BRIEF REPORT



How orthographic-specific characteristics shape letter position coding: The case of Thai script

Manuel Perea¹ · Heather Winskel² · Pablo Gomez³

Published online: 12 April 2017 © Psychonomic Society, Inc. 2017

Abstract A central question for any model of visual word identification is the representation of the position at which letters are encoded (e.g., calm vs. clam). In this article, we examine whether the orthographic-specific characteristics of a writing system-namely, Thai-shape the process of letter position coding. Thai is an alphabetic script that lacks interword spaces and has an orthographic order that does not necessarily correspond to the phonological order for initial vowels. This implies that the initial letter position coding in Thai needs to be flexible enough that readers can successfully encode the letter positions of words. To compare letter position coding in Thai to that in English, we conducted an experiment that paralleled Experiment 3 in Gomez, Ratcliff, and Perea (Psychological Review, 115, 577-600, 2008), including 23 conditions (single-letter replacements, letter transpositions, letter migrations, and a corresponding control). We obtained fits from Gomez et al.'s overlap model, which is a model that has been shown to account for letter position coding in the Roman alphabet across this variety of letter manipulations. The overlap model was found to successfully fit the Thai data. Our results revealed that the position encoding was better for the first letter than for the rest of the positions in both languages; however, in English the position uncertainty grows as a function of letter order quite abruptly, whereas in Thai it grows gradually. Thus, the orthographic-specific

Manuel Perea mperea@uv.es

² Southern Cross University, Coffs Harbour, New South Wales, Australia

³ DePaul University, Chicago, IL, USA

characteristics of the Thai writing system do play a role in shaping the process of letter position coding.

Keywords Visual word recognition · Modeling · Letter position coding

A central question for any model of visual word identification is the representation of the position at which letters are encoded. Consequently, all recently proposed models offer an explanation of this process in the course of orthographic processing (e.g., the LTRS model: Adelman, 2011; spatialcoding model: Davis, 2010; overlap model: Gomez, Ratcliff, & Perea, 2008; overlap open-bigram model: Grainger, Granier, Farioli, Van Assche, & van Heuven, 2006; Bayesian reader model: Norris, Kinoshita, & van Casteren, 2010; and SERIOL model: Whitney, 2001). Not surprisingly, most of the evidence that those models have tried to account for has been obtained using readers of languages with the Roman alphabet.

Expanding the empirical record to include the process of letter position coding in non-Roman alphabets is a worthwhile endeavor, beyond the goals of inclusion and diversity, because it encourages theory development; namely, the orthographic features of different writing systems could shape the process of letter position coding. Hence, the research agenda while studying different writing systems is to explore both the differences (the changes in the processes as a function of writing systems) and the invariances (the parts of the processes that stay the same across writing systems) in letter position coding (see Wiley, Wilson, & Rapp, 2016, for an examination of the similarities/differences in letter identity coding between the Arabic vs. Roman alphabets).

In this article, we examine whether the orthographic-specific characteristics of the Thai writing system shape the process of

¹ Departamento de Metodología, Universitat de València, Av. Blasco Ibáñez, 21, 46010 Valencia, Spain

letter position coding. Prior research using the Roman alphabet has offered remarkably similar patterns of letter position coding across a variety of languages (e.g., across the Germanic [English: Perea & Lupker, 2003], Romance [Spanish: Perea & Lupker, 2004], pre-Indo-European [Basque: Perea & Carreiras, 2006], Semitic [Maltese: Perea, Gatt, Moret-Tatay, & Fabri, 2012], and Uralic [Hungarian: Tóth & Csépe, 2016] families). This could be taken to suggest that letter position coding processes are to some degree language-independent, at least in alphabetic languages using the Roman alphabet.¹

