OPEN

Edited by:

INTRODUCTION

To group words into meaningful strings with a hierarchical structure, the human language system employs sophisticated rules and constraints at di erent levels of representation, such as syntactic and semantic agreements (Frazier and Rayner, 1982; MacDonald et al., 1994; McRae et al., 1998). Violations of these constraints lead to impairments in comprehension, resulting in the slowing of reading speed (Van Gompel et al., 2001; Swets et al., 2008; Wu et al., 2014), defective memory of the sentence (Cohen et al., 2001), and ambiguous, unintended or even unrecognized interpretations (Ferreira et al., 2002; Sanford and Sturt, 2002; Ferreira, 2003). Prosody, i.e., the supra-segmental information conveyed in language, is a type of constraint in the language system. Usually recognized in terms of acoustic variations such as pitch, intensity, and duration, prosodic properties constitute a hierarchical structure, i.e., the prosodic structure, which is believed to be independent of, but related to, the surface syntactic structure of a sentence (Nespor and Vogel, 1986; Selkirk, 2011; for more details of the prosodic structure in Chinese, please see Supplementary Material). Although prosodic constraints are considered to play a fundamental role in speech production and comprehension (Scherer et al., 1984; Fodor, 2002), their functions in written language processing have generally been overlooked in the past.

Recent studies have shown that prosodic and especially metrical features at the word level are represented and utilized during sentence reading (Ashby and Rayner, 2004; Lukatela

brain responses to a Chinese V-O phrase positioned at the end of a sentence, which could be of either the correct [1 + 1] or the incorrect [2 + 1] pattern. While the V-O was visually presented as a whole phrase with two or three characters on the screen in Experiment 1, the verb and the object were seen consecutively in Experiment 2. Results reliably showed that the [2 + 1] pattern elicited both a larger frontocentral negativity and a larger posterior positivity as compared with the [1 + 1] pattern in the N400 time window after the onset of the whole phrase (Experiment 1) or after the onset of the one-character object (Experiment 2), thus ruling out a visual complexity account for this e ect. Moreover, these e ects were unlikely to be the carryover of distinct lexical processing of the mono- and disyllabic verbs because there was no remarkable di erence in ERP responses to these verbs. Instead, these e ects indicated a rapid utilization of the rhythmic pattern during the processing of written sentences: when words are combined to form a phrase, prosody (the rhythmic pattern in this case) proactively imposes constraints on the expectation and selection/access of words that can enter the combination. The violation of the prosodic constraints could initiate reanalysis in a later time window to engender a coherent representation, as reflected by the increased late positivity for the [2 + 1] pattern, relative to the [1 + 1] pattern, 700 ms after the onset of the whole phrase. The findings of N400-like and late positivity e ects for the abnormal rhythmic pattern, i.e., the phrasal structure with an inappropriate combination of syllable numbers, were also observed in studies of speech in other languages (Knaus et al., 2007; Magne et al., 2007; Domahs et al., 2008, 2015; Schmidt-Kassow et al., 2015). In these studies, the placement of word stress or syllabic length was manipulated, resulting in correct or abnormal metrical rhythm at the word level. Thus, cross-linguistic evidence seems to indicate common cognitive processing of rhythmic information in languages with divergent default prosodic structures.

Extending our previous work, the aim of the current study was to further investigate how the rhythmic pattern constrains the build-up of phrases or compounds during Chinese sentence reading. Instead of using the ERP measures, here we recorded oculomotor activities while participants read fully visible sentences that contained the critical constituents. The rhythmic pattern was manipulated on the critical V-O with either [2 + 2] or [2 + 1] pattern, and the combination was positioned at the middle of a sentence to avoid the potential influence of sentence-final integration process on local processing. Moreover, such a V-O combination was used to modify a disyllabic noun (e.g., 基地

pattern or word order violation; this detection could take place with different time courses. Readers were supposed to be aware of the prosodic anomaly when they finished reading the verb and its object, i.e., the first two words of the compound (Luo and Zhou, 2010), before viewing the head noun. In comparison, the anomaly of word order was unlikely to be detected before the head noun was encountered because the alteration of word order occurred only under the specific circumstance of constructing a compound. Therefore, reanalysis was expected to take place at the second word of the three-word compound for the violation of rhythmic pattern but at the third word (head noun) for the violation of word order. Moreover, we expected that the subsequent reanalyses triggered by these two types of violations would be reflected in temporally dissociated oculomotor measures, given that the effects for prosodic violation (i.e., intonation mismatch) and syntactic violation (i.e., word category or word order violation) were obtained in different time windows in the previous ERP studies (Eckstein and Friederici, 2005, 2006). An alternative prediction would be that the reanalysis of either type of anomaly would mainly affect the late stage of processing according to the findings in Breen and Clifton (2013). Accordingly, we would specifically expect differences in regressive eye movements in response to our manipulations.

In particular, we explored the reanalysis process at a relatively late stage, which has generally been overlooked in the past, i.e., after the reader has finished viewing the sentence at least once. The reader may still entertain the demand to re-read the sentence or to go over its constituents even when he/she finishes reading the last word of the sentence (Meseguer et al., 2002; Sturt, 2007). That is, reanalysis can last through the sentence re-reading stage. Malsburg and Vasishth (2011) pointed out that over half of the sentences engender regressions from the end of the sentence no matter whether the reader was garden-pathed or not. These regressive eye movements tend to illustrate a long-lasting integration process in which all the lexical information has been accessed. Here we not only used the conventional analysis (see Clifton et al., 2007, for a review) to examine the total reading durations for re-reading but also adopted a newly developed method, scanpath analysis, to analyze the pattern of re-reading (Malsburg and Vasishth, 2011). A scanpath refers to a sequential pattern of eye-fixations. Distinct from conventional eye-tracking measures, scanpath analysis integrates both the spatial and the temporal structure of the eye-movements during reading, providing a global view of eye movements' dynamic changes within a space. It would particularly make contributions in studies that have multiple regions of interest and a large amount of long saccades (Mitchell et al., 2008; Malsburg et al., 2014), and be beneficial to the exploration of re-reading. For example, inbound or outbound regressions and second-pass reading times for the pre-defined region(s) are usually measured to reflect the regressive eye movements occurring at the late stage of sentence processing. But these measures, by definition, have the deficits in quantitatively depicting the sequence of serial saccadic movements and the start-and-end of each of them. This information might be peripheral for investigating the initial reading of the sentence because readers are inclined to read

