Perception and Production of a Feature Correlation in Chinese Characters

James Myers
National Chung Cheng University

May 9, 2025

[forthcoming in *Written Language and Literacy*]

Abstract

How detailed is the implicit knowledge of regularities in written form? To address this question experimentally in Chinese, we examined a subtle feature correlation in which left stroke curving is obligatory in narrow arched-shaped components but not in wide ones, despite neither feature being lexically contrastive. While width did not directly affect curving classifications in a perception experiment, a statistically significant subset of participants gave significantly more "curved" responses and/or significantly increased them for narrow arches. In a handwriting experiment, the contrast in written curving degree was significantly greater for wide arches, and stroke speeds were also significantly more similar, as if they were planned separately. Together these results confirm that even very subtle formal patterns may become mentally active through experience with a writing system, and also suggest that stroke planning in handwriting shares similarities with the effects of prosody on articulation in speech and signing.

Keywords: Chinese, formal patterns, strokes, perception experiment, handwriting experiment, prosody

1. Introduction

A key function of writing is to represent speech, so it is unsurprising that the mapping between written form and spoken language dominates the literature on writing systems. Graphemes represent phonemes, morphemes, or various other linguistic units (Meletis 2019), and even subgraphemic regularities can provide important information, most famously in the shapes of Korean letters (Lee 2009), but even in the Roman alphabet, lowercase consonant letters tend to extend above or below the line while vowel letters tend not to (Primus 2004). Components internal to Chinese characters also give useful clues about the meanings or pronunciations of the monosyllabic morphemes represented by the characters as wholes (Chang, Hsu, Tsai, Chen & Lee 2016; Sze, Rickard Liow & Yap 2014).

Nevertheless, all writing systems also contain purely formal patterns. English, for example, has numerous phonetically vacuous spelling rules, like the alternation between <y> and <i> in <happy>/<happiness> and <baby>/<babies> (Albrow 1972; McCawley 1994; Venezky 1970). Many writing systems also have context-dependent allography (McCawley 1994; Meletis 2020), such as initial capitalization or the position-specific variants in Arabic script (Friedmann & Haddad-Hanna 2012). Chinese characters show formal regularities of these types as well, such as components that take on a different allographic form in certain positions (e.g., 心 *xīn* 'heart' must appear as 忄 on the left, as in 恨 *hèn* 'hate'). Moreover, in every writing system where it has been tested experimentally, readers and writers generalize formal patterns to never-before-seen test cases (see review in Treiman & Kessler 2022), and indeed, acquisition of these patterns does not necessarily depend on overt instruction (Pacton, Perruchet & Fayol 2001; Tsai & Nunes 2003).

This sort of implicit learning is often conflated with statistical learning (again see Treiman & Kessler 2022), but the concepts are distinct. All learning depends on sampling, but this doesn't mean that what ends up represented in the learner's head must be coded in statistical rather than categorical terms, that the representations must be superficial rather than abstract, or even that the sampling procedure itself has no innate biases (Lidz & Gagliardi 2015). The pattern examined in this study is learned implicitly in the sense that it is never taught in school, but whether it is best understood as mere statistical learning of sensorimotoric correlations or as something more abstract requires further discussion (a point we return to at the end of the paper).

In any case, implicit knowledge of regularities in written form remains understudied, particularly for non-alphabetic writing systems. Here we report two experiments on an extremely subtle pattern in Chinese character form that has received virtually no research attention at all, in order to explore just how rich implicit written knowledge can be and what might motivate it.

1.1 The importance of formal linguistic patterns

Formal patterns also have functions. In spoken language, allophonic and phonotactic patterns aid perception and recognition by adding redundancies that enhance the cues needed to identify lexical items. This conclusion, with a solid basis in information theory (Goldsmith 2000; Hall, Hume, Jaeger & Wedel 2018; Hockett 1953), has repeatedly been demonstrated empirically (Beddor, McGowan, Boland, Coetzee & Brasher 2013; Massaro & Cohen 1983; Penney, Cox & Szakay 2020). Formal phonological patterns also ease language production by making some aspects of lexical form predictable. This is why speakers underarticulate more frequently used words and sounds, both because they expect that listeners can fill in the familiar details themselves (Cohen Priva 2015), and because well-practiced gestures can be produced more efficiently (Tomaschek, Tucker, Fasiolo & Baayen 2018). Speakers may also hyperarticulate sounds with canonical (i.e., prototypical) status, even if they are not the most frequently encountered (Lu and Lee-Kim 2021; for canonicity effects in speech perception, see Sumner & Samuel 2005).

These same phenomena are found in reading and writing. Chetail (2015) reviews evidence for the benefits of redundancy in the reading of alphabetic scripts, and according to the somewhat controversial word shape effect (Beech & Mayall 2005; Lavidor 2011; Perea & Rosa 2002), readers of alphabetic scripts may even be able to make use of the outlines of printed words to aid lexical access. Similarly, Chinese readers recognize individual characters faster and more accurately the more frequent their components (Taft & Zhu 1997; Wu, Mo, Tsang & Chen 2012), and they also use the overall layout of components (e.g., horizontal vs. vertical) to help identify characters (Yeh, Li, Takeuchi, Sun & Liu 2003). Underarticulation is also a major influence on handwriting, which helps explain the tendency for frequent graphemes to reduce (Koshevoy, Miton & Morin 2023). Canonicity effects are found as well. In a charming study, Wong, Wadee, Ellenblum, and McCloskey (2018) found that despite looptail <g> being the dominant variant of *g* in English print, most of their participants expressed no awareness of its existence, were unable to handwrite it correctly, and failed to detect errors in modified forms. Instead, they seemed to treat the opentail <g> used in handwriting as the canonical form.

While formal regularities affect both input and output processing, there is an important difference. In principle, a spoken word can be recognized successfully as soon as enough of it has been heard to confirm that it cannot be any other word (Allopenna, Magnuson & Tanenhaus 1998; Ernestus & Cutler 2015; Marslen-Wilson 1987). Similarly, studies suggest that visually discriminating among Roman letters seems to depend more on stroke endings (e.g., four in <X> vs. none in <O>) than any of the many other available visual features (Fiset, Blais, Ethier-Majcher, Arguin, Bub & Gosselin 2008). By contrast, whether in writing, speech, or signing, producing a lexical item requires one not just to make it distinct from all other lexical items but also to generate accurate motor plans for all of its salient features, whether or not they are redundant. In Chinese character processing, this asymmetry between input and output is seen in so-called character amnesia (Huang, Zhou, Du, Wang & Cai 2021), whereby characters that are recognized without effort may not be so easily written by hand.

1.2 A feature correlation in Chinese characters

The specific formal pattern examined in this study involves a correlation between the features of width and curving within Chinese character components (often called radicals, though this term is traditionally restricted to dictionary indexing components). Width is contrastive in only one minimal pair: high-frequency 日 *rì* ‘day’ versus low-frequency 曰 *yuē* ‘say (classical Chinese).’ Two other minimal pairs, both relatively frequent, contrast only in stroke length: 土 *tǔ* ‘soil’ versus 士 *shì* ‘gentleman,’ and 未 *wèi* ‘not yet’ versus 末 *mò* ‘final.’ Yang and Wang (2018) explain the paucity of such contrasts by noting that width and length are relative features that require context to detect, which makes them harder to learn (whether explicitly or implicitly) and process. Nevertheless, these features still play an active role in the system because their distribution is highly predictable: character components enlarge on the right and bottom edges of a character, as shown by the larger size of 馬 *mǎ* 'horse' in 媽 *mā* 'mother' versus 騎 *qí* 'ride,' and of 大 *dà* 'big' in 尖 *jiān* 'sharp' versus 奇 *qí* 'strange' (we cite traditional characters here because our experiments were run in Taiwan, but all the generalizations we review also hold of the simplified characters used in China and Singapore). In the same way, strokes are consistently lengthened on the bottom and right of components, as in 三 *sān* 'three' and 川 *chuān* 'river' (Wang 1983). Acceptability judgment experiments on nonce characters demonstrate that both of these patterns are productive (Myers 2019).