The Thai writing system has an alphabetic orthography, but it has two idiosyncratic features that make it particularly interesting to compare with the Roman alphabet in relation to letter position encoding. First, Thai has some commonly used vowels (i.e., ι /e:/, ι / ϵ :/, $\tilde{\iota}$ /o:/, l /aj/, l /aj/) that are written before the consonant but articulated after the consonant in speech (e.g., the word unu "flat" is not pronounced as /ɛ:bn/ but as $b\epsilon:n/-i.e.$, the written vowel u is misaligned relative to the spoken vowel $(\varepsilon:/)$; however, not all vowels follow this pattern, since other vowels are pronounced in the order that they are written (e.g., the word unn 'Baht' is pronounced /ba:t/-i.e., the written vowel 1 is aligned; see Winskel & Iemwanthong, 2010). Therefore, letter position coding in Thai needs to be flexible enough to deal with the processing of words containing either aligned or misaligned vowels. Second, the Thai script does not have interword spaces (e.g., the sentence นักศึกษาไปชื่อตะไคร้มาจากตลาด "The student goes to the market to buy lemon grass"). As Winskel, Perea, and Ratitamkul (2012) indicated, this implies that during sentence reading "there is a degree of ambiguity in relation to which word a given letter belongs to" (p. 1523) (e.g., the boundaries of the word ตะไคร้ "lemon grass" are difficult to discern in the sentence above).

Our tool for comparing letter position coding in the Thai and Roman writing systems is a mathematical/ computational model: the overlap model (Gomez et al., 2008). The basic assumption of this model is that perceptual uncertainty is associated with locating objects in space (i.e., letters in words), so the locations of letters in a letter string should be considered distributions along a dimension rather than exact points (see Logan, 1996). Each letter position has a different degree of "perceptual uncertainty," which is treated as a free parameter in the model (*s* parameter)—other models of visual word identification have incorporated similar mechanisms (e.g., Adelman, 2011; Davis, 2010; Grainger et al., 2006; Norris et al., 2010). As was shown by Gomez et al., the overlap model can account for letter position encoding in the Roman alphabet across a variety of manipulations: letter replacement, letter transposition, letter addition/deletion, and letter migration.

One of the critical benchmarks in the experiments and fits reported by Gomez et al. (2008) is that the overlap model captured the first letter as being more important for letter position coding than were interior letters in the Roman alphabet. According to the model, this occurs because the degree of perceptual noise in the internal letter positions is larger than the perceptual noise in the first letter position (see Gomez et al., 2008, Fig. 14). In other words, jugde resembles judge much more closely than *ujdge*—note that in the Roman alphabet, transposedletter effects tend not to occur when the initial letter is involved (see White, Johnson, Liversedge, & Rayner, 2008). The presence of some misaligned vowels in Thai words and the absence of interword spaces may lead to a more flexible process of letter position coding than occurs in languages that employ the Roman alphabet. Indeed, some evidence from Thai supports the view that letter position coding is quite flexible, even in relation to the initial letter position (see Perea, Winskel, & Ratitamkul, 2012, and Winskel et al., 2012, for evidence from masked priming and sentence reading).

To compare how letter position coding is attained in Thai in comparison with English (i.e., a language that uses the Roman alphabet), we conducted an experiment that paralleled Experiment 3 in Gomez et al. (2008). In this experiment they examined how letter positions are encoded in English across 23 conditions that involved letter replacements, letter transpositions, and letter migrations. The present experiment with Thai words and Thai participants allowed us to compare the fits obtained by Gomez et al. in English with the fits obtained in the present Thai experiment.

Therefore, on the basis of previous research and the characteristics of Thai, we explored two questions: (1) Would the position uncertainty model (the overlap model) be able to account for the data obtained with Thai readers? (2) If the model could account for the data, would the configuration of the parameter values nonetheless be related to some of the idiosyncrasies of the Thai writing system? The present experiment involved 23 conditions that included (a) a single-letter replacement condition, (b) an adjacent-letter transposition condition, (c) a letter migration condition, and (d) an orthographic control for the letter migration condition (see Table 1 for further details). These are the same conditions as in Gomez et al.'s (2008) Experiment 3, in which they found an abrupt increase in perceptual uncertainty from the initial position to the other letter positions.

