquisition of novel words during  
 1274 Chinese children's reading. *Journal of Cognitive Psychology*, 27(5), 594–608.  
 1275 <https://doi.org/10.1080/20445911.2014.1000918>
- 1276 Liang, F., Gao, Q., Li, X., Wang, Y., Bai, X., & Liversedge, S. P. (2023). The importance of  
 1277 the positional probability of word final (but not word initial) characters for word  
 1278 segmentation and identification in children and adults' natural Chinese reading. *Journal*  
 1279 *of Experimental Psychology: Learning, Memory, and Cognition*, 49(1), 98–115.  
 1280 <https://doi.org/10.1037/xlm0001116>
- 1281 Liang, F., Ma, J., Bai, X., & Liversedge, S. P. (2021). Initial landing position effects on  
 1282 Chinese word learning in children and adults. *Journal of Memory and Language*, 116,  
 1283 104183. <https://doi.org/10.1016/j.jml.2020.104183>
- 1284 McCauley, S. M., & Christiansen, M. H. (2019). Language learning as language use: A cross-  
 1285 linguistic model of child language development. *Psychological Review*, 126(1), 1–51.  
 1286 <https://doi.org/10.1037/rev0000126>
- 1287 McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context  
 1288 effects in letter perception: Part 1. An account of basic findings. *Psychological Review*,  
 1289 88(5), 375–407. <https://doi.org/10.1037/0033-295X.88.5.375>
- 1290 Meylan, S. C., Frank, M. C., Roy, B. C., & Levy, R. (2017). The emergence of an abstract  
 1291 grammatical category in children's early speech. *Psychological Science*, 28(2), 181–192.

- 1292 <https://doi.org/10.1177/0956797616677753>
- 1293 Minier, L., Fagot, J., & Rey, A. (2016). The temporal dynamics of regularity extraction in  
 1294 non-human primates. *Cognitive Science*, 40(4), 1019–1030.  
 1295 <https://doi.org/10.1111/cogs.12279>
- 1296 Mirault, J., Snell, J., & Grainger, J. (2018). You that read wrong again! A transposed-word  
 1297 effect in grammaticality judgments. *Psychological Science*, 29(12), 1922–1929.  
 1298 <https://doi.org/10.1177/0956797618806296>
- 1299 Miyashita, Y. (2004). Cognitive memory: cellular and network machineries and their top-  
 1300 down control. *Science*, 306, 435-440.
- 1301 Monaghan, P., Chater, N., & Christiansen, M. H. (2003). Inequality between the classes:  
 1302 Phonological and distributional typicality as predictors of lexical processing.  
 1303 *Proceedings of the Annual Meeting of the Cognitive Science Society*, 25, 810–815.
- 1304 Mosse, E. K., & Jarrold, C. (2008). Hebb learning, verbal short-term memory, and the  
 1305 acquisition of phonological forms in children. *Quarterly Journal of Experimental*  
 1306 *Psychology*, 61(4), 505–514. <https://doi.org/10.1080/17470210701680779>
- 1307 Muggeo, V. M. R. (2008). Segmented: An R package to fit regression models with broken-  
 1308 line relationships. *R News*, 8(1), 20–25. <https://cran.r-project.org/doc/Rnews/>
- 1309 New, B., & Pallier, C. (2020). *Lexique 3.83*. <http://www.lexique.org>
- 1310 Norris, D., Page, M. P. A., & Hall, J. (2018). Learning nonwords: the Hebb repetition effect  
 1311 as a model of word learning. *Memory*, 26(6), 852–857.  
 1312 <https://doi.org/10.1080/09658211.2017.1416639>
- 1313 Northbrook, J., Allen, D., & Conklin, K. (2022). ‘Did you see that?’—The role of repetition  
 1314 and enhancement on lexical bundle processing in English learning materials. *Applied*  
 1315 *Linguistics*, 43(3), 453–472. <https://doi.org/10.1093/applin/amab063>
- 1316 Ordonez Magro, L., Mirault, J., Grainger, J., & Majerus, S. (2022). Sequential versus