The second feature examined in this study is stroke curving. Even though straight and curved vertical strokes are traditionally treated as distinct (豎 *shù* 丨 vs. 豎撇 *shù piē* 丿), there are no minimal pairs at all, and again the distribution is predictable: curved vertical strokes can only appear on the left edge, as in 川, 月 *yuè* 'moon,' 片 *piàn* 'slice,' and 爿 *qiáng* 'piece of wood.' The left-edge restriction usually applies at the component level, as in 所 *suǒ* 'place' (戶 + 斤), though a handful of components are lexically marked to undergo curving at the left edge of the whole character, as in 辣 *là* 'spicy' (辛 + 束). Again, acceptability judgment experiments in Myers (2019) confirm that Chinese readers generalize the positional restriction on curving to nonce components.

As Wang (1983) suggests and Myers (2019) confirms in quantitative corpus analyses, the features of width and curving are themselves correlated: stroke curving is obligatory in relatively narrow characters like 月, 丹 *dān* 'cinnabar,' and 拜 *bài* 'honor,' but straight strokes are common in characters like 冊 *cè* 'volume (book)' and 而 *ér* 'and,' where the arch is wide relative to its height. The only near-minimal pairs in curving involve wider components like curved 周 *zhōu* 'cycle' versus straight 同 *tóng* 'same,' though as Wang (1983) points out, the extra strokes in 周 relative to 同 may make it appear proportionally taller and thus narrower. Myers (2019) provides several other observational arguments for the historical productivity of this width/curving correlation, including the replacement of wide and straight 肉 *ròu* 'flesh' with a narrow and curved allograph in the narrowing left-edge position in characters like 胖 *pàng* 'fat,' and the statistically significant tendency for straight components to disfavor the left-edge position, like straight 甬 *yǒng* 'corridor' on the top in 勇 *yǒng* 'brave' versus curved 角 *jiǎo* 'horn' on the left in 解 *jiě* 'separate.'

While not explicitly recognizing this correlation, Unicode implicitly encodes it by having separate entries for the two forms of the "arch" structure in the above characters: the narrow and curved ⺆ versus the wide and straight 冂. While arches never form full characters and thus have no meanings or pronunciations, they are always written as wholes, with the simple left stroke (丿 or 丨) written immediately before the complex stroke on the right (㇆). They are also quite common, having a combined lognormed type frequency 1.58 standard deviations above the mean for the 441 distinct traditional character components identified by Chuang and Teng (2009). Nevertheless, the contrast between them is not taught as part of ordinary childhood literacy education.

Building on previous linguistically motivated analyses of writing systems (including Evertz 2018; McCawley 1994; Wang 1983; Watt 1975), Myers (2019) attempts to account for these formal asymmetries in terms of a written analog to prosody. Very briefly, component enlargement and stroke lengthening are argued to be analogous to stress, and curving is argued to be analogous to segmental correlates of the lack of stress, like vowel reduction. The analogies are not just physical but structural, since as we have seen, prominence and curving are both predictable from relative position, just as stress and vowel reduction are. Specifically, the rightmost component in a character and the rightmost stroke in a component are enlarged, as if they lie in prosodically strong positions, whereas curving only occurs in the leftmost position of a character or component, as if it lies in a prosodically weak position. In essence, the proposal is that characters have hierarchical prosodic structure, including a foot-like weak-strong constituent [WS] (see Myers 2021 for more formal details on what these feet are built on, and Myers 2024 for how prosody interacts with recursive component combination). The width difference between characters like 月 and 冊 can then be encoded in terms of the number of these "feet," respectively as one-foot [WS] and two-foot [S][S], analogous to the contrast between the one-foot English word *háppy* [SW] versus the two-foot *ábstràct* [S][S]. Following standard theories of stress (Hayes 1995), each foot is obliged to have a strong head (making *[W] ungrammatical). The upshot is that in this analysis, curving is disfavored in wide character components like 冊 because the leftmost position is strong, given its structure [S][S].

1.3 The present study

Despite such ambitious theories, the width/curving correlation itself has not yet been tested experimentally. Thus the first goal of this study is simply to establish whether this correlation has any effect whatsoever on reading or writing. It is not obvious that it must; not only is the pattern not taught in school, but anecdotally some Chinese writers claim that they always write arch components with a straight left stroke, regardless of the character. Moreover, the pattern involves features rather than full components (let alone full graphemes), the features are not lexically contrastive, and it is the feature correlation that matters, not each feature on its own; even in spoken language, feature correlations are hard to learn (see Warker & Dell 2006 on the effects of implicit learning on speech errors). If readers or writers show any sign of such an extremely subtle correlation, it would be striking evidence that the implicit knowledge of written form can be quite detailed indeed.

If the width/curving correlation does affect processing, our next question is whether the effects differ for input (perception) and output (handwriting). As noted earlier, while perceiving formal patterns may be easier than producing them, only writers are obliged to produce arch components as wholes, and so writing may be more likely than reading to reflect this pattern. Our third and final question is whether there is any empirical support for the hypothesis that the width/curving correlation involves something analogous to prosodic structure. In particular, if the manual production of narrow curved arches is timed using one prosodic unit [WS] while that of wide straight arches is timed using two [S][S], this should influence the relative speeds of the two strokes that comprise the arches.

To address these questions, we ran two experiments on arch characters. The first built on the categorical perception task applied to Chinese characters by Peng, Minett, and Wang (2010) and Yang and Wang (2018), in order to test if character width influences the identification of stroke curving. The second experiment was a repetition production task that tested if character width modulates the degree of stroke curving in handwriting. The results of the second experiment were then further examined for evidence that the relative timing of the arch's two strokes depended on width.

2. Perception experiment

2.1 Method

2.1.1 Participants

We recruited 40 native Chinese-speaking traditional character readers at National Chung Cheng University in Taiwan (mean age 21 years, range 18 to 25). None reported impairments in language, vision, hearing, or learning. Participants signed a consent form approved by the university ethics review board and were paid a nominal fee.

2.1.2 Materials and Design

Stimuli consisted of the Unicode character ⺆ (the narrow and curved arch component), modified to form distinct wide and narrow sets, each with gradiently varying degrees of left stroke curving, using the Character Description Language (CDL) tool of Wenlin (https://wenlin.com/; Bishop & Cook 2007). The CDL code described each character as points in a 128 × 128 square, with the topmost horizontal stroke extending 101 points along the *x*-axis in wide characters and 56 points in narrow ones. As in all modern typefaces, curving in Wenlin's character set is computed using Bézier splines, here linking a vertical straight line to a cubic curve in the leftmost stroke of the arch character. In our stimuli, the inflection point was always 42 points from the top of the character, but curving degree was gradiently adjusted by changing the *x* coordinate of the stroke's bottom point at ten 3-point increments, from no displacement relative to the stroke's top point (coded in the analyses as curving degree 1, i.e., fully straight) to a leftward displacement of 27 points (curving degree 10). These displacements were identical regardless of character width, including for the practice items with intermediate widths (76 and 81 points, alternating across increasing degrees of curving). Wenlin's plain typeface was used instead of its Song typeface (the serif family used in most printed text) because the latter has differently shaped tips for straight and curved strokes.

The 30 characters (narrow, wide, and intermediate for practice, each with 10 curving degrees) were exported from Wenlin as SVG files with stroke widths of 4% and converted to white lines on a black background for more rapid perception, using the magick package (Ooms 2024) in R (R Core Team 2024). The topmost horizontal stroke was always centered along the *x*-axis of the image, so that within each experimental block (one for each character width), the top of the leftmost stroke would remain in the same retinal location even as curving varied. This meant that the display field had to be rectangular in order to accommodate the maximally curved characters. A mask consisting entirely of white-on-black Xs was created separately. As presented to participants, the rectangular field containing the stimuli had a width of 1.6° vertically and 2.1° horizontally when viewed from a distance of about 80 cm (about the size of a viewer's thumb held out at arm's length: O'Shea 1991), so that each stimulus could be taken in at a glance. The 20 experimental stimuli are shown in Figure 1.