region-by-region, from the beginning toward the end. By "initial reading," we mean the reading processes that occur before the eyes reach the end of the sentence. It nevertheless is critical for re-reading since saccadic span and moving directions are more flexible and optional at such a late stage, and as a consequence, should be more vulnerable to experimental manipulations. In other words, investigating eye movements for sentence re-reading requires approaches with better integration of temporal and spatialalalcd.385642(r)0.1392953353175-0.306.398(i)-0.159932(n)0.0988 reas a(n)0.221841(t)-0.247636(e)-1.01118(g)0.149614(r)0.13932502tcage-49.

congruent object noun being monosyllabic (e.g., 蒜, *suan*, garlic) or disyllabic (e.g., 大蒜, *dasuan*, garlic). Different forms of each pair were synonyms, expressing the same meanings and having the same V-O syntactic relationship; moreover, the character (morpheme) of the monosyllabic word was a constituent of its disyllabic counterpart. The word frequencies of the monosyllabic nouns are generally higher than those of the disyllabic forms, 60 vs. 29 per million according to the Lancaster Corpus of Mandarin Chinese (McEnery and Xiao, 2004) or 42 vs. 17 per million according to SUBTLEX-CH (Cai and Brysbaert, 2010). We selected three kinds of combinations in each set as the experimental stimuli: a monosyllabic verb with a monosyllabic noun, a disyllabic verb with a monosyllabic noun, and a disyllabic verb with a disyllabic noun. For the latter two, word order of the combination was manipulated as either V-O or O-V. Thus, there were five types of pairs, each of which was then combined with a disyllabic noun (e.g., 基地, *jidi*, farm, base) to form a compound (see Table 1). This disyllabic noun could only act as the head noun of the compound and could not be viewed as an object of the verb because of the selectional restriction of the verb.

Note that, apart from the four critical conditions, the experiment also included a fifth, unanalyzed condition in which sentences had the same structure and words as the other critical sentences except that the compound took a monosyllabic verb, which was not used in the 4 critical conditions, with a monosyllabic noun as the modifier in a V-O sequence. These correct sentences, 19 in each test list, were taken as fillers in the five lists constructed using a Latin-square procedure. Thus, each test list had 95 critical sentences, 19 for each experimental condition. Another 80 sentences were also added into each list as fillers. They were structurally similar to the critical sentences except that the critical compounds were replaced by constructions like "V-O + *de* + noun," "O-V + *de* + noun," and "object + *d*e+ verb + noun" with various types of acceptable rhythmic patterns and/or by compounds with a monosyllabic head. All the fillers were well-formed. Stimuli in each list were pseudo-randomized such that no more than 3 sentences from the same condition would appear consecutively.

Apparatus

Eye movements were recorded with an EyeLink 2K system at a sampling rate of 2000 Hz. Each sentence was presented in one line at the middle vertical position of a 21-inch CRT screen (1024*768 resolution; frame rate 100 Hz). The font *Fangsong*-29 was used, with one character subtending 1 degree of visual angle. Participants read each sentence with their head positioned on a chin rest 78 cm from the screen. All recordings and calibrations were based on the left eye but viewing was binocular.

Procedures

Participants were calibrated with a nine-point grid. A fixation cross was presented at the position on the screen where the first character of the sentences would appear. The fixation was presented for 500 ms, followed by visual presentation of the whole sentence. Participants were required to silently read the sentence and to press a button on a joystick when finished reading. The button pressing caused the sentence to disappear

and a verification sentence to present in about one-third of the trials in the list, including fillers. Participants were instructed to answer, by pressing the "yes" button with their left index finger and the "no" button with their right one, whether the verification sentence was semantically congruent with the preceding sentence. Half of the trials required a "yes" answer and half required a "no" answer. The content of the verification sentence could be related to the overall meaning or to any part of the target sentence; this was to minimize any potential influence on eye movements (particularly the re-reading pattern) by task demand. Specifically, 30 critical trials were followed by the verification sentences. In 18 trials, the head noun of the critical compound, but not the V-O combination, was mentioned in the probe sentence (e.g., "The farm's irrigation system will be changed" for the exemplar in Table 1). In another 12 trials, the probe sentence concerned the comprehension of the V-O combination (e.g., "The planting of garlic needs watering."). Each trial ended with a rating task on the well-formedness of the sentence. A 2-by-2 grid filled with the numbers 1, 2, 3, 4 respectively would appear on the screen, and participants were instructed to press one of the four corresponding buttons on the joystick to assess the well-formedness of each sentence, with "4" representing that the sentence was not well-formed or the expression was unnatural and "1" representing that the sentence was well-formed and the sentence was expressed in a conventional way. While the verification task required memory and comprehension of the sentence overall as well as parts of the sentence, the well-formedness rating could boost the sensitivity to the unnatural regions. Participants underwent a practice block of 15 trials before the formal experiment.

