What's in a Name? Electrophysiological Differences in Processing Proper Nouns in Mandarin Chinese

Bernard Jap[1], Yu-Yin Hsu[2], Lavinia Salicchi[2], and Yuxi Li[2]
University of Macau[1], The Hong Kong Polytechnic University[2]
bernardjap@um.edu.mo, yyhsu@polyu.edu.hk, {lavinia.salicchi, yu-xi.li}@connect.polyu.hk

Abstract

The current study examines how proper names and common nouns in Chinese are cognitively processed during sentence comprehension. EEG data was recorded when participants were presented with neutral contexts followed by either a proper name or a common noun. Proper names in Chinese often consist of characters that can function independently as words or be combined with other characters to form words, potentially benefiting from the semantic features carried by each character. Using cluster-based permutation tests, we found a larger N400 for common nouns when compared to proper names. Our results suggest that the semantics of characters do play a role in facilitating the processing of proper names. This is consistent with previous behavioral findings on noun processing in Chinese, indicating that common nouns require more cognitive resources to process than proper names. Moreover, our results suggest that proper names are processed differently between alphabetic languages and Chinese language.

Keywords: proper names, common nouns, ERP, Mandarin Chinese

1. Introduction

A longstanding philosophical and linguistic debate concerns the definition and distinction between proper names (PNs) and common nouns (CNs). CNs are widely accepted as words that denote *classes* of real-world concrete and abstract entities, while PNs specify a particular *individual* entity within a class (Yasuda et al., 2000). Their difference is clear in clinical settings and daily life. For example, aphasic patients may struggle to recall either CNs or PNs (Warrington and McCarthy, 1987). Systematic difficulties in recalling PNs has also been related to early stages of Alzheimer's disease (Mueller et al., 2020; Semenza et al., 2003). These findings suggest that PNs pose more processing challenges (Brown, 1991). Adorni et al. (2014) propose that one key processing difference between the two categories is the relationship between a word and its reference. CNs are associated with a wider range of real-world entities and contain richer semantic features, whereas PNs have a direct connection to individual entities and are thus mostly associated with episodic memory.

Nonetheless, most of this research on PNs and CNs has been conducted in languages that use the Latin alphabet, such as English and Italian. These languages follow precise orthographic rules to distinguish words in the two categories, i.e., capitalizing the first letter for PNs (Sulpizio and Job, 2018). Considering this, we think Chinese language, a typologically distinct language, provides a useful context for further understanding the differences in processing PN and CNs. Chinese writing is mostly logographic, offering no orthographic or typographical cues to distinguish nouns' subcategories. Moreover, Chinese characters do not exclusively represent sounds; they also convey the semantics of the concepts they symbolize. For example, the two characters in the name 老周(Lao Zhou), 老and 周, can each be used independently as a standalone word, e.g., 'old' and 'week', respectively, while they may have other meanings in specific contexts. Even in PNs where the characters are not independent words, the fact that these characters are used in other words and may preserve some senses from those words potentially makes the processing of Chinese PNs different from that of PNs in alphabetic languages, in which individual letters of a PN (like *John*) do not have semantics.

Nonetheless, previous studies on processing Chinese PN and CNs, adopting either ERP (Wang et al., 2016) or behavioral method (Yen and Müller, 2003), do not reach a consensus. The question remains as to whether a language that does not impose orthographic constraints in distinguishing PNs from CNs, like Chinese, exhibits different processing patterns during the language comprehension of CNs and PNs, in which the latter type also contains semantically meaningful characters. Therefore, to address this question, we conducted an ERP experiment to investigate the processing of PNs and CNs in Mandarin during language comprehension.

2. Related Work

Lu and Bai (2023) examined whether CNs and PNs are processed differently in the left and right hemispheres by Chinese speakers. Their results sug-

The Workshop on Cognitive Aspects of the Lexicon, pages 79–85
20 May, 2024. © 2024 ELRA Language Resource Association: CC BY-NC 4.0

gest a lateralization of CN processing, while PNs did not show the same hemispheric advantage. However, these findings are partially disproved by Desai et al. (2023), which showed that in fMRI both PNs and CNs activated a wide network, across both hemispheres, with several overlapped active areas. These differences concerned the level of activation of these areas, with PNs leading to greater involvement of the right hemisphere.