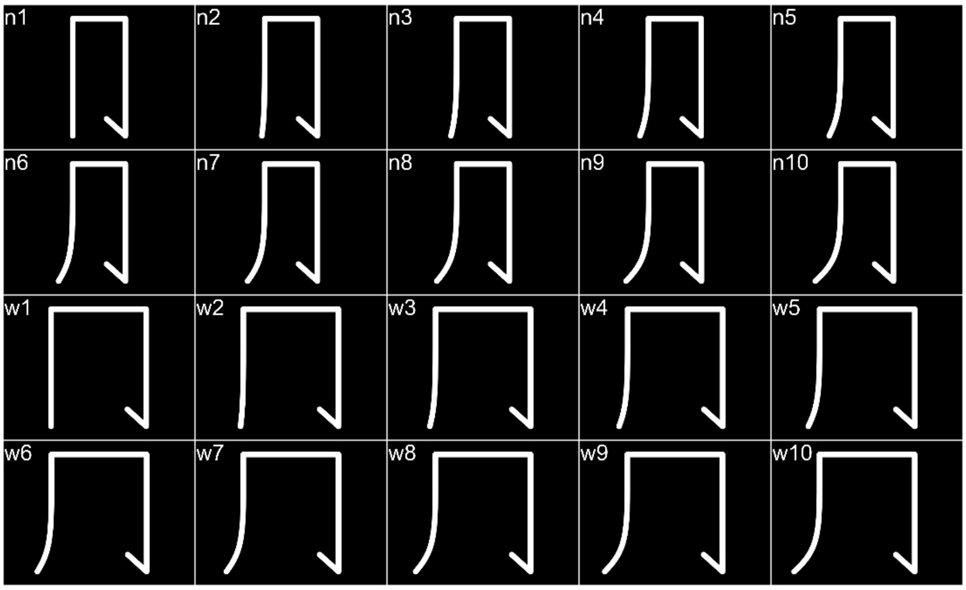

Figure 1. **Stimuli used in Experiment 1**, with 10 degrees of curving in narrow (n1-n10) and wide (w1-w10) arch characters.

2.1.3 Procedure

The experiment was run on Windows 10 computers using PsychoPy (v2022.2.5) (Peirce et al. 2019; Peirce et al. 2022). After 10 practice trials with arch characters of intermediate width, participants were presented with the wide and narrow character sets in separate blocks in counterbalanced orders across participants, and within each block, the ten items of each width were presented in random order ten times. Participants were asked to judge if the leftmost stroke was straight (豎 *shù* 丨) or not (撇 *piē* 丿) as quickly as possible. In both practice and experimental trials, labels (curved vs. straight) for the response keys on the computer keyboard (“S” on left vs. “L” on right) were also counterbalanced across participants, forming four participant groups in all (two block orders × two arrangements of response key labels). After a 1000 ms fixation symbol (“o” in white on a black background) and a 500 ms black screen, characters were flashed for 50 ms and then followed by a very long visual mask (2000 ms), during which time responses were recorded; trials ended with a button press or with a black screen for up to 600 ms, for a maximum trial duration of 4,150 ms (see Figure 2). As Peng et al. (2010) note, the rapid display and backwards masking in this procedure make it difficult for participants to become consciously aware of character form. The experiment took approximately 10 minutes per participant.

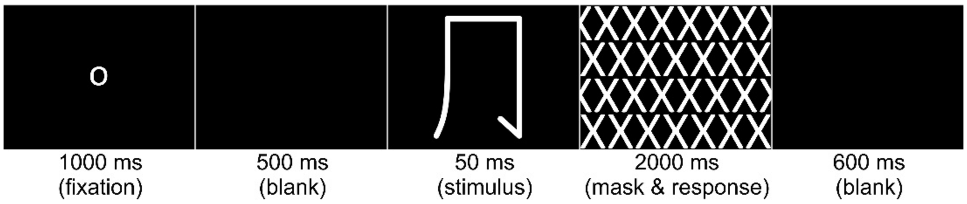

Figure 2. **Sample trial from Experiment 1**, illustrated with a practice trial with a target character of intermediate width and curving.

2.2 Results

199 trials without responses and 69 trials with reaction times faster than 100 ms were dropped (3% of the total), leaving 7,732 datapoints (40 participants × 20 stimuli × 10 repetitions – 268). The response choices were analyzed using logistic generalized additive mixed-effects modeling (GAMM), implemented using the gam function in R's mgcv package (Wood 2017). This class of model was chosen in order to capture nonlinear categorical perception without assuming that it must necessarily take a sigmoid shape, unlike the logistic regression models used by Yang and Wang (2018). Response choices were predicted from character width (narrow vs. wide), degree of stroke curving (1-10), and stimulus repetition (1-10). The two numerical independent variables and their interactions were treated as smooth terms, with the "wiggliness" of their effects expressed via the estimated degrees of freedom (*edf*) of their basis functions, where *edf* = 1 implies a linear effect. All smooth terms were based on thin plate regression splines, which, by analogy to real metal plates, simulate curving with resistance by finding a set of functions that collectively capture the observed curving in the most efficient way. Following Winter and Wieling (2016), we reduced overfitting by setting the number of basis functions to five, half the number of points in the curving continuum.

Our primary interest was in how character width modulated the influence of the degree of curving on responses, but as the word "additive" in its name suggests, GAMM does not model interactions in the way familiar from other types of regression. In order to compare the effects of curving across character widths, we followed van Rij (2015) in coding width as an ordinal factor, so that the curving effect was first estimated in narrow characters only and then in the residuals, thereby testing for a difference in the curving effect between narrow and wide characters. The GAMM analogs of the interaction between the two numerical variables (curving degree and repetition), and both with width, were modeled as tensor product smooths that factored out the effects of the individual variables. There were thus GAMM analogs for all three main effects and all two-way and three-way interactions. Finally, by exploiting the mathematical relation between smooths and random effects structure, the model also included by-participant random intercepts and random linear slopes for width, curving, repetition, and their interactions, each of which, in this type of model, yields its own *p*-value that tests the null hypothesis of cross-participant identity. The R code for this and all other analyses in this paper are given in the supplementary materials (https://osf.io/9f7pq/).

The resulting model is illustrated in Figure 3, which reconfirms the sigmoid shape expected of categorical perception, with the category boundary reflected in the sharp increase in "curved" judgments around curving degree 4. However, width did not have an overall effect on judgments ($p > .9$), nor did it modulate the curving effect. That is, while curving did matter in narrow characters (*edf* = 3.840, $\chi^2 = 68.51$, $p < .0001$), its effect in wide characters was no different from that in narrow ones ($p > .7$). Overall, "curved" judgments decreased with repetition (*edf* = 3.619, $\chi^2 = 25.40$, $p < .0001$), and as shown by a tensor product interaction between repetition and stimulus curving (*edf* = 3.307, $\chi^2 = 19.78$, $p < .001$), this decrease was more rapid on the straight end of the continuum. There were no other significant effects or interactions (*p*s > .7).

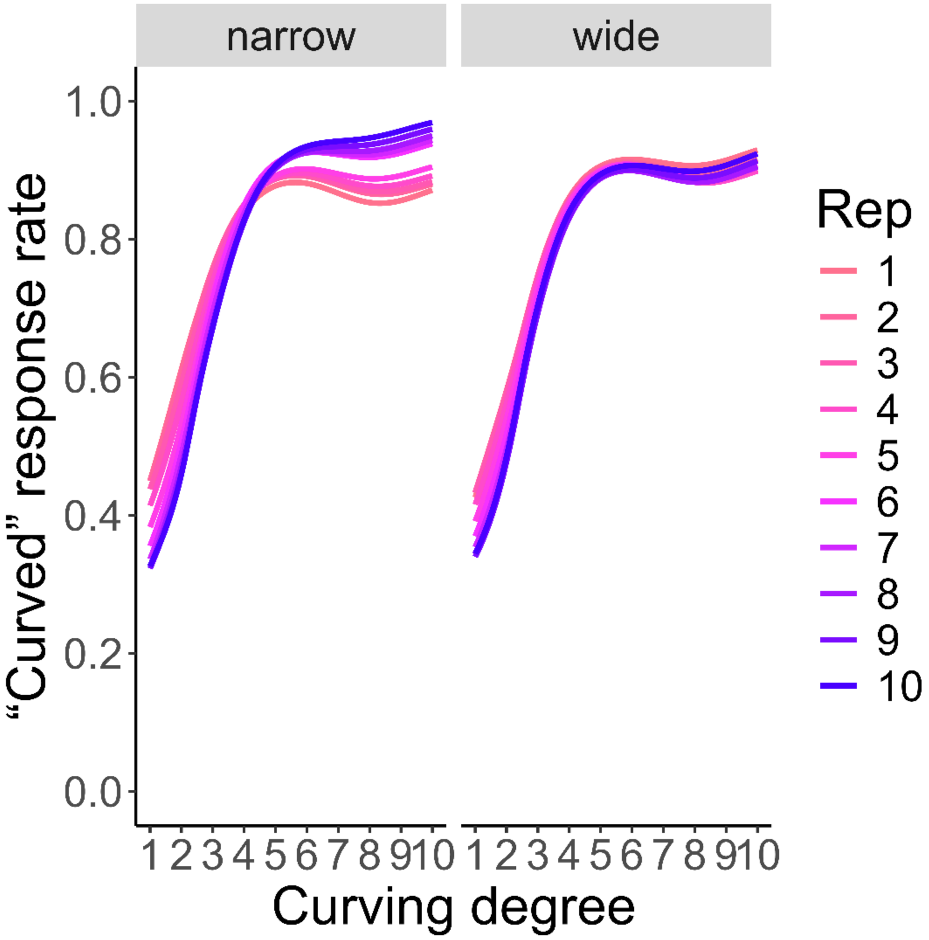

Figure 3. Proportion of "Curved" responses in Experiment 1 as a function of curving degree, width, and repetition (Rep).

