Worlds shaped by words: a cross-linguistic investigation into the neural mechanisms of lexico-syntactic feature production

Wang, J.

CHAPTER 3

Processing of visual shape information in Chinese classifier-noun phrases

This is an Accepted Manuscript of an article published by Taylor & Francis in the Journal of Cognitive Neuropsychology on April 1, 2025, available online: https://doi.org/10.1080/02643294.2025.2485974.

Abstract: Previous studies have demonstrated that classifiers associated with nouns are activated during lexical access. Shape classifiers, a specific type, incorporate visual shape information. This study examined how visual shape information in classifiers is processed during the production of classifier-noun phrases by native Chinese speakers. Participants performed a picture-naming task using the blocked cyclic naming paradigm, where classifier congruency and shape similarity were manipulated. Behavioural results revealed a classifier congruency effect, with slower reaction times for classifier-incongruent conditions, and a shape interference effect, where classifiers with similar shapes slowed responses. EEG analysis showed that classifier-incongruent conditions elicited more positive voltage amplitudes than congruent ones, while shape-dissimilar conditions produced more negative amplitudes compared to shape-similar conditions after 300 ms post-stimulus. These findings indicated that classifiers were activated when producing noun phrases in a blocked cyclic naming paradigm. Moreover, visual shape information embedded in classifiers was processed during the production of classifier-noun phrases.

Keywords: speech production, lexical access, blocked cyclic naming, shape interference effect, Mandarin Chinese

3.1. Introduction

Different languages adopt various means to categorise nouns, with the main nominal categorisation systems being grammatical gender and classifier systems. In many Indo-European languages, grammatical gender of nouns is primarily assigned according to morpho-phonological rules, with semantic parameters seldom serving as constraints. For instance, the word "*table*" can be categorized as a masculine noun in German (*der Tisch*) and as a feminine noun in French (*la table*), with little relation to its semantics. In contrast, the classifier in Chinese (Adams & Faires Conklin, 1973; Allan, 1977; Kilarski, 2013), demonstrates a high degree of semantic constraints. For example, in Mandarin Chinese, the word "*table*" is collocated with the classifier "*zhang*", which denotes flat objects and corresponds to the flat shape of table surfaces. The role of semantic information in classifier production has been scarcely explored in the literature. This study aimed to investigate how classifiers and the semantic information they contain are processed during noun phrase production.

3.1.1. The activation of lexico-syntactic features during speech production

Various language production models have proposed hypotheses regarding the mental representation and processing mechanisms of grammatical gender (for a review, see Wang & Schiller, 2019), which have been extended to classifiers (Wang et al., 2019). The *WEAVER++ model* assumes that word production is a serial and staged process that traverses from conceptual preparation via lemma retrieval to word-form encoding (Levelt, 1992, 1999; Levelt et al., 1999a, 1999b; Roelofs, 1992, 1993, 2005; Roelofs & Meyer, 1998). Grammatical gender and classifier are stored at the so-called "lemma" level in the mental lexicon as lexico-syntactic features (Levelt et al., 1999a; Nickels et al., 2015; Wang et al., 2019). The lexico-syntactic features of a word are automatically activated during lemma access and are selected when the corresponding gender-marked elements (e.g., determiners; Roelofs, 1992, 1993, 2018) or classifiers (Huang & Schiller, 2021; Wang et al., 2019) are required to be produced. In contrast, *the* "*interactive*" *spreading-activation model* (Dell, 1986, 1988; Dell & O'Seaghdha, 1991, 1992) and the "*Independent-Network*" *model* (Caramazza, 1997) challenge the seriality and discreteness of activation flow, allowing bypassing the retrieval of lexico-syntactic features during bare noun naming. Nevertheless, lexico-syntactic features are retrieved when the production of their lexical forms is required.

The models were empirically tested in studies using the *picture-word interference* (PWI) paradigm. Participants are required to name a target picture while ignoring a superimposed distractor word. It has been found that naming latencies are shorter when the grammatical gender of the target picture is congruent with the distractor word, compared to the incongruent condition, suggesting that grammatical gender is automatically activated during lexical retrieval. This phenomenon is interpreted as a *gender congruency effect* (Schriefers, 1993) or a *determiner congruency effect* (Schiller & Caramazza, 2003), which has been verified across several Germanic languages, such as Dutch (La Heij et al., 1998; Schiller & Caramazza, 2003, 2006; Schriefers, 1993; Starreveld & La Heij, 2004; Van Berkum, 1997) and German (Bürki et al., 2016; Heim et al., 2009; Schiller & Caramazza, 2003;

Schiller & Costa, 2006; Schriefers & Teruel, 2000). Moreover, an analogous effect was observed in the study of Chinese classifiers, i.e., the *classifier congruency effect*. Participants named pictures faster when the distractor and target word shared the same classifier than when they did not (Huang & Schiller, 2021). This finding indicates that classifiers are automatically activated during lexical access and compete for selection when the classifiers are incongruent (Li et al., 2006; Wang et al., 2019).

The processing mechanism of lexico-syntactic features has also been tested using the *blocked cyclic naming* (BCN) paradigm. In this paradigm, all target pictures are grouped in various combinations to form either homogeneous or heterogeneous blocks. Participants repeatedly name a small set of pictures (e.g., three to six pictures) within a block. Each repetition is called a cycle. In a Dutch adjective-noun phrase production task, Vigliocco et al. (2002) found that participants were faster at naming blocks of pictures with the same grammatical gender (homogeneous block) than blocks of pictures with different grammatical gender (heterogeneous block). This result has been interpreted as an effect similar to syntactic priming (Hartsuiker & Kolk, 1998; Pickering, 1999). When previously accessed syntactic features are re-accessed in subsequent trials, the enhanced activation or availability facilitates the process of re-accessing (Bock, 1986; Levelt et al., 1999a). Bi et al. (2010) used the BCN paradigm in a Chinese naming task to investigate the differences in the processing of semantic and visual shape information between bare noun naming and classifier-noun phrase naming. Their findings revealed that the inclusion of classifiers increased the activation levels of both semantic and visual shape information. However, this study did not directly address the process of classifier activation.

3.1.2. Semantic information in classifiers

Although both classifiers and grammatical gender are lexico-syntactic features, classifiers are comparatively more semantically transparent and are subject to semantic constraints in their selection (Kemmerer, 2017; Kilarski, 2013; Seifart, 2010). In noun phrases of the type "number/demonstrative/quantifier + classifier + noun", it is mandatory for the selected classifier to be consistent with the head noun with respect to semantic features. For example, the classifier "*fu*" in the noun phrase "*zhe fu hua*" (this + classifier + painting) must be collocated with noun entities that indicate textiles or pictures; the classifier "*ba*" in "*yi ba dao*" (one + classifier + knife) implies objects with handles. The semantic constraints between classifiers and nouns therefore enable readers to predict the upcoming noun based on the preceding classifier when reading a sentence (Chou et al., 2014), even if there is a long distance between the two (Hsu et al., 2014).

The semantic functions of classifiers are realized through the retention of semantic features such as animacy, shape, function, consistency, and size to varying degrees (Shi, 1996; Tai, 1994; Tai & Chao, 1994; Tai & Wang, 1990). One example is that of shape classifiers, which are closely associated with the external shape of physical objects. They can be systematically categorized from a cognitive perspective based on spatial dimensionality. For instance, the classifier "*tiao*" indicates a long-shaped object that extends in a one-dimensional space; "*zhang*" refers to a flat-shaped object in a two-dimensional space; "*kuai*" denotes a three-dimensional-shaped object

(Shao, 1993; Shi, 2001; Tai, 1994; Tai & Wang, 1990). It was observed that native Chinese speakers rely more on shape features when categorizing nouns compared to speakers of languages without a classifier system (Kuo & Sera, 2009; Schmitt & Zhang, 1998; Sera et al., 2013; Zhang & Schmitt, 1998). However, how shape information is processed when producing classifiers remains to be investigated.

3.1.3. Processing of visual shape information

The processing mechanisms of shape information have been elucidated in some bare noun naming tasks. De Zubicaray et al. (2018) investigated the role of visual form in lexical access through a series of picture-naming experiments using the PWI and BCN paradigms. The results of the PWI paradigm showed that visual form-related distractors caused longer naming latencies than unrelated distractors. The authors concluded that it was a *visual form interference effect*. In the BCN paradigm, a significant visual form interference effect emerged from the second cycle onward, mirroring the *semantic interference effect* (Abdel Rahman & Melinger, 2009; Damian et al., 2001; Janssen et al., 2011, 2015; Kroll & Stewart, 1994; Python et al., 2018; Belke & Stielow, 2013) observed when manipulating semantic categorical relations in this paradigm. De Zubicaray et al. (2018) propose that these two effects may reflect analogous mechanisms, namely the *incremental learning mechanism* (Oppenheim et al., 2010). The lexical nodes link to semantic features (e.g., colour, shape) via excitatory and inhibitory connections. Naming target pictures strengthens the connection between lexical nodes and semantic features. Concurrently, the connections between semantic features and lexical nodes of co-activated non-target words are weakened, which lengthens their retrieval time. Gauvin et al. (2019) investigated the visual form interference effect by utilizing functional magnetic resonance imaging (fMRI). A reduced BOLD signal was observed in the left posterior middle temporal gyrus (pMTG) for visually similar distractors compared to visually dissimilar distractors, an area identified as being associated with lexical processing during speech production. The authors, therefore, concluded that the visual form interference effect reflects lexical processing and not pre-lexical or post-lexical perceptual evaluation.