- 1317 simultaneous presentation of memoranda in verbal working memory: (How) does it  
 1318 matter? *Memory and Cognition*. <https://doi.org/10.3758/s13421-022-01284-4>
- 1319 Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb  
 1320 repetition effect and the learning of phonological word forms. *Philosophical*  
 1321 *Transactions of the Royal Society B: Biological Sciences*, 364(1536), 3737–3753.  
 1322 <https://doi.org/10.1098/rstb.2009.0173>
- 1323 Page, M. P. A., Cumming, N., Norris, D., McNeil, A. M., & Hitch, G. J. (2013). Repetition-  
 1324 spacing and item-overlap effects in the Hebb repetition task. *Journal of Memory and*  
 1325 *Language*, 69(4), 506–526. <https://doi.org/10.1016/j.jml.2013.07.001>
- 1326 Pellicer-Sánchez, A. (2017). Learning L2 collocations incidentally from reading. *Language*  
 1327 *Teaching Research*, 21(3), 381–402. <https://doi.org/10.1177/1362168815618428>
- 1328 Pellicer-Sánchez, A., Siyanova-Chanturia, A., & Parente, F. (2022). The effect of frequency  
 1329 of exposure on the processing and learning of collocations: A comparison of first and  
 1330 second language readers' eye movements. *Applied Psycholinguistics*, 43(3), 1–30.  
 1331 <https://doi.org/10.1017/S014271642200011X>
- 1332 Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by  
 1333 8-month-old infants. *Child Development*, 80(3), 674–685. [https://doi.org/10.1111/j.1467-](https://doi.org/10.1111/j.1467-8624.2009.01290.x)  
 1334 [8624.2009.01290.x](https://doi.org/10.1111/j.1467-8624.2009.01290.x)
- 1335 Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of*  
 1336 *Memory and Language*, 39(2), 246–263. <https://doi.org/10.1006/jmla.1998.2576>
- 1337 Pinker, S. (1991). Rules of language. *Science*, 253(5019), 530–535.  
 1338 <https://doi.org/10.1126/science.1857983>
- 1339 Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive*  
 1340 *Sciences*, 6(11), 456–463. [https://doi.org/10.1016/S1364-6613\(02\)01990-3](https://doi.org/10.1016/S1364-6613(02)01990-3)
- 1341 Ramisch, C. (2015). *Multiword Expressions Acquisition*. Springer.

- 1342 <https://doi.org/10.1007/978-3-319-09207-2>
- 1343 Rey, A., Bogaerts, L., Tosatto, L., Bonafos, G., Franco, A., & Favre, B. (2020). Detection of  
 1344 regularities in a random environment. *Quarterly Journal of Experimental Psychology*,  
 1345 73(12), 1–13. <https://doi.org/10.1177/1747021820941356>
- 1346 Rey, A., Fagot, J., Mathy, F., Lazartigues, L., Tosatto, L., Bonafos, G., Freyermuth, J.-M., &  
 1347 Lavigne, F. (2022). Learning higher-order transitional probabilities in nonhuman  
 1348 primates. *Cognitive Science*, 46(4), e13121. Doi: 10.1111/cogs.13121
- 1349 Rey, A., Minier, L., Malassis, R., Bogaerts, L., & Fagot, J. (2019). Regularity extraction  
 1350 across species: Associative learning mechanisms shared by human and non-human  
 1351 primates. *Topics in Cognitive Science*, 11(3), 573–586.  
 1352 <https://doi.org/10.1111/tops.12343>
- 1353 Robinet, V., Lemaire, B., & Gordon, M. B. (2011). MDLChunker: A MDL-Based cognitive  
 1354 model of inductive learning. *Cognitive Science*, 35, 1352–1389.  
 1355 <https://doi.org/10.1111/j.1551-6709.2011.01188.x>
- 1356 Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old  
 1357 infants. *Science*, 274(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>
- 1358 Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental  
 1359 language learning: Listening (and learning) out of the corner of your ear. *Psychological*  
 1360 *Science*, 8(2), 101–105. <https://doi.org/10.1111/j.1467-9280.1997.tb00690.x>
- 1361 Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition  
 1362 effects in lexical memory. *Journal of Experimental Psychology: Human Perception and*  
 1363 *Performance*, 3(1), 1–17. <https://doi.org/10.1037/0096-1523.3.1.1>
- 1364 Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori  
 1365 contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, 110,  
 1366 104038. <https://doi.org/10.1016/j.jml.2019.104038>