The figure nevertheless suggests a trend towards an interaction between curving and width, with an overall higher proportion of curving in narrow characters, as well as an apparent three-way interaction of these two variables with repetition, whereby "curved" judgments increased across repetitions only for narrow characters. The by-participant random effect associated with the curving by width interaction was marginal (*edf* = 31.146, $\chi^2 = 39.00$, $p = .08$) and that associated with the three-way interaction was significant (*edf* = 23.073, $\chi^2 = 38.12$, $p < .01$), suggesting that these trends were driven by a subset of participants. Indeed, when GAM models were built separately for each participant (which converged in all but two of the 40 participants), 23.7% of the converging participants (9/38) showed $p < .05$ for one, the other, or both interactions in the predicted directions (i.e., more "curved" responses and/or an increasing number of "curved" responses for narrow stimuli). An additional two participants showed significant interactions in the opposite directions (i.e., a greater or increasing number of "curved" judgments in wide characters). The chance probability of so many participants showing the expected pattern is less than .003, as crudely estimated in a two-sided binomial test with success per trial set at .05 for seven successes (9 significant results in the predicted direction - 2 significant results in the opposite direction) in the 38 converging participants. A similarly significant result (two-tailed $p = .0001$) was found in a resampling analysis that better accounted for the two- and three-way interactions and for the expected and unexpected directions (see supplementary materials).

2.3 Discussion

Judgments of the curving of the leftmost stroke in the arch character were, unsurprisingly, categorical. Moreover, given that there is only one way to be truly straight but many ways to be curved, it took only a small increase in the degree of curving for participants to switch from usually judging the stroke as straight to almost always judging it as curved.

However, the results did not reveal a robust effect of character width on judgments of stroke curving, with the judgments for narrow and wide stimuli in Figure 3 both showing an inflection around the fourth point in the 10-point curving continuum. Nevertheless, motivation to keep looking for productivity in the width/curving interaction comes from trends in the same figure, driven by almost one-fourth of the participants, whereby across repetitions, narrow characters tended to yield a greater and/or an increasing proportion of "curved" judgments than wide characters. These trends are consistent with the hypothesis that a statistically significant subset of participants had, in some sense, internalized the regularity in printed text whereby the leftmost stroke is always curved in narrow arches.

While the width/curving correlation had at most weak effects on perception, this need not be the case for production. As noted in the introduction, lexical representations can afford to be less precise for processing inputs, where the goal is merely to detect lexical contrast, than when processing outputs, where all features, even predictable ones, must be produced accurately. Thus in Experiment 2 we turn to handwriting.

3. Experiment 2: Handwriting

3.1 Method

3.1.1 Participants

A new group of 41 native Chinese-speaking traditional character readers was recruited at the same university in Taiwan (mean age 21 years, range 20 to 30). None reported impairments in language, vision, hearing, or learning. Participants signed a consent form approved by the university ethics review board and were paid a nominal fee. Data from two participants were set aside because they self-reported as being left-handed, which may affect implementation of the conventional left-to-right stroke order of Chinese, leaving 39 participants for analysis.

3.1.2 Materials and Design

Four arch characters were chosen from the materials used in Experiment 1, two narrow and two wide, with curving degrees across, but close to, the identification boundary determined in Experiment 1 (stimuli n2, n5, w2, and w5 in Figure 1). These continuum points were chosen as perceptually unambiguous representatives of the four ways to combine width and curving, without however providing precise writing targets for prototypically straight or curved strokes.

3.1.3 Procedure

The experiment was run via PsychoPy (v2022.2.5) on a Windows 10 computer, with the display of instructions and stimuli and the collection of writing data handled by a Wacom® One tablet and stylus. PsychoPy treated the tablet and stylus as a mouse, recording the time, position, and stylus contact status (treated as a left mouse button press) approximately every 20 ms.

After eight practice trials with two arch characters of intermediate width and curving degrees 2 and 5, participants were presented with 40 experimental trials consisting of a random order of the four experimental stimuli shown 10 times each. Stimuli were displayed in white on a black 5 cm by 5 cm square, which was then replaced in the same location by a black-bordered white square where participants were asked to write the character they had just seen. More precisely, each trial began with a 1000 ms fixation symbol ("o" in white on a gray background) and a 500 ms gray screen, followed by the target character for 500 ms and a brief mask of 100 ms (see Figure 4). Trials ended only when participants used the stylus to tap a button labeled "OK" located below the writing area. Since the focus of this experiment was on writing and not perception, a relatively long display was used, and the mask was used only to prevent participants from tracing retinal afterimages. The experiment took around 10 minutes per participant.

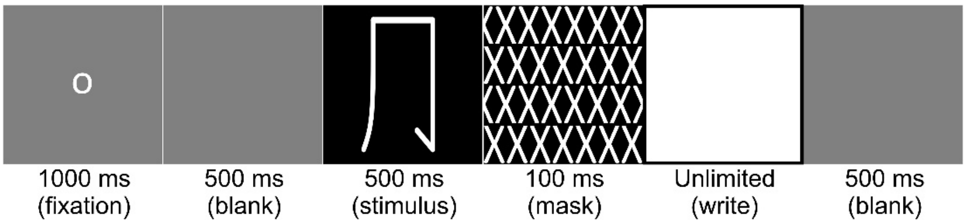

Figure 4. **Sample trial from Experiment 2**, illustrated with a practice trial with a target character of intermediate width and curving.

3.2 Results

The hypotheses tested in this experiment relate to the curvature of the first (leftmost) stroke in the two-stroke arches, and the relative amount of resources devoted to it in production. To calculate these two dependent variables, the times, locations, and stylus contacts were extracted automatically from the raw PsychoPy results files, with strokes defined in terms of the start and end of stylus contact (see Myers 2023, for more on how this was done). This information was then used to identify and remove trials with fewer than two strokes (42 trials, via writing the arch like an upside down U), trials with more than two strokes (25 trials, via false starts on one or both strokes), and trials in which at least one stroke crossed the border of the writing area (9 trials, including 2 with missing or extra strokes). These removals left a data set with 1,486 trials; see Figure 5.