Pecher and Zwaan (2017) propose that shape information may be activated automatically if the language possesses a shape-related classifier system. However, limited research has tested this hypothesis, and the findings remain inconclusive. Bi et al. (2010) observed a visual shape interference effect only in the classifier-noun phrase naming task, not in the bare noun naming task. Participants named pictures more slowly when the shape classifiers within each block had similar shape features compared to when the shapes were dissimilar. Bi and colleagues propose two explanations: (1) that the production of classifiers may lead participants to allocate more attention to shape information, and (2) that the activation level of classifiers is increased in blocks with homogeneous shape information, thereby eliciting the visual shape interference effect, which manifests as prolonged naming in the classifier-noun phrase naming task. The former interpretation suggests that the effect occurs at the pre-lexical level, whereas the latter implies that it takes place at the lexical level. The specific stage of lexical production at which shape information in classifiers is processed, therefore, remains undetermined.

3.1.4. Electrophysiological correlates of processing classifiers

By combining classifier production with concurrent EEG measurements, the locus of activation of semantic features stored in classifiers can be more precisely studied.

Levelt's model suggests that lemma retrieval occurs around 200 ms after the stimulus onset, followed by lemma selection at approximately 270 - 290 ms (Indefrey, 2011; Indefrey & Levelt, 2004). Wang et al. (2019), therefore, postulate that the activation of the classifier should not be earlier than 300 ms post-stimulus onset. Consistent with their hypothesis, the congruency of the classifier elicited an N400 effect (Kutas & Federmeier, 2011) emerging around 350 ms post-stimulus onset in the PWI paradigm (Huang & Schiller, 2021; Wang et al., 2019). In the context of the current study, if shape information in classifiers is also activated following lemma selection (and therefore also has a lexical locus), it is expected to occur no earlier than 300 ms after stimulus onset, and it should fall within the time window associated with the N400 effect.

Studies on the processing of classifiers have also reported P600 effects. The P600 is typically associated with processing grammatical anomalies or incongruities and has a maximum over the centro-parietal regions of the scalp (Gouvea et al., 2010). When reading phrases or sentences, mismatched classifier-noun pairs can elicit both N400 and P600 effects, indicating difficulties in integrating semantic and syntactic information (Chan, 2019; Hsu et al., 2014; Zhou et al., 2010). These findings suggest that classifier processing involves both semantic and morphosyntactic processing. Building on previous work, this study primarily focused on the N400 and P600 components to observe the processing of semantic and syntactic information in shape classifiers.

3.1.5. The present study

Previous studies have provided some insight into the processing mechanisms of Chinese classifiers. However, it remains unclear whether shape information in classifiers is processed, and at which stage of classifier-noun phrase production this processing occurs. The present study aims to elucidate this question by investigating the production process of shape classifiers. Using ERPs, we further investigate the locus of the visual shape interference effect observed by Bi et al. (2010).

The classifier congruency and shape similarity were manipulated in a blocked cyclic naming task. We hypothesize that if classifier nodes are activated when using the BCN paradigm, the naming latency of classifier-homogeneous blocks will be shorter than in classifier-heterogeneous blocks, similar to previous findings regarding grammatical gender. Drawing from the observed P600 effect induced by syntactic violations in prior studies, we hypothesize that activating different classifiers within a block will impede the integration of syntactic information in syntactically heterogeneous contexts. Therefore, we expect to observe the elicitation of a larger positive amplitude around 600 ms post onset of the target picture (i.e., a P600 effect) in the classifier heterogeneous block compared to the classifier homogeneous block.

Regarding the processing of shape information, we first hypothesize that if shape information in classifiers is activated, a visual shape interference effect should occur. Specifically, we expect longer naming latencies for blocks with homogeneous shapes compared to those with heterogeneous shapes. Furthermore, if shape information is indeed processed during classifier activation (i.e., at the lexical level), we expect to observe an N400 effect. Specifically, shape-heterogeneous blocks should elicit more negative voltage amplitudes around 400 ms post-stimulus onset compared to shape-homogeneous blocks, reflecting the increased difficulty in integrating semantic information.

3.2. Methods

3.2.1. Participants

Thirty-six (ten males and twenty-six females) right-handed native Mandarin Chinese speakers were recruited. The average age of the participants was 25.81 years (SD = 3.25). All participants had normal or corrected-to-normal vision and no history of psychological or neurological impairments or language disorders at the time of testing. Their basic information, including age, gender, handedness, education and language background, was collected via a questionnaire. A consent form, as well as an information form in compliance with the Ethics Code for linguistic research in the Faculty of Humanities at Leiden University, were provided before and after the experiment, respectively. Participants received a monetary compensation for their participation.

3.2.2. Materials

Thirty-six black-and-white line drawings of common objects were selected from the databases of Liu et al. (2011) and Severens et al. (2005). The objects corresponded to monosyllabic (19%) or disyllabic (81%) nouns in Mandarin Chinese. Half of the nouns are associated with classifiers denoting long shapes, and the other half with classifiers denoting flat shapes. Each shape type included three shape-similar classifiers: "*gen*", "*tiao*", and "*zhi*" for long-shape classifiers, and "*mian*", "*pian*", and "*zhang*" for flat-shape classifiers. Collocations of classifiers and nouns were retrieved from the BCC corpus (Xun et al., 2016) and filtered based on their frequency of occurrence, ensuring the selection of high-frequency collocations. Nouns sharing the same shape classifiers were grouped into twelve triplets, forming the classifier-congruent (C+) and shape-similar (S+) conditions (see Table 3.1). These triplets were then rearranged to create the classifier-incongruent (C-) and shape-similar (S+) conditions and the classifier-incongruent (C-) and shape-dissimilar (S-) conditions. Items within each triplet were phonologically, orthographically, and semantically unrelated. Additionally, a rating test was conducted to control for the potential confounding effect of high visual similarity between target pictures. Fifteen native Mandarin Chinese speakers who did not participate in the subsequent experiment rated the visual similarity of paired pictures on a scale of 1 to 7. The average visual similarity scores for the C+S+, C-S+, and C-S- conditions were 1.9, 1.91, and 1.74, respectively. To determine whether the condition significantly predicted the visual similarity scores, mixed-effects ordinal regression models were used in RStudio Version 2024.4.1.748 (R Core Team, 2023) with the *clmm()* function from the R package "*ordinal*" (Christensen, 2023). A top-down model selection procedure was followed, testing the significance of each factor and performing model comparisons using the *anova()* function. The final model included a fixed effect of Condition and random effects of subject and item, which indicated that the condition did not significantly predict the visual similarity scores (see Appendix 3.A for details of the best-fitting model). Three additional unrelated pictures were selected for the training session.

Table 3.1

Example of triplets used in the experimental session in all three conditions.

Condition	Target noun (*English translation, "classifier - shape"*)
Classifier congruent & shape similar	骨头 (bone, "gen – long shape classifier")
	火柴 (match, "gen - long shape classifier")
	香蕉 (banana, "gen - long shape classifier")
Classifier incongruent & shape similar	骨头 (bone, "gen - long shape classifier")
	项链 (necklace, "tiao - long shape classifier")
	枪 (gun, "zhi - long shape classifier")
Classifier incongruent & shape dissimilar	骨头 (bone, "gen - long shape classifier")
	毛巾 (towel, "tiao - long shape classifier")
	地图 (map, "zhang – flat shape classifier")

3.2.3. Design and procedure

This experiment employed a within-subjects design, with classifier congruency (C) and shape similarity of the classifier (S) as the fixed factors, resulting in three conditions: C+S+, C-S+, and C-S-. The C+S- condition was not involved, as identical classifiers inherently share the same shape features. For the picture naming task, each triplet was repeated eight times in a pseudo-random order to form a block. The first cycle served as a familiarization session, the second as a practice session, and the subsequent six cycles constituted the experimental session. Each of the three conditions contained twelve blocks, yielding a total of 648 experimental trials for the entire naming task. Participants were given a training block without experimental materials before the formal naming task to familiarize themselves with the process. The order of blocks and trials was pseudo-randomized using the Windows program Mix (Van Casteren & Davis, 2006), ensuring that no identical pictures or conditions appeared consecutively.

The experiment was conducted using E-prime 3.0 software (Psychology Software Tools, Pittsburgh, PA). For each block, three target pictures were first presented simultaneously for 2,000 ms and then shown individually for 1,200 ms with their target names in the form of classifier-noun phrases (see Figure 3.1). Following this familiarisation session, participants practised naming the three pictures using their designated names, with any deviant responses corrected by the experimenter. In the subsequent six experimental cycles, target pictures were presented as follows: a fixation point "+" for 500 ms, a blank screen for 300 ms, two or three identical target objects for 2,000 ms, and another blank screen for 200 ms. Participants were instructed to name the pictures as quickly and accurately as possible in noun phrases ("number 2/3 + classifier + noun"), with the number 2 or 3 appearing alternately. Each slide advanced to the next according to the set time, regardless of the participant's response. Vocal responses and electroencephalogram (EEG) data were recorded simultaneously. A short break was provided between each block.