Further proof of the processing differences between PNs and CNs can be found in EEG studies. Dehaene (1995) tested the neural correlates of five sub-categories of nouns and found a stronger N400 in temporal regions associated with PNs. Adorni et al. (2014) recorded in a lexical decision task the same late negativity with a P300 linked, again, to PNs only. Early correlates were recorded as well: in Müller and Kutas (1996) and Proverbio et al. (2001) a stronger P200 and N100 in the left anterior temporal and left fronto-central cerebral areas were found in association with PNs. Proverbio et al. (2009) focused on the evaluation of pairs of words being either name/surname or compounds of CNs, both existent and nonexistent while reaction time (RT) and brain activities were recorded. In contrast with previous studies, it revealed longer RTs and a stronger N400 in association with CNs.

Most previous research focused on PNs and CNs has been conducted in languages that use the Latin alphabet and require capital first letter for PNs. Sulpizio and Job (2018) studied the influence of orthographic variations on noun processing and found that N100 and P200 are associated with early processing of the form-category typicality. This indicates that studies on alphabet-based languages are heavily influenced by orthographic rules.

Wang et al. (2016) presented to Chinese participants PNs and CNs that were (in-)congruent with a given context. The ERP analysis showed that the N400 elicited by incoherent sentences was stronger in front of PNs, especially in the left hemisphere. Incongruent CNs, however, led to a stronger P600, suggesting a later, and more challenging, processing. Finally, Yen and Müller (2003), a behavioral study on Chinese nouns, found that CNs were harder to process than PNs, leading to longer RTs.

Generally speaking, there is no consensus on the processing of CNs and PNs, and most of them studied nouns independent of contexts. Previous studies do not agree on the brain networks involved in processing the two noun categories, whether or not lateralization takes place, what the neural correlates associated with PNs and CNs are, and the timing of their process. However, PNs seem to be more challenging in alphabetic languages.

In this study, we examined the processing patterns of PN and CN in Mandarin sentence comprehension, to simulate real-world language processing as both noun types are often encountered in reading or listening. We hypothesize that if semantic access to PNs through individual characters indeed facilitates their reading, processing PNs should require similar or even less cognitive effort compared to CNs, as suggested by Yen and Müller (2003). Specifically, we examine early potentials (N100, P200), and the N400. Were there no observable ERPs *OR* observable N100, P200, or N400 for the CN stimuli, it can be interpreted as evidence for semantic facilitation in processing Chinese PNs. Conversely, if observable N100, P200, or N400 for the PN stimuli occur, it would suggest that Chinese PNs are processed similarly to PNs in Indo-European languages, where the sub-components typically do not correspond to a lexical entry.

Table 1: An example set of the target nouns. *Note: PN means PN, and AN means animate nouns.*

	Context sentence	Target
PN	在学校组织的郊游途中， 'In a trip organized by the school,'	小婷.... 'Xiaoting'
AN	在学校组织的郊游途中， 'In a trip organized by the school,'	妹妹.... 'sister'

3. Method

3.1. Stimuli

As each experimental item started with a neutral sentence context, we divided nouns into 24 PNs and animate 24 CNs, with animacy effect in control. Before the experiment, a naturalness judgment task of the experimental items was conducted on a five-point Likert scale by 30 native speakers of Mandarin who did not participate in the experiment. All items used in this study were rated as 3 points or above in the judgment task. While we did not set a specific parameter for the selection of PNs and CNs, all PNs and CNs are two-character Chinese words/names, and their linguistic characteristics are delineated in subsection 3.3. Each set of experimental sentences (Table 1) involves two types of target nouns serving as the subject of a sentence and was introduced by the same context sentence in each set. Sentences were pseudorandomized and organized into two sets; each set had 48 trials and 72 fillers. The materials were presented in simplified Chinese characters, and word-by-word, each for 600ms, with a 500ms blank screen between each word, except for the context sentence displayed as a unit for 2000ms. Digital triggers were manually inserted at the relevant time point in every sentence, which is at the onset of the noun.