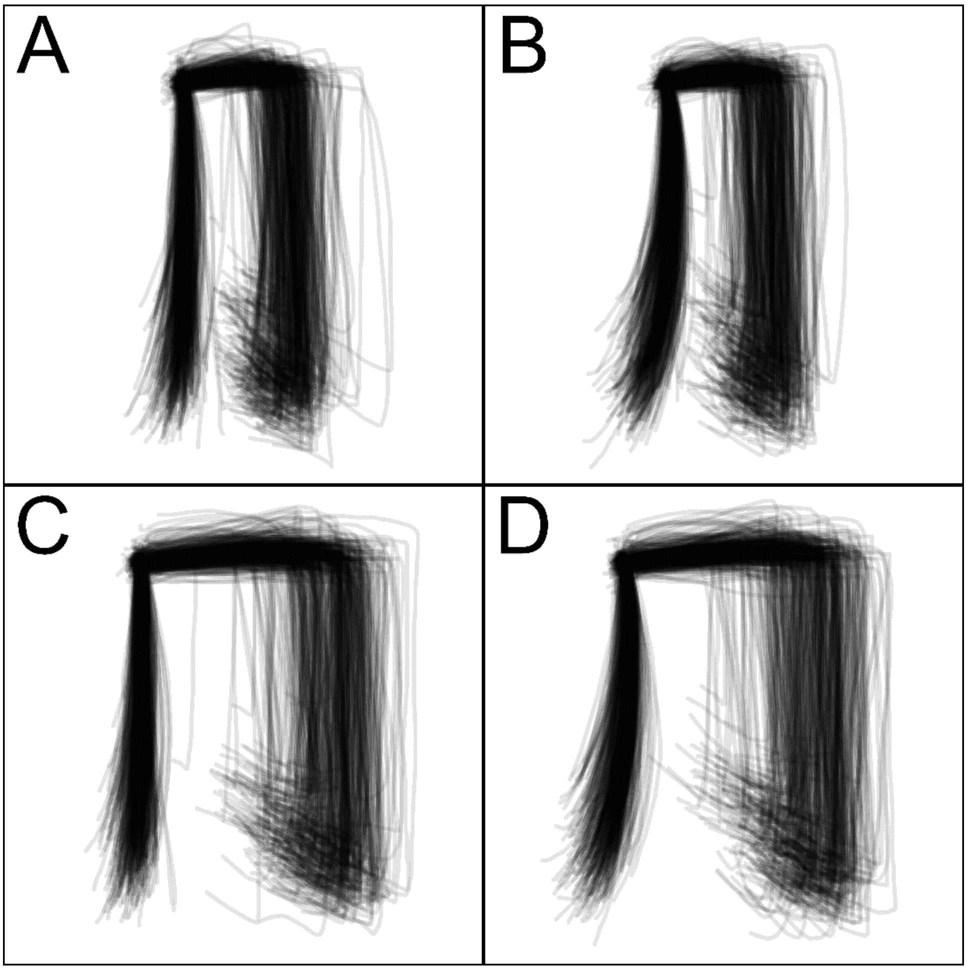

Figure 5. Stylus paths for written characters in Experiment 2. Only responses with exactly two strokes entirely within the writing area are included. Images are aligned to the position of first stylus contact. A, B, C, D show responses to stimuli n2, n5, w2, and w5, respectively. See Myers (2023) for information on how this figure was created.

The dependent variable of leftmost handwritten stroke curving was computed using the localCurvature function in the EBImage package (Pau, Fuchs, Sklyar, Boutros & Huber 2010) in R. This fits a parabola tangent to each point on the contour of an area, yielding a value that has greater magnitude the narrower the parabola (i.e., greater curvature). Positive values indicate concavity relative to the area's contour, while negative values imply convexity. Because the function processes contours in a clockwise direction, it assigns negative values to left-curving strokes. The means of these values were then negated so that higher values would represent greater mean curvature.

To estimate the resources devoted to each stroke, we quantified their relative speeds. We could not follow the methods used for comparing syllable stresses in spoken and signed languages, where relative duration is used (e.g., Reinisch, Jesse & McQueen 2011; Van der Kooij & Crasborn 2008), because in arch characters the simple first stroke, whether straight or curved, always has a shorter length and duration than the complex second stroke. Speed eliminates these confounds by dividing length by duration, and it is also known that faster gestures have less accurate gestural plans

(Latash 2012). Thus we expect “unstressed” strokes to be produced more quickly than “stressed” ones. Stroke duration was simply the difference between the first and last time measurements for each stroke, and stroke length was the sum of all Euclidean distances between each pair of points in each stroke. From these we calculated each stroke speed (length divided by duration) and finally the stroke speed ratio (the first stroke speed divided by the second stroke speed). The closer this ratio to one, the more equal the resources devoted to the two strokes.

In separate analyses, we modeled each of these two dependent variables as a function of first stroke type in the stimuli (below vs. above the perceptual threshold for curving determined in Experiment 1), stimulus width type (narrow vs. wide), repetition (1 through 10), and all of their interactions. All analyses used mixed-effects linear regression with by-participant random intercepts and slopes, except for repetition and its interactions, since unlike Experiment 1, the number of observations in Experiment 2 was too low to permit such a complex random effects structure. Moreover, due to the models’ complexity, we found it necessary to assume weakly informative Bayesian priors to avoid singular fits (a kind of overfitting), which were computed using the blmer function in R’s blme package (Chung, Rabe-Hesketh, Dorie, Gelman & Liu 2013). Likelihood ratio tests showed that the models with repetition did not give a significantly better fit than the models without it when predicting handwritten curving degree ($\chi^2(4) = 3.06$, $p > .5$) or stroke speed ratio ($\chi^2(4) = 6.59$, $p > .1$) and so the models reported here drop this variable. Because p values are not well defined in this partially Bayesian approach, we estimated them using Type II Wald F tests (which factor out main effects before interactions) with Kenward-Roger degrees of freedom, via the Anova function in R’s car package (Fox & Weisberg 2019). For completeness, we also report the t values given by the blmer function (as desired, we found that generally $F \approx t^2$).

The model predicting handwritten curving degree is illustrated in Figure 6. Stimulus stroke type significantly affected handwritten curving in the expected way ($B = 0.75$, $SE = 0.06$, $t = 11.90$, $F(1, 37.00) = 137.69$, $p < .0001$), while stimulus width had no significant effect on its own ($B = -0.03$, $SE = 0.02$, $t = -1.20$, $F(1, 36.98) = 0.94$, $p > .3$). Crucially, however, there was also a significant interaction between these two variables ($B = -0.06$, $SE = 0.02$, $t = -2.77$, $F(1, 36.97) = 7.65$, $p < .01$), whereby the difference in handwritten curving was greater in wide characters than in narrow ones.

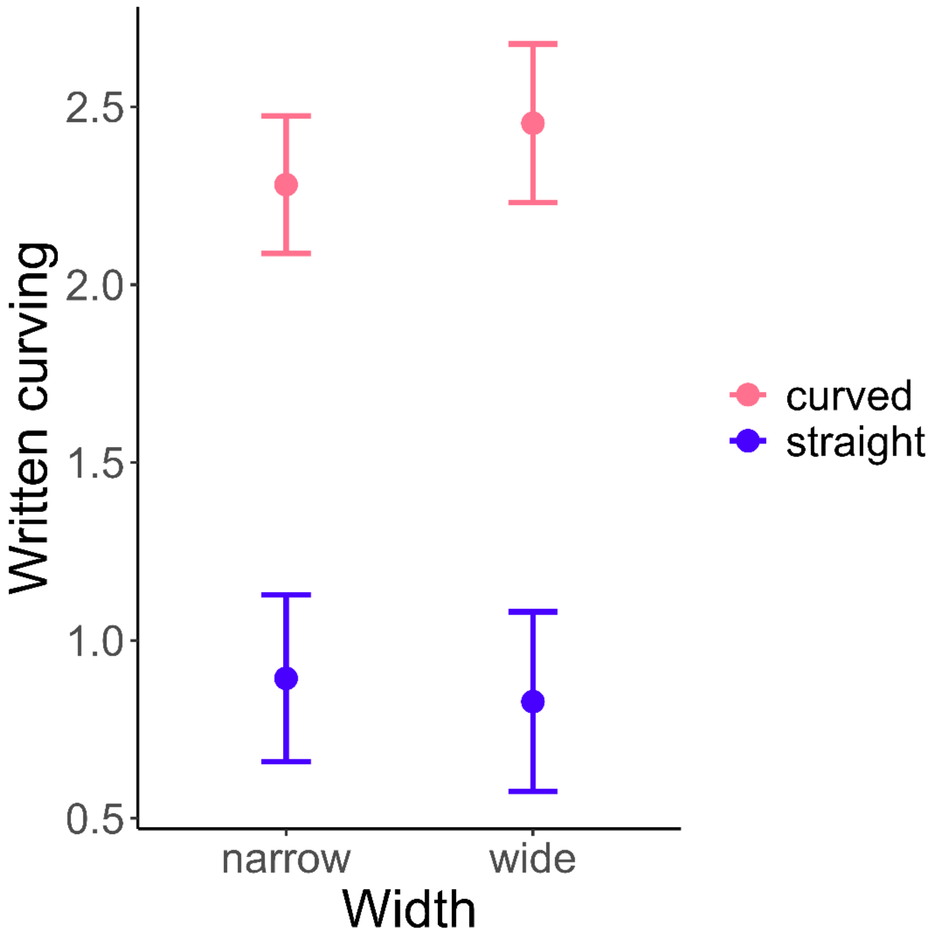

Figure 6. The degree of written curving as a function of stimulus curving degree. Written curving is the negation of the mean curvature at each point in the first stroke. Error bars represent 95% confidence intervals.