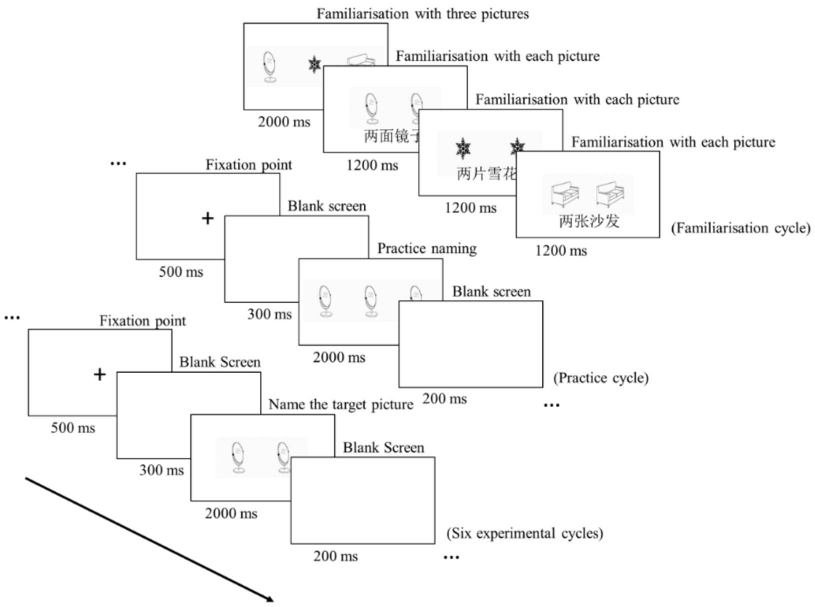


Figure 3.1. Order of stimulus presentation in each block.

3.2.4. EEG Recording and pre-processing

The EEG signals were recorded using 32 Ag/AgCl active electrodes arranged according to the international 10/20 system layout with an Active-Two Biosemi system (BIOSEMI, Amsterdam), as detailed in Appendix 3.B. Additionally, six external electrodes were utilised: two electrodes affixed to the outer canthi of the left and right eyes to measure the horizontal electrooculogram (HEOG), two electrodes placed above and below the left eye to measure the vertical electrooculogram (VEOG), and two electrodes attached to the mastoid bones behind the left and right ears for offline re-referencing. All channels were online referenced to the Common Mode Sense (CMS), with the Driven Right Leg (DRL) used to capture ground circuit noise. EEG recordings were continuously digitised using ActiView software (ActiView806-Lores) by BioSemi at a sampling rate of 512 Hz with a 0.16–100 Hz band-pass filter during the task.

After data acquisition, the EEG signals were pre-processed and analysed offline using Brain Vision Analyzer 2.2 (Brain Products GmbH, 2013). The EEG data were re-referenced to the average of both mastoid electrodes and filtered with a 0.1 Hz high-pass and 30 Hz low-pass filter. Noisy channels comprised less than 2% of electrodes per participant (mean = 0.0139%) and were corrected through spherical spline interpolation. Artefacts, including eye blinks, eye movements and muscle activity, were identified and corrected through visual inspection and independent component analysis (ICA). Epochs with voltage fluctuations exceeding 50 μV within a 200 ms interval or with voltage amplitudes outside the range of [-200, 200] μV were automatically rejected. Additionally, the maximum absolute amplitude difference within a 100 ms interval was limited to 200 μV. After pre-processing, EEG segments around the picture onset markers were extracted, encompassing a temporal range from -200 ms before to 800 ms after the picture onset. Baseline correction was applied using the 200 ms period preceding picture onset. Five participants were excluded from subsequent behavioural and ERP analyses because their number of valid trials fell below 60% after artefact rejection, resulting in a final data set of 31 participants.

3.3. Results

3.3.1. Behavioural data results

Naming accuracy and reaction times were extracted using Praat 6.3.08 (Boersma, 2001). Data exclusion for the behavioural analyses was based on the following criteria: incorrect naming responses (1.18%), reaction times shorter than 200 ms or longer than 2,000 ms (0.99%), reaction times exceeding three standard deviations from the participant's mean (1.27%), and trials containing artefacts in the EEG data (12.28%). Consequently, 17,036 trials (84.81%) out of the initial 20,088 recorded trials remained for the final analysis.

Reaction times were analysed using generalised linear mixed-effects models (GLMMs) with the *lme4* package (Bates et al., 2015) in RStudio Version 2024.4.1.748 (R Core Team, 2023). The *glmer()* function with a gamma distribution and identity link function was employed to model the positively skewed reaction times. The multi-level regression model included fixed-effect predictors for Condition (treatment-coded; baseline C−S+) and Cycle (mean-centred continuous numeric variable). The maximal random effects structure supported by the data based on model comparison included intercepts for participants and items, along with random slopes for the fixed-effect predictors by participant and item. The random effect structure was simplified when the model fitted singularly or failed to converge. Models were compared using the *anova()* function based on Akaike's Information Criterion (AIC; Akaike, 1974), Bayesian Information Criterion (BIC; Neath & Cavanaugh, 2012), and the log-likelihood ratio (Lewis et al., 2011). Model fitness was assessed by plotting residuals against predicted values.

Table 3.2

The best-fitting model for reaction times.

Formula: reaction times ~ 1 + condition + cycle + (1 + condition + cycle || subject) + (1 + cycle || item)

Predictors	*Estimates*	*std. Error*	*95%CI*	*Pr(>|z|)*
(Intercept)	588.166	2.072	584.104 – 592.229	**<0.001**
Condition [C+S+]	-33.011	2.348	-37.613 – -28.409	**<0.001**
Condition [C-S-]	-13.995	1.698	-17.322 – -10.668	**<0.001**
Cycle	-4.844	1.816	-8.403 – -1.285	**0.008**
Random Effects				
σ^2	0.04			
τ_{00} Item.1	116.34			
τ_{00} Subject.2	1,008.61			
τ_{11} Item.Cycle	23.14			
τ_{11} Subject.Cycle	13.53			
τ_{11} Subject.Condition [C+S+]	70.99			
τ_{11} Subject.Condition [C-S+]	113.71			
τ_{11} Subject.Condition [C-S-]	47.92			
ρ_{01}				
ρ_{01}				
ICC	1.00			
N Subject	31			
N Item	36			
Observations	17,036			
Marginal R^2 / Conditional R^2	0.704 / 1.000			

The baseline C-S+ condition was compared to the C+S+ and C-S- conditions, respectively. Results (see Table 3.2) showed that reaction times were significantly shorter in the C+S+ condition (*M* = 538 ms, *SD* = 158) compared to the C-S+ condition (*M* = 568 ms, *SD* = 164), β = -33.011, *95% CI* [-37.613, -28.409], *SE* = 2.348, $p < 0.001$. Additionally, reaction times in the C-S+ condition were significantly longer than those in the C-S- condition (*M* = 556 ms, *SD* = 160), β = -13.995, *95% CI* [-17.322, -10.668], *SE* = 1.698, $p < 0.001$. The main effect of Cycle was also significant, showing that reaction times decreased with every cycle, β = -4.844, *95% CI* [-8.403, -1.285], *SE* = 1.816, $p < 0.01$. See Figure 3.2 for a graphical representation of the reaction time effects.