3.2. EEG data

3.2.1. Participants and Procedure

47 adult native speakers (mean age 22.21 $\pm$ 2.35, 26 female) of Mandarin Chinese from the Northern provinces of China participated in our study. Data of 9 participants were excluded due to low trial counts that remain after artifact rejection; thus, 38 participants' data were used for analyses.

In each experiment, participants were seated in front of a monitor presented with the sentences using E-Prime. The monitor showed written instructions that were explained orally by the experimenter. Participants were instructed to minimize head movements and keep their eyes open during the experiment, but blinking was allowed. During the experiment, a fixation cross was shown between trials and sets. Each trial began with a 500ms blank screen, followed by the phrase "准备好了吗？" (ready?)" shown on the screen until participants pressed any key. To prevent fatigue, there were breaks after every block (10 blocks in total), allowing the participants to read at their own pace. After every 3-8 trials (randomized), a comprehension prompt was given to ensure that the participants remained focused and to provide a measure of comprehension performance. Each session lasted about one hour, including cap and electrode preparation. Participants received US$25.

3.2.2. Measurements and Preprocessing

The participants' EEGs were recorded using a 64-channel and then preprocessed using EEGLAB (Delorme and Makeig, 2004). FieldTrip (Oostenveld et al., 2011) was used for statistical analysis. The EEG data was re-referenced to the two mastoid electrodes, and bad channels were interpolated. The remaining data was filtered using a 0.1Hz high-pass filter. ERPs were calculated for each participant, electrode, and condition in an interval from 200ms before onset to 1000ms after onset for each time-locked trigger. These epochs were then demeaned per channel and subjected to independent component analysis. Components associated with blinks, saccades, and muscle artifacts were removed. After this step, baseline correction was applied to the data using a 200 ms pre-stimulus onset baseline. Then, a threshold rejection function was used to detect and reject artifacts. Finally, data was filtered using a 40.0Hz low-pass filter.

3.2.3. Statistical analysis

For the statistical analysis, we used cluster-based permutation tests (Maris and Oostenveld, 2007) on all the scalp electrodes and the specific time window within the epoch. The tests compared the PN and CN ERPs at each channel and each sample, identifying clusters of spatiotemporally adjacent data points where the difference between the two conditions exceeds a threshold of $p < .1$ in a t-test for that time window. Given the previous ERP results, we conducted two cluster tests to capture potential differences between PNs and CNs: (1) a two-tailed test on 0-300ms to measure early potentials that have been reported, specifically N100 and P200. (2) a two-tailed test on 300-500ms to measure the N400, which has been the most common ERP observed in the study of PN processing. In (2) we ran a two-tailed test despite testing for an ERP of negative polarity because the direction of the effect is unclear, as we reviewed in Section 2.

3.3. Extraction of linguistic features

For each target noun and its initial and final characters, we extracted four types of linguistic features: frequency, stroke counts, word status of the characters, and finally the orthographic neighbor density of the noun. The purpose of extracting these features is to better interpret the ERP results in light of the differences in linguistic features between CN and PN: were we to find effects resembling, for example, a frequency effect in a direction that fits the word frequency profile of the stimuli, we would be equipped to avoid misinterpreting specific ERPs as waveforms unique to processing that is associated with common or proper nouns.

We calculated **frequencies** based on the corpus 'Chinese Web 2017 (zhTenTen17) Simplified' in Sketch Engine (Kilgarriff et al., 2014). We then normalized the frequencies on the basis of 10,000 words. We retrieved the **stroke number** of characters from hanziDB [1].The **word status** of characters was determined by two of the authors of this study who are native speakers of Mandarin, and specialized in Chinese linguistics. The criteria were the characters' meaningfulness and independence. Following Xiong et al. (2021), we calculated the **orthographic neighbors**, and its **density** (as the ratio of orthographic neighbors to the number of total word types). Table 2 (Appendix A) shows the means for these features.

