Original Article

QJEP

Quarterly Journal of Experimental Psychology
2021, Vol. 74(9) 1542–1552

DOI: 10.1177/17470218211012960
qjep.sagepub.com
SAGE

Are better young readers more likely to confuse their mother with their mohter?

Pablo Gomez[1], Ana Marcet[2] and Manuel Perea[2,3]

Abstract

One of the most replicated effects in the contemporary word recognition literature is the transposed-letter effect (TL effect): pseudowords created by the transposition of two letters (e.g., MOHTER) are often misread as the real word. This effect ruled out those accounts that assume that letter position is encoded accurately and led to more flexible coding schemes. Here, we examined whether reading skill modulates this effect. The relationship between reading skill and the TL effect magnitude is a contentious issue both empirically and theoretically. The present lexical decision experiment was designed to shed some light on the relationship between reading skill and the TL effect magnitude with a large sample of Grade 6 children. To that end, we conducted both multiple regression and path analyses. Results showed that a specific aspect of reading skills (pseudoword reading) negatively correlates with the TL effect's magnitude in the error data (i.e., MOHTER is less wordlike for better readers). This finding highlights the need for a comprehensive visual-word recognition model that includes individual variability and the multidimensional character of reading in school-age children.

Keywords

Letter position coding; lexical decision; individual variability; word recognition

Received: 7 May 2020; revised: 10 February 2021; accepted: 14 February 2021

One of the critical manipulations to crack the orthographic code during word recognition is the use of transposed-letter (TL) stimuli (see Grainger, 2008, 2018). As first shown by Bruner and O'Dowd (1958), Chambers (1979), and O'Connor and Forster (1981), pseudowords created by the transposition of two letters (JUGDE) are often misread as their base word. Indeed, JUGDE and JUDGE produce a very similar N400 component (i.e., a marker of lexical-semantic processing; see Vergara-Martínez et al., 2013). Furthermore, the TL similarity effect is also very robust during sentence reading. Rayner et al. (2006) found that TL pseudowords inserted in sentences (e.g., "he saw a young jugde at the court today" in which *jugde* is the TL pseudoword) barely disrupted eye movement control when compared with intact sentences, especially for internal transpositions (see also Blythe et al., 2014, for further evidence).

The robustness of the TL effect across paradigms rules out those models of visual-word recognition that assume that letter position is encoded with precision (e.g., McClelland & Rumelhart, 1981, interactive activation model, and its descendants). This effect is highly replicable and has been found in different families of languages that use the Latin script: Romance language like Spanish, Italian, Portuguese, and French (e.g., Schoonbaert & Grainger, 2004); Germanic languages like German and English (e.g., Perea & Lupker, 2004); Semitic languages like Maltese (e.g., Perea et al., 2012); and pre-Indo-European languages like Basque (e.g., Perea & Carreiras, 2006a). Indeed, the TL phenomenon led to implementing quantitative models of visual-word recognition with a flexible scheme when encoding letter position within words (see Grainger, 2018, for review).[1] Most empirical evidence for the TL effect comes from two paradigms. First, single-presentation techniques in which latency and accuracy for TL pseudowords (e.g., JUGDE) are typically compared with orthographic controls (e.g., replacement-letter [RL]

[1]California State University San Bernardino, Palm Desert, CA, USA
[2]Universitat de València, València, Spain
[3]Universidad Nebrija, Hoyo de Manzanares, Spain

Corresponding author:
Pablo Gomez, Department of Psychology, California State University San Bernardino, Ringgold Standard Institution, 37500 Cook, Palm Desert, CA 92211, USA.
Email: Pablo.Gomez@csusb.edu

pseudowords, as JUPTE). The TL effect is defined as the difference in performance (i.e., latency and error rates) between pseudowords created by transposing two consonants (TL pseudowords) and pseudowords created by replacing two consonants (RL pseudowords). In previous research, TL pseudowords (e.g., JUGDE) have shown both longer latencies and more errors (i.e., incorrect "word" responses) than RL pseudowords (e.g., JUPTE). Second, priming paradigms (e.g., masked priming; boundary technique), in which the relationship between a masked prime (or parafoveal preview) and a target word is manipulated (e.g., the TL pair jugde-JUDGE vs. the RL pair jupte-JUDGE). The TL priming effect is defined as the difference between the response times (RTs) (or error rates) on the target word when preceded by an RL prime and when preceded by a TL prime. In prior studies, RTs to a target word (e.g., JUDGE) are faster when preceded by TL pseudoword prime (e.g., judge) than when preceded by an RL pseudoword prime (e.g., jupte). Although these methodologies are related, they answer different questions (Andrews, 1997). Single-presentation techniques directly answer the question of "how much lexical activity is generated by the transposed-letter pseudoword?" In contrast, priming paradigms answer the question of "once the orthographic representation of the transposed-letter prime/preview is pre-activated, how does the processing of the target word proceed?" Here we focus on the first question, but not to imply that the second question is less valid.

To fully comprehend the nature of letter position coding in word recognition, we should examine how reading ability may modulate this process. There are two competing formulations. On one hand, better reading skills could be associated with a better ability to encode precise letter positions. Castles et al. (2007) proposed that, as reading abilities develop, letter position coding becomes more accurate (lexical tuning account). Similarly, Perfetti's lexical quality hypothesis (Perfetti & Hart, 2002) assumes that better-skilled readers have more stable orthographic representations than less-skilled readers—this includes the encoding of letter position (see Perfetti, 2017). Thus, in lexical tuning and lexical quality accounts, better readers can encode orthographic representations with higher precision than less-skilled readers, thus yielding smaller TL effects. This mechanism of lexical tuning/quality could be modelled as a reduction of the uncertainty in the locations of letters in computational models of visual-word recognition (e.g., *s* parameter in the overlap model, Gomez et al., 2008; *sigma* parameter in the spatial coding model, Davis, 2010).

On the other hand, better reading skills could be associated with a less precise letter position coding. In Grainger and Ziegler's (2011) dual-rote model of visual-word recognition, there are two primary routes to the lexicon: phonological and orthographic. The phonological route would emerge first in reading development, requiring serial letter processing (i.e., precise letter position coding). In contrast, the orthographic route would enable more parallel and faster processing of words via coarser letter position coding. The TL effect would increase with reading development in the dual-route model because of the greater reliance of better readers on the more efficient orthographic route (see Ziegler et al., 2014). However, note that these accounts (lexical tuning and dual-route models in particular) are concerned mostly with the *development* of reading skills rather than on individual differences per se.