¹ Transposed-letter effects are typically small in Arabic and Hebrew (Perea, Abu Mallouh, & Carreiras, 2010; Velan & Frost, 2011). Note, however, that these two languages have two distinctive features: (i) a rigid morphological structure, and (ii) vowels that are not regularly written down. Indeed, transposed-letter effects are robust in languages that use Arabic script in which vowel information is written down (e.g., Uyghur; see Yakup, Abliz, Sereno, & Perea, 2015).

Table 1 Types of stimuli in the experiment

	1	
Type of Probe	Thai Stimuli	Roman Equivalent
One-letter replacement	ยาบอก, ลเบอก, ลายอก, ลาบเก, ลาบอย	XBCDE, AXCDE, ABXDE, ABCXE, ABCDX
Adjacent letter transposition	าสบอก, ลบาอก, ลาอบก, ลาบกอ	BACDE, ACBDE, ABDCE, ABCED
Letter migration	าบลอก, ลบอาก, บลาอก, ลาอกบ, ลอาบก, ลกาบอ, ลากบอ	BCADE, ACDBE, CABDE, ABDEC, ADBCE, ABECD, AEBCD
Migration + replacement	าบยอก, ลบอเก, ยลาอก, ลาอกย, ลเวบก, ลยาบอ, ลายบอ	BCXDE, ACDXE, XABDE, ABDEX, AXBCE, ABXCD, AXBCD

Assuming that the target is anuan (ABCDE), the table shows the 23 types of distractors. For ease of reading, we present Roman alphabet equivalents of the Thai stimuli in the final column. Letters that change position are indicated in *italic* font, and replaced letters are indicated in **bold** font.

Method

Participants

Twenty students and staff from Chulalongkorn University, Bangkok, participated in the experiment. All of them were native speakers of Thai and had normal or corrected-to-normal vision.

Materials

In this experiment, sets of similar stimuli were formed by rearranging the letters in Thai pseudowords. The creation of the stimuli was parallel to that employed by Gomez et al. (2008, Exp. 1) with the Roman alphabet in English readers. We created 1,196 pseudowords by substituting one letter within five-letter Thai words obtained from the Thai National Corpus (Aroonmanakun, 2007). The base words included both alignedand misaligned-vowel words. Tone diacritics were not included in the experimental pseudoword stimuli. Each target stimulus (or any of the items generated from it) was presented only once. During the generation of items, vowels were always substituted for vowels, and consonants were always substituted for consonants in the replacement conditions. There were 23 conditions altogether, which can be divided into four categories: (a) a single-letter replacement condition, (b) an adjacent-letter transposition condition, (c) a letter migration condition, and (d) a letter migration + replacement condition (refer to Table 1). Overall, the experimental session was composed of 13 blocks of 92 trials each (1,196 trials).

Procedure

Participants were tested individually in a quiet room at the Center for Research on Speech and Language Processing. Stimuli were presented in the 24-point Courier Proportional Thai font. The DMDX software (Forster & Forster, 2003) was used to display the stimuli and record the participants' responses. We used a two-alternative forced choice paradigm that mimicked the procedure used by Gomez et al. (2008). On each trial, a fixation point was initially presented on the computer screen for 500 ms, and then a target stimulus was

presented for 83 ms in the center of the screen, which was subsequently masked with segments of Thai letters. Note that Gomez et al. employed a duration of 50 ms for the target stimulus; however, a pilot study in Thai revealed that accuracy was substantially lower than in the Gomez et al. experiment, so the target duration was increased to 83 ms in order to produce more comparable results for the two scripts.² The participant had to choose between two alternatives that were presented simultaneously below the mask (one to the right and the other to the left of the mask, as can be seen in Gomez et al.'s, 2008, Fig. 2). The alternatives were the pseudoword and a letter string from one of the four conditions outlined below. Participants were asked to indicate which alternative was the letter string that had been presented briefly. Each alternative was the correct response on an equal number of trials. The order of trials was randomized for each participant. Prior to the experimental trials, 14 practice trials were given. The experiment took approximately 1 h to complete.