Word frequency and visual complexity have been shown to impact early ERPs. A meta-study of 1100 English words and pseudowords by Dufau et al. (2015) observed frequency effects starting from 100ms, while visual complexity affects begin as early as 30-50ms, with another time window at 100-150ms also showing such effect. In the case of Chinese characters, visual complexity may be assessed by stroke count of a word or a character. Characters with more strokes elicited larger P200 and smaller N200, and similar levels of N400

[1] http://hanzidb.org/

(Yang et al., 2016). Similarly, words with many orthographic neighbors generate larger N400 than words with fewer ones (Müller et al., 2010).

In our stimuli, the word frequency for CNs is much higher than for PNs, which is expected when extracting generic names from corpora. However, PNs have higher individual character frequency. The stroke count is comparable, with the first character of PNs having a lower stroke count. The orthographic neighborhood shows that PN stimuli have a higher orthographic neighborhood compared to the CN stimuli. There is strong evidence that factors like frequency are task-dependent (Fischer-Baum et al., 2014), and not all tasks will elicit ERPs that correspond to these factors. Therefore, it is possible that these effects will not be observed in our study, despite the items not being perfectly balanced for these factors.

4. Results

Figure 1 shows grand average waveforms at electrodes starting from the baseline (200ms before the stimuli onset) to 1000ms. The contrast between CNs and PNs generated a clear negative shift at around the time window of the N400, peaking at almost precisely 400ms. Additionally, a minor contrast seems to occur at around the 250ms mark with the PN condition having a higher peak amplitude.

Specific channels were chosen to present the results when the channels corresponded to significant differences in the cluster test (refer to the raster plot in Appendix B). Figure 2 shows topographic maps for the waveforms of CN minus PN between 300ms and 500ms. The color bars on the right side of the topographic plots show the amplitude of the channel from -2μV (blue) to 2μV (yellow). The effect's topographical distribution intensifies between 350-450ms, with the negativity ranging from centroparietal to occipital channels. The ERP is centered in the midline channels, distributed along both sides of the hemisphere. The latency at around 300-500ms, the peak of the ERP at around 400ms, and the central to posterior distribution of the effect suggest that this is an N400. Although the topography seems to be distributed in a slightly more posterior area compared to the typical centroparietal N400, significant differences were found at the centroparietal channels and even some of the central electrodes (Figures 1-3).

Statistical analyses confirmed these observations. In the early potentials time window (0-300ms) for the N100 and P200, we find no effect, but a significant (p=.028) effect at the N400 time window (300-500ms) across 28 channels from central to occipital electrodes with the negative shift for CNs starting from 332ms to 484ms. We also found the clusters in the permutation test for the 300-500ms time window to correspond with the ERP and topographic plots (see Figure 3 in Appendix B). Our results show an N400 with a mainly central to posterior distribution for CNs when compared to PNs.

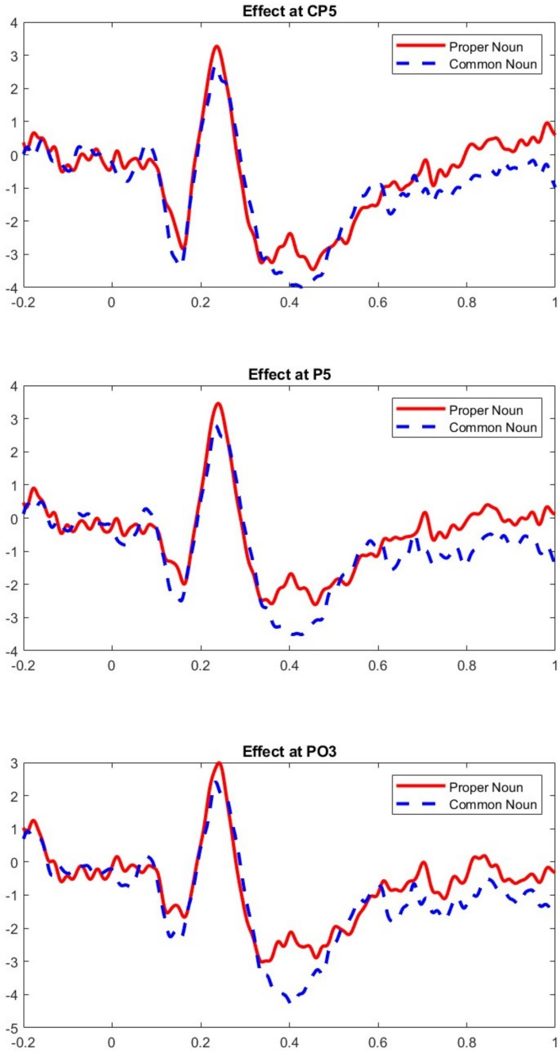

Figure 1: ERPs at three individual channels.