The empirical examination of whether reading skills (as measured by standardised reading tests) modulate letter position coding in word recognition tasks has been scarce. To our knowledge, no previous studies have directly tackled this issue with a single-presentation technique. Indirect evidence was obtained in a lexical decision experiment conducted by Perea et al. (2016). They compared the size of the TL effect of individuals with outstanding orthographic-lexical capabilities (competitive Scrabble players) and a control group of readers with no expertise in Scrabble. Perea et al. (2016) found a substantially smaller TL effect for the Scrabble players. They interpreted these findings as favouring the lexical tuning and lexical quality accounts: competitive Scrabble players likely acquired highly stable orthographic representations that allowed them to encode letter position in word recognition tasks more precisely than non-players. However, Perea et al. (2016) did not obtain a standardised reading measure of the participants, limiting their findings' implications. Direct evidence of the role of reading abilities in letter position coding has been obtained in two masked priming experiments (Andrews & Lo, 2012; Ziegler et al., 2014). Andrews and Lo (2012) conducted a masked priming experiment with university students and collected reading proficiency measures. Each target word could be preceded by a TL prime, a one-letter different neighbour prime, or an unrelated prime. They also manipulated the primes' lexical status: word primes (*colt-CLOT*, *plot-CLOT*, *punt-CLOT*) versus pseudoword primes (*crue-CURE*, *cire-CURE*, *gine-CURE*). Andrews and Lo (2012) found facilitation from TL primes for those readers with low levels of lexical expertise. They concluded that readers with "lower quality representations do not accurately code letter order" (p. 159). The lack of a similar effect for those readers with high levels of lexical expertise was interpreted as a more precise letter position coding, offering some support to Castles et al.'s (2007) lexical tuning account and Perfetti and Hart's (2002) lexical quality hypothesis. However, this work has limitations: the experiment lacked the appropriate control condition for TL primes (i.e., an RL condition [*cnoe-CRUE*] or an identity condition [*cure-CURE*]; see Kezilas et al., 2017). Also, the overall difference between *crue-CURE* and *gine-CURE* was only 8 ms, making it challenging to establish conclusions on if (or how) a variable might modulate such effect.

More recently, Ziegler et al. (2014) conducted a masked priming experiment with readers from Grades 1 to 5 and examined the relationship between scores in a standardised reading test and letter position coding using the sandwich variant of masked priming (Lupker & Davis, 2009): the prime was not immediately presented after a forward mask, but it was briefly preceded by a 27 ms target (500 ms mask, 27 ms target, 57 ms prime, TARGET). The TL priming effect was measured as the difference in latencies between a TL pseudoword condition (*cousre-COURSE*) and an RL pseudoword condition (*coufpe-COURSE*). Latencies differed widely among grades (around 2 s for Grade 1 children and around 700–800 ms for Grade 5 children); hence, Ziegler et al. computed *z*-scores of the RT differences between the TL and RL conditions. They found a positive relationship between the size of the TL priming effect and reading ability ($r = .304$). That is, better readers showed larger masked TL priming effects, thus favouring Grainger and Ziegler's (2011) dual-route model. However, it may be difficult to make inferences between the TL effect in beginning readers (Grade 1) and intermediate, independent readers (Grade 5). Performance in a lexical decision task (LDT) is a convolution of several components like motor processes (executing the keypress), core processes (lexical, semantic), strategic considerations (emphasis on speed vs. accuracy), and encoding processes (mapping the retinotopic input onto abstract representations of letters/words). The processes underlying lexical decision responses in Grade 1 and Grade 5 children may be quite different. It would have been desirable to examine whether the correlation between TL priming effect and reading ability occurs on a Grade-by-Grade basis.

In sum, whereas the Perea et al. (2016) and the Andrews and Lo (2012) experiments with adult readers produced evidence favouring the lexical tuning account, the Ziegler et al. (2014) experiment with developing readers (children of Grades 1–5) had evidence favouring the dual-route model. Likewise, the experiments that examined the magnitude of TL effects as a function of grade—under the assumption that younger children have less developed reading abilities than older children—have not produced a consistent pattern either (see Supplemental Appendix A for a summary of previous experiments with single-presentation lexical decision and masked priming, where the inconsistency is particularly noticeable in the RT data and less so in the accuracy data). In fact, Paterson et al. (2014) found that letter position coding in children around 9 years old is remarkably similar to adult readers. While the term *grade-level* reading as an index of reading skill might imply that a poor sixth-grade reader is just like an average third-grade reader (Davidson & Myhre, 2000), examining the variability within specific age groups is a critical piece of the reading development puzzle.

In the present study, we sought to shed light on whether reading abilities play a role in letter position coding by focusing on a large sample of children in Grade 6 to directly explore variability within one age group instead of comparing across different ages. We selected this group of developing readers because of two reasons: first, previous research suggests that, at around Grade 6, the uncoupling of general intelligence and reading skill becomes evident (Ferrer et al., 2009); also, in Spain, there is close to 100% schooling of that age group as opposed to the university level, in which the coverage is about 37% of individuals. Furthermore, there is a broader range in reading ability among Grade 6 students than among university students (i.e., university students are assumed to be in the mid/upper reading ability segment).

As in the Grainger et al. (2012) and the Colombo et al. (2017) experiments with developing readers, we employed a single-presentation LDT comparing TL versus RL pseudowords. As indicated earlier, this task allows us to directly answer how wordlike the TL pseudowords are compared with their appropriate controls. Furthermore, this task produces much greater effect sizes than a masked priming LDT (Comesaña et al., 2016). Moreover, it also allows us to examine not only the RTs but also the error rates. As Perea and Lupker (2004) and Marcet et al. (2019) showed, error rates can be more sensitive to subtle manipulations on TL effects (e.g., consonant/vowel status, visual-spatial manipulation) than the RTs in this task—of course, RT effects are still most important for skilled readers in experiments where accuracies are near ceiling levels. For each participant, we measured reading ability via a standardised test of word reading and pseudoword reading in Catalan (PROLEC-R test; Cuetos et al., 2007). Then, we examined the relationships between the magnitude of the TL effect and reading ability. As a secondary goal, we also examined whether this relationship could be modulated by visual perception speed as measured by the visual search test.