Results

The empirical findings are described in this section, and the model fits will be described and evaluated in the Model Fits subsection.³ Given the two-alternative forced choice procedure presented in this article, our analysis focused on accuracy rates (the RT analyses are available in the online appendix). The average accuracy rates for each condition are shown in Table 2 in the columns labeled *Thai*. The results were rather straightforward: The transposition conditions were more difficult than the letter replacement conditions. To examine the evidence for and against differences among the different conditions within this experiment, we utilized Bayes factors.

² As can be seen in Table 1, the results showed that the overall accuracies across conditions were comparable for Thai and English participants (.749 vs. .718). One reason that we needed to increase the stimulus duration was to keep the overall accuracy rate similar to that in English, which was probably related to the fact that Thai letters are visually complex and share many features (e.g., $\eta - \eta$, $\mu - \mu - \eta$, $\mu - \mu - \eta$, $\mu - \omega - \eta$, $\mu - \mu - \eta$, $\mu - \omega - \psi - \psi$, $\mu - \omega - \psi - \psi$, $\mu - \omega - \psi - \psi$, $\mu - \psi - \psi - \psi$, $\mu - \psi - \psi$

 $^{^3}$ We have included the analyses presented in this article, along with every other analysis carried out, in the online Appendix, available at https://osf.io/ n8hkr/

 Table 2
 Proportions of correct responses for the present experiment, as well as the results from Gomez et al. (2008) in English, for comparison purposes

Conditions	English (G	English (Gomez et al., 2008)			Thai		
	Data	2 <i>SE</i> s	Model	Data	2 SEs	Model	BF ₁₀
Replacement 1	.872	.042	.848	.832	.069	.764	0.46
Replacement 2	.688	.050	.642	.750	.064	.713	0.76
Replacement 3	.696	.052	.651	.694	.052	.660	0.31
Replacement 4	.637	.046	.646	.646	.056	.658	0.32
Replacement 5	.618	.054	.639	.653	.063	.632	0.41
Transposition 1 and 2	.825	.052	.807	.781	.069	.754	0.47
Transposition 2 and 3	.603	.040	.564	.695	.041	.640	12.55
Transposition 3 and 4	.589	.040	.566	.642	.046	.589	0.98
Transposition 4 and 5	.561	.040	.559	.577	.053	.571	0.34
Migration Letter 1 to 3	.897	.038	.885	.839	.067	.858	0.74
Migration Letter 2 to 4	.668	.052	.658	.767	.049	.757	5.81
Migration Letter 3 to 1	.867	.040	.895	.838	.067	.877	0.38
Migration Letter 3 to 5	.710	.058	.651	.682	.045	.671	0.39
Migration Letter 4 to 2	.649	.058	.655	.765	.061	.782	5.47
Migration Letter 5 to 2	.698	.058	.723	.802	.057	.823	320.7
Migration Letter 5 to 3	.625	.054	.660	.684	.058	.691	0.32
Mig. 1 to 3 + Replacement	.878	.036	.918	.866	.067	.888	0.32
Mig. 2 to 4 + Replacement	.730	.062	.713	.801	.065	.796	0.82
Mig. 3 to 1 + Replacement	.885	.048	.895	.853	.068	.882	0.39
Mig. 3 to 5 + Replacement	.710	.060	.708	.711	.057	.724	0.31
Mig. 4 to 2 + Replacement	.704	.062	.711	.791	.065	.799	1.30
Mig. 5 to 2 + Replacement	.713	.060	.739	.820	.063	.824	3.27
Mig. 5 to 3 + Replacement	.684	.062	.713	.728	.067	.733	0.44

Bayes factors are ratios of the probabilities of the data given two competing models (H_0 vs. H_1 , in this case; see Rouder, Morey, Speckman, & Province, 2012).