5. Discussion and Conclusion

The representational model proposed by Cohen (1990) postulates that the difficulty in retrieving PNs stems from processing rather than storage, as names are typically semantically neutral and offer little semantic clues for retrieval. However, this may not be fully apply to all languages. In Chinese, many names' characters can function as a standalone word. In our stimuli, the majority of the PNs (83 percent) contain such characters. Thus, Chinese readers are likely to use the semantic information of such characters to facilitate PN processing. However, if this is the case, we should expect *no* differences between PNs and CNs, since both types allow semantic facilitation in Chinese. This

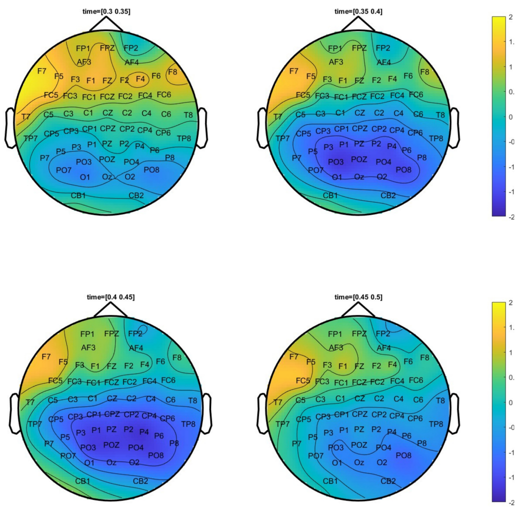

Figure 2: Topographic plot for CNs - PNs from 300ms to 500ms in 50ms intervals.

raises the question of why the differences still exist.

Our results differ from the ERP study by Wang et al. (2016) that reported larger N400s for incongruent PNs compared to CNs; yet, it is important to note that the differences in that study were elicited from a violation-based design, making comparisons between the two studies inappropriate. While limited evidence directly supports the idea that each of the various factors (e.g., frequency and imageability) facilitates the processing of either type of nouns, some studies suggest that most PNs' lack of semantic features is the primary factor to PNs' increased processing efforts (see Adorni et al. (2014) for a summary). Our finding is consistent with results in Yen and Müller (2003), showing that CNs are more difficult than PNs.There are several possibilities as to why PNs can be easier to process than CNs in Chinese: a) PNs are usually more imageable than CNs (Proverbio et al., 2009); b) PNs may evoke more emotional and sensory activations than CNs (Gorno-Tempini et al., 1998; Douville et al., 2005) with the left temporal cortex playing an important role in PN retrieval; c) the retrieval of PNs generates visual representation in brain areas involved in processing of visual images, even when not required by the tasks (Campanella et al., 2001), and in this case, visual imagery generation may assist in the processing of PNs.

Limitations

The main limitation is that there are linguistic characteristic differences between the two noun groups: compared to CNs, the PNs in the stimuli have a lower frequency and more orthographic neighbors, but the two groups have somewhat comparable stroke counts. In this case, according to previous research, we would expect to find early effects from 100ms (Dufau et al., 2015) which corresponds to low frequency, and a larger N400 amplitude (Müller et al., 2010) which corresponds to higher orthographic density for the PN. Nonetheless, in our data, we did not find ERPs that correspond with word frequency and neighborhood density. We did not find any early potentials, and instead, we observed an N400 for the CNs, which are words with higher frequency. As such, it is rather unlikely that the effects found in the analysis are driven by differences in the listed linguistic features, given that their corresponding ERP effects are not present.