To sum up, if better readers have a better-tuned system for letter position coding (lexical tuning account; Castles et al., 2007; lexical quality hypothesis, Perfetti, 2017), there would be a negative relationship between reading ability and the size of the TL effect. Alternatively, if better readers use a coarser orthographic coding of letter position—in the spirit of the dual-route model of word recognition (Grainger & Ziegler, 2011), one would expect a positive relationship between reading ability and the size of the TL effect. We examined the likelihood of these two accounts with the observed data using Bayesian linear regression analyses and path analyses.

Method

Participants

In total 87 sixth-grade children (45 boys; mean age in months = 142.2 months [$SD = 4.5$]; range in years: 11–12) from three schools in Barcelona's metropolitan area took part in the experiment. They were native speakers of Catalan with normal/corrected-to-normal vision. Catalan had been their instruction language at all levels of school. Their parents signed informed consent forms before the experiment. Seven individuals were excluded because they had been diagnosed with learning disabilities or dyslexia.

Materials

The base words for the pseudoword stimuli in the lexical decision experiment were 80 nouns extracted from the Catalan word-frequency database (Rafel i Fontanals, 1998). The mean word-frequency per million was 128.4 (range: 25–1,102), the average length was 5.8 letters (range: 5–7), and the average Coltheart's N was 4.4 (range: 0–17). For each word, we created two pseudowords: (1) a TL pseudoword in which two adjacent middle consonants were transposed (JUGTE [base word: JUTGE, the Catalan for JUDGE]); and (2) an RL pseudoword in which two adjacent middle consonants were replaced (JUDLE). The average mean log-bigram frequencies and the mean Coltheart's N were similar for the two types of pseudowords (1.46 [range: 0.47–2.53] and 0.56 [range: 0–4] for the TL pseudowords and 1.44 [range: 0.34–2.29] and 0.69 [range: 0–4] for the RL pseudowords, both ps > .34). We constructed two counterbalanced sets of materials so that if the TL pseudoword JUGTE appeared in one set, its corresponding RL pseudoword (JUDLE) would appear in the other set. We also selected 80 Catalan words (mean word-frequency per million = 122; range: 26–1,076; average length in letters = 5.8; range: 5–7) to act as words in the LDT. Each participant was presented with 160 experimental trials (80 words, 40 TL pseudowords, and 40 RL pseudowords). We also included a practice phase composed of 16 trials with the same characteristics as the experimental trials. The stimuli are presented in Supplemental Appendix B.

Procedure

Participants in the lexical decision experiment were tested in groups of four on computers running DMDX (Forster & Forster, 2003). Each trial started with the presentation of a fixation point ("+") for 500 ms, which was replaced by a letter string in uppercase that remained on the screen until the participant responded or after 2 s had passed. Each letter string was presented in black (Courier New 14-pt) font on a white background. Participants were told to indicate if the string presented was a Catalan word or not by pressing—as fast and as accurately as possible—the "yes" or "no" keys. To measure reading ability, participants were evaluated with the PROLEC-R test in Catalan (Cuetos et al., 2007); this test offers a measure of the accuracy/time when reading words (word subtest) and pseudowords (pseudoword subtest). Finally, to obtain a global measure of visual scanning and processing speed, we employed the Symbol Search subset of the Wechsler Intelligence Scale for Children (WISC; Wechsler, 2001); this was done to explore if the TL effect might be related to visuo-attentional processes.

Results

Incorrect responses and RTs shorter than 200 ms were excluded from the latency analyses. Seven premature responses were removed, 2% of trials for words, and 5.7% of nonword trials timed out at 2,000 ms. We computed the mean RT and percent error of each participant in each condition. We excluded four participants with more than 40% of errors for words or RL pseudowords (the final sample was $n = 76$). Our inferential methods were based on model comparison using Bayes factors (Rouder et al., 2009), which allow us to quantify the support for the null/alternate hypothesis. For instance, a BF_{10} of 12 would indicate that the current data set is 12 times more likely under the alternative than under the null hypothesis.

As a requisite for any further analysis in this type of experiment, we believe that it is important to establish, as a form of quality control, that the data have three features: (1) performance is above chance, (2) there is a TL effect, and (3) the scores in the standardised test in our sample are reasonably close to those from the populations of reference.

As a first step, we checked that the error rates were well below chance (see Perea et al., 2016). Indeed, our participants were able to carry out the task with levels of error rates similar to other LDT studies. The average (by participant) error rate for words was only 9.9% ($SE = 0.6\%$). The RL pseudowords also had an error rate that demonstrated that our group of participants could accurately carry out the task: the mean percent error was 12.5% ($SE = 1.2\%$). In contrast, the average percent error for TL pseudowords was 45.6% ($SE = 2.3$); while this percent of error is almost four-times larger than for RL pseudowords ($BF_{10} = 1.63e28$), it is in line with other single-presentation lexical decision experiments with TL pseudowords (e.g., Perea & Lupker, 2004, Experiment 4; Perea & Carreiras, 2006b, Experiment 3), and overall performance is well above chance: the d' between words and TL nonwords is 1.45, and between words and replaced letter nonwords is 2.49.

Having established the above-chance performance, we can explore the latency data, which also shows a sizeable TL effect. The mean RTs were longer for TL pseudowords (1,264 ms) than for RL pseudowords (1,123 ms); averaged across participants, the TL effect had a mean of 141 ms ($SE = 12$ ms; $BF_{10} = 2.7e11$).

We obtained three scores from standardised tests: the first two were reading measures from the Catalan standardised test PROLEC-R (word- and pseudoword-reading scores; Cuetos et al., 2007; see Perea et al., 2014, for a similar procedure), and the third one was the score from