As might be expected, the overall the accuracy rates were higher for replacement conditions than for transposition conditions, and when comparing replacements versus transpositions at specific letter positions, the comparisons that provide the most support for the H₁ are replacement of the 1st letter versus transposition of Letters 1 and 2 (t = 3.89, BF₁₀ = 36.9) and replacement of the 4th letter versus transposition of Letters 4 and 5 (t = 3.46, BF₁₀ = 16).⁴ Table 3 shows the BF₁₀s for other comparisons.

In addition, we compared the accuracies for the present experiment to those of the parallel experiment in English. Notably, the conditions with the greatest evidence for a cross-language difference are those that involve the second letter: transposition of 2 and 3 (t =3.16, BF₁₀ = 12.55), migration of 2 to 4 (t = 2.791, BF₁₀ = 5.81), migration of 4 to 2 (t = 2.761, BF₁₀ = 5.47), migration of 5 to 2 (t = 4.459, BF₁₀ = 320.67), and migration of 5 to 2 (t = 2.488, BF₁₀ = 3.27). For the rest of the conditions, the Bayes factors are less than 1, meaning that there is in fact support for the null model (see Table 3).

To summarize, the data show some relevant qualitative patterns: (a) There is a transposed-letter effect, since the transposed-letter conditions are more difficult than the single-replacement conditions; (b) although it is numerically smaller than in the English data, there is also an advantage for first-letter manipulations, since the conditions in which the first letter are manipulated are easier than the conditions in which the first letter was not manipulated; and finally, (c) performance is quite similar in this experiment versus its English counterpart, except for conditions in which the second letter was different between the target and the lure, in which case the performance of Thai readers was superior to that of English readers.

⁴ Jeffreys (1961) provides a scale for the interpretation of Bayes factors. For example, a BF of 1:1 to 3:1, is "barely worth mentioning," whereas a BF of 100:1 is "decisive." We prefer not to use an arbitrary cutoff that could be construed as a critical value; however, we will use Jeffreys's wording to present the outcomes of our analyses.

Table 3 Comparisons among conditions in the experiment

	<i>df</i> = 19		p(correct)		
Conditions	t Value	BF (H ₁ /H ₀)	Cond. 1	Cond. 2	Difference
R1–TL12	-3.89	36.9	.832	.781	051
R2-TL23	-2.11	1.4	.750	.695	055
R3-TL34	-2.19	1.6	.694	.642	052
R4-TL45	-3.46	16.0	.646	.577	069
RML-ML 1 to 3	-1.65	0.7	.866	.839	027
RML-ML 2 to 4	-1.68	0.8	.801	.767	034
RML-ML 3 to 1	-0.83	0.3	.853	.838	014
RML-ML 3 to 5	-1.78	0.9	.711	.682	029
RML-ML 4 to 2	-2.27	1.8	.791	.765	026
RML-ML 5 to 2	-1.27	0.5	.820	.802	018
RML-ML 5 to 3	-2.34	2.1	.728	.684	044
TL12 vs. TL23	-2.94	5.1	.781	.695	086
TL23 vs. TL34	-2.38	2.2	.695	.642	053
TL34 vs. TL35	-1.96	1.1	.642	.577	065
R1 vs. R2	-3.82	32.4	.832	.750	082
R2 vs. R3	-2.47	2.5	.750	.694	056
R3 vs. R4	-1.53	0.6	.694	.646	048
R4 vs. R5	-0.35	0.3	.646	.653	.007

For graphical displays of the differences between languages, refer to Fig. A18 in the online appendix, in which we show a scatterplot of the performance in the English experiment against the performance in the Thai experiment. For visual displays on the quality of the model fits, see Fig. A4 in the online appendix. RX, Replace letter X; TLXY, Transpose letters X and Y; RML–ML X to Y, Migrate + replace letter X to Y vs. Migrate letter X to Y