Acknowledgment

This study was approved by the University's Ethics Review Board (HSEARS20211223003), and was supported by the Departmental General Research Fund (4-ZZRX) funded by the Hong Kong Polytechnic University. We would like to thank the anonymous reviewers for their feedback and suggestions. Special thanks to Deran Kong, Wenxi Fei, and Zhihong Chen for assisting in part of the experimental preparation.

Bibliographical References

Roberta Adorni, Mirella Manfredi, and Alice Mado Proverbio. 2014. Electro-cortical manifestations of common vs. proper name processing during reading. 135:1–8.

Alan S Brown. 1991. A review of the tip-of-the-tongue experience. *Psychological bulletin*, 109(2):204.

Salvatore Campanella, Frédéric Joassin, Bruno Rossion, Anne De Volder, Raymond Bruyer, and Marc Crommelinck. 2001. Association of the distinct visual representations of faces and names: A pet activation study. *NeuroImage*, 14(4):873–882.

Gillian Cohen. 1990. Why is it difficult to put names to faces? *British Journal of Psychology*, 81(3):287–297.

Stanislas Dehaene. 1995. Evidence for category-specific word processing in the normal human brain. *NeuroReport*, 6(2):2153–2157.

Arnaud Delorme and Scott Makeig. 2004. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21.

Rutvik H. Desai, Usha Tadimeti, and Nicholas Riccardi. 2023. Proper and common names in the semantic system. 228(1):239–254.

Kelli Douville, John L Woodard, Michael Seidenberg, Sarah K Miller, Catherine L Leveroni, Kristy A Nielson, Malgorzata Franczak, Piero Antuono, and Stephen M Rao. 2005. Medial temporal lobe activity for recognition of recent and remote famous names: an event-related fmri study. *Neuropsychologia*, 43(5):693–703.

Stéphane Dufau, Jonathan Grainger, Katherine J Midgley, and Phillip J Holcomb. 2015. A thousand words are worth a picture: Snapshots of printed-word processing in an event-related potential megastudy. *Psychological science*, 26(12):1887–1897.

Simon Fischer-Baum, Danielle S Dickson, and Kara D Federmeier. 2014. Frequency and regularity effects in reading are task dependent: evidence from erps. *Language, cognition and neuroscience*, 29(10):1342–1355.

Maria Luisa Gorno-Tempini, Cathy J Price, Oliver Josephs, Rik Vandenberghe, Stefano F Cappa, Narinder Kapur, Richard S Frackowiak, and ML Tempini. 1998. The neural systems sustaining face and proper-name processing. *Brain: a journal of neurology*, 121(11):2103–2118.

Adam Kilgarriff, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. The sketch engine: ten years on. *Lexicography*, 1:7–36.

Zijia Lu and Xuejun Bai. 2023. The processing differences between chinese proper nouns and common nouns in the left and right hemispheres of the brain. 13(3):424.

Eric Maris and Robert Oostenveld. 2007. Nonparametric statistical testing of eeg-and meg-data. *Journal of neuroscience methods*, 164(1):177–190.

Kimberly D Mueller, Rebecca L Koscik, Lianlian Du, Davide Bruno, Erin M Jonaitis, Audra Z Koscik, Bradley T Christian, Tobey J Betthauser, Nathaniel A Chin, Bruce P Hermann, et al. 2020. Proper names from story recall are associated with beta-amyloid in cognitively unimpaired adults at risk for alzheimer's disease. *Cortex*, 131:137–150.

Horst M Müller and Marta Kutas. 1996. What's in a name? electrophysiological differences between spoken nouns, proper names and one's own name. *NeuroReport*, 8(1):221–225.

Oliver Müller, Jon Andoni Duñabeitia, and Manuel Carreiras. 2010. Orthographic and associative neighborhood density effects: What is shared, what is different? *Psychophysiology*, 47(3):455–466.

Robert Oostenveld, Pascal Fries, Eric Maris, and Jan-Mathijs Schoffelen. 2011. Fieldtrip: open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational intelligence and neuroscience*, 2011:1–9.

Alice Mado Proverbio, Stefania Lilli, Carlo Semenza, and Alberto Zani. 2001. ERP indexes of functional differences in brain activation during proper and common names retrieval. 39(8):815–827.