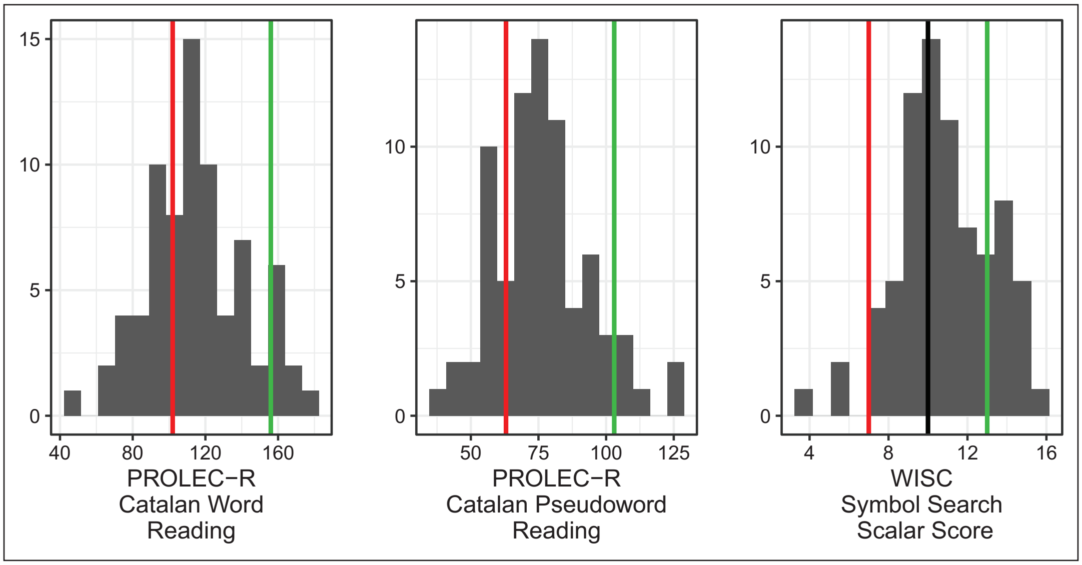

Figure 1. The panels show the distribution of scores in the three standardised tests.
The first and second panels show the PROLEC scores; the vertical lines represent the cut-off points for the low performance, middle performance, and high-performance groups as described in the test's interpretation manual. The right-most panel shows the participant's scalar scores for the WISC-IV symbol search test; the vertical lines represent the normed −1 *SD*, mean, and +1 *SD* scores.

symbol search task of the WISC (a measure of perceptual processing speed). Figure 1 shows the distribution of scores in our sample; as can be seen, the scores in our sample have appropriate variation and seem to be consistent with the norming data.

Given these large TL effects both in the accuracy and the latency measurements, and that the scores in the standardised tests are well behaved, we can now explore the central question: if reading ability and processing speed are related to the magnitude of the TL effect.

Our data analysis strategy consisted of three complementary methods: first, we performed Bayesian multiple regression analyses on the TL effect in percent error (or in RT) as the dependent variable and the scores from the standardised test as independent variables; second, to get a holistic picture of the covariance structure in our data, we implemented path models with the standardised tests as the exogenous variables and performance in the experiment (i.e., RT and percentage of errors) as the endogenous variables.

Bayesian multiple regression

TL effects as measured by percent error. We tested a regression model in which the accuracy TL effect for each participant was the dependent variable. The regressors were the three standardised test scores (i.e., PROLEC-word reading, PROLEC-pseudoword reading, and the WISC search task). We used the BayesFactor R package to compare the full model with all the simpler (nested) models; the model with the highest Bayes factor (i.e., the model with the highest likelihood relative to other models) included only the PROLEC-pseudoword-reading score as a regressor, with a *BF* of 146 over the null model. The standardised coefficient for PROLEC pseudoword was −0.42, $t(74)=8.00$, $p<.001$: this indicates that the TL effect was smaller for the better pseudoword readers.

TL effects as measured by RT. We tested the same regression model as in the analysis of the error data with the TL effect on the RT as the dependent variable. The best model was the intercept-only model, with a $BF_{0F}=21$ (i.e., these data are 21 times more likely under the intercept-only model than under the full model). The best of the nested models included PROLEC-word reading as a regressor, but compared with the null model (intercept-only), the *BF* was close to 1; recall that the *BF* quantifies the likelihood of the data given competing models, and hence a $BF=1$ means that there is not enough evidence to favour either of the two models.

Performance in word trials. While this work's focus is on the performance to nonwords, we also tested the same regression models using the RT and the percent error to words as the outcome variables. In both cases, the preferred model had a single regressor: nonword reading. Compared with the intercept-only model, for the RT, the *BF* was only 6.8, and, for percent error, it was 286.

Path analyses

Path modelling is a way to describe the dependencies between exogenous variables and dependent variables. Our dataset includes six dependent variables: two task-related

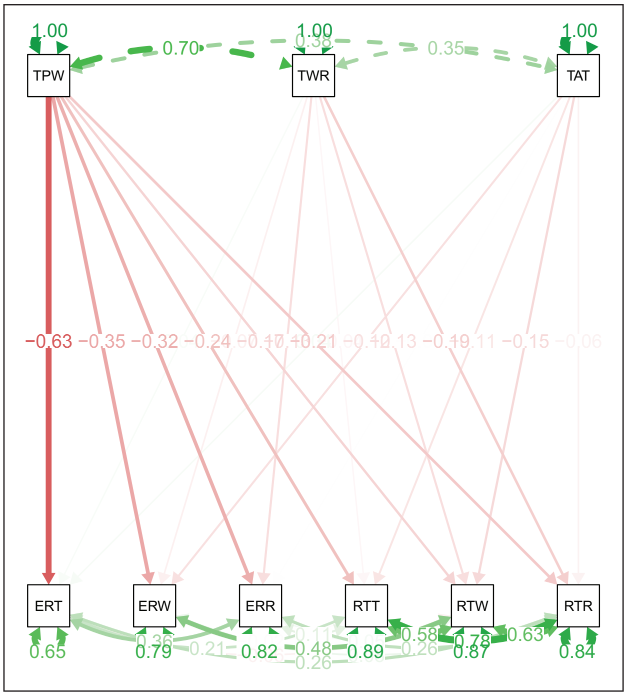

Figure 2. The path model. The model has the three exogenous variables labelled as TPW: test of pseudoword reading; TW: test of word reading; TAT: WEIS Search subtest. It also has six dependent variables labelled as ERW: error rate for words; ERT: error rate for transposed-letter nonwords; ERR: error rate for replacement-letter nonwords; RTW: RT for words; RTT: RT for transposed-letter nonwords; RTR: RT for replacement-letter nonwords (all RTs are for correct responses only). The thickness of the line represents the strength of dependency between variables.

measurements (error rate and latency) for three types of stimuli (words, RL nonwords, and TL nonwords); and three exogenous variables: test scores for PROLEC-words, PROLEC-nonword, and WISC-symbol search.

