A Meta-Analysis on the N1 Print Tuning Effect in Early and Late N1 Time Window

Alex X. Ye[1], Kexin Huang[1], Urs Maurer[2,3,4]

[1] *Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, , Hong Kong, China*

[2] *Department of Psychology, The Chinese University of Hong Kong, Hong Kong, China*

[3] *Centre for Developmental Psychology, The Chinese University of Hong Kong, Hong Kong, China*

[4] *Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong, China*

Introduction

In electroencephalography (EEG), N1 print tuning is indicated by greater negativity for words than symbols or false fonts at occipito-temporal electrodes at around 120-200 ms after the stimulus onset (Maurer et al., 2006; Xue, 2008; Brem et al., 2010; Wang and Maurer, 2020). It is associated with reading expertise since longitudinal studies found an increased N1 print tuning effect in children as they acquire reading skills (Maurer et al., 2006).

Previous studies that investigated this effect, however, reported variation in how print tuning unfolds in time within the N1 component. To look into the temporal details, previous studies used two methods to analyze their data: subdivision of the N1 based on Global Field Power (GFP) and peak latency analysis. By identifying the minima and maxima of the grand average GFP curves, some studies segmented the N1 component into early (minima of N1 onset to peak) and late (peak to minima of N1 offset) parts (Wang and Maurer, 2017; Wang and Maurer, 2020). While the early N1 consistently presented the N1 print tuning effect, a study combining EEG and fMRI reported more pronounced difference between word and symbol conditions (Brem, 2006), and another study reported weaker effect compared to that occurred in early N1 (Eberhard-Moscicka at al., 2016). Peak latency analyses also had mixed findings: some studies found a consistent earlier peak for word condition and later peak for control condition (Shirahama, 2004; Wang and Maurer, 2017; Wang and Maurer, 2020), other studies did not find a significant difference between the two (Eulitz, 2000; Eberhard-Moscicka et al., 2016), even though print tuning was larger in the early than late N1 in one of these studies (Eberhard-Moscicka et al., 2016).

A further look at previous studies suggested that for studies that did not find significant latency difference (Eulitz, 2000; Eberhard-Moscicka et al., 2016), they used false-fonts instead of symbols in previous studies and kept using alphabetic words in the word condition. As false-fonts might share more visual similarity with real words when compared with symbols, the findings might suggest that visual similarity between word and control conditions could potentially influence the N1 print tuning effect.

This influence, however, might also be due to the writing system, as studies that found inconsistent results used word from alphabetic writing system (Eulitz, 2000; Brem, 2006; Eberhard-Moscicka et al., 2016), while studies using stimuli in logo-graphic system found consistent difference in early and late N1 (Wang and Maurer, 2017; Wang and Maurer, 2020).

To understand the time course of tuning tuning it would be important to systematically investigate the influence of visual similarity on the difference of the N1 in response to words and the N1 in response to visual control stimuli by contrasting the early and late stage of N1.

This meta-analysis study hypothesizes that: 1) Across the selected studies, there is a variation in the processing time window of N1 print tuning effect, and 2) visual similarity shows a moderation effect on such a difference. If the visual similarity does not show a significant moderation effect, an alternative hypothesis is that the writing system might serve as a moderator.

Methods

We conducted a meta-analysis of ERP print tuning studies using visual complexity of the stimuli as a moderator variable. As we needed to partially reanalyze the ERP data for the N1 onset-offset analysis and extract visual complexity measures from the stimulus material, we could only include those studies, where we had access to the EEG data and the stimulus material therefore restricting the studies to those conducted by the corresponding author.

After selecting the studies and the stimulus material, an image analysis to quantify the visual similarity of stimuli used in the experiments was conducted based on Chang et al.'s quantification of visual complexity, where they used 4 parameters to measure the visual complexity (Chang et al., 2015). By adding an additional parameter, object number, the analysis used 5 parameters to quantify visual complexity: perimetric complexity (PC), object number, disconnected stroke, stroke sum, and junctions. Object number is added for that it is more tailored to measure visual complexity of logographic scripts such as Chinese. The visual similarity is defined as the difference between the visual complexity of the stimuli, and a smaller difference between conditions indicates greater visual similarity. This visual similarity index is later added into the meta-analysis as a factor to test whether it shows moderation effect.

This meta-analysis is based on the EEG data collected from 7 previous studies. Specifically, these studies used German, English, Chinese as stimuli for the familiar word condition, while symbols, false-fonts and Korean were used for unfamiliar word or symbol condition. We selected these studies because we had access to the data through the corresponding author of this paper, Dr. Urs Maurer.