Alice Mado Proverbio, Serena Mariani, Alberto Zani, and Roberta Adorni. 2009. How are 'barack obama' and 'president elect' differentially stored in the brain? an ERP investigation on the processing of proper and common noun pairs. 4(9):e7126.

Carlo Semenza, Sara Mondini, F Borgo, M Pasini, and MT Sgaramella. 2003. Proper names in patients with early alzheimer's disease. *Neurocase*, 9(1):63–69.

Simone Sulpizio and Remo Job. 2018. Early and multiple-loci divergency of proper and common names: An event-related potential investigation. 119:107–117.

Lin Wang, Rinus G. Verdonschot, and Yufang Yang. 2016. The processing difference between person names and common nouns in sentence contexts: an ERP study. 80(1):94–108.

Elizabeth K Warrington and Rosaleen A McCarthy. 1987. Categories of knowledge: Further fractionations and an attempted integration. *Brain*, 110(5):1273–1296.

Jianping Xiong, Yujie Zhang, and Ping Ju. 2021. The effects of orthographic neighborhood size and the influence of individual differences in linguistic skills during the recognition of chinese words. *Frontiers in Psychology*, 12:727894.

Shasha Yang, Shunmei Zhang, and Quanhong Wang. 2016. P2 and behavioral effects of stroke count in chinese characters: Evidence for an analytic and attentional view. *Neuroscience letters*, 628:123–127.

Kiyoshi Yasuda, Bobbie Beckmann, and Tetsuo Nakamura. 2000. Brain processing of proper names. 14(11):1067–1089.

Huei-Ling Yen and Horst M. Müller. 2003. Processing of proper names in mandarin chinese. In Franz Schmalhofer, Richard Young, and Graham Katz, editors, *Proceedings of EuroCogSci 03*, 1 edition, pages 450–450. Routledge.

Appendix A. Summary of Linguistic Features

	Freq	C1F	C2F	C1S	C2S	C1R	C2R	ON*
PN	0.01	5.11	0.65	5.29	7.96	0.58	0.46	128.62
AN	0.25	0.59	0.26	7.16	7.25	0.92	0.83	13.19

Table 2: Means of the targets' frequency (Freq), characters' frequency (C F), stroke count (S), word ratio (R), and orthographic neighbor (ON) (* in thousands)

Appendix B. Raster Plot

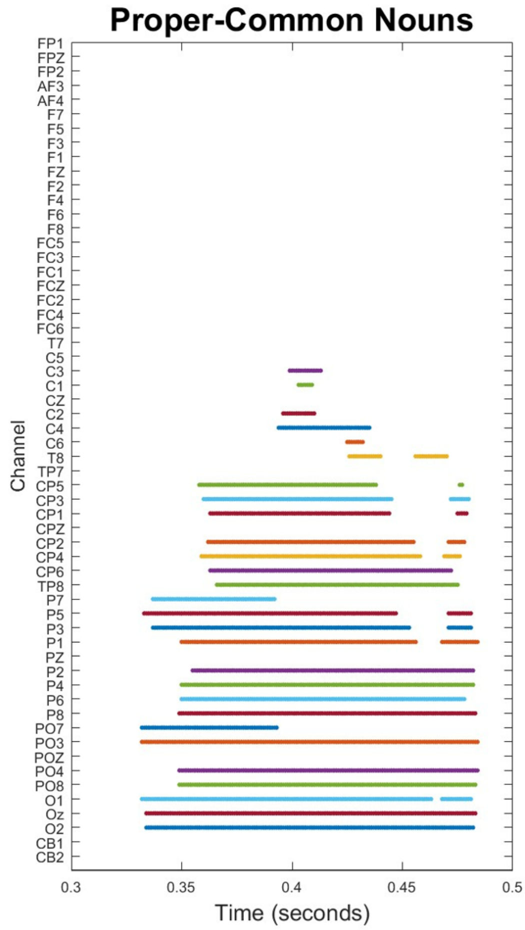

Figure 3: Raster plot showing data point included in the cluster tested with the permutation test for the 300-500ms time window