Conventional Analysis

Five regions were selected as the regions of interest, as shown in Table 1. Region 1 contained the first component of the V-O, composed of 1–2 characters (e.g., the word "to plant" in the example); Region 2 contained the second component of the V-O, which was composed of 1–2 characters (e.g., "garlic"); Region 3 contained the head noun of the critical compound, composed of 2 characters (e.g., "district"); Region 4 contained the component (a verb phrase or an adverbial of the predicate structure in the clause) after the compound, composed of 2–3 characters (e.g., "will introduce"); Region 5 contained the last 3 characters of the sentence (e.g., the last three character of "irrigation system"), except for a few sentences in which only the last two characters were included. The latter was because there were only two characters left after Region 4. Alternatively, defining the last 2 characters as Region 5 for all sentences yielded the same pattern of effects as the one reported in this article; but given that the last two characters were easily parafoveally processed and skipped without fixations, we choose to report the analysis using the 3-character definition for Region 5.

Regression Path Duration (RPD) and the probability of Regression Out (REG) for each region were the measures of instant reanalysis during first reading. RPD was the summed fixation duration from when the region was first fixated until the eyes first moved past the region. REG included the percentage of trials in which at least one regression was made from a given

region to previous parts of the sentence prior to leaving that region in a forward direction. First-fixation durations shorter than 60 ms or longer than 800 ms, or Gaze Duration (GD, i.e., the sum of fixation duration from the eyes first entered the region until the eyes moved out) shorter than 60 ms or longer than 1000 ms were excluded from duration and regression analyses, leaving 96% of observations across the five defined regions for statistical analyses.

Estimates were from a linear mixed model (LMM) for durations and a generalized linear mixed model (GLMM) for percent regressions (Baayen et al., 2008), with crossed random e ects for participants and items using the *lmer* program of the *lme4* package (Bates et al., 2008) in the R environment for statistical computing (R-Development Core Team, 2009). Due to the large number of trials, the t-distribution approximated the normal distribution, and estimates larger than 2 SE, i.e., absolute t-values (for LMM) or z-values (for GLMM) > 1.96 were interpreted as being significant.

Scanpath Analysis

We generally followed the method introduced in Malsburg and Vasishth (2011) to perform scanpath analysis. Scanpath analysis (Cristino et al., 2010; Malsburg and Vasishth, 2011) first quantifies the dissimilarities between every two scanpaths. Similar scanpaths are then clustered and, as a consequence, a prototype of each cluster can be extracted. In this study, each prototype depicted one specific regressive pattern of re-reading. Thus, we were able to both explore what kinds of regressive patterns were mainly triggered for a particular condition and to perform comparisons between conditions by examining the distribution (i.e., percentages) of their scanpaths in each cluster.

The dissimilarities between scanpaths can be measured with "scasim" (Malsburg and Vasishth, 2011), which is a type of global distance between any two fixation sequences. The core idea of this distance is like the edit distance (Levenshtein, 1966), quantifying the dissimilarity of two sequences as the overall penalty to transform one sequence into the other. The pre-defined penalty is a function of locations and durations of fixations in the two sequences:

$$d(f, g) = |dur(f) - dur(g)| \times m^{distance(f,g)} + |dur(f) + dur(g)| \times (1 - m^{distance(f,g)}).$$

In this definition, f and g refer to any of the fixations from two scanpaths, respectively. The function $dur()$ defines the duration of the fixation, the function $distance()$ defines the distance in the visual field between f and g, and m is a constant which approximates the drop of visual acuity as the distance increases. As there could be at maximum C_{m+n}^{n} ways to align the fixations given that the two sequences have m and n fixations respectively, at maximum C_{m+n}^{n} overall penalty values could be generated. The minimal value among them was defined as the scasim between these two sequences.

For the further clustering, it is recommended to map all the scanpaths into a multi-dimensional space while keeping the distances between them undistorted because a coordinate space

RESULTS

Accuracy and Rating

On average, participants correctly answered 94.9% ($SD = 6\%$) of all the probe questions, indicating that they read the sentences carefully. For the experimental sentences, 30 of them were followed by a probe question referring to the interpretation of the noun compound: the accuracy rate was 95.48% ($SD = 6\%$) for the 18 questions concerning the head noun and was 94.30% ($SD = 9\%$) for the remaining 12 questions concerning the meaning of the V-O combination. The high accuracy suggests that the critical compounds were well comprehended across all the experimental conditions (see Table 2), although the readers showed the tendency of having higher accuracy for sentences with the abnormal rhythmic pattern than for sentences with the normal rhythmic pattern (by 4.5%, $t = 1.9$). As shown on Table 2, the rating scores of all conditions were less than 2, which may imply that the readers were adopting a relatively loose criterion in judging the well-formedness with respect to the violations of rhythmic pattern or word order. But there were distinctions between the four critical conditions, as confirmed by the statistical analysis with the LMM, which included rhythmic pattern and word order as two within-participant factors. Sentences with abnormal rhythmic pattern were generally rated as more odd than those with normal rhythmic pattern (1.48 for sentences with abnormal rhythmic pattern and 1.69 for sentences with normal rhythmic pattern), $t = -6.51$. Although there was no significant main effect of word order, an interaction between the two factors was found, $t = 2.3$. Further comparisons showed that the abnormal rhythmic pattern led to worse comprehensibility not only when the word order was correct (by 0.27, $t = 6.65$) but also when the word order was incorrect (by 0.13, $t = 3.26$). However, the incorrect word order resulted in worse comprehensibility only when the rhythmic pattern was abnormal (by 0.18, $t = 4.55$), not when the rhythmic pattern was normal ($t < 1.2$).

Conventional Analysis

Measures for all the five regions are shown in Table 3.

Region 1. Words in this region were disyllabic verbs in the RHY+ORD− and RHY−ORD− conditions and were monosyllabic or disyllabic nouns in the RHY+ORD+ and RHY−ORD+ conditions. Since comparing verbs and nouns

---

TABLE 2 | Grand means and standard errors of accuracy rate and

correct word order, 17% vs. 14%, $b = 0.29$, $SE = 0.13$, $z = 2.2$, $p = 0.028$. The same pattern was also obtained in GD analysis.