Model fits

The main objective of the present experiment in Thai was not to establish that transpositions are harder to detect than replacements. Instead, the goal was to assess whether the overlap model could account for the differences and similarities in performance across the 23 conditions in our experiment. The model assumes that in our experimental paradigm the briefly presented target would be compared to the two alternatives; this comparison yields a measurement of orthographic similarity, which is computed from the overlap between the flashed string and the alternatives. The overlap is calculated by the model as follows: For each letter position, the area under the curve of the letter in the slot within the target string is multiplied by the area of the same letter in the corresponding slot within the study string, and these products are summed over all the slots, as described by Eq. 1 in Gomez et al. (2008):

overlap =
$$\sum_{i=1}^{L} \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} f_1(x) dx \int_{i-\frac{1}{2}}^{i+\frac{1}{2}} \int f_2(x) dx$$
,

where *i* is the center of the position slot, $f_1(x)$ is the distribution of the first stimulus centered on *i*, $f_2(x)$ is the distribution of the second stimulus, and *x* is the position along the word order dimension (horizontal axis).

Given the use of two alternatives in our experiment, the model computes the two overlap measurements (i.e., the target stimulus with the correct alternative $[overlap_t]$ and the target stimulus with the foil $[overlap_f]$). The *overlaps* between the target stimulus and the two alternatives are transformed into correct-response proportions using a power function (Eq. 2 in Gomez et al., 2008):

$$p(\text{correct}) = \frac{(overlap_t)^a}{(overlap_t)^a + (overlap_f)^a},$$

where *a* is a scaling parameter greater than 1; this allows small differences in overlap to produce larger effects in accuracy.

As in the Gomez et al. (2008) study, to fit the overlap model to the data, we utilized the general-purpose optimization method in R based on the Nelder-Mead algorithm (Nocedal & Wright, 1999), which adjusts the values of the six parameters of the model. The data entered into the minimization routine were the response proportions for the 23 conditions. These data could be averages across participants, but they could also be the data for each individual participant. We performed calculations both ways: When comparing the model fits, we present the fits to the averaged data (as in Table 2 and Fig. 1). However, when performing statistical inferencesfor example, when comparing the values of parameters for the present experiment against the parameter values in Gomez et al. (2008)—we use the parameters obtained from fitting the model to each participant's data as our dependent variable. For these comparisons (see Table 4), we utilized Bayes factors, which revealed substantial evidence of a cross-language difference only for the second letter (with a larger s_2 value for English), and anecdotal evidence of a difference for the final letter position (with larger s5 values for Thai).

We can compare the parameters of the model for this experiment to the parameters in the parallel experiment with English-speaking participants. As can be observed in Fig. 1, there are qualitative differences between the parameters for the two languages. Namely, the parameters that describe the position uncertainty have higher values as the letter position increases in both languages; however, the growth in value is more pronounced in the English than in the Thai experiment.

Gomez et al. (2008) argued that the five *s* parameters could be described with a simple exponential growth-to-asymptote function over letter positions:

$$S_i = d\left(1 - e^{-(i-0.5)/r}\right).$$

2



Fig. 1 Values of the *s* parameter of the overlap model for each of the letter positions: s_1 to s_5 . For comparison purposes, we plot the parameter values for the experiment in English from Gomez et al. (2008) as the "E" points, and the parameter values for the present experiment as the "T" points. Note that these are parameters obtained from fitting the model to the averaged data. The solid line represents the exponential growth-to-asymptote function that best fits the E points (*s* parameters for the experiment in English), whereas the thick dashed line represents a linear function that fits the T points (*s* parameters for the experiment in Thai). The thin dotted line represents the exponential growth-to-asymptote function that best fits the T points (which misses the points badly). Refits of the empirical data with the simplified models can be seen in the online appendix: for English (Figs. A12 to A17) and for Thai (Figs. A4 to A9).