The model predicting stroke speed ratio is illustrated in Figure 7. Stroke type did not have a significant effect ($|t| < 1$, $p > .3$) but width did ($B = 0.053$, $SE = 0.01$, $t = 5.25$, $F(1, 36.99) = 27.42$, $p < .0001$); as shown in the figure, the stroke speed ratio was higher for narrow characters. There was also a marginal interaction between these two variables ($B = 0.014$, $SE = 0.007$, $t = 2.03$, $F(1, 36.97) = 4.1$, $p = .05$); as the figure suggests, the stroke speed ratio was higher when the first stroke was curved, but only in narrow characters.

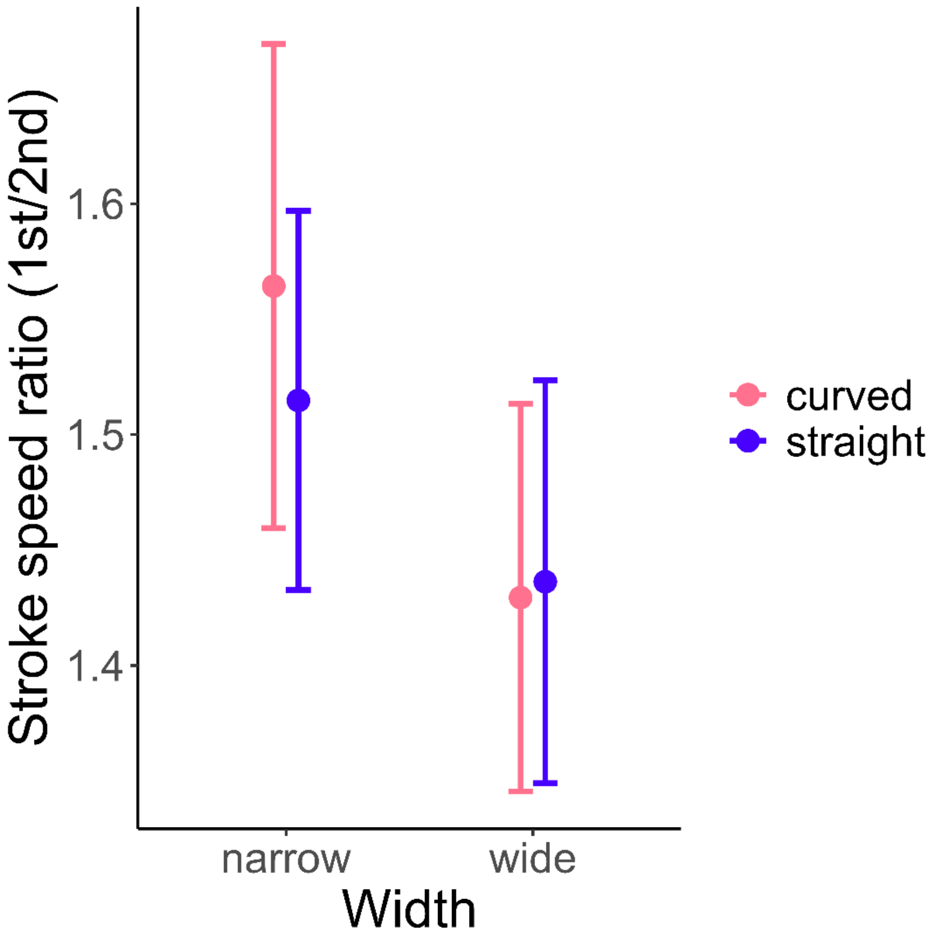

Figure 7. The ratio of first to second stroke speeds as a function of stimulus curving degree and width. Stroke speed ratio has no units. Error bars represent 95% confidence intervals.

3.3 Discussion

Handwritten arch characters consistently reflected the curving and width categories exemplified by the arch character stimuli. At the very least, then, this experiment shows that Chinese writers are generally able to write stroke curving distinctly in the <⺆> and <冂> variants of the arch form, even if anecdotal reports suggest that not all writers make this distinction in natural writing.

More importantly, stimulus width influenced the production of curving: the difference in the degree of handwritten curving for straight and curved stimuli was significantly greater for wide than for narrow characters. This means that the participants produced the left stroke in narrow arches similarly regardless of stimulus curving, consistent with their experience that this stroke is always curved in real narrow characters like 月, whereas they made a greater curving distinction in wide arches, a context in which lexical distinctions occur in real characters, as in 周 versus 同. Moreover, since the left stroke was written first, the effect of character width must have occurred during mental planning, before the width was physically realized via the second stroke. The results thus provide unambiguous evidence for the productive implementation of the width/curving correlation in Chinese arch components.

The ratio of first to second stroke speed was also significantly closer to one in wide than in narrow characters. This suggests that writers devoted a more equal amount of resources to the strokes in the wide context, consistent with the “prosodic” explanation proposed by Myers (2019) for the width/curving correlation, whereby wide character components have two foot-like groupings [S][S] rather than just one as in narrow components [WS]. Also consistent with the prosodic analysis is the finding that only in narrow characters was the stroke speed ratio (marginally) higher in curved-stroke arches, as if only in [WS] structures was the curved stroke “unstressed.”

4. General Discussion

Chinese character components, including the common arch component, show a quite subtle pattern in which the leftmost vertical stroke is always curved in narrow components but may be straight in wide components. Despite this pattern’s subtlety, a perception experiment suggested that some readers generalize the width/curving correlation to isolated arch components, and a production experiment provided robust evidence of this correlation in handwriting. Moreover, further analysis of the handwriting results suggests that the width/curving correlation may be mentally encoded in prosody-like terms, with narrow arches planned as a single unit while wide ones are planned as two separate units.

4.1 The productivity of formal written feature correlations

The answer to our first research question is thus yes: the width/curving correlation has productive effects on the processing of written forms. The evidence from the categorical perception task in Experiment 1 is admittedly weak: only around one fourth of the participants gave more curving judgments for narrow arches than for wide ones, and/or increased their curving judgments across repetitions more for narrow than for wide arches. One may prefer to interpret these results as pessimistically as Wong et al. (2018) did in their study of looptail <g> and opentail <g>, where their first experiment found that only one sixth of their participants expressed awareness of the two *g* forms (6/36 in their Table 1, p. 1326). However, in our experiment the number of participants whose responses reflected the width/curving correlation was significantly greater than would have been expected under the null hypothesis that the correlation had no effect at all, and our study differed even more dramatically from Wong et al. (2018) when it came to handwriting. While they found that only 28% of the participants in their second experiment could accurately write looptail *g* by hand (their Figure 4, p. 1330), the handwriting task in our Experiment 2 provided statistically significant evidence for the width/curving correlation in arch components, that is, reason to believe that it should generalize beyond our sample.

It may be that formal written patterns within a single system (the width/curving correlation) are much easier to learn, at least implicitly, than patterns that only exist across systems, as with the two *g* variants (though Wong et al. 2018 only tested for explicit awareness of them). Moreover, the width/curving correlation is not only not taught in school but is not a writing system universal either (the Roman alphabet does not have it), so despite the weak results in the perception task, the most likely source of the patterns observed in handwriting is reading experience.

4.2 Perception versus production

The fact that the width/curving correlation was so much more robust in handwriting than in perception answers our second research question: yes, modality does matter. After all, the arch component has quite a different status in handwriting than in reading. Because it is merely a stroke group rather than a full component, for readers its shape is obscured by the strokes within, around, and across it, whereas in handwriting it is always produced in an unbroken two-stroke sequence. This means that writers, unlike readers, do need to treat arches as planning units, making them what the Chinese psycholinguistics literature calls “logographemes” (Chen & Cherng 2013; Law & Leung 2000). The new discovery here is that writers can productively generalize the width/curving correlation to nonce characters, demonstrating that they do not merely memorize by rote which form of the arch is lexically specified to appear in each character.

4.3 Explaining formal written patterns

Our final research question asked whether the width/curving correlation could be explained via a written analog of prosody. The results showed that the speeds were more similar across the two strokes in wide arches, as if they were planned in separate and equally “stressed” units, while in narrow arches the first stroke was produced much more quickly than the second one, as if they were planned in a single asymmetric weak-strong “foot.”

While this analysis may seem overly abstract, alternative explanations that appeal to universal psychophysics just do not fare as well. We have already mentioned that no such width/curving correlation has been noted in any other writing system. Consider also the proposal by Yang and Wang (2018) that Chinese has virtually no minimal pairs in stroke length because relative features are difficult to detect. While initially plausible, this psychophysical explanation is immediately challenged by signed and spoken languages, where lexical contrasts in length are quite common, as in geminates versus singleton consonants and long versus short vowel phonemes. What is actually rare is lexical contrast in another relative feature, namely stress, which is either predictable (including in sign languages: Van der Kooij & Crasborn 2008) or not used phonologically at all (Goedemans & van der Hulst 2013). Even in English, with its highly lexicalized stress system, there are no true minimal stress pairs (Cutler 2015), since within a given dialect or register, stress contrasts always correlate with syntactic category (nominal *áddress* vs. verbal *addréss*). Similarly, in German, the apparent minimal stress pair *umfáhren* ‘bypass’ versus *úmfahren* ‘to knock over’ actually involves morphologically distinct affixes (respectively, inseparable vs. separable *um-*) (these examples were suggested by an anonymous reviewer, albeit for a different purpose). The explanation for this universal is that prosody is a grammar-governed regularity, and this regularity makes it unavailable for lexical contrast. If regular stroke lengthening in Chinese characters is an analog of stress, as argued in the introduction, then the paucity of stroke length contrasts can be ascribed to the same abstract grammatical principle.

Prosodic theory has already provided important insights into the structure and processing of writing systems. The correlation between phoneme type and lowercase Roman letter shape, as noted in the introduction, actually reflects sonority, which results in written syllables being roughly demarcated by “taller” letters (e.g., <or.tho.gra.phic>). The alternation between <y> and <i> in English spelling is motivated by a constraint against word-final <i> (<baby> vs. *<babi>), which interacts with constraints against overly short content words (<die> vs. *<dy>) and identical adjacent vowel letters (<dying> vs. *<diing>), all of which have parallels in prosodic phonology, with its restrictions on word-final segments, minimal word sizes, and onsetless syllables. Similar conclusions apply to the written analogs of syllables in scripts otherwise quite different from that of English, including Korean hangul and Maldivian Thaana (Gnanadesikan 2023). Evertz (2018) also provides experimental evidence for analogs of metrical feet in written English and German. Reinken (2023) extends these observations to handwritten German as well, noting among other things that letters tend to be less reduced in syllable onsets and more reduced in unstressed syllables. The present study adds to this literature by showing that an analog of prosody may affect handwriting even in a logographic script, independent of the prosody of the corresponding spoken language.

Such considerations suggest a possible answer to our earlier question about the nature of implicit written knowledge: if it is influenced by prosody, it may go beyond mere statistical learning of sensorimotor correlations. However, the status of prosody itself remains somewhat controversial even for spoken language; see, for example, the debates reviewed in Ladd (2008) about whether prosody is truly phonological. In any case, we cannot expect the first study on a hitherto neglected phenomenon to answer all the questions it may raise. Following Myers (2019), then, we merely report an intriguing similarity we have observed among writing, speech, and signing, and leave its full interpretation for future research.

4.4 Conclusions

Chinese readers and writers implicitly know a highly subtle feature correlation in character form that is not explicitly taught and has no direct communicative function. This implicit knowledge seems to be particularly active in handwriting, which requires richer mental representations than perception. While the pattern itself seems to be unique to Chinese script, it may reflect the same principles that govern the prosodic structures of spoken and signed languages.

Funding

Financial support for this study was provided by Taiwan’s National Science and Technology Council grant MOST109-2410-H-194-096-MY3.

Acknowledgements

This study was made possible by my lab assistants Wenchi Lin, Yangru Chen, Meichun Liu, and Yuchu Chang, as well as all of the experimental participants. A very early version of this study was presented at the 34th North American Conference on Chinese Linguistics. Tom Bishop provided essential help in understanding the inner workings of Wenlin. Suggestions on argumentation came from Tsungying Chen and two anonymous reviewers. I hereby absolve of all of the above for any responsibility for my misdeeds.

References

Albrow, Kenneth H. (1972). *The English writing system: Notes toward a description*. London: Longman.

Allopenna, Paul D., James S. Magnuson & Michael K. Tanenhaus (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language* 38.4: 419-439.

Beddor, Patrice Speeter, Kevin B. McGowan, Julie E. Boland, Andries W. Coetzee &

Anthony Brasher (2013). The time course of perception of coarticulation. *The Journal of the Acoustical Society of America* 133.4: 2350-2366.

Beech, John R. & Kate A. Mayall (2005). The word shape hypothesis re-examined: evidence for an external feature advantage in visual word recognition. *Journal of Research in Reading* 28.3: 302-319.

Bishop, Tom & Richard Cook (2007). A character description language for CJK. *Multilingual* 18.7: 62-68.

Chang, Ya-Ning, Chun-Hsien Hsu, Jie-Li Tsai, Chien-Liang Chen & Chia-Ying Lee (2016). A psycholinguistic database for traditional Chinese character naming. *Behavior Research Methods* 48: 112-122.

Chen, Jenn-Yeu & Rong-Ju Cherng (2013). The proximate unit in Chinese handwritten character production. *Frontiers in Psychology* 4: 517.

Chetail, Fabienne (2015). Reconsidering the role of orthographic redundancy in visual word recognition. *Frontiers in Psychology* 6: Article 645.

Chuang, Te-Ming & Hsien-Ying Teng (2009). *Research and development of Chinese characters information database and its application* [in Chinese]. Taipei: Academia Sinica.

Chung, Yeojin, Sophia Rabe-Hesketh, Vincent Dorie, Andrew Gelman & Jingchen Liu (2013). A nondegenerate penalized likelihood estimator for variance parameters in multilevel models. *Psychometrika* 78.4: 685-709.

Cohen Priva, Uriel (2015). Informativity affects consonant duration and deletion rates. *Laboratory Phonology* 6.2: 243-278.

Cutler, Anne (2015). Lexical stress in English pronunciation. In Marnie Reed & John M. Levis (eds.), *The handbook of English pronunciation*, 106-124. Hoboken, NJ: John Wiley & Sons.

Ernestus, Mirjam & Anne Cutler (2015). BALDEY: A database of auditory lexical decisions. *The Quarterly Journal of Experimental Psychology* 68.8: 1469-1488.

Evertz, Martin (2018). *Visual prosody: The graphematic foot in English and German*. Berlin: Walter de Gruyter.

Fiset, Daniel, Caroline Blais, Catherine Ethier-Majcher, Martin Arguin, Daniel Bub & Frédéric Gosselin (2008). Features for identification of uppercase and lowercase letters. *Psychological Science* 19.11: 1161-1168.

Fox, John & Sanford Weisberg (2019). *An R companion to applied regression*, third edition. Thousand Oaks CA: Sage.

Friedmann, Naama & Manar Haddad-Hanna (2012). Letter position dyslexia in Arabic: From form to position. *Behavioural Neurology* 25.3: 193-203.

Gnanadesikan, Amalia E. (2023). Segments and syllables in Hangeul and Thaana: A comparison and optimality theoretic analysis. *Written Language & Literacy* 26.2: 238-265.

Goedemans, Rob & Harry van der Hulst (2013). Fixed stress locations. In Matthew S. Dryer & Martin Haspelmath (eds.), WALS Online (v2020.4) [Data set]. Zenodo. Retrieved November 16, 2024, from https://wals.info/chapter/14

Goldsmith, John (2000). On information theory, entropy, and phonology in the 20th century. *Folia Linguistica* 34.1: 85-100.

Hall, Kathleen Currie, Elizabeth Hume, T. Florian Jaeger & Andrew Wedel (2018). The role of predictability in shaping phonological patterns. *Linguistics Vanguard* 4.s2: 20170027.

Hayes, Bruce (1995). *Metrical stress theory: Principles and case studies*. Chicago:

University of Chicago Press.

Hockett, Charles F. (1953). Review of *The mathematical theory of communication* by Claude L. Shannon and Warren Weaver. *Language* 29.1: 69-93.