- 1367 Schimke, E. A. E., Angwin, A. J., Cheng, B. B. Y., & Copland, D. A. (2021). The effect of  
 1368 sleep on novel word learning in healthy adults: A systematic review and meta-analysis.  
 1369 *Psychonomic Bulletin & Review*, 28(6), 1811–1838. [https://doi.org/10.3758/s13423-021-](https://doi.org/10.3758/s13423-021-01980-3)  
 1370 01980-3
- 1371 Sereno, J. A. (1999). Hemispheric differences in grammatical class. *Brain and Language*,  
 1372 70(1), 13–28. <https://doi.org/10.1006/brln.1999.2137>
- 1373 Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in  
 1374 statistical learning: Current pitfalls and possible solutions. *Behavior Research Methods*,  
 1375 49(2), 418–432. <https://doi.org/10.3758/s13428-016-0719-z>
- 1376 Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic  
 1377 entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, 177,  
 1378 198–213. <https://doi.org/10.1016/j.cognition.2018.04.011>
- 1379 Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2018). Redefining “learning” in  
 1380 statistical learning: What does an online measure reveal about the assimilation of visual  
 1381 regularities? *Cognitive Science*, 42, 692–727. <https://doi.org/10.1111/cogs.12556>
- 1382 Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An  
 1383 eye-tracking study of idiom processing by native and non-native speakers. *Second*  
 1384 *Language Research*, 27(2), 251–272. <https://doi.org/10.1177/0267658310382068>
- 1385 Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. J. B. (2011). Seeing a phrase “time  
 1386 and again” matters: The role of phrasal frequency in the processing of multiword  
 1387 sequences. *Journal of Experimental Psychology: Learning Memory and Cognition*,  
 1388 37(3), 776–784. <https://doi.org/10.1037/a0022531>
- 1389 Skarabela, B., Ota, M., O’Connor, R., & Arnon, I. (2021). ‘Clap your hands’ or ‘take your  
 1390 hands’? One-year-olds distinguish between frequent and infrequent multiword phrases.  
 1391 *Cognition*, 211, 104612. <https://doi.org/10.1016/j.cognition.2021.104612>



- 1392 Smalle, E. H. M., Bogaerts, L., Simonis, M., Duyck, W., Page, M. P. A., Edwards, M. G., &  
1393 Szmalec, A. (2016). Can chunk size differences explain developmental changes in  
1394 lexical learning? *Frontiers in Psychology*, *6*, 1–14.  
1395 <https://doi.org/10.3389/fpsyg.2015.01925>
- 1396 Smith, S. A., & Murphy, V. A. (2015). Measuring productive elements of multi-word phrase  
1397 vocabulary knowledge among children with English as an additional or only language.  
1398 *Reading and Writing*, *28*(3), 347–369. <https://doi.org/10.1007/s11145-014-9527-y>
- 1399 Snell, J., & Theeuwes, J. (2020). A story about statistical learning in a story: Regularities  
1400 impact eye movements during book reading. *Journal of Memory and Language*, *113*,  
1401 104127. <https://doi.org/10.1016/j.jml.2020.104127>
- 1402 Snell, J., Leipsig, S. Van, & Grainger, J. (2018). OB1-Reader: A Model of Word Recognition  
1403 and Eye Movements in Text Reading. *Psychological Review*, *125*(6), 969–984.
- 1404 Solan, Z., Horn, D., Ruppin, E., & Edelman, S. (2005). Unsupervised learning of natural  
1405 languages. *Proceedings of the National Academy of Sciences*, *102*(33), 11629–11634.  
1406 <https://doi.org/10.1073/pnas.0409746102>
- 1407 Sonbul, S., El-Dakhs, D. A. S., Conklin, K., & Carrol, G. (2022). “Bread and butter” or  
1408 “butter and bread”? Nonnatives’ processing of novel lexical patterns in context. *Studies*  
1409 *in Second Language Acquisition*, 1–23. <https://doi.org/10.1017/S0272263122000237>
- 1410 Szmalec, A., Duyck, W., Vandierendonck, A., Mata, A. B., & Page, M. P. A. (2009). The  
1411 Hebb repetition effect as a laboratory analogue of novel word learning. *Quarterly*  
1412 *Journal of Experimental Psychology*, *62*(3), 435–443.  
1413 <https://doi.org/10.1080/17470210802386375>
- 1414 Szmalec, A., Page, M. P. A., & Duyck, W. (2012). The development of long-term lexical  
1415 representations through Hebb repetition learning. *Journal of Memory and Language*,  
1416 *67*(3), 342–354. <https://doi.org/10.1016/j.jml.2012.07.001>