This function represents the idea that the value of the s parameter rises across letter positions (i) at a rate r until it reaches an asymptotic value d.

Importantly, the *s* parameters do not behave in this manner for the Thai experiment. Indeed, the *s* parameters for Thai rise in a linear manner. For comparison purposes, in Fig. 1 we show the parameters and the exponential growth to asymptote from Gomez et al.'s (2008) article, as well as both the linear and the exponential growth-to-asymptote fits for the *s* parameters from the present experiment with Thai. We compared the sums of squares of the residuals and found that the linear function had a much better fit (SS = 0.02) than the exponential growth function (SS = 0.07). Conversely, when we reanalyzed the Gomez et al. data, the exponential growth-to-asymptote

Table 4Parameters of the overlap model for Thai, again with paralleldata from Gomez et al. (2008) in English for comparison purposes

Parameter	English	Thai	t Score	BF (Alt/Null)
<i>s</i> ₁	0.566	0.924	-1.421	0.683
<i>s</i> ₂	2.264	1.02	4.691	539.803
<i>s</i> ₃	1.412	1.414	-0.006	0.309
<i>s</i> ₄	2.132	1.827	1.066	0.484
<i>s</i> ₅	1.108	1.916	-3.474	25.204
а	3.53	2.964	1.374	0.649

function fared better (SS = 0.07, vs. SS = 13 for the linear function). Note that both the linear and exponential approaches to a limit have two parameters.

Discussion

The main aim of the present study was to uncover differences in the encoding of letter position due to the nature of the script in Thai (i.e., an unspaced orthographic system in which some of the letters may be misaligned). We used the overlap model as a medium to explore this question. The conclusions are straightforward: Whereas the general perceptual-uncertainty mechanisms might be the same in Thai and in English—as deduced by the good fits in both Thai and English—the specific characteristics of the languages do seem to modulate this mechanism. In both scripts, the degrees of perceptual uncertainty are not the same across all letter positions, with a growth in the value of the uncertainty parameter as a function of condition. Critically, the forms of the function are different for Thai and for English.

Although for English it can be described as an exponential growth to a limit, in Thai it seems to be a linear function. This means that the position encoding for the first letter is better than those for the rest of the positions in both languages, but not as dramatically different in Thai as in English. These results support the prediction that position uncertainty for initial letter position is somewhat greater in Thai than in English. Thus, these results also support the view that orthographicspecific characteristics of the Thai writing system play a role in shaping the process of letter position coding. Specifically, two differences between Thai and English writing might be at play: (1) the lack of interword spacing might make the first letter less salient for Thai readers, which might then make the decay of the position coding accuracy less pronounced than for English readers; and (2) the misaligned vowels might deemphasize the role of the first letter, since it is not necessarily the one that needs to articulated first during reading. We acknowledge that determining the relative importance of these two factors is beyond the scope of our work, and we do hope that these findings might result in further research with Thai readers.

Although the present study focused on the fits from the overlap model of letter position coding, we acknowledge that other models that have employed the principle of "perceptual uncertainty" to encode letter position can also capture the obtained effects (e.g., the spatial-coding model, overlap open-bigram model, Bayesian reader model, or LTRS model).

In sum, the present findings are consistent with wellstudied phenomena related to the learned aspects of perception and attention. Indeed, previous research in our laboratory revealed that the degree of letter position coding during visualword recognition is modulated by expertise in orthographic– lexical processing (Perea, Marcet, & Gomez, 2016). Although the task will be challenging, future directions of research may aim to uncover the possible trajectories of learning letter encoding (see Wiley et al., 2016, for evidence from letter identity coding).

Author note The research reported in this article was partially supported by Grant No. PSI2014-53444-P from the Spanish Ministry of Economy and Competitiveness. We thank Sudaporn Luksaneeyanawin, Wirote Aroonmanakun, and Theeraporn Ratitamkul, at the Centre for Research in Speech and Language Processing (CRSLP) and the Linguistics Department, Chulalongkorn University, Bangkok, for advice and assistance as well as the use of their laboratory facilities. We also thank Chalong Saengsirivijam for assistance with participant recruitment. Finally, we thank two anonymous reviewers for their helpful comments and suggestions.