A significant interaction between rhythmic pattern and word order was found in this region on RPD, $b = 97.7$, $SE = 33.22$, $t = 2.94$, but not on regression probability measures, $p > 0.1$. Further analysis showed that, while RPD showed only a tendency of being longer for sentences with incorrect word order than for sentences with correct word order when rhythmic pattern was normal, $t < 1.5$ or $p > 0.1$, the difference between the two conditions was highly significant when the rhythmic pattern was abnormal, 627 ms for the RHY-ORD- condition and 527 ms for the RHY−ORD+ condition, $b = 114.13$, $SE = 27.53$, $t = 4.15$. This interaction could also be interpreted in terms of the effect of rhythmic pattern as a function of word order. When the word order was correct, sentences with abnormal rhythmic pattern yielded longer RPD (527 ms) than sentences with normal rhythmic pattern (400 ms), $b = 125.03$, $SE = 23.46$, $t = 5.33$; when the word order was incorrect, the difference was even larger (627 ms for the RHY-ORD- condition and 402 ms for the RHY+ORD− condition), $b = 221.29$, $SE = 22.93$, $t = 9.65$.

To summarize, on the head noun, the rhythmic pattern effect emerged regardless of whether the word order was correct or not; however, the size of the effect was larger in sentences with incorrect word order.

Region 4. In this region, the abnormal rhythmic pattern led to not only longer RPD (611 ms for sentences with abnormal rhythmic pattern and 555 ms for sentences with normal rhythmic pattern), $b = 56.21$, $SE = 20.06$, $t = 2.8$, but also a higher REG (17% for sentences with the abnormal rhythmic pattern and 14% for sentences with the normal rhythmic pattern), $b = 0.29$, $SE = 0.13$, $z = 2.22$, $p = 0.026$. On the other hand, the violation of word order also resulted in longer RPD (608 ms for sentences with the incorrect word order and 559 ms for sentences with the correct word order), $b = 46.77$, $SE = 20.06$, $t = 2.33$, as well as a higher REG (18% for sentences with incorrect word order and 13% for sentences with the correct order), $b = 0.51$, $SE = 0.13$, $z = 3.32$, $p < 0.001$. Interaction between rhythmic pattern and word order did not reach significance, $t < 1.0$. No significant results were observed on GD, $t < 0.5$.

Region 5. Compared with sentences with acceptable rhythmic pattern, sentences with abnormal rhythmic pattern induced *shorter* RPD (1059 ms for sentences with the abnormal rhythmic pattern and 1134 ms for sentence with the normal rhythmic pattern), $b = -74.69$, $SE = 31.21$, $t = -2.39$, and a *reduced* REG (45% for sentences with the abnormal pattern and 50% for sentences with the normal pattern), $b = -0.24$, $SE = 0.1$, $z = -2.43$, $p = 0.015$. Neither the main effect of word order nor the interaction between word order and rhythmic pattern was significant, $ts < 1$. Again, no significant results were observed on GD, $t < 1.4$.

Scanpath Analysis

Participants executed regressions from the last word in 1292 trials of all 2250 trials: 327 for the RHY+ORD+ condition, 335 for the RHY−ORD+ condition, 333 for the RHY+ORD− condition, and 297 for the RHY-ORD- condition. Although 51 trials (3.9%) seemed unusual since their converted distances to

all other scanpaths were over three standard deviations larger than those of the other scanpaths, they were kept in the following analysis as removing them did not change the pattern of results.

We first built a 2-dimension map due to the simplicity and visibility of this model. The stress of this map was 13.32%, indicating that this map was good enough for our purpose (c.f., Kruskal, 1964; Malsburg and Vasishth, 2011). Thirteen clusters were detected on this map using the mixture of Gaussian modeling, which is able to identify the clusters even if they intersect or overlap. **Figure 1** shows the map and **Figure 2** shows the prototype of each cluster. To better illustrate the locations of the fixations in the scanpaths in terms of sentence structure, we matched the coordinates of location and the critical regions, as shown in **Figure 2**. Clusters were sorted according to the complexity of constitutional scanpaths.

Scanpath patterns in **Figure 2** could be roughly classified into three groups in terms of durations and complexity. Clusters such as Cluster I, II, and III contained simple scanpaths with a single fixation shorter than 400 ms. Clusters IV, V, VI, and VII manifested relatively efficient regressions targeting the compound region, as the scanpaths consisted of 2–4 fixations on the compound region for a total of 600–1000 ms. The other clusters (Clusters VIII–XIII) could be viewed as complex regression patterns, with durations longer than 1000 ms and diversified spatial trajectories. Detailed descriptions of clusters according to the prototypes are as follows.

Simple Cluster I: Regressing from the end of the sentence, with only a fixation located on the beginning of the sentence, including Region 1 and the words before Region 1.

Simple Cluster II: The single fixation on the scanpath mainly dwelled on Region 2, 3, or 4, which was within or next to the critical compound in the middle of the sentence. But the whole scanpath lasted no more than 200 ms, relatively shorter than the single fixations compared with other simple clusters.

Simple Cluster III: The single fixation located on the middle and the later part of the sentence, including Region 4 and words between Region 4 and 5.