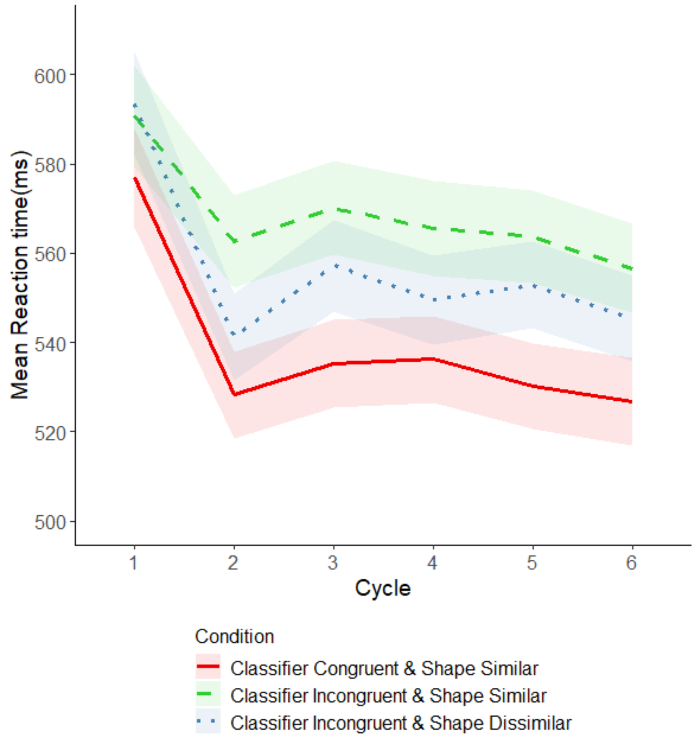
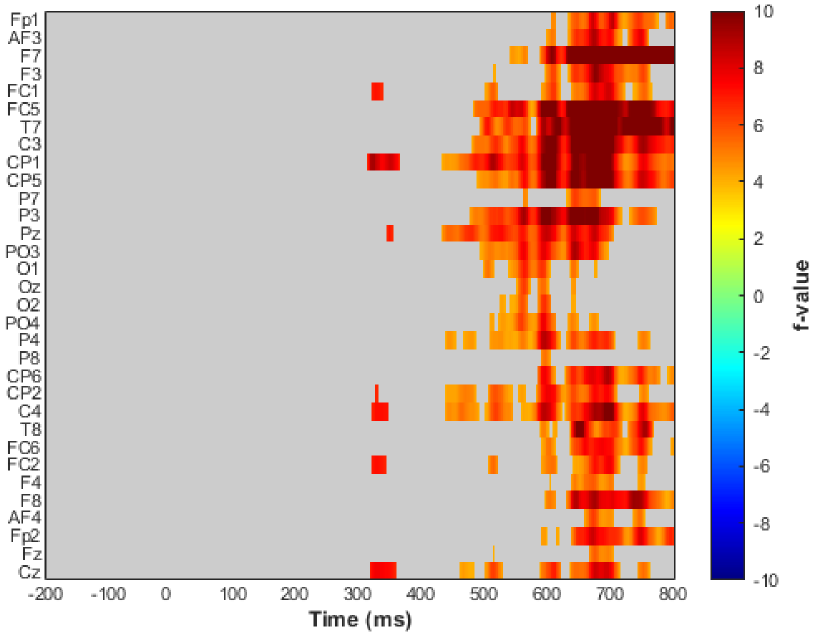


Figure 3.2. Mean and 95% confidence interval of reaction time per condition and per cycle (in ms).

3.3.2. EEG data results

The data sets excluded for behavioural analysis were also removed from the EEG data. A mass univariate linear mixed-effects model analysis (lmeEEG, Visalli et al., 2024) was performed across all electrodes in MATLAB Version R2022b (The MathWorks Inc., 2022) to identify the temporal dynamics of the classifier effect and shape effect. The lmeEEG analysis followed a three-step procedure: (1) a linear mixed model (LMM) was conducted for each channel and timepoint combination on epoched EEG data vectors, comprising all trials from all subjects. The LMM formula used was Voltage amplitude ~ Condition (three levels) + (1 | Subject) + (1 | Item); (2) mass univariate linear regression models (LM) were conducted on marginal EEG data (mEEG), producing a channel-by-timepoint map of the observed F-values for the fixed effect. The mEEG was obtained by removing the random effects contributions estimated in step (1); (3) permutation testing (using 2000 permutations) was performed, and threshold-free cluster enhancement (TFCE) was applied (E = 0.66, H = 2; Smith & Nichols, 2009) with a family wise error of 5%. The results of the mass univariate analysis revealed significant differences in voltage amplitudes across conditions at the Cz, CP1, CP2, Pz, FC1, FC2, and C4 electrodes in the central area from 300 to 400 ms post-stimulus onset (see Figure 3.3). Additionally, potential effects were observed in a broader array of electrodes (anterior: F3, F7, FC1, FC5; central: C3, C4, Cz, CP1, CP2, CP5, CP6; posterior: P3, P4, PO3, Pz) from



approximately 500 to 800 ms post-stimulus onset. These electrodes and time windows were selected for further analysis using linear mixed-effects models (LMMs).

Figure 3.3. The raster diagram shows significant effects elicited by the Condition predictor across all electrodes from -200 to 800 ms post-stimulus onset. The rectangles in warm colours indicate significant channel/time-point pairs, while grey rectangles indicate no significant modulations.

The modelling and model selection process for EEG data was similar to that for behavioural data. An additional covariate was included for the 500 - 800 ms time window: Anteriority (sum-coded with three levels: anterior, central, posterior), which indicates the brain region where the electrodes were located.

In the time window of 300 - 400 ms, the best-fitting model was Voltage amplitudes ~ 1 + Condition + Cycle + (1 + Cycle || Subject) + (1 + Cycle || Item); see

Table 3.3 for the model details. Voltage amplitudes were significantly more positive in the C-S+ condition (M = 3.41, SD = 14.00) than in the C+S+ condition (M = 2.78, SD = 13.52), β = -0.512, *95% CI* [-0.536, -0.487], SE = 0.012, t = -41.319, p < 0.001. Additionally, voltage amplitudes were significantly more negative in the C-S- condition (M = 3.37, SD = 13.52) than in the C-S+ condition, β = 0.064, *95% CI* [0.039, 0.088], SE = 0.012, t = 5.145, p < 0.001. No significant effect of Cycle was observed, β = -0.013, *95% CI* [-0.223, 0.198], SE = 0.107, t = -0.119, p = 0.905.

In the time window of 500 - 800 ms, the best-fitting model (see Table 3.4) was Voltage amplitudes ~ 1 + Condition + Cycle + Anteriority + Condition: Anteriority + (1 | Subject) + (1 | Item). Voltage amplitudes were significantly more positive in the C-S+ condition (M = 4.71, SD = 15.78) than in the C+S+ condition (M = 3.94, SD = 15.43), β = -0.812, *95% CI* [-0.823, -0.800], SE = 0.006, t = -136.961, p < 0.001. Furthermore, voltage amplitudes were significantly more negative in the C-S- condition (M = 4.66, SD = 15.43) than in the C-S+ condition, β = -0.021, *95% CI* [-0.033, -0.010], SE = 0.006, t = -3.617, p < 0.001. A significant effect of Cycle was observed, β = -0.189, *95% CI* [-0.192, -0.186], SE = 0.001, t = -137.727, p < 0.001. The effect of Anteriority indicated that differences in amplitudes were significant at central (β = 0.715, *95% CI* [0.693, 0.736], SE = 0.011, t = 65.756, p < 0.001) and posterior electrodes (β = -0.989, *95% CI* [-1.013, -0.965], SE = 0.012, t = -79.834, p < 0.001). Figure 3.4 presents the temporal dynamics of all conditions for representative electrodes Cz and Pz across each cycle.

Table 3.3

The best-fitting model for voltage amplitudes in the 300 - 400 ms time window for channels Cz, CP1, CP2, Pz, FC1, FC2, and C4 (n = 31).

Formula: voltage amplitudes ~ 1 + condition + cycle + (1 + cycle || subject) + (1 + cycle || item)

Predictors	*Estimates*	*std. Error*	*95% CI*	*t-value*	*Pr(>|t|)*
(Intercept)	3.147	0.886	1.410 – 4.884	3.551	**<0.001**
Condition [C+S+]	-0.512	0.012	-0.536 – -0.487	-41.319	**<0.001**
Condition [C-S-]	0.064	0.012	0.039 – 0.088	5.145	**<0.001**
Cycle	-0.013	0.107	-0.223 – 0.198	-0.119	0.905

Random Effects

σ^2	160.80
$\tau_{00\ \mathrm{Item.1}}$	0.68
$\tau_{00\ \mathrm{Subject.1}}$	23.76

$\tau_{11\ \text{Item.Cycle}}$	0.17
$\tau_{11\ \text{Subject.Cycle}}$	0.22
ρ_{01}	
ρ_{01}	
ICC	0.01
N Subject	31
N Item	36
Observations	6,320,356
Marginal R^2 / Conditional R^2	0.000 / 0.007

Table 3.4

The best-fitting model for voltage amplitudes in the 500 - 800 ms time window for channels F3, F7, FC1, FC5, C3, C4, Cz, CP1, CP2, CP5, CP6, P3, P4, PO3, and Pz (n = 31).