References

- Adelman, J. S. (2011). Letters in time and retinotopic space. *Psychological Review, 118,* 570–582. doi:10.1037/a0024811
- Aroonmanakun, W. (2007). Creating the Thai National Corpus. Manusaya, 13, 4–17.
- Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, 117, 713–758. doi:10.1037/a0019738
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods*, *Instruments*, & Computers, 35, 116–124. doi:10.3758/bf03195503
- Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, 115, 577–600. doi: 10.1037/a0012667
- Grainger, J., Granier, J. P., Farioli, F., Van Assche, E., & van Heuven, W. J. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 865–884. doi:10.1037/0096-1523.32.4.865
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). New York, NY: Oxford University Press.
- Logan, G. D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, 103, 603–649. doi:10.1037/0033-295X.103.4.603
- Nocedal, J., & Wright, S. J. (1999). Numerical optimization. New York, NY: Springer.
- Norris, D., Kinoshita, S., & van Casteren, M. (2010). A stimulus sampling theory of letter identity and order. *Journal of Memory and Language*, 62, 254–271. doi:10.1016/j.jml.2009.11.002
- Perea, M., Abu Mallouh, R., & Carreiras, M. (2010). The search of an input coding scheme: Transposed-letter priming in Arabic. *Psychonomic Bulletin & Review*, 17, 375–380. doi:10.3758/pbr.17. 3.375
- Perea, M., & Carreiras, M. (2006). Do transposed-letter effects occur across lexeme boundaries? *Psychonomic Bulletin and Review*, 13, 418–422. doi:10.3758/bf03193863

- 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. doi:10.3758/ s13423-012-0273-3
- Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter confusability effects in masked associative priming. *Memory & Cognition*, 31, 829–841. doi:10.3758/bf03196438
- 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. doi: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. doi:10. 7334/psicothema2015.167
- Perea, M., Winskel, H., & Ratitamkul, T. (2012). On the flexibility of letter position coding during lexical processing: The case of Thai. *Experimental Psychology*, 59, 68–73. doi:10.1027/1618-3169/ a000127
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56, 356–374. doi:10.1016/j.jmp.2012. 08.001
- Tóth, D., & Csépe, V. (2016). Adaptive specialization in position encoding while learning to read. *Developmental Science*. doi:10. 1111/desc.12426
- Velan, H., & Frost, R. (2011). Words with and without internal structure: What determines the nature of orthographic and morphological processing? *Cognition*, 118, 141–156. doi:10.1016/j.cognition.2010.11. 013
- White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of wordbeginning letters. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1261–1276. doi:10.1037/0096-1523.34.5.1261
- Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, 8, 221–243. doi:10.3758/ bf03196158
- Wiley, R. W., Wilson, C., & Rapp, B. (2016). The effects of alphabet and expertise on letter perception. *Journal of Experimental Psychology: Human Perception and Performance*, 42, 1186–1203. doi:10.1037/ xhp0000213
- Winskel, H., & Iemwanthong, K. (2010). Reading and spelling acquisition in Thai children. *Reading and Writing*, 23, 1021–1053. doi:10. 1007/s11145-009-9194-6
- Winskel, H., Perea, M., & Ratitamkul, T. (2012). On the flexibility of letter position coding during lexical processing: Evidence from eye movements when reading Thai. *Quarterly Journal of Experimental Psychology*, 64, 1522–1536. doi:10.1080/17470218.2012.658409
- Yakup, M., Abliz, W., Sereno, J., & Perea, M. (2015). Extending models of visual-word recognition to semicursive scripts: Evidence from masked priming in Uyghur. *Journal of Experimental Psychology: Human Perception and Performance*, 41, 1553–1562. doi:10.1037/ xhp0000143