Moderate Cluster IV: This cluster consisted of a backward-moving pattern with the first fixation on Region 1 or 2 (i.e.,
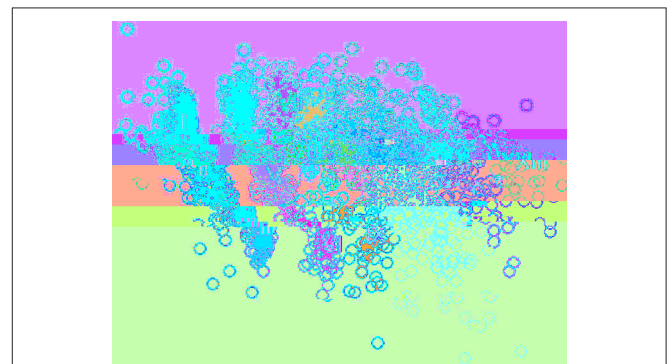
FIGURE 1 | Map of all regression patterns in the data set originating from region 5. Colors indicate clusters that were found using mixture of Gaussian modeling. The roman numbers mark the positions on the center of these clusters.
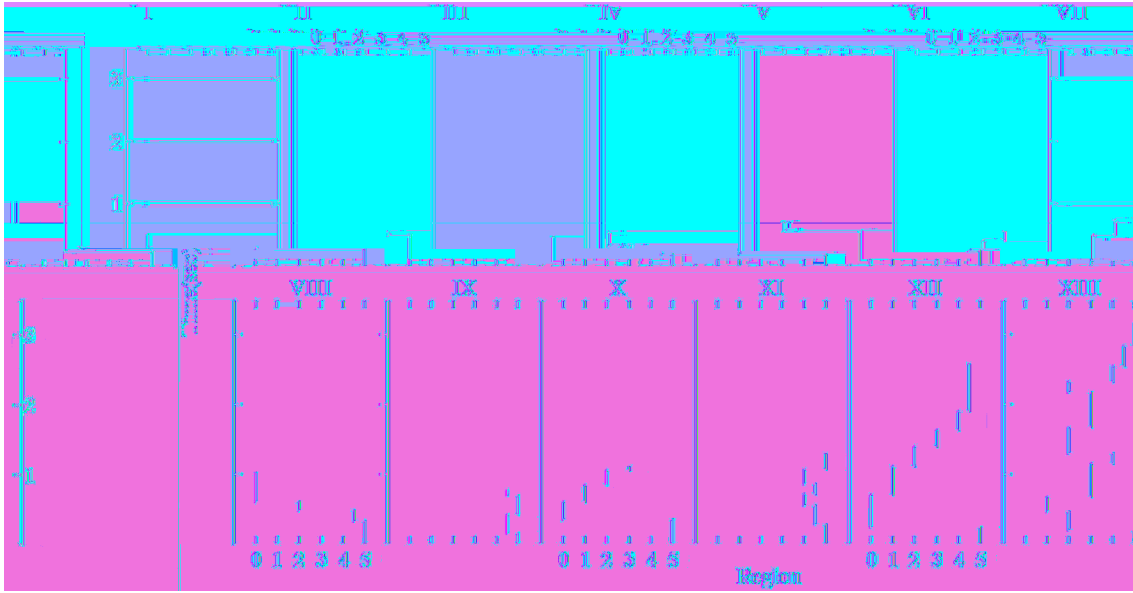
FIGURE 2 | The regression patterns that were closest to the gravity center of the clusters identified on the 2-dimensional map of all regressions from the data set (see Figure 1), called the prototypical regress

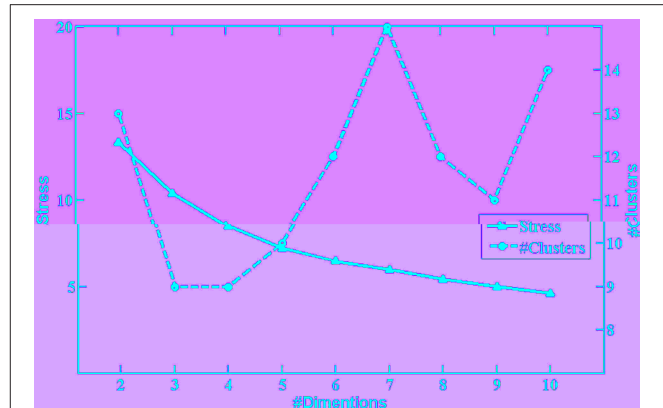**TABLE 4 | Count of scanpaths by cluster and condition (2-dimensional map).**

	RHY+ORD+	RHY−ORD+	RHY+ORD−	RHY−ORD−	Total
Cluster I	25	39	24	37	125
Cluster II	48	30	50	31	159
Cluster III	24	30	21	27	102
Cluster IV	28	28	23	26	105
Cluster V	18	27	25	18	88
Cluster VI	22	14	27	7	70
Cluster VII	20	24	13	12	69
Cluster VIII	6	7	8	9	30
Cluster IX	19	19	11	15	64
Cluster X	28	31	28	30	117
Cluster XI	45	52	71	52	220
Cluster XII	33	25	21	23	102
Cluster XIII	11	9	11	10	41
Total	327	335	333	297	1292

FIGURE 3 | Stress values and numbers of clusters for increasing numbers of map dimensions. As the number of dimensions goes up, the stress of maps decreases, i.e., more variance is explained by higher-dimensional maps.

normal sentences, sentences with the abnormal rhythmic pattern were associated with more scanpaths of simple patterns and with fewer scanpaths of complex patterns during re-reading. Neither the main effect of word order nor the interaction of rhythmic pattern and word order were significant[1].

In order to test the reliability of the scanpath classification and to validate the result found on the 2-dimensional space, we also fitted maps for 2–10 dimensions and calculated clusters models for each of them. **Figure 3** shows the stress of those maps and the number of clusters obtained as a function of the number of dimensions. The stress (or variance not represented by the map) decreased as the dimension of the map increased, while the number of clusters reached a plateau of around 10 after dimension of the map became large enough. To contrast the 2-dimensional model with a more complex one, we chose the clustering on the 5-dimensional map for further analysis since the 5-dimensional model approximated to the knee in the stress curve which methodologically (or structurally) indicated the possible dimensionality of the data. **Figure 4** shows the prototypical scanpaths of the clusters and **Table 5** shows the count of scanpaths by cluster and condition.

Multinomial logistic regression with Cluster X as the baseline category yielded a significant difference between the numbers of trials with the normal and abnormal rhythmic patterns, $z = 2.03$, $p = 0.04$ for Cluster I, and $z = 1.89$, $p = 0.06$ for Cluster III, respectively. As these two clusters reflected simple scanpath patterns, this result indicated that re-reading sentences with the abnormal rhythmic pattern induced more, rather than less, simple regressive eye movements as depicted by cluster I and III. This finding is consistent with the findings of the analysis with the 2-dimensional map which showed increased trials of

simple patterns but fewer trials of complex patterns for prosodic violation.

DISCUSSION

By manipulating the rhythmic pattern and word order of the compound, the present study showed that, during the initial reading of the sentence, both types of violations elicited immediate reanalysis locally within the compound, as reflected by longer reading times and more regressions. These effects also extended to the reading of the region right next to the critical compound. But the interaction, which suggested more demanding reanalysis for double violations than for single ones, only occurred on the head noun of the compound, not on the spillover region. After all words of the sentence had been accessed, however, readers tended to initiate less reanalysis for sentences with the abnormal rhythmic pattern than for sentences with the normal pattern, as shown by shorter re-reading times on the sentence-final words as well as fewer and simpler regressive eye movements. In the following discussion, we start with the issue of processing word order information, and then focus on the immediate use and the late influence of rhythmic pattern information during sentence reading. We note especially the similarities and dissimilarities between the processing of different types of information.

The effect of word order violation was in line with a great number of previous studies reporting prolonged viewing durations and more regressions for sentences with ambiguities or errors (for a review, see Rayner, 1998; for the study on Chinese, see Shi et al., 2000; Zhang et al., 2000, 2002; Hsieh et al., 2009). In particular, the violation of word order was immediately detected at the head noun, resulting in an increase of gaze duration, which is commonly considered to indicate the difficulty of lexical access during sentence reading (Rayner, 2009; Yan et al., 2014). More regressive saccades were then launched toward the preceding regions, reflecting the attempt to integrate the current word into the upstream context (Spivey and

---

[1]We also performed the multinomial logistic regression model with cluster XIII as the baseline category since this cluster had the most similar distribution to the overall distribution of trials according to the result of chi-square test (with the minimum $\chi 2$ value 0.33). Cluster VI showed a robust effect, $z = 2.06$, $p < 0.05$. Test on Cluster I was not significant though, $p = 0.29$.
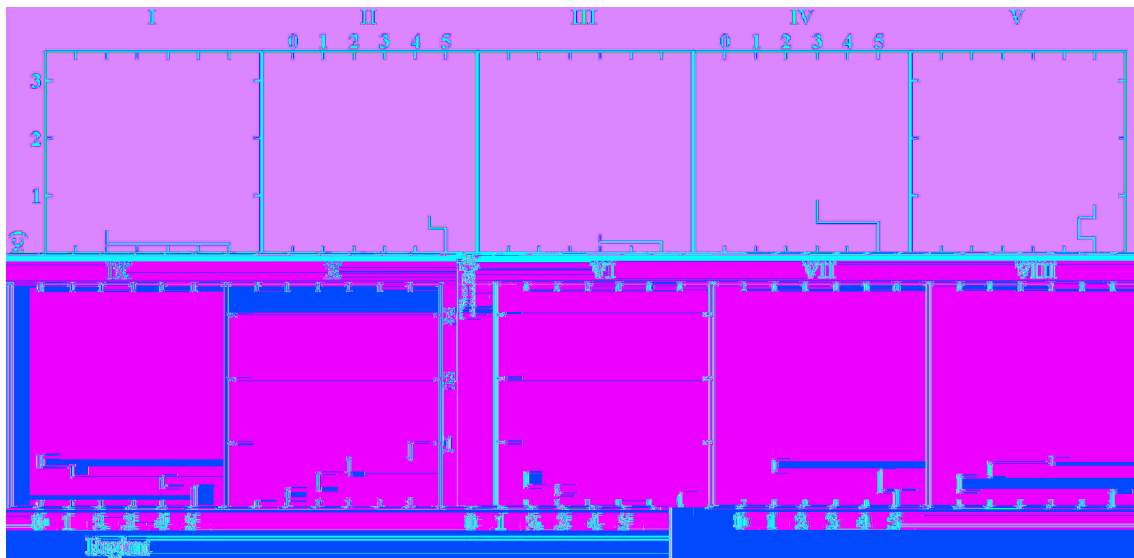
FIGURE 4 | Prototypical regressive patterns of the clusters on the 5-dimensional map.

TABLE 5 | Count of scanpaths by cluster and condition (5-dimensional map).

	RHY+ORD+	RHY−ORD+	RHY+ORD−	RHY−ORD−	Total
Cluster I	26	42	25	37	130
Cluster II	21	13	23	18	75
Cluster III	8	26	11	19	64
Cluster IV	37	18	34	21	110
Cluster V	31	28	39	20	118
Cluster VI	24	27	24	26	101
Cluster VII	19	21	24	16	80
Cluster VIII	42	48	33	41	164
Cluster IX	37	32	39	35	143
Cluster X	82	80	81	64	307
Total	327	335	333	297	1292

Tanenhaus, 1998; Boland and Blodgett, 2001). Such reanalysis continued even to the next region, which can be interpreted as the spillover of integration demands for the previous anomaly (Rayner et al., 1989; Rayner, 1998). However, for the abnormal word order, there seemed to be neither lexical access difficulty nor enhanced reanalysis at the end of the sentence and on re-reading, where sentences with the word order violation were read as fast and smooth as sentences with no word order violation. Results of the offine tasks also showed no impairment of comprehensibility by the single violation of word order, implying that the rapid local reanalysis at the compound was very

Note that, e ects for the abnormal rhythmic pattern cannot be simply explained away as being due to the absence of co-occurrence between the disyllabic verb and the monosyllabic object noun. Abnormal and normal rhythmic patterns used essentially the same words, di ering only on one morpheme (see the RHY+ and RHY− conditions in Table 2). More importantly, if the e ects were due to the absence of co-occurrence, one would expect to observe similar e ects for di erent types of violations in either oculomotor or ERP responses, an expectation not confirmed by the findings in either the present study or Luo and Zhou (2010). Orthographical or segmental di erences between the disyllabic noun in the RHY+ORD+ condition and the monosyllabic noun in the RHY−ORD+ condition could not provide a tenable account for the rhythmic pattern e ects either, since the e ects were observed on the verb and regions downstream, which were visually and phonologically undi erentiated between conditions. Moreover, this account would predict prolonged viewing times for the combinations with normal rhythmic pattern (2-character noun) than those with abnormal rhythmic pattern (1-character noun) during the first pass reading (Rayner and Raney, 1996; Wang et al., 1999) or for regions (Regions, 2, 3, and 4) following the lower frequency 2-character noun than following the higher frequency 1-character noun, apparently contradicting what we observed in this study.

Local and immediate reanalysis of the abnormal rhythmic pattern also manifested itself as more regressive eye movements launched from Region 2 before reading on, presumably for further confirmation of the perceived information and for repair of the mismatching prosodic structure. Similar patterns were further observed on the subsequent head noun of the compound regardless of whether the word order was correct or not. The prosodic violation caused di culty in lexical access for the unambiguous head noun of the compound, suggesting that the expectation toward the target word based on rhythmic pattern, which would normally facilitate the processing of the upcoming word, was disrupted. But after carrying out the reanalysis for the whole compound, readers seemed to no longer su er from failing to generate the incremental lexical expectation, as indicated by the null e ect on gaze duration in post-compound Region 4, although increased regressive eye movements remained for reanalysis of contextual integration.

Unexpectedly, when readers continued on and approached the end of the sentence for the first time, the preceding rhythmic pattern violation seemed to facilitate, rather than interfere with the later stage of the sentence comprehension. This observation was confirmed by both approaches of data analysis: the conventional measures revealed o5(u)0ntafr o

Given these unexpected findings for rhythmic pattern, how should we compare the current study with previous research? Do they merely suggest the unique processing for a specific prosodic property, or do they point to more general but subtle mechanisms in language comprehension? We believe that our findings may provide a possible perspective for interpreting the similarities and dissimilarities between different types of information, i.e., the recoverability when a certain constraint is violated. Indeed, manipulations of syntactic information usually have significant impacts upon the buildup of the main predicate structure of the sentence and even cause harm to the comprehension (Frazier and Rayner, 1982; MacDonald et al., 1994; McRae et al., 1998; Christianson et al., 2001; Ferreira et al., 2001; Sanford and Sturt, 2002). Similarly, the prosodic properties that were investigated in the previous reading studies could also lead to a syntactic garden-path if mistakenly represented (Kentner, 2012; Breen and Clifton, 2013). By contrast, the rhythmic pattern here does not remarkably affect the representation at the sentential-meaning level in the given context, presumably because the relatively intact visual cues, with only one semantically redundant morpheme missing from the original, correct V-O combination, would suffice for lexical access and syntactic parsing. The "Above and Beyond" approach is assumed to be applied to recover from the rhythmic pattern anomaly which is less damaging to parsing than some other prosodic properties studied so far. On the other hand,

Eckstein, K., and Friederici, A. D. (2006). It's early: event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *J. Cogn. Neurosci.* 18, 1696–1711. doi: 10.1162/jocn.2006.18.10.1696

Feng, S. (1997). *Han yu de yun lü, ci fa yu ju fa (The Rhythm, Morphology, and Grammar of Chinese) (in Chinese).* Beijing: The Peking University Press.

Feng, S. (2004). Verb-object inversion and prosodic morphology. (in Chinese). *Ling. Sci.* 3, 12–20. Available online at: http://www.linsci.com/EN/article/downloadArticleFile.do?attachType=PDF&id=371

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cogn. Psychol.* 47, 164–203. doi: 10.1016/S0010-0285(03)00005-7

Ferreira, F., Bailey, K. G. D., and Ferraro, V. (2002). Good-enough representations in language comprehension. *Curr. Dir. Psychol. Sci.* 11, 11–15. doi: 10.1111/1467-8721.00158

Ferreira, F., Christianson, K., and Hollingworth, A. (2001). Misinterpretations of garden-path sentences: implications for models of sentence processing and reanalysis. *J. Psycholinguist. Res.* 30, 3–20. doi: 10.1023/A:1005290706460

Fodor, J. D. (1998). Learning to parse? *J. Psycholinguist. Res.* 27, 285–319. doi: 10.1023/A:1023258301588

Fodor, J. D. (2002). "Prosodic disambiguation in silent reading," in *Proceedings of the North East Linguistic Society 32* (Amherst, MA: GLSA), 113–132.

Fraley, C., and Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *J. Am. Stat. Assoc.* 97, 611–631. doi: 10.1198/016214502760047131

Fraley, C., and Raftery, A. E. (2007). *MCLUST Version 3 for R: Normal Mixture Modeling and Model-Based Clustering.* (Tech. Rep. No. 504), Department of Statistics, University of Washington, Seattle, WA.

Frazier, L., and Rayner, K. (1982). Making and correcting errors during sentence comprehension: eye movements in the analysis of structurally ambiguous sentences. *Cogn. Psychol.* 14, 178–210. doi: 10.1016/0010-0285(82)90008-1

Hirose, Y. (2003). Recycling prosodic boundaries. *J. Psycholinguist. Res.* 32, 167–195. doi: 10.1023/A:1022448308035

Hsieh, Y., Boland, J. E., Zhang, Y., and Yan, M. (2009). Limited syntactic parallelism in Chinese ambiguity resolution. *Lang. Cogn. Proc.* 24, 1227–1264. doi: 10.1080/01690960802050375

Huestegge, L. (2010). Effects of vowel length on gaze durations in silent and oral reading. *J. Eye Mov. Res.* 3, 1–18. Available online at: www.jemr.org/online/3/5/5

Hwang, H., and Schafer, A. J. (2009). Constituent length affects prosody and processing for a dative NP ambiguity in Korean. *J. Psycholinguist. Res.* 38, 151–175. doi: 10.1007/s10936-008-9091-1

Hwang, H., and Steinhauer, K. (2011). Phrase length matters: the interplay between implicit prosody and syntax in Korean "garden path" sentences. *J. Cogn. Neurosci.* 23, 3555–3575. doi: 10.1162/jocn_a_00001

Jun, S. A. (2003). Prosodic phrasing and attachment preferences. *J. Psycholinguist. Res.* 32, 219–249. doi: 10.1023/A:1022452408944

Jun, S., and Kim, S. (2004). "Default phrasing and attachment preferences in Korean," in *Proceedings of INTERSPEECH-ICSLP (International Conference on Spoken Language Processing)* (Jeju).

Kentner, G. (2012). Linguistic rhythm guides parsing decisions in written sentence comprehension. *Cognition* 123, 1–20. doi: 10.1016/j.cognition.2011.11.012

Kentner, G. (2015). "Stress clash hampers processing of noncanonical structures in reading," in *Rhythm in Cognition and Grammar: A Germanic Perspective, eds R. Vogel and R. Vijver* (Berlin: Walter de Gruyter GmbH and Co KG), 111–136.

Knaus, J., Wiese, R., and Janßen, U. (2007). "The processing of word stress: EEG studies on task-related components," in *Proceedings of the International Congress of Phonetic Sciences* (Saarbrücken), 709–712.

Kruskal, J. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29, 1–27. doi: 10.1007/BF02289565

Levenshtein, V. (1966). Binary codes capable of correcting deletions and insertions and reversals. *Sov. Phys. Dokl.* 10, 707–710.

Lukatela, G., Eaton, T., Sabadini, L., and Turvey, M. T. (2004). Vowel duration affects visual word identification: evidence that the mediating phonology is phonetically informed. *J. Exp. Psychol. Hum.* 30, 151–162. doi: 10.1037/0096-1523.30.1.151

Luo, Y., Yan, M., Yan, S., Zhou, X., and Inhoff, A. W. (2015). Syllabic tone articulation influences the identification and use of words during Chinese

sentence reading: evidence from ERP and eye movement recordings. *Cogn. Affect. Behav. Neurosci.* doi: 10.3758/s13415-015-0368-1. [Epub ahead of print].

Luo, Y., Yan, M., and Zhou, X. (2013). Prosodic boundaries delay the processing of upcoming lexical information during silent sentence reading. *J. Exp. Psychol. Learn.* 39, 915–930. doi: 10.1037/a0029182

Luo, Y., and Zhou, X. (2010). ERP evidence for the online processing of rhythmic pattern during Chinese sentence reading. *Neuroimage* 49, 2836–2849. doi: 10.1016/j.neuroimage.2009.10.008

MacDonald, M. C., Pearlmutter, N. J., and Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychol. Rev.* 101, 676–703. doi: 10.1037/0033-295X.101.4.676

Magne, C., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., and Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken French: behavioral and electrophysiological evidence. *Cereb. Cortex* 17, 2659–2668. doi: 10.1093/cercor/bhl174

Malsburg, T., Kliegl, R., and Vasishth, S. (2014). Determinants of scanpath regularity in reading. *Cogn. Sci.* 39, 1675–1703. doi: 10.1111/cogs.12208

Malsburg, T., and Vasishth, S. (2011). What is the scanpath signature of syntactic reanalysis? *J. Mem. Lang.* 65, 109–127. doi: 10.1016/j.jml.2011.02.004

McEnery, T., and Xiao, Z. (2004). "The lancaster corpus of mandarin Chinese: a corpus for monolingual and contrastive language study," in *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC) 2004* (Paris: ELRA/ELDA), 1175–1178.

McRae, K., Spivey-Knowlton, M. J., and Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *J. Mem. Lang.* 38, 283–312. doi: 10.1006/jmla.1997.2543

Meseguer, E., Carreiras, M., and Clifton, C. Jr. (2002). Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Mem. Cogn.* 30, 551–561. doi: 10.3758/BF03194956

Mitchell, D. C., Shen, X., Green, M. J., and Hodgson, T. L. (2008). Accounting for regressive eye-movements in models of sentence processing: a reappraisal of the Selective Reanalysis hypothesis. *J. Mem. Lang.* 59, 266–293. doi: 10.1016/j.jml.2008.06.002

Nespor, M., and Vogel, I. (1986). *Prosodic Phonology.* Dordrecht: Foris.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124, 372–422. doi: 10.1037/0033-2909.124.3.372

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Q. J. Exp. Psychol.* 62, 1457–1506. doi: 10.1080/17470210902816461

Rayner, K., and Raney, G. E. (1996). Eye movement control in reading and visual search: effects of word frequency. *Psychon. Bull. Rev.* 3, 245–248. doi: 10.3758/BF03212426

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., and Clifton, C. Jr. (1989). Eye movements and on-line language comprehension processes. *Lang. Cogn. Proc.* 4, SI21–SI49. doi: 10.1080/01690968908406362

R-Development Core Team (2009). *R: A Language and Environment for Statistical Computing.* ISBN 3-900051-07-0. Vienna: R Foundation for Statistical Computing. Available online at: http://www.R-project.org

Rothermich, K., Schmidt-Kassow, M., and Kotz, S. A. (2012). Rhythm's gonna get you: regular meter facilitates semantic sentence processing. *Neuropsychologia* 50, 232–244. doi: 10.1016/j.neuropsychologia.2011.10.025

Rothermich, K., Schmidt-Kassow, M., Schwartze, M