- 1417 Tomasello, M. (2003). *Constructing a language: A usage-based theory of language*  
 1418 *acquisition* (1st ed.). Harvard University Press.
- 1419 Tosatto, L., Bonafos, G., Melmi, J.-B., & Rey A. (2022). Detecting non-adjacent  
 1420 dependencies is the exception rather than the rule. *PLoS ONE*, *17*(7), e0270580.  
 1421 Doi: 10.1371/journal.pone.0270580
- 1422 Tovar, Á. E., Westermann, G., & Torres, A. (2018). From altered synaptic plasticity to  
 1423 atypical learning: A computational model of Down syndrome. *Cognition*, *171*, 15-24.
- 1424 Treiman, R., Gordon, J., Boada, R., Peterson, R. L., & Pennington, B. F. (2014). Statistical  
 1425 learning, letter reversals, and reading. *Scientific Studies of Reading*, *18*(6), 383–394.  
 1426 <https://doi.org/10.1080/10888438.2013.873937>
- 1427 Tyler, L. K., Richard, R., Fadili, J., & Moss, H. E. (2001). The neural representation of nouns  
 1428 and verbs: PET studies. *Brain*, *124*(8), 1619–1634.  
 1429 <https://doi.org/10.1093/brain/124.8.1619>
- 1430 Ullman, M. T., Pancheva, R., Love, T., Yee, E., Swinney, D., & Hickok, G. (2005). Neural  
 1431 correlates of lexicon and grammar: Evidence from the production, reading, and judgment  
 1432 of inflection in aphasia. *Brain and Language*, *93*(2), 185–238.  
 1433 <https://doi.org/10.1016/j.bandl.2004.10.001>
- 1434 Wilson, B., Spierings, M., Ravignani, A., Mueller, J. L., Mintz, T. H., Wijnen, F., vander  
 1435 Kant, A., Smith, K., & Rey, A. (2018). Non- adjacent dependency learning in humans  
 1436 and other animals. *Topics in Cognitive Science*, *12*(3), 843-858.
- 1437 Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.
- 1438 Wray, A. (2012). What do we (think we) know about formulaic language? An evaluation of  
 1439 the current state of play. *Annual Review of Applied Linguistics*, *32*, 231–254.  
 1440 <https://doi.org/10.1017/S026719051200013X>
- 1441 Zang, C., Fu, Y., Du, H., Bai, X., Yan, G., & Liversedge, S. P. (2023). Processing

1442 multiconstituent units: Preview effects during reading of Chinese words, idioms, and  
1443 phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.  
1444 <https://doi.org/10.1037/xlm0001234>

1445

1446

1447 Appendix A. Participants' scores and standard deviations on the LexTALE task for each  
 1448 experiment.  
 1449

Experiment	Vocabulary score	<i>SD</i>
Unrelated words (1)	86.53%	5.76
Pseudowords (2)	85.13%	7.08
Semantically related words (3)	85.08%	6.37
Idioms (4)	86.18%	6.19