Huang, Shuting, Yacong Zhou, Menglin Du, Ruiming Wang & Zhenguang G. Cai (2021). Character amnesia in Chinese handwriting: A mega-study analysis. *Language Sciences* 85: 101383.

Koshevoy, Alexey, Helena Miton & Olivier Morin (2023). Zipf's law of abbreviation holds for individual characters across a broad range of writing systems. *Cognition* 238: 105527.

Ladd, D. Robert. (2008). *Intonational phonology*, second edition. Cambridge, UK: Cambridge University Press.

Latash, Mark L. (2012). *Fundamentals of motor control*. London: Academic Press.

Lavidor, Michal (2011). Whole-word shape effect in dyslexia. *Journal of Research in Reading* 34.4: 443-454.

Law, Sam Po & Mak Tak Leung (2000). Structural representations of characters in Chinese writing: Evidence from a case of acquired dysgraphia. *Psychologia* 43: 67-83.

Lee, Sang-Oak (2009). The Korean alphabet: An optimal featural system with graphical ingenuity. *Written Language & Literacy* 12.2: 202-212.

Lidz, Jeffrey and Annie Gagliardi (2015). How nature meets nurture: Universal grammar and statistical learning. *Annual Review of Linguistics* 1.1: 333-353.

Lu, Yu-An & Sang-Im Lee-Kim (2021). The effect of linguistic experience on perceived vowel duration: Evidence from Taiwan Mandarin speakers. *Journal of Phonetics* 86: 101049.

Marslen-Wilson, William D. (1987). Functional parallelism in spoken word recognition. *Cognition* 25: 71-102.

Massaro, Dominic W. & Michael M. Cohen (1983). Phonological context in speech perception. *Perception & Psychophysics* 34.4: 338-348.

McCawley, James D. (1994). Some graphotactic constraints. In William C. Watt (ed.), *Writing systems and cognition: Perspectives from psychology, physiology, linguistics, and semiotics*, 115-127. Berlin: Springer.

Meletis, Dimitrios (2019). The grapheme as a universal basic unit of writing. *Writing Systems Research* 11.1: 26-49.

Meletis, Dimitrios (2020). Types of allography. *Open Linguistics* 6.1: 249-266.

Myers, James (2019). *The grammar of Chinese characters: Productive knowledge of formal patterns in an orthographic system*. Abingdon-on-Thames, UK: Routledge.

Myers, James (2021). Levels of structure within Chinese character constituents. In Yannis Haralambous (ed.), *Grapholinguistics in the 21st Century: From graphemes to knowledge, Part II*, 645-582. Paris: Fluxus Editions.

Myers, James (2023). Nonstatistical things that linguists can do with R: A review with a case study on Chinese handwriting. In Alex Cherici, Bihua Chen & Chien-Jer Charles Lin (eds.), *Proceedings of the 34th North American Conference on Chinese Linguistics* (NACCL-34). Department of East Asian Languages and Cultures, Indiana University Bloomington.

Myers, James (2024). Interactions among patterns in Chinese character form. In James Myers & James H-Y. Tai (eds.), *Looking at language from all side*s, 129-152. Taipei: Crane Publishing.

O'Shea, Robert P. (1991). Thumb's rule tested: Visual angle of thumb's width is about 2 deg. *Perception* 20: 415-418.

Ooms, Jeroen (2024). magick: Advanced Graphics and Image-Processing in R. R package version 2.8.4. Retrieved November 16, 2024, from https://CRAN.R-project.org/package=magick

Pacton, Sébastien, Pierre Perruchet, Michel Fayol & Axel Cleeremans (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General* 130.3: 401-426.

Pau, Grégoire, Florian Fuchs, Oleg Sklyar, Michael Boutros & Wolfgang Huber (2010). EBImage—an R package for image processing with applications to cellular phenotypes. *Bioinformatics* 26.7: 979-981.

Peirce, Jonathan, Rebecca Hirst & Michael MacAskill (2022). *Building experiments in PsychoPy*, 2nd edition. London: Sage.

Peirce, Jonathan, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman & Jonas Kristoffer Lindeløv (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* 51: 195-203.

Peng, Gang, James W. Minett & William S-Y. Wang (2010). Cultural background influences the liminal perception of Chinese characters: An ERP study. *Journal of Neurolinguistics* 23.4: 416-426.

Penney, Joshua, Felicity Cox & Anita Szakay (2021). Effects of glottalisation, preceding vowel duration, and coda closure duration on the perception of coda stop voicing. *Phonetica* 78.1: 29-63.

Perea, Manuel & Eva Rosa (2002). Does "whole-word shape" play a role in visual word recognition? *Perception & Psychophysics* 64: 785-794.

Primus, Beatrice (2004). A featural analysis of the Modern Roman Alphabet. *Written Language & Literacy* 7.2: 235-274.

Reinken, Niklas (2023). *Die Grammatik der Handschriften*. Heidelberg: Winter.

R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved November 16, 2024, from https://www.R-project.org/.

Reinisch, Eva, Alexandra Jesse & James M. McQueen (2011). Speaking rate affects the perception of duration as a suprasegmental lexical-stress cue. *Language and Speech* 54.2: 147-165.

Sumner, Meghan & Arthur G. Samuel (2005). Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language* 52: 322-338.

Sze, Wei Ping, Susan J. Rickard Liow & Melvin J. Yap (2014). The Chinese Lexicon Project: A repository of lexical decision behavioral responses for 2,500 Chinese characters. *Behavior Research Methods* 46: 263-273.

Taft, Marcus & Xiaoping Zhu (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23.3: 761-775.

Tomaschek, Fabian, Benjamin V. Tucker, Matteo Fasiolo & R. Harald Baayen (2018). Practice makes perfect: The consequences of lexical proficiency for articulation. *Linguistics Vanguard* 4.s2: 20170018.

Treiman, Rebecca & Brett Kessler (2022). Statistical learning in word reading and spelling across languages and writing systems. *Scientific Studies of Reading* 26.2: 139-149.

Tsai, Kuan-Chan & Terezinha Nunes (2003). The role of character schema in learning novel Chinese characters. In Hsuan-Chih Chen & Catherine McBride-Chang (eds.), *Reading development in Chinese children*, 109-125. Westport, CT: Praeger.

Van der Kooij, Els & Onno Crasborn (2008). Syllables and the word-prosodic system in Sign Language of the Netherlands. *Lingua* 118.9: 1307-1327.

van Rij, Jacolien (2015). Overview GAMM analysis of time series data. Retrieved November, 2024, from https://jacolienvanrij.com/Tutorials/GAMM.html

Venezky, Richard L. (1970). *The structure of English orthography*. The Hague: Mouton.

Wang, Jason Chia-Sheng (1983). *Toward a generative grammar of Chinese character structure and stroke order*. Ph.D. thesis, University of Wisconsin-Madison.

Warker, Jill A. & Gary S. Dell (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32.2: 387-398.

Watt, William C. (1975). What is the proper characterization of the alphabet? I: Desiderata. *Visible Language* 9.4: 293-327.

Winter, Bodo & Martijn Wieling (2016). How to analyze linguistic change using mixed models, growth curve analysis and generalized additive modeling. *Journal of Language Evolution* 1.1: 7-18.

Wong, Kimberly, Frempongma Wadee, Gali Ellenblum & Michael McCloskey (2018). The devil's in the g-tails: Deficient letter-shape knowledge and awareness despite massive visual experience. *Journal of Experimental Psychology: Human Perception and Performance* 44.9: 1324-1335.

Wood, Simon N. (2017) *Generalized additive models: An introduction with R* (2nd edition). Chapman and Hall/CRC Press.

Wu, Yan, Deyuan Mo, Yiu-Kei Tsang & Hsuan-Chih Chen (2012). ERPs reveal sub-lexical processing in Chinese character recognition. *Neuroscience Letters* 514.2: 164-168.

Yang, Ruoxiao & William Shi Yuan Wang (2018). Categorical perception of Chinese characters by simplified and traditional Chinese readers. *Reading and Writing* 31.5: 1133-1154.

Yeh, Su-Ling, Jing-Ling Li, Tatsuto Takeuchi, Vincent Sun & Wen-Ren Liu (2003). The role of learning experience on the perceptual organization of Chinese characters. *Visual Cognition* 10.6: 729-764.