We implemented a path model in which the LDT performance (RT and error rate) depended on the test scores. This model is depicted in Figure 2. The arrows in the graph represent dependencies, and the coefficients on the lines are measurements of statistical relationship (also graphically illustrated by the width of the line). The results of these analyses are clear if we examine the paths that start from each of the test scores: for both the PROLEC-word reading and the visual attention test (WISC-Search), there are only negligible dependencies for any of the dependent variables (all standardised coefficients are less than .|19| [$zs < |1.6|$, $ps > .19$]); on the contrary, for PROLEC-pseudoword test, the dependencies are much stronger: the coefficients for the effects of the PROLEC-pseudoword readings on the error rate variables are all significant: for words, standard coefficient $= -.346$, $z = -2.473$, $p = .13$; for TLs, standard coefficient $= -0.630$, $z = -5.524$, $p < .001$; for RL, standard coefficient $= -0.316$, $z = -2.197$, $p = .028$. Even though none of the coefficients for the RT measurements are significant (all $ps > .1$), the largest standardised coefficients are related to the PROLEC-pseudoword reading (RTs for words: -0.167, $z = -1.10$; RTs for RL: -0.210, $z = -1.42$; and RTs for TL: -0.244, $z = -1.61$).

Table 1 presents the covariance structure and selected estimates and fit measures; the complete output and code for this analysis are available in the online appendix. Also, Figure 3 displays the pairs-plot for all the variables used in this analysis. In the pairs-plot, we can show the distribution of all the variable scores (in the diagonal), along with the scatterplots and the pairwise correlations. As can be seen, there is some relationship among most variables, and the test scores are highly related to each other ($r > .37$), while the RT measurements are also highly related to each other ($r > .62$). Notably, the two reading scores have an $r = .7$, but only the nonword reading score seems to be an important predictor of the TL effect. This pattern of results suggests that the TL effect is not primarily driven by lexical competition.

As a summary, the analyses using Bayesian multiple regression and path modelling give us a clear picture of our experimental data (LDT) and the standardised testing (PROLEC and WISC-Search). These two methods have slightly different goals: the Bayesian multiple regression allows us to use Bayes factors, which provide a measurement of the likelihood of the data given competing models, and hence can quantify the support for null versus alternative hypotheses; the path model, on the contrary, provides with an overall account of the covariances among all the variables in the study. The top-line conclusions are two: (1) Better scores in the pseudoword-reading test is associated with lower error rates in all types of items (including words) and weakly associated (non-significantly, but indeed numerically) to shorter RTs to all categories of items and (2) while the TL effects occur in both the error rate and the latency data, only the accuracy effects are related to the standardised test scores. Specifically, better pseudoword readers have smaller TL effects.

Discussion

Researchers have proposed several models to explain how readers encode letter position during visual-word recognition (see Grainger, 2018, for review). Most of this research employed the size of the TL effect as a signature of letter position coding. To fully understand how letter position is encoded when reading, it is critical to know whether reading ability modulates letter position coding and, if so, in which direction: whether better readers are more precise at encoding letter position, as espoused by Castles et al.'s (2007) lexical tuning model and Perfetti's (2017) lexical quality hypothesis; or, on the contrary, whether better readers use a coarser encoding of letter position, in the spirit of Grainger and Ziegler's (2011) dual-route model.

Table 1. The covariance matrix, means, and *SD*s for the path model variables ($N = 76$).

	TestWRD	TestPW	TestATT	RTRL	RTTL	RTWD	RTEffect	ERRL	ERTL	ERWD	EREffect
TestWRD	748.51										
TestPW	333.76	303.56									
TestATT	24.01	16.71	6.32								
RTRL	−1,702.12	−1,109.20	−90.40	30,392.95							
RTTL	−1,094.60	−881.96	−87.00	22,734.69	26,442.19						
RTWD	−1,110.79	−745.93	−88.03	16,229.83	13,910.62	18,742.55					
RTEffect	607.52	227.23	3.41	−7,658.26	3,707.50	−2,319.21	11,365.77				
ERRL	−104.35	−76.96	−4.80	716.15	394.74	315.58	−321.42	115.48			
ERTL	−209.74	−201.48	−9.19	1,377.57	1,090.18	341.55	−287.39	105.39	85.43		
ERWD	−52.11	−41.54	−3.78	239.42	121.24	416.36	−118.18	14.82	26.89	29.60	
EREffect	−105.39	−124.52	−4.39	661.42	695.45	25.98	34.03	−10.09	280.04	12.07	290.13
Mean	114.605	76.132	10.763	1,122.724	1,263.911	966.265	141.187	12.514	45.612	9.912	33.098
SD	27.359	17.423	2.513	174.336	162.611	136.903	106.610	10.746	19.632	5.441	17.033

Selected model fit measurements

Npar	χ^2	df	p value	RFI	NFI	AIC	RMSEA
37	0.594	2	.743	0.965	0.998	6,120.004	<0.001

Selected parameter estimates

Lhs	Op	Rhs	Est.std	SE	z	p value
ERTL	~	TestPW	−0.630	0.114	−5.524	.000
ERTL	~	TestWRD	0.036	0.131	0.279	.781
ERTL	~	TestATT	0.042	0.101	0.413	.679
ERWD	~	TestPW	−0.346	0.140	−2.473	.013
ERWD	~	TestWRD	−0.065	0.144	−0.453	.650
ERWD	~	TestATT	−0.122	0.110	−1.103	.270
ERRL	~	TestPW	−0.316	0.144	−2.197	.028
ERRL	~	TestWRD	−0.130	0.146	−0.889	.374
ERRL	~	TestATT	−0.012	0.113	−0.106	.915
RTTL	~	TestPW	−0.244	0.152	−1.609	.108
RTTL	~	TestWRD	−0.038	0.153	−0.248	.804
RTTL	~	TestATT	−0.106	0.117	−0.907	.364
RTWD	~	TestPW	−0.167	0.152	−1.100	.271
RTWD	~	TestWRD	−0.129	0.150	−0.858	.391
RTWD	~	TestATT	−0.148	0.115	−1.283	.199
RTRL	~	TestPW	−0.210	0.148	−1.416	.157
RTRL	~	TestWRD	−0.189	0.147	−1.288	.198
RTRL	~	TestATT	−0.060	0.115	−0.526	.599

Model

```
ERTL  ~  TestPW  +  TestWRD  +  TestATT
ERWD  ~  TestPW  +  TestWRD  +  TestATT
ERRL  ~  TestPW  +  TestWRD  +  TestATT

RTTL  ~  TestPW  +  TestWRD  +  TestATT
RTWD  ~  TestPW  +  TestWRD  +  TestATT
RTRL  ~  TestPW  +  TestWRD  +  TestATT
ERWD  ~~  0*  ERRL
ERWD  ~~  0*  ERTL
```

RFI: relative fit index; NFI: normed fit index; AIC: Akaike information criterion; RMSEA: root mean square error of approximation.