Formula: voltage amplitudes ~ 1 + condition + cycle + anteriority + condition: anteriority + (1 | subject) + (1 | item)

Predictors	*Estimates*	*std. Error*	*95% CI*	*t-value*	*Pr(>|t|)*
(Intercept)	4.507	0.924	2.696 – 6.317	4.879	**<0.001**
Condition [C+S+]	-0.812	0.006	-0.823 – -0.800	-136.961	**<0.001**
Condition [C-S-]	-0.021	0.006	-0.033 – -0.010	-3.617	**<0.001**
Cycle	-0.189	0.001	-0.192 – -0.186	-137.727	**<0.001**
Anteriority [Central]	0.715	0.011	0.693 – 0.736	65.758	**<0.001**
Anteriority [Posterior]	-0.989	0.012	-1.013 – -0.965	-79.834	**<0.001**
Condition [C+S+] * Anteriority [Central]	0.006	0.015	-0.024 – 0.036	0.390	0.697

Condition [C-S-] * Anteriority [Central]	0.059	0.015	0.029 – 0.089	3.854	**<0.001**
Condition [C+S+] * Anteriority [Posterior]	0.320	0.017	0.286 – 0.354	18.362	**<0.001**
Condition [C-S-] * Anteriority [Posterior]	0.180	0.017	0.146 – 0.214	10.326	**<0.001**
Random Effects					
σ^2	214.15				
$\tau_{00\ \mathrm{Item}}$	0.70				
$\tau_{00\ \mathrm{Subject}}$	25.84				
ICC	0.11				
N_{Subject}	31				
N_{Item}	36				
Observations	39,353,160				
Marginal R^2 / Conditional R^2	0.001 / 0.112				

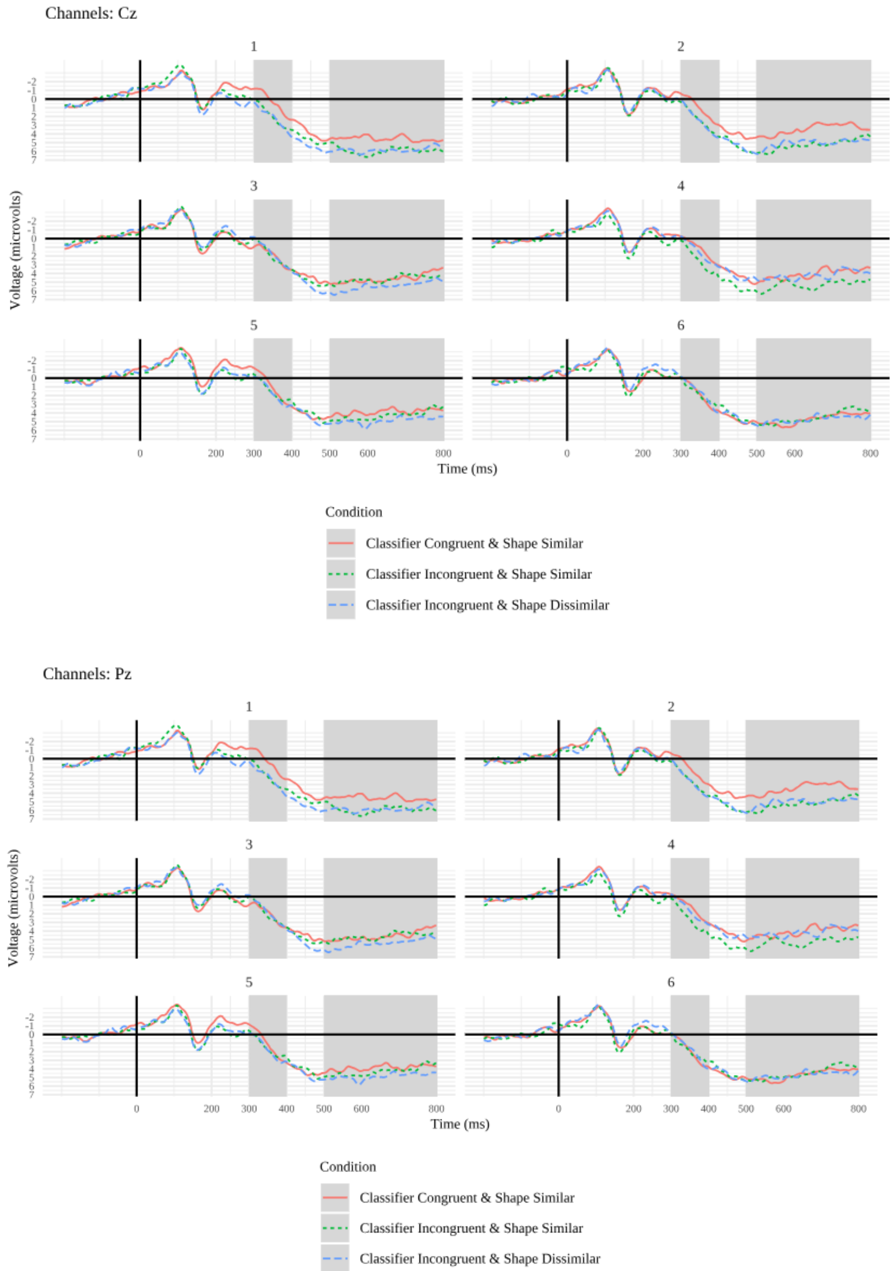


Figure 3.4. Mean voltage amplitudes of all conditions as a function of time and presentation cycle for electrode Cz and Pz. The selected time windows are highlighted in grey.

3.4. Discussion

The present study demonstrated a large classifier congruency effect, with participants naming classifier-congruent blocks faster than incongruent ones. Corresponding EEG data revealed that classifier-incongruent conditions elicited more positive waveforms in the central-parietal regions within the 300-400 ms and 500-800 ms time windows compared to congruent conditions. Furthermore, we observed a shape interference effect where participants named blocks with classifiers sharing similar shapes more slowly than those with dissimilar shapes. EEG data showed that shape-dissimilar conditions elicited more negative voltage amplitudes in the central-parietal regions during the 300-400 ms and 500-800 ms time windows compared to shape-similar conditions. The classifier effects were much larger than the shape effects.

3.4.1. Activation of classifier nodes during noun phrase production in the BCN paradigm

This present study unveiled a large classifier congruency effect in naming latencies. This observation aligns with the findings by Vigliocco et al. (2002) concerning grammatical gender in Dutch, wherein naming speed was observed to be faster when the grammatical genders of the target words were congruent in each block, compared to gender-incongruent blocks. Consistent with behavioural results, our EEG data indicate that the classifier-incongruent condition elicited more positive waveforms starting 300 ms post-stimulus onset, compared to the classifier-congruent condition. This ERP effect occurred after the lemma retrieval stage (Levelt et al., 1999a), which is consistent with conclusions on the temporal dynamics of classifier activation in previous studies with the PWI paradigm (Huang & Schiller, 2021; Li et al., 2006; Wang et al., 2019).

The amplitude differences between classifier-congruent and incongruent conditions were primarily localised in the central-parietal areas during the 300-400 ms and 500-800 ms time windows, showing positive polarity. Therefore, we attribute this ERP component to the late positive complex (LPC). Given that the 500-800 ms time window involves more electrodes and exhibits more pronounced amplitude differences than the 300-400 ms window, the LPC component may consist of two subcomponents. On a more specific note, this LPC component may contain a P300 with a partially overlapping positive Slow Wave (Squires et al., 1975; Sutton & Ruchkin, 1984).

A P300 is typically observed in the oddball paradigm and its variants (Sutton et al., 1965; for a review, see Polich, 2011). It is often elicited by the presentation of novel or infrequent stimuli among a sequence of frequent stimuli, primarily distributed over the fronto-central region of the brain (Katayama & Polich, 1998; Polich, 2003, 2007; Squires et al., 1975) and has been suggested to reflect updating of working memory (Donchin, 1981). The P300 was also observed in a Chinese study using the BCN paradigm (Wang et al., 2018), where unpredictable stimuli within heterogeneous blocks require the updating of context formed by blocks, compared to homogeneous blocks. In the current experiment, we propose that classifier nodes were activated during naming. In classifier-incongruent blocks, the activation of different unpredictable classifiers required more attentional and cognitive resources for context updating compared to classifier-congruent blocks, leading to a P300.

The positive Slow Wave, co-occurring with the P300, indexes further processing activity on the novel stimulus after its detection (Ruchkin et al., 1980; Sutton & Ruchkin, 1984), such as the process of re-analysing and integrating the stimulus with the context. Previous literature has classified ERP components within similar time windows and brain regions as P600 (Frisch et al., 2002; Kaan et al., 2000; Osterhout & Holcomb, 1992), semantic P600 (Kuperberg et al., 2003, 2020; Thornhill & Van Petten, 2012), or post-N400 positivity (PNP; DeLong et al., 2014; Juottonen et al., 1996; Van Petten & Luka, 2012), reflecting a nuanced distinction of the underlying cognitive processes (Leckey & Federmeier, 2020). In general terms, the late positive component reflects the difficulty in integrating information when encountering unexpected stimuli or stimuli that are semantically or syntactically inconsistent with the preceding context. This difficulty in information integration may result in the cost of reanalysing and repairing the conflict (DeLong et al., 2014; Juottonen et al., 1996; Leckey & Federmeier, 2020; Pijnacker et al., 2010; Van de Meerendonk et al., 2010; Van Petten & Luka, 2012). Such positive potentials were observed previously in experiments using mismatched classifier-noun pairs, possibly reflecting the integration and reanalysis of mismatched information between classifiers and nouns or contexts (Chan, 2019; Hsu et al., 2014; Li et al., 2021; Li & Xu, 2023; Zhou et al., 2010). Given that in the present study, the classifiers in the classifier-incongruent blocks were not always consistent with the context, we hypothesise that similar cognitive processes explain the results, i.e., repeated naming maintains the activation of the nouns and their corresponding classifiers at a higher level, while in classifier-incongruent blocks, the activated classifier does not match the preceding context, requiring participants to expend additional cognitive resources to resolve the conflict through reanalysis.

Together, we posit that lexico-syntactic features associated with the target word's lemma, such as grammatical gender or classifier in our study, are activated when participants named target pictures in noun phrases under the BCN paradigm, as has been suggested in previous work using the PWI paradigm (Huang & Schiller, 2021; Li et al., 2006; Wang et al., 2019).

3.4.2. Processing visual shape information in shape classifiers

A shape interference effect was found, although much smaller than the classifier congruency effect. This may be because the shape factor only concerns a single semantic feature of classifiers, whereas the classifier factor manipulates the entire classifier.

The shape interference effect that was observed in the behavioural results is consistent with the findings of Bi et al. (2010). Furthermore, using EEG, we observed that the shape-dissimilar condition elicited more negative ERP responses than the shape-similar condition in the centro-parietal region of the brain after 300 ms post-stimulus onset. These more negative waveforms have been observed to be elicited by semantically unrelated items in studies with the BCN paradigm (Janssen et al., 2011, 2015; Python et al., 2018; Wang et al., 2018). Janssen et al. (2015) interpreted the negative waveforms as requiring more cognitive resources to integrate semantic information with the context or retrieve semantic information from memory in heterogeneous blocks. Moreover, for Chinese, it has been found that mismatched classifier-noun stimuli in sentence comprehension experiments elicited an N400 component, reflecting the difficulty in integrating the semantic information of the classifiers and nouns (Chan, 2019; Hsu et al., 2014; Li et al., 2021; Li & Xu, 2023; Zhou et al., 2010). Based on these findings, we hypothesise that the more negative waveforms we observed in the shape-dissimilar blocks might reflect a similar cognitive process of integrating shape information from classifiers into the context or retrieving shape information of classifiers from memory.

Oppenheim et al.'s (2010) incremental learning mechanism accounts for the semantic interference effect in the BCN paradigm during noun production, which assumes the dynamic adjustment of the connections between semantic features and lexical items. Breining et al. (2016) propose that this mechanism, which applies to semantic-lexical mappings, can also affect lexical-segmental mappings. They suggest that the interference effect observed when comparing blocks with and without segmental overlap also reflects incremental learning. During target word production (e.g., *cat*), the connection between the word and its segments (/k/, /æ/, /t/) is strengthened, while the connection to segments (/m/, /æ/, /t/) from co-activated lexical candidates (e.g., *mat*) is weakened. Consequently, when a word with segmental overlap (e.g., *mat*) is later named within the same block, lexical retrieval is hindered. The shape interference effect we observed may similarly reflect incremental learning during the activation of shape information in classifier nodes during lemma retrieval. For instance, when producing the target phrase "*yi gen gutou*" (one + classifier + bone), the long-shape feature associated with the classifier "*gen*" is activated. This feature, in turn, spreads activation to other classifiers with the long-shape feature (e.g., *tiao*). Successful naming of the target phrase strengthens the connection between the long-shape feature and "*gen*" while weakening its connection with "*tiao*". Consequently, when the next target phrase requires the classifier "*tiao*", its activation becomes more effortful than a classifier with a dissimilar shape feature (e.g., *zhang*).

As naming latencies alone are insufficient to capture changes at different stages of the production process, Bi et al. (2010) proposed two possible loci for the shape interference effect. They suggested that it may occur either at the lexical level, as we have proposed, or at the pre-lexical level. Specifically, they argued that the task requirement to name shape classifiers prompted participants to focus more on the shape features of the target word, leading to difficulties in picture recognition. However, EEG results from our experiment indicate that the differences in voltage amplitude between shape-similar and dissimilar conditions manifested approximately 300 ms post-stimulus onset, indicating that the effect occurred during lexical retrieval (Levelt et al., 1999a). De Zubicaray et al. (2018) found that the shape interference effect observed in the picture naming task disappeared when participants performed a semantic classification task. Since the semantic classification does not involve lexicalisation, this result also suggests that the shape interference effect occurs at the lexical level rather than the pre-lexical level.

3.5. Conclusions

To conclude, this study observed a classifier congruency effect and a shape interference effect through a noun phrase naming experiment using the BCN paradigm. Consistent with previous research concerning the processing mechanisms of classifiers, our findings indicate that the classifiers associated with nouns are activated during the lexical retrieval stage in the BCN paradigm, similar to grammatical gender. Moreover, the semantic information retained in the classifiers is also activated at the lexical level.

Acknowledgements

The research received partial funding from the China Scholarship Council (CSC) under grant number 202006200006. We are grateful to Yufang Wang for providing valuable advice on the analysis of the EEG data in this study. We would like to express our gratitude to all the members of the LUCL Experimental Linguistics Lab. Lastly, we extend our gratitude to the anonymous reviewers and participants for their valuable contributions.

Declaration of interest statement

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported.

Author contributions

Jin Wang: conceptualisation, methodology, programming, data acquisition, formal analysis, data curation, writing-original draft, writing-review and editing, funding acquisition. **Jurriaan Witteman**: formal analysis, writing-review and editing, supervision. **Niels O. Schiller**: conceptualisation, methodology, writing-review and editing, supervision, funding acquisition.

Appendix

3.A Model parameters: visual similarity rating scores

Table 3.A.1 Model of best fit for visual similarity rating scores (n = 15).

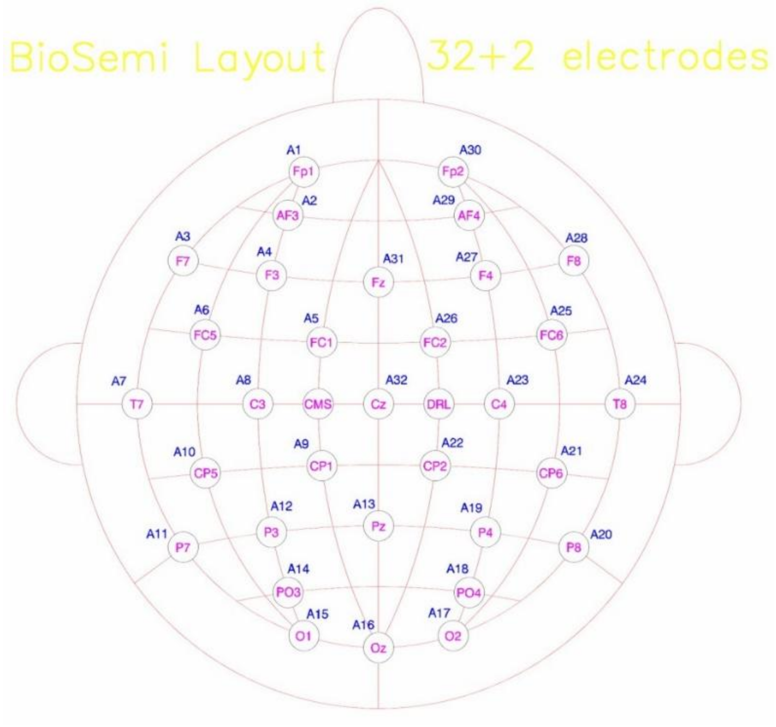
Formula: rating scores ~ condition + (1 + condition | subject) + (1 + condition | item)

Predictors	*Log-Odds*	*std. Error*	*95% CI*	*z-value*	*Pr(>|z|)*
1|2	1.057	0.581	-0.082 – 2.196	1.819	0.069
2|3	2.510	0.585	1.363 – 3.658	4.289	**<0.001**
3|4	3.573	0.590	2.416 – 4.730	6.053	**<0.001**
4|5	3.710	0.591	2.551 – 4.868	6.275	**<0.001**
5|6	4.849	0.602	3.669 – 6.029	8.055	**<0.001**
6|7	6.484	0.644	5.222 – 7.746	10.070	**<0.001**
Condition [C+S+]	-0.529	0.338	-1.192 – 0.133	-1.567	0.117
Condition [C-S+]	0.076	0.356	-0.621 – 0.774	0.214	0.831
Random Effects					
σ^2	3.29				
$\tau_{00\ \text{Noun}}$	0.66				
$\tau_{00\ \text{Subject}}$	5.41				
$\tau_{11\ \text{Noun.Condition [C-S+]}}$	1.82				
$\tau_{11\ \text{Noun.Condition [C+S+]}}$	1.90				
$\tau_{11\ \text{Subject.Condition [C-S+]}}$	0.17				
$\tau_{11\ \text{Subject.Condition [C+S+]}}$	0.23				
ρ_{01}	-0.29				
	-0.44				
	-0.84				
	-0.99				

ICC	0.65
N $_{Subject}$	15
N $_{Noun}$	36
Observations	1,570
Marginal R^2 / Conditional R^2	0.008 / 0.651

3.B EEG montage

Figure 3.B.1: 32 channels with 10/20 system layout including CMS and DRL (www.biosemi.com/headcap.htm).



References

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes*, *24*(5), 713–734. https://doi.org/10.1080/01690960802597250

Adams, K. L., & Faires Conklin, N. (1973). Toward a theory of natural classification. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, *9*(1), 1–10.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Allan, K. (1977). Classifiers. *Language*, *53*(2), 285–311. https://doi.org/10.1353/lan.1977.0043

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *ArXiv Preprint ArXiv:1506.04967*.

Belke, E., & Stielow, A. (2013). Cumulative and non-cumulative semantic interference in object naming: Evidence from blocked and continuous manipulations of semantic context. *Quarterly Journal of Experimental Psychology*, *66*(11), 2135–2160. https://doi.org/10.1080/17470218.2013.775318

Bi, Y., Yu, X., Geng, J., & Alario, F. X. (2010). The role of visual form in lexical access: Evidence from Chinese classifier production. *Cognition*, *116*(1), 101–109. https://doi.org/10.1016/j.cognition.2010.04.004

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341–347.

Breining, B., Nozari, N., & Rapp, B. (2016). Does segmental overlap help or hurt? Evidence from blocked cyclic naming in spoken and written production. *Psychonomic Bulletin & Review*, *23*(2), 500–506. https://doi.org/10.3758/s13423-015-0900-x

Bürki, A., Sadat, J., Dubarry, A. S., & Alario, F. X. (2016). Sequential processing during noun phrase production. *Cognition*, *146*, 90–99. https://doi.org/10.1016/j.cognition.2015.09.002

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Chan, S. hui. (2019). An elephant needs a head but a horse does not: An ERP study of classifier-noun agreement in Mandarin. *Journal of Neurolinguistics*, *52*(June), 100852. https://doi.org/10.1016/j.jneuroling.2019.100852

Chou, C., Huang, H., & Lee, C. (2014). Effects of semantic constraint and cloze probability on Chinese classifier-noun agreement. *Journal of Neurolinguistics*, *31*, 42–54. https://doi.org/https://doi.org/10.1016/j.jneuroling.2014.06.003

Christensen, R. H. B. (2023). *ordinal---Regression models for ordinal data* (R package version 2023.12-4). https://cran.r-project.org/package=ordinal

Damian, M. F., Vigliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*(3), 77–86. https://doi.org/10.1016/S0010-0277(01)00135-4

De Zubicaray, G. I., McLean, M., Oppermann, F., Hegarty, A., McMahon, K., & Jescheniak, J. D. (2018). The shape of things to come in speech production: Visual form interference during lexical access. *Quarterly Journal of Experimental Psychology*, *71*(9), 1921–1938. https://doi.org/10.1080/17470218.2017.1367018

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, *27*(2), 124–142. https://doi.org/10.1016/0749-596X(88)90070-8

Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al (1991). *Psychological Review*, *98*(4), 604–614. https://doi.org/10.1037/0033-295X.98.4.604

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, *42*(1–3), 287–314. https://doi.org/10.1016/0010-0277(92)90046-K

DeLong, K. A., Quante, L., & Kutas, M. (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia*, *61*(1), 150–162. https://doi.org/10.1016/j.neuropsychologia.2014.06.016

Donchin, E. (1981). Surprise!...Surprise? *Psychophysiology*, *18*(5), 493–513. https://doi.org/10.1111/j.1469-8986.1981.tb01815.x

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition*, *85*(3), B83–B92. https://doi.org/10.1016/S0010-0277(02)00126-9

Gauvin, H. S., McMahon, K. L., Meinzer, M., & De Zubicaray, G. I. (2019). The shape of things to come in speech production: A functional magnetic resonance imaging study of visual form interference during lexical access. *Journal of Cognitive Neuroscience*, *31*(6), 913–921. https://doi.org/10.1162/jocn_a_01382

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, *25*(2), 149–188. https://doi.org/10.1080/01690960902965951

Hartsuiker, R. J., & Kolk, H. H. J. (1998). Syntactic facilitation in agrammatic sentence production. *Brain and Language*, *62*(2), 221–254. https://doi.org/10.1006/brln.1997.1905

Heim, S., Friederici, A. D., Schiller, N. O., Rüschemeyer, S., & Amunts, K. (2009). The determiner congruency effect in language production investigated with functional MRI. *Human Brain Mapping*, *30*(3), 928–940. https://doi.org/10.1002/hbm.20556

Hsu, C. C., Tsai, S. H., Yang, C. L., & Chen, J. Y. (2014). Processing classifier-noun agreement in a long distance: An ERP study on Mandarin Chinese. *Brain and Language*, *137*, 14–28. https://doi.org/10.1016/j.bandl.2014.07.002

Huang, S., & Schiller, N. O. (2021). Classifiers in Mandarin Chinese: Behavioral and electrophysiological evidence regarding their representation and processing. *Brain and Language*, *214*(March 2020), 104889. https://doi.org/10.1016/j.bandl.2020.104889

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*(OCT), 1–16. https://doi.org/10.3389/fpsyg.2011.00255

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1–2), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001

Janssen, N., Carreiras, M., & Barber, H. A. (2011). Electrophysiological effects of semantic context in picture and word naming. *NeuroImage*, *57*(3), 1243–1250. https://doi.org/10.1016/j.neuroimage.2011.05.015

Janssen, N., Hernández-Cabrera, J. A., Van Der Meij, M., & Barber, H. A. (2015). Tracking the time course of competition during word production: Evidence for a post-retrieval mechanism of conflict resolution. *Cerebral Cortex*, *25*(9), 2960–2969. https://doi.org/10.1093/cercor/bhu092

Juottonen, K., Revonsuo, A., & Lang, H. (1996). Dissimilar age influences on two ERP waveforms (LPC and N400) reflecting semantic context effect. *Cognitive Brain Research*, *4*(2), 99–107. https://doi.org/10.1016/0926-6410(96)00022-5

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201. https://doi.org/10.1080/016909600386084

Katayama, J., & Polich, J. (1998). Stimulus context determines P3a and P3b. *Psychophysiology*, *35*(1), 23–33. https://doi.org/10.1111/1469-8986.3510023

Kemmerer, D. (2017). Some issues involving the relevance of nominal classification systems to cognitive neuroscience: Response to commentators. *Language, Cognition and Neuroscience*, *32*(4), 447–456. https://doi.org/10.1080/23273798.2016.1274414

Kilarski, M. (2013). *Nominal classification* (Vol. 121). John Benjamins Publishing Company. https://doi.org/10.1075/sihols.121

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kuo, J. Y. C., & Sera, M. D. (2009). Classifier effects on human categorization: The role of shape classifiers in Mandarin Chinese. *Journal of East Asian Linguistics*, *18*(1), 1–19. https://doi.org/10.1007/s10831-008-9036-6

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience*, *32*(1), 12–35. https://doi.org/10.1162/jocn_a_01465

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*(1), 117–129. https://doi.org/10.1016/S0926-6410(03)00086-7

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

La Heij, W., Mak, P., Sander, J., & Willeboordse, E. (1998). The gender-congruency effect in picture-word tasks. *Psychological Research*, *61*(3), 209–219. https://doi.org/10.1007/s004260050026

Leckey, M., & Federmeier, K. D. (2020). The P3b and P600(s): Positive contributions to language comprehension. *Psychophysiology*, *57*(7), 1–15. https://doi.org/10.1111/psyp.13351

Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, *42*(1–3), 1–22. https://doi.org/10.1016/0010-0277(92)90038-J

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*(6), 223–232. https://doi.org/10.1016/S1364-6613(99)01319-4

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999a). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(01), 1–38. https://doi.org/10.1017/S0140525X99001776

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999b). Multiple perspectives on word production. *Behavioral and Brain Sciences*, *22*(1). https://doi.org/10.1017/S0140525X99451775

Lewis, F., Butler, A., & Gilbert, L. (2011). A unified approach to model selection using the likelihood ratio test. *Methods in Ecology and Evolution*, *2*(2), 155–162. https://doi.org/10.1111/j.2041-210X.2010.00063.x

Li, F., Hong, X., & Wang, Y. (2021). The N400 and Post-N400 positivity effect in Mandarin classifier-noun congruence: An ERP study. *Journal of Neurolinguistics*, *57*(May 2019), 100958. https://doi.org/10.1016/j.jneuroling.2020.100958

Li, F., & Xu, X. (2023). The effects of lexical representation on the dynamic process of phrase comprehension. *Journal of Neurolinguistics*, *68*(April). https://doi.org/10.1016/j.jneuroling.2023.101145

Li, W., Jia, G., Yanchao, B., & Hua, S. (2006). Classifier congruency effect in the production of noun phrases. *Studies of Psychology and Behavior*, *4*(1), 34. https://doi.org/https://psybeh.tjnu.edu.cn/EN/Y2006/V4/I1/34

Liu, Y., Hao, M., Li, P., & Shu, H. (2011). Timed picture naming norms for Mandarin Chinese. *PLoS ONE*, *6*(1). https://doi.org/10.1371/journal.pone.0016505

Neath, A. A., & Cavanaugh, J. E. (2012). The Bayesian information criterion: Background, derivation, and applications. *WIREs Computational Statistics*, *4*(2), 199–203. https://doi.org/10.1002/wics.199

Nickels, L., Biedermann, B., Fieder, N., & Schiller, N. O. (2015). The lexical-syntactic representation of number. *Language, Cognition and Neuroscience*, *30*(3), 287–304. https://doi.org/10.1080/23273798.2013.879191

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z

Pecher, D., & Zwaan, R. A. (2017). Flexible concepts: a commentary on Kemmerer (2016). *Language, Cognition and Neuroscience*, *32*(4), 444–446. https://doi.org/10.1080/23273798.2016.1274413

Perniss, P., Vinson, D., Seifart, F., & Vigliocco, G. (2012). Speaking of shape: The effects of language-specific encoding on semantic representations. *Language and Cognition*, *4*(3), 223–242. https://doi.org/10.1515/langcog-2012-0012

Pickering, M. (1999). Syntactic priming in language production. *Trends in Cognitive Sciences*, *3*(4), 136–141. https://doi.org/10.1016/S1364-6613(99)01293-0

Pijnacker, J., Geurts, B., Van Lambalgen, M., Buitelaar, J., & Hagoort, P. (2010). Exceptions and anomalies: An ERP study on context sensitivity in autism. *Neuropsychologia*, *48*(10), 2940–2951.

Polich, J. (2003). Theoretical overview of P3a and P3b. In *Detection of Change* (pp. 83–98). Springer US. https://doi.org/10.1007/978-1-4615-0294-4_5

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Polich, J. (2011). Neuropsychology of P300. In *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0089

Python, G., Fargier, R., & Laganaro, M. (2018). ERP evidence of distinct processes underlying semantic facilitation and interference in word production. *Cortex*, *99*, 1–12. https://doi.org/10.1016/j.cortex.2017.09.008

R Core Team. (2023). *R: A language and environment for statistical computing* (2024.4.1.748). R Foundation for Statistical Computing. https://www.r-project.org/

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1–3), 107–142. https://doi.org/10.1016/0010-0277(92)90041-F

Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: Retrieval of verbs. *Cognition*, *47*(1), 59–87. https://doi.org/10.1016/0010-0277(93)90062-z

Roelofs, A. (2005). Spoken word planning, comprehending, and self-monitoring: Evaluation of WEAVER++. In R. J. Hartsuiker, R. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (1st ed., pp. 42–63). Psychology Press. https://doi.org/10.4324/9780203506196

Roelofs, A. (2018). A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition*, *172*(December 2017), 59–72. https://doi.org/10.1016/j.cognition.2017.12.007

Roelofs, A., & Meyer, A. S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(4), 922–939. https://doi.org/10.1037/0278-7393.24.4.922

Ruchkin, D. S., Sutton, S., Kietzman, M. L., & Silver, K. (1980). Slow wave and P300 in signal detection. *Electroencephalography and Clinical Neurophysiology*, *50*(1–2), 35–47. https://doi.org/10.1016/0013-4694(80)90321-1

Schiller, N. O., & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language*, *48*(1), 169–194. https://doi.org/10.1016/S0749-596X(02)00508-9

Schiller, N. O., & Caramazza, A. (2006). Grammatical gender selection and the representation of morphemes: The production of Dutch diminutives. *Language and Cognitive Processes*, *21*(7–8), 945–973. https://doi.org/10.1080/01690960600824344

Schiller, N. O., & Costa, A. (2006). Different selection principles of freestanding and bound morphemes in language production. *Journal of Experimental Psychology: Learning Memory and Cognition*, *32*(5), 1201–1207. https://doi.org/10.1037/0278-7393.32.5.1201

Schmitt, B. H., & Zhang, S. (1998). Language structure and categorization: A study of classifiers in consumer cognition, judgment, and choice. *Journal of Consumer Research*, *25*(2), 108–122. https://doi.org/10.1086/209530

Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 841–850.

Schriefers, H., & Teruel, E. (2000). Grammatical gender in noun phrase production: The gender interference effect in German. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1368–1377. https://doi.org/10.1037/0278-7393.26.6.1368

Seifart, F. (2010). Nominal Classification. *Language and Linguistics Compass*, *4*(8), 719–736. https://doi.org/10.1111/j.1749-818X.2010.00194.x

Sera, M. D., Johnson, K. R., & Kuo, J. Y. (2013). Classifiers augment and maintain shape-based categorization in Mandarin speakers. *Language and Cognition*, *5*(1), 1–23. https://doi.org/10.1515/langcog-2013-0001

Severens, E., Lommel, S. Van, Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, *119*(2), 159–187. https://doi.org/10.1016/j.actpsy.2005.01.002

Shao, J. (1993). The semantic analysis of classifiers and the bilateral selection between the classifier and the noun [量词的语义分析及其与名词的双向选择]. *Studies of the Chinese Language*, *234*(3), 181–188.

Shi, Y. Z. (2001). The cognitive foundations of the shape-based classifiers in modern Chinese [表物体形状的量词的认知基础]. *Language Teaching and Linguistic Studies*, *27*(1), 34–41. https://doi.org/CNKI:SUN:YYJX.0.2001-01-006.

Shi, Y. Z. (1996). Proportion of extensional dimensions: The primary cognitive basis for shape-based classifiers in Chinese. *Journal-Chinese Language Teachers Association*, *31*, 37–60.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*(1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*(4), 387–401. https://doi.org/10.1016/0013-4694(75)90263-1

Starreveld, P. A., & La Heij, W. (2004). Phonological facilitation of grammatical gender retrieval. *Language and Cognitive Processes*, *19*(6), 677–711. https://doi.org/10.1080/01690960444000061

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187–1188. https://doi.org/10.1126/science.150.3700.1187

Sutton, S., & Ruchkin, D. S. (1984). The late positive complex. *Annals of the New York Academy of Sciences*, *425*(1), 1–23. https://doi.org/10.1111/j.1749-6632.1984.tb23520.x

Tai, J. H. Y. (1994). Chinese classifier systems and human categorization. *In Honor of William S.-Y. Wang: Interdisciplinary Studies on Language and Language Change*, 479–494.

Tai, J. H. Y., & Chao, F. (1994). A semantic study of the classifier zhang. *Journal of the Chinese Language Teachers Association*, *29*(3), 67–78.

Tai, J., & Wang, L. (1990). A semantic study of the classifier tiao. *Journal of the Chinese Language Teachers Association*, *25*(1), 35–56.

The MathWorks Inc. (2022). *MATLAB version: 9.13.0 (R2022b)*. Natick, Massachusetts: The MathWorks Inc.

Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology*, *83*(3), 382–392. https://doi.org/10.1016/j.ijpsycho.2011.12.007

Tsang, C., & Chambers, C. G. (2011). Appearances aren't everything: Shape classifiers and referential processing in Cantonese. *Journal of Experimental Psychology: Learning Memory and Cognition*, *37*(5), 1065–1080. https://doi.org/10.1037/a0023601

Van Berkum, J. J. A. (1997). Syntactic processes in speech production: The retrieval of grammatical gender. *Cognition*, *64*(2), 115–152. https://doi.org/10.1016/S0010-0277(97)00026-7

Van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, *38*(4), 584–589. https://doi.org/10.3758/BF03193889

Van de Meerendonk, N., Kolk, H. H. J., Vissers, C. Th. W. M., & Chwilla, D. J. (2010). Monitoring in language perception: Mild and strong conflicts elicit different ERP patterns. *Journal of Cognitive Neuroscience*, *22*(1), 67–82. https://doi.org/10.1162/jocn.2008.21170

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015

Vigliocco, G., Lauer, M., Damian, M. F., & Levelt, W. J. M. (2002). Semantic and syntactic forces in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 46–58. https://doi.org/10.1037/0278-7393.28.1.46

Visalli, A., Montefinese, M., Viviani, G., Finos, L., Vallesi, A., & Ambrosini, E. (2024). lmeEEG: Mass linear mixed-effects modeling of EEG data with crossed random effects. *Journal of Neuroscience Methods*, *401*. https://doi.org/10.1016/j.jneumeth.2023.109991

Wang, M., Chen, Y., & Schiller, N. O. (2019). Lexico-syntactic features are activated but not selected in bare noun production: Electrophysiological evidence from overt picture naming. *Cortex*, *116*(SI), 294–307. https://doi.org/10.1016/j.cortex.2018.05.014

Wang, M., & Schiller, N. O. (2019). A review on grammatical gender agreement in speech production. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02754

Wang, M., Shao, Z., Chen, Y., & Schiller, N. O. (2018). Neural correlates of spoken word production in semantic and phonological blocked cyclic naming. *Language, Cognition and Neuroscience*, *33*(5), 575–586. https://doi.org/10.1080/23273798.2017.1395467

Xun, E., Rao, G., Xiao, X., & Zang, J. (2016). The construction of the BCC Corpus in the age of Big Data. *Corpus Linguistics*, *3*(1), 93–109.

Zhang, S., & Schmitt, B. (1998). Language-dependent classification: The mental representation of classifiers in cognition, memory, and ad evaluations. *Journal of Experimental Psychology: Applied*, *4*(4), 375–385. https://doi.org/10.1037/1076-898X.4.4.375

Zhou, X., Jiang, X., Ye, Z., Zhang, Y., Lou, K., & Zhan, W. (2010). Semantic integration processes at different levels of syntactic hierarchy during sentence comprehension: An ERP study. *Neuropsychologia*, *48*(6), 1551–1562. https://doi.org/10.1016/j.neuropsychologia.2010.02.001