The data was preprocessed separately for each study. After artifact rejection, offline filtering, baseline correction, selecting target epoch, and ICA (Maurer & Brandeis et al., 2005; Maurer & McCandliss, 2005; Maurer & McCandliss, 2008; Eberhard-Moscicka et al., 2016; Wang & Maurer, 2017), the GFP was calculated for each condition within each study individually. The N1 segmentation was conducted based on the same criteria as previous studies. The raw data collected for this meta-analysis contains the average amplitude within early N1 (minima of N1 onset to peak) or late N1 (peak to minima of N1 offset) of a specific participant in each condition of the same task and different electrodes.

As previous studies found that the left hemisphere in temporo-occipital lobe showed the most robust N1 print tuning effect (Maurer et al., 2005; Maurer et al., 2006; Xue, 2008; Brem et al., 2010; Wang and Maurer, 2020), this study limited electrode selection to the left hemisphere. Finally, the electrode in the left hemisphere with the largest difference in amplitudes between word and control conditions was selected. Among included studies, the first study used a different EEG montage from others. To unify all studies, reference is made to a technical document from EGI (Luu & Ferree, 2005) to help identify which electrodes are located in nearest locations in different systems. For the first study, the electrode showing the most robust N1 print tuning effect is PO9, which is the closest to E64 in the other system.

After selecting the target electrode, the average amplitude across all participants for specific electrode (PO9 or E64), task type (repetition detection or script decision) was calculated for each time window (N1 onset and N1 offset), word condition (word condition and control condition) separately for each study. The standard deviation was calculated in a similar manner.

The 5 parameters used in this study were defined as the following. As Watson (2011) suggested, PC is a shape descriptor that quantifies the complexity of an object's perimeter relative to its area. The image was first loaded using the Python Imaging Library (PIL) and converted to a grayscale representation, then calculated through the following steps:

Perimeter Extraction: Inside the image, the object's perimeter in the image was determined through a series of operations involving the application of the ImageChops module from PIL.

Ink Area Calculation: The ink area of the image was computed by inverting the original image, summing the intensities of all pixels, and normalizing the values by dividing them by 255.

Perimetric Complexity Calculation: Based on the calculated perimeter and ink area, the perimetric complexity was derived using the equation [PC_equa].

$$PC = \left(perimeter^2\right)/(inkArea * 4 * \pi)$$

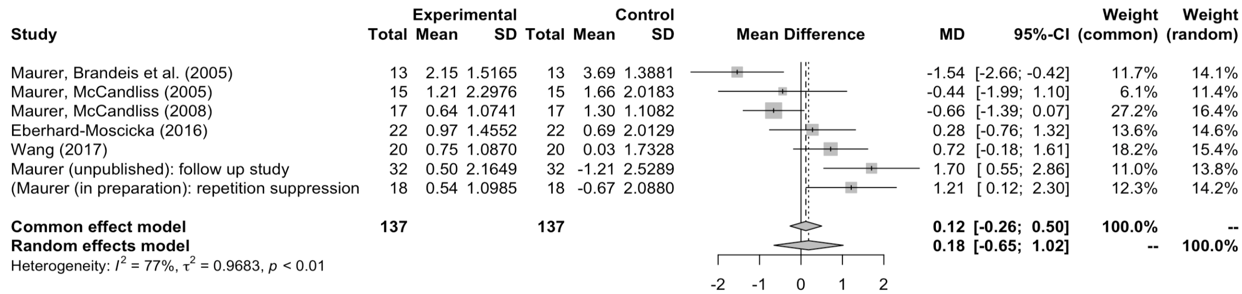
This method facilitates the quantification of an image's perimetric complexity, serving as a valuable shape descriptor for a variety of image analysis applications.

For the stroke sum (SS), disconnected strokes (DS), and junctions (JC), we counted based on the images of each stimulus presented in the experiment. Here, instead of the definition of stroke in Chinese, we defined strokes according to standards used in alphabetic languages and Korean: a mark in one direction. Therefore, we re-calculated the stroke sum of Chinese characters according to this standard.

For object number (ON), we define an object as a unit of several connected strokes that is not connected to any other stroke. For example, the character "指"has 3 objects. Specifically, if a single stroke is not connected to any other stroke, it is a disconnected stroke, not an object. By this, we developed an image analysis algorithm to help us find the number of objects in each stimulus automatically.