1450  
 1451

1452

Appendix B. List of triplets per participant and fillers in Experiment 1

Participant	Triplet
1	armure gilet nectar
2	armure nectar gilet
3	gilet armure nectar
4	gilet nectar armure
5	nectar armure gilet
6	nectar gilet armure
7 ... 12	mule proie noeud
13 ... 18	virage rasoir festin
19 ... 24	livret tirage cloche
25 ... 30	graine tronc berger
31 ... 36	rappel noyau palace
37 ... 42	calcul balai volcan

**Filler words**

angle armée assaut bague balle base câble canard chute cible  
coton dent doute fuite grève hasard jardin lion moteur musée  
neige nuage offre orage ours papier parole pause perte pilote  
pluie porc prince rideau roue signe site soupe tarte tenue tigre  
trafic valise violon voisin

**Filler pseudowords**

abréne acogne acrule acun agarte ainte alire alme altace anet  
anide antôt appome arine artal arti arut arêl asode atinle  
augard ausi autoce blose boce borté bugue bune carc caruce  
catail catire cepame cerson cetir chamir charde charon chefet  
cheler chouï chroid chume ciroir claint cluise coble cocère  
colmel counai crupe cumite cuse damade danfum degite derler  
ditrer double drugue dévede elsir engade ensime falber farbe  
fauf flerse folde forni frone fube fule garsu gitre glac granal  
gretro grode grort grupe gumble hatour heudit hoire hontôt  
hougue humice inssir iple ipéth iseau jada japite jutand laitôt  
latace lavec lieune lipin léble léfile lémece mevail miman  
molome moléce monner mèlui môvec naille natéme noil nomsi  
nopore noxe omante oufage ounite pachet paseau pecran pertif  
piale plaze pleité pordée preper preur preuro psat puif pérée  
ranu relle renchi renre renume rerile rerise retave ricit rocèle  
roulip sabe sarie satu sepoce sesin soite sona sorat sounir spho  
spile suine sule supe taivec taute touet touge toutôt trounu  
tulque tyle uant ucun unate vecote ventin visise vosé voule  
vrande vèpel âleur ésale êder îcun

1453

1454

1455

Appendix C. List of triplets per participant and fillers in Experiment 2

<b>Participant</b>	<b>Triplet</b>
1	acré mouffe brague
2	acré brague mouffe
3	mouffe acré brague
4	mouffe brague acré
5	brague acré mouffe
6	brague mouffe acré
7 ... 12	étrin rompte flais
13 ... 18	charpe crêle joine
19 ... 24	souffe prompe dige
25 ... 30	bloue sprère vauve
31 ... 36	fitre plou boge
37 ... 42	relet harte loude

**Filler words**

acier acteur adieu agence aile angle anneau armée assaut auteur  
 bague balle ballon banque barbe base beurre bible bière blague  
 bombe bonté bourse bouton bruit canard carte cesse change  
 chat chaîne choc chute cible cirque client code coffre copain  
 corde coton course crème cuir câble dent destin dette devoir  
 disque douche doute drôle défi départ dîner empire emploi  
 ennui ferme fleuve forme forêt four fuite fumée genou golf  
 grâce grève gâteau génie hasard herbe hiver huile humour  
 image indice jardin joie lame ligne lion loup lycée magie  
 mairie mardi milieu moteur mouche musée nature navire neige  
 neveu nuage objet offre ombre orage ours page pain papier  
 parfum pari parole pause perte pilote plage pluie poche poil  
 pomme porc port poste potion pouce poème presse preuve  
 prince prénom prêtre pêche radio rideau robe rocher rose roue  
 rythme régime salade salon sauce savon scène signe site siège  
 soirée souci soupe sport statue style tabac tante tarte tasse  
 temple tenue texte tigre toile tombe trafic trou troupe trésor  
 tuyau vague valise vallée ventre veste violon vitre voeu voie  
 voisin vote vélo écoute épée

**Filler pseudowords**

ainte alire altace antôt arut atinle bune charon chefet choui  
 ciroir cocère counai cuse engade ensime falber farbe fauf flerse  
 fube gretro iple iseau jada latace lieune mèlui naille ounite  
 paseau puif renre retave rocèle sepoce sesin soral suine trounu  
 vecote voule vrande âleur ésale