In the present experiment, we examined the relationship between TL effects and reading ability within a large sample of sixth graders who participated in a lexical decision experiment with TL/RL pseudowords undertook standardised reading tests. Results showed—using both Bayesian multiple regression and path analyses—that individuals with high scores in the pseudoword-reading test tended to have smaller TL effects in the error data. Significantly, the error rate in TL pseudoword trials was negatively associated with the pseudoword-reading score.

Thus, our data showed that more proficient young readers are unlikely to confuse their MOTHER with their MOHTER. This pattern supports the hypothesis that the components of reading measured by the pseudoword-reading test have a stronger relationship to the TL effect, as measured by the percent error in a single-presentation

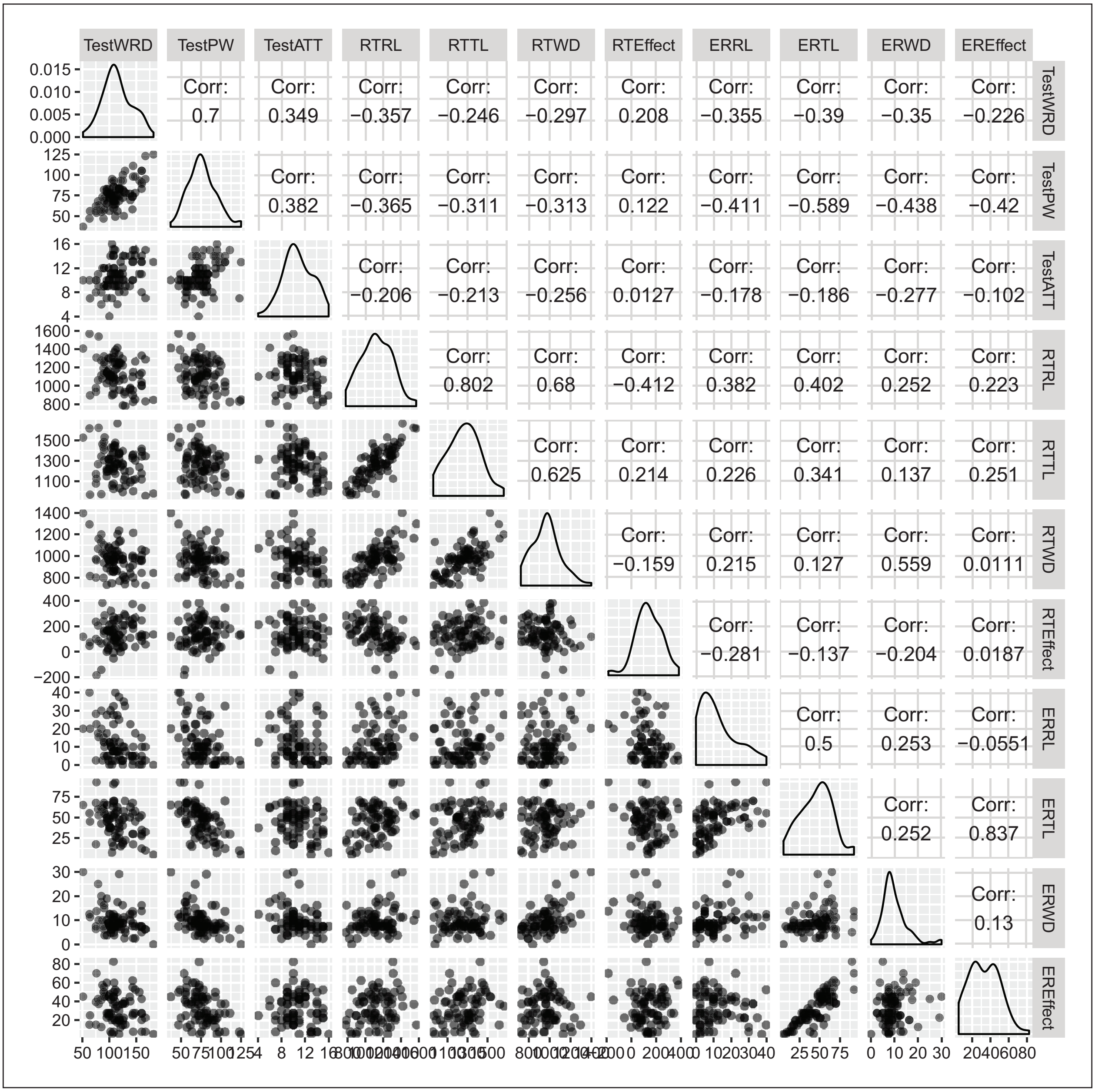

Figure 3. This graph allows us to visualise all the variables used in this study: each variable's distributional shape is displayed in the diagonal panels.
The scatterplots are on the bottom panels and the numerical values of the correlations on the top panels.

LDT, than word reading or visual attention. It is worth noting that these findings occur within the context of the language studied. The experiment was conducted in Catalan, a relatively transparent orthography in which it is difficult to disentangle orthographic and phonological processing—to dissociate the role of phonology it would be necessary to run a parallel study using a deep orthographic system. Of note, a recent reading-like experiment with Grades 3 to 4 children in English, Pagán et al. (2021) also reported that better readers encoded letter position of TL pseudowords more precisely than less-skilled readers.

Overall, our findings—together with the Pagán et al. (2021) findings—are more consistent with the lexical tuning account (Castles et al., 2007) and the lexical quality account (Perfetti, 2017) than with the dual-route account (Ziegler et al., 2014). In fairness to Ziegler et al. (2014), their model focuses on reading development in the initial years of literacy rather than on individual differences per se. The dual-route model focuses on the process where the orthographic route becomes preponderant in reading development. Critically, neither of the above accounts has been implemented as a model that predicts the relationship between standardised reading tests and the TL effects within a single age group (but see Figure 1 of Kohnen & Castles, 2013). Indeed, these theories could be implemented as models in many ways (see McElreath, 2016), and comparing such implementations would be a fruitful future endeavour.

It is important to address a critical question: why is reading skill correlated with TLs' effects on accuracy but not on latency? This dissociation can be explained in two ways: one could conceive a processing explanation in which the component of processing that relates to the latency effects is different than the component of processing that relates to accuracy effects, and only the latter is related to reading skill; on the contrary, the lack of correlation between reading skill and TL effect in latency could be simply because of excessive measurement noise (i.e., the trial-by-trial variability is simply too large). We do not take a position on this issue at the moment. Still, it is important to point out that a similar issue is currently under debate in the cognitive control literature, where classic effects (e.g., Stroop and flanker) have near-zero correlation across participants. Some researchers interpret this lack of correlation as evidence of cognitive control being comprised of several independent skills (e.g., Rey-Mermet et al., 2018), while others see it as a consequence of trial noise attenuating any actual correlation (e.g., Whitehead et al., 2019). Furthermore, together with the question of individual differences (i.e., subject variability), there is another issue of relevance to understand the processes underlying letter position coding fully. We would also need to examine in-depth, with a large set of items, which lexical/sublexical elements (e.g., consonant/vowel status, letter position, bigram frequency, morphological boundaries, phonology) make a TL pseudoword more or less wordlike (i.e., item variability).

While both single-presentation experiments and masked priming experiments coexist in the study of letter position coding, one might argue that the masked priming technique would have been a better choice to minimise the potential role of participant's strategies. We chose the single-presentation task because effect sizes are noticeably larger than in masked priming and also because—unlike masked priming—it directly answers the question of how wordlike a TL pseudoword like MOHTER is.

To sum up, the present experiment showed, using both Bayesian linear regression and path analyses, that letter position coding is modulated by reading ability in sixth graders: better readers are less likely to confuse their MOHTER with their MOTHER. This pattern is consistent with the reduced TL effects found with adult individuals that excel in orthographic abilities (competitive Scrabble players; Perea et al., 2016). The empirical evidence on the relationship between reading skills and letter position coding is emerging—but still scarce—with a diverse set of age groups, populations, and tasks. For example, Friedmann and Rahamim (2014) and Kezilas et al. (2014) found that there can be letter position coding deficits that cannot be attributed to phonological decoding abilities (letter position dyslexia; Friedmann & Rahamim, 2007). Thus, the take-home message is that future implementations of visual-word recognition models need to consider individual differences to capture the intricacies of letter position coding accurately—and likely other important phenomena underlying visual-word recognition.

Declaration of conflicting interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This paper was funded by Grant PSI2017-86210-P from the *Spanish Ministry of Science, Innovation, and Universities* to MP, by Grant 0115/2018 (*Estades d'investigadors convidats*) from the *Universitat de València* to PG, and Grant GV/2020/074 from the Department of Innovation, Universities, Science and Digital Society of the Valencian Government to AM.

ORCID iDs

Pablo Gomez https://orcid.org/0000-0003-4180-1560

Ana Marcet https://orcid.org/0000-0001-8755-5903

Manuel Perea https://orcid.org/0000-0002-3291-1365

Data accessibility statement

The data and materials from the present experiment are publicly available at the Open Science Framework website: DOI 10.17605/OSF.IO/WGKPC.

Supplemental material

The supplementary material is available at qjep.sagepub.com.

Note

1. Theorists have proposed two non-mutually exclusive mechanisms. One explanation is that, because of the visual system's limitations, there is some noise at encoding letter positions in a string (Gomez et al., 2008). As a result, the letter H in the visual input MOHTER would activate not only the third position but it would also activate—to some degree—neighbouring letter positions, thus explaining that MOHTER and MOTHER are orthographically very similar. Another explanation is via the activation of open bigrams (Grainger & van Heuven, 2004). The presentation of MOHTER would activate several open bigrams that are

common with MOTHER (e.g., MO, MH, MT, ME; MR, OH, OT, OE, OR, HE, ER, TE, TR, ER). Thus, the transposed-letter pseudoword MOHTER would share most of its open bigrams with MOTHER (i.e., all except TH), hence explaining why MOHTER is easily misperceived as MOTHER.

References

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*, *4*, 439–461. https://doi.org/10.3758/bf03214334

Andrews, S., & Lo, S. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 152–163. https://doi.org/10.1037/a0024953

Blythe, H. I., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2014). Reading transposed text: Effects of transposed letter distance and consonant-vowel status on eye movements. *Attention, Perception, & Psychophysics*, *76*, 2424–2440. https://doi.org/10.3758/s13414-014-0707-2

Bruner, J. S., & O'Dowd, D. (1958). A note on the informativeness of parts of words. *Language and Speech*, *1*, 98–101.

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology*, *97*, 165–182. https://doi.org/10.1016/j.jecp.2007.01.006

Chambers, S. M. (1979). Letter and order information in lexical access. *Journal of Verbal Learning and Verbal Behavior*, *18*, 225–241. https://doi.org/10.1016/s0022-5371(79)90136-1

Colombo, L., Sulpizio, S., & Peressotti, F. (2017). Serial mechanism in transposed letters effects: A developmental study. *Journal of Experimental Child Psychology*, *161*, 46–62. https://doi.org/10.1016/j.jecp.2017.04.002

Comesaña, M., Soares, A. P., Marcet, A., & Perea, M. (2016). On the nature of consonant/vowel differences in letter position coding: Evidence from developing and adult readers. *British Journal of Psychology*, *107*, 651–674. https://doi.org/10.1111/bjop.12179

Cuetos, F., Rodríguez, B., Ruano, E., & Arribas, D. (2007). *PROLEC-R: Català—Bateria d'avaluació dels processos lectors* [PROLEC-R: Catalan—Battery of assessment of reading processes]. TEA Edicíciones.

Davidson, M., & Myhre, O. (2000). Measuring reading at grade level. *Educational Leadership*, *57*, 25–28.

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*, 713–758. https://doi.org/10.1037/a0019738

Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K., & Shaywitz, S. E. (2009). Uncoupling of reading and IQ over time: Empirical evidence for a definition of dyslexia. *Psychological Science*, *21*, 93–101. https://doi.org/10.1177/0956797609354084

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116–124. https://doi.org/10.3758/bf03195503

Friedmann, N., & Rahamim, E. (2007). Developmental letter position dyslexia. *Journal of Neuropsychology*, *1*, 201–236. https://doi.org/10.1348/174866407X204227

Friedmann, N., & Rahamim, E. (2014). What can reduce letter migrations in letter position dyslexia? *Journal of Research in Reading*, *37*, 297–315. https://doi.org/10.1111/j.1467-9817.2011.01525.x

Gómez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*, 577–601. https://doi.org/10.1037/a0012667

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes*, *23*, 1–35. https://doi.org/10.1080/01690960701578013

Grainger, J. (2018). Orthographic processing: A "mid-level" vision of reading. *Quarterly Journal of Experimental Psychology*, *71*, 335–359. https://doi.org/10.1080/17470218.2017.1314515

Grainger, J., Lété, B., Bertand, D., Dufau, S., & Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition*, *123*, 280–292. https://doi.org/10.1016/j.cognition.2012.01.003

Grainger, J., & van Heuven, W. J. B. (2004). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *Mental lexicon: Some words to talk about words* (pp. 1–23). Nova Science Publishers.

Grainger, J., & Ziegler, J. C. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology*, *2*, Article 54. https://doi.org/10.3389/fpsyg.2011.00054

Kezilas, Y., Kohnen, S., McKague, M., & Castles, A. (2014). The locus of impairment in English developmental letter position dyslexia. *Frontiers in Human Neuroscience*, *8*, Article 356. https://doi.org/10.3389/fnhum.2014.00356

Kezilas, Y., McKague, M., Kohnen, S., Badcock, N. A., & Castles, A. (2017). Disentangling the developmental trajectories of letter position and letter identity coding using masked priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 250–258. https://doi.org/10.1037/xlm0000293

Kohnen, S., & Castles, A. (2013). Pirates at parties: Letter position processing in developing readers. *Journal of Experimental Child Psychology*, *115*, 91–107. https://doi.org/10.1016/j.jecp.2012.08.011

Lupker, S. J., & Davis, C. J. (2009). Sandwich priming: A method for overcoming the limitations of masked priming by reducing lexical competitor effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 618–639. https://doi.org/10.1037/a0015278

Marcet, A., Perea, M., Baciero, A., & Gomez, P. (2019). Can letter position encoding be modified by visual perceptual elements? *Quarterly Journal of Experimental Psychology*, *72*, 1344–1353. https://doi.org/10.1177/1747021818789876

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I—An account of basic findings. *Psychological Review*, *88*, 375–407. https://doi.org/10.1037/0033-295x.88.5.375

McElreath, R. (2016). *Chapman & Hall/CRC texts in statistical science: Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC Press.

O'Connor, R. E., & Forster, K. I. (1981). Criterion bias and search sequence bias in word recognition. *Memory & Cognition*, *9*, 78–92. https://doi.org/10.3758/bf03196953

Pagán, A., Blythe, H. I., & Liversedge, S. P. (2021). The influence of children's reading ability on initial letter position encoding during a reading-like task. *Journal of Experimental Psychology: Learning, Memory and Cognition*. Advance online publication. https://doi/org/10.1037/xlm0000989

Paterson, K. B., Read, J., McGowan, V. A., & Jordan, T. R. (2014). Children and adults both see "pirates" in "parties": Letter-position effects for developing readers and skilled adult readers. *Developmental Science*, *18*, 335–343. https://doi.org/10.1111/desc.12222

Perea, M., & Carreiras, M. (2006a). Do transposed-letter effects occur across lexeme boundaries? *Psychonomic Bulletin & Review*, *13*, 418–422. https://doi.org/10.3758/BF03193863

Perea, M., & Carreiras, M. (2006b). Do transposed-letter similarity effects occur at a prelexical phonological level? *Quarterly Journal of Experimental Psychology*, *59*, 1600–1613. https://doi.org/10.1080/17470210500298880

Perea, M., Gatt, A., Moret-Tatay, C., & Fabri, R. (2012). Are all Semitic languages immune to letter transpositions? The case of Maltese. *Psychonomic Bulletin & Review*, *19*, 942–947. https://doi.org/10.3758/s13423-012-0273-3

Perea, M., Jiménez, M., & Gomez, P. (2016). Does location uncertainty in letter position coding emerge because of literacy training? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 996–1001. https://doi.org/10.1037/xlm0000208

Perea, M., Jiménez, M., Suárez-Coalla, P., Fernández, N., Viña, C., & Cuetos, F. (2014). Ability for voice recognition is a marker for dyslexia in children. *Experimental Psychology*, *61*, 480–487. https://doi.org/10.1027/1618-3169/a000265

Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language*, *51*, 231–246. https://doi.org/10.1016/j.jml.2004.05.005

Perea, M., Marcet, A., & Gomez, P. (2016). How do Scrabble players encode letter position during reading? *Psicothema*, *28*, 7–12. https://doi.org/10.7334/psicothema2015.167

Perfetti, C. A. (2017). Lexical quality revisited. In E. Segers & P. van den Broek (Eds.), *Developmental perspectives in written language and literacy: In honor of Ludo Verhoeven* (pp. 51–68). John Benjamins.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–214). John Benjamins.

Rafel i Fontanals, J. (1998). *Diccionari de Freqüencies* [Dictionary of frequencies]. Institut d'Estudis Catalans.

Rayner, K., White, S. J., Johnson, R. L., & Liversedge, S. P. (2006). Raeding wrods with jubmled lettres. *Psychological Science*, *17*, 192–193. https://doi.org/10.1111/j.1467-9280.2006.01684.x

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(4), 501–526. https://doi.org/10.1037/xlm0000450

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225–237. https://doi.org/10.3758/PBR.16.2.225

Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes*, *19*(3), 333–367. https://doi.org/10.1080/01690960344000198

Vergara-Martínez, M., Perea, M., Gomez, P., & Swaab, T.Y. (2013). ERP correlates of letter identity and letter position are modulated by lexical frequency. *Brain and Language*, *125*, 11–27. https://doi.org/10.1016/j.bandl.2012.12.009

Wechsler, D. (2001). *Wechsler Intelligence Scale for Children–Revised* (Spanish adaptation). TEA Ediciones.

Whitehead, P. S., Brewer, G. A., & Blais, C. (2019). Are cognitive control processes reliable? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(5), 765–778. https://doi.org/10.1037/xlm0000632

Ziegler, J. C., Bertrand, D., Lété, B., & Grainger, J. (2014). Orthographic and phonological contributions to reading development: Tracking developmental trajectories using masked priming. *Developmental Psychology*, *50*, 1026–1036. https://doi.org/10.1037/a0035187