After collecting the data of all 5 parameters from each stimulus, we averaged each condition, normalized across all conditions, and compared the difference between word and control conditions. These 5 parameters represent different aspects of visual complexity in stimuli, and they were added individually or together into the meta-analysis as moderators, to test whether they showed moderation effect in the EEG data. To investigate which parameters influenced the N1 print tuning effect, a stepwise multiple meta-regression was conducted and a model that described the EEG data the best was chosen. Step-wise meta-regression with forward selection was used to explore sources of heterogeneity among studies. This approach began with an empty model and sequentially added predictor variables based on their statistical significance, seeking to create a model that best explained the variations in effect sizes across studies.

Results



According to the random-effect model shown in figure 1, no significant general effect size ((word_late N1 - symbol_late N1) - (word_early N1 - symbol_early N1)) was found in the k = 7 studies. The Hedges' g was -0.2427, and the p-value was 0.6660. Significant heterogeneity was shown as $I^2 = 77.1\%$. Figure 1 shows individual study and overall effect size, along with the 95% confidence interval.

Forest plot comparing early and late N1 print tuning effect

The Egger's regression-based test (Egger et al., 1997) did not show the overall publication bias (z = 0.0977, p = 0.9222). In addition, the subgroup analysis displayed much larger p-value within the subgroups ($z_{alphabetic} = 0.0187$, $p_{alphabetic} = 0.9851$ and $z_{logographic} = 0.9299$, $p_{logographic} = 0.3524$), which further proves the symmetries of the selected papers.

	Model 1		Model 2		Model 3		Model 4		Model 5	
	estimate	p value	estimate	p value	estimate	p value	estimate	p value	estimate	p value
PC	2.6583	0.7240	-9.2747	0.2875	-30.8616	0.0033	-30.6845	0.0075	-42.0121	0.0211
ON			8.7323	0.0621	20.0045	0.0004	20.0073	0.0004	21.1050	0.0003
DS					-10.1781	0.0117	-10.1129	0.0210	-8.0435	0.1139
SS							-0.2076	0.9695	-0.3908	0.9427
JC									-8.9467	0.4235
R^2	0.00%		37.65%		100.00%		100.00%		100.00%	

PC: Perimetric Complexity, ON: Object Number, DS: Disconnected Stroke, SS: Stroke Sum, JC: Junction.

Table 1: Stepwise Random-Effect Moderator Analysis

As shown in table [table1], step-wise random-effect moderator analysis (k = 7; estimator: restricted maximum-likelihood estimator) suggested that Model 3 accounted for all variability with least parameters. Therefore it was the model that described the data the best. The parameters in this model were PC, ON, and DS.

The random-effect moderator analysis (k = 7; estimator: restricted maximum-likelihood estimator) showed a significant result (QM(df = 2) = 10.8292, p = 0.0045), while the test for residual heterogeneity was not significant (QE (df = 4) = 4.1641, p = 0.3842).

Additionally, we did a multiple meta-regression analysis involving writing system, and it only raised the p-value rather than lowering it. However, when writing system served as the only moderator, the random-effect model showed a significant moderation effect (QM(df = 1) = 10.0814, p = 0.0015), while the test for residual heterogeneity was not significant (QE(df = 5) = 3.9813, p = 0.5521).

The post-hoc analysis compared word and control conditions separately in the early and late N1 time windows. In the early N1, the random-effect model showed a significant general effect size (MD = -1.0816, p = 0.0007) and no significant heterogeneity (QE (df = 6) = 3.17, p = 0.7871). In the late N1, the random-effect model did not show a significant general effect size (MD = -0.8259, p = 0.1732), but significant heterogeneity was present (QE (df = 6) = 23.24, p = 0.0007). As shown in table [table2] The post-hoc meta-regression found a significant correlation between the difference among studies in the late N1 and 3 visual complexity parameters: PC, ON, and DS.

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-0.4947	0.5010	-0.9875	0.3234	-1.4766	0.4872

	estimate	se	zval	pval	ci.lb	ci.ub
PC	-42.3097	14.9328	-2.8333	0.0046	-71.5773	-13.0420
obj_num	27.4348	7.6107	3.6048	0.0003	12.5181	42.3515
disc_strk	-15.3988	5.7660	-2.6706	0.0076	-26.6999	-4.0978

Conclusion

Effect Sizes

The random-effect model showed significant heterogeneity, supporting our hypothesis that there is a difference in the N1 print tuning effect processing time window among the selected studies. Study 1-3 had a more pronounced early N1 print tuning effect when compared with study 4-7, where no obvious difference in N1 print tuning effect between early and late N1 was observed. In other words, the N1 print tuning effect started similarly for all studies, but lasted longer and was still significant in the late N1 in study 1-3, whereas in study 4-7 the effect was shorter. Post-hoc analysis confirmed this finding.