1456

1457

1458

## Appendix D. List of triplets per participant and fillers in Experiment 3

Participant	Triplet
1	banane cerise fraise
2	banane fraise cerise
3	cerise banane fraise
4	cerise fraise banane
5	fraise banane cerise
6	fraise cerise banane
7 ... 12	coco figue poire
13 ... 18	citron tomate oignon
19 ... 24	saumon truite requin
25 ... 30	pigeon dinde hibou
31 ... 36	coyote renard lièvre
37 ... 42	cobaye tortue taupe

**Filler words**

angle armée assaut bague balle base câble canard chute cible  
coton dent doute fuite grève hasard jardin lion moteur musée  
neige nuage offre orage ours papier parole pause perte pilote  
pluie porc prince rideau roue signe site soupe tarte tenue tigre  
trafic valise violon voisin

**Filler pseudowords**

abréne acogne acrule acun agarte ainte alire alme altace anet  
anide antôt appome arine artal arti arut arêle asode atinle  
augard ausi autoce blose boce borté buge bune carc caruce  
catail catire cepame cerson cetir chamir charde charon chefet  
cheler chouï chroid chume ciroir claint cluise coble cocère  
colmel counai crupe cumite cuse damade danfum degite derler  
ditrer double drugue dévede elsir engade ensime falber farbe  
fauf flerse folde forni frone fube fule garsu gitre glac granal  
gretro grode grort grupe gumble hatour heudit hoire hontôt  
hougue humice inssir ipele ipéth iseau jada japite jutand laitôt  
latace lavec lieune lipin léble léfile lémece mevail miman  
molome moléce monner mèlui môvec naille natéme noil nomsi  
nopore noxe omante oufage ounite pachet paseau pecran pertif  
piale plaze pleité pordée preper preur preuro psat puif pérée  
ranu relle renchi renre renume rerile rerise retave ricit rocèle  
roulip sabe sarie satu sepoce sesin soite sona soral sounir spho  
spile suine sule supe taivec taute touet touge toutôt trounu  
tulque tyle uant ucun unate vecote ventin visise vosé voule  
vrande vèpel âleur ésale êder îcun

1459

1460

1461

Appendix E. List of triplets per participant and fillers in Experiment 4

<b>Participant</b>	<b>Triplet</b>
1 ... 6	être mauvais joueur
7 ... 12	faire fausse route
13 ... 18	qui dort dîne
19 ... 24	faire chou blanc
25 ... 30	faire profil bas
31 ... 36	faire grise mine
37 ... 42	montrer patte blanche

**Fillers**

angle armée assaut bague balle base câble canard chute cible  
coton dent doute fuite grève hasard jardin lion moteur musée  
neige nuage offre orage ours papier parole pause perte pilote  
pluie porc prince rideau roue signe site soupe tarte tenue tigre  
trafic valise violon voisin

**Filler pseudowords**

abréne acogne acrule acun agarte ainte alire alme altace anet  
anide antôt appome arine artal arti arut arêle asode atinle  
augard ausi autoce blose boce borté buge bune carc caruce  
catail catire cepame cerson cetir chamir charde charon chefet  
cheler chouï chroid chume ciroir claint cluise coble cocère  
colmel counai crupe cumite cuse damade danfum degite derler  
ditrer double drugue dévede elsir engade ensime falber farbe  
fauf flerse folde forni frone fube fule garsu gitre glac granal  
gretro grode grort grupe gumble hatour heudit hoire hontôt  
hougue humice inssir iple ipéth iseau jada japite jutand laitôt  
latace lavec lieune lipin léble léfile lémece mevail miman  
molome moléce monner mèlui môvec naille natéme noil nomsi  
nopore noxe omante oufage ounite pachet paseau pecran pertif  
piale plaze pleité pordée preper preur preuro psat puif pérée  
ranu relle renchi renre renume rerile rerise retave ricit rocèle  
roulip sabe sarie satu sepoce sesin soite sona sorat sounir spho  
spile suine sule supe taivec taute touet touge toutôt trounu  
tulque tyle uant ucun unate vecote ventin visise vosé voule  
vrande vêpel âleur ésale êder îcun

1462

1463

1464