Interestingly, this finding echoes with the results from previous studies as studies using symbols in control conditions found more pronounced word vs. control condition difference in a later time window (Shirahama, 2003; Brem, 2006), while those using false-fonts or unfamiliar words only found an earlier difference (Eberhard-Moscicka et al., 2016; Wang & Maurer, 2017).

Moderation Effect

The step-wise meta-regression analysis revealed that the 3-parameter model (PC, ON, and DS) best fits the EEG data. This suggests that visual similarity in these aspects of stimuli affects how our brain processes reading. PC and DS are negatively correlated, while ON is positively correlated with the EEG data. This means that the increase in PC and DS shifts the N1 print tuning effect towards early N1, while the increase in ON shifts it towards late N1.

Limitation

We did not examine the difference between alphabetic and logographic groups in our analysis.

No previous studies have done analysis of visual complexity in objects. More detailed investigation is required to better understand how it influences perception of logographic languages.

Reference

[1] Maurer, U., Brandeis, D., McCandliss, B.D., 2006. Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behav. Brain Funct*. 1 (1), 13. https://doi.org/10.1186/1744-9081-1-13.

[2] Xue, G., Jiang, T., Chen, C., & Dong, Q. (2008). Language experience shapes early electrophysiological responses to visual stimuli: The effects of writing system, stimulus length, and presentation duration. *NeuroImage*, 39(4), 2025–2037. https://doi.org/10.1016/j.neuroimage.2007.10.021

[3] Brem, S. Brem S, Bach S, Kucian K, Guttorm TK, Martin E, Lyytinen H, Brandeis D, Richardson U.(2010). Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proc Natl Acad Sci USA* 107 (17): 7939-44.

[4] Wang, F., & Maurer, U. (2020). Interaction of top-down category-level expectation and bottom-up sensory input in early stages of visual-orthographic processing. *Neuropsychologia.*

[5] Wang, F., & Maurer, U. (2017). Top-down modulation of early print-tuned neural activity in reading. *Neuropsychologia*, 102, 29–38. https://doi.org/10.1016/j.neuropsychologia.2017.05.028 137(December 2019). https://doi.org/10.1016/j.neuropsychologia.2019.107299

[6] Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E., & Brandeis, D. (2006). Evidence for developmental changes in the visual word processing network beyond adolescence. *NeuroImage*, 29(3), 822–837. https://doi.org/10.1016/j.neuroimage.2005.09.023

[7] Eberhard-Moscicka, A. K., Jost, L. B., Fehlbaum, L. v., Pfenninger, S. E., & Maurer, U. (2016). Temporal dynamics of early visual word processing – Early versus late N1 sensitivity in children and adults. *Neuropsychologia*, 91, 509–518. https://doi.org/10.1016/j.neuropsychologia.2016.09.014

[8] Shirahama, Y., Ohta, K., Takashima, A., Matsushima, E., & Okubo, Y. (n.d.). Magnetic brain activity elicited by visually presented symbols and Japanese characters. https://doi.org/10.1097/01.wnr.0000122485.43641.4b

[9] Eulitz, C., Eulitz, H., Maess, B., Cohen, R., Pantev, C., & Elbert, T. (2000). Magnetic brain activity evoked and induced by visually presented words and nonverbal stimuli. Psychophysiology, 37(4), 447-455.

[10] Chang, L. Y., Plaut, D. C., & Perfetti, C. A. (2016). Visual complexity in orthographic learning: Modeling learning across writing system variations. *Scientific Studies of Reading*, *20*(1), 64–85. https://doi.org/10.1080/10888438.2015.1104688

[11] Maurer, U., Brem, S., Bucher, K., Brandeis, D., 2005. Emerging neurophysiological specialization for letter strings. J. Cogn. *Neurosci*. 17 (10), 1532–1552. https://doi. org/10.1162/089892905774597218.

[12] Luu, P., & Ferree, T. (2005). Determination of the HydroCel Geodesic Sensor Nets' average electrode positions and their 10–10 international equivalents. *Inc, Technical Note*, 1-11.
MATLAB. (2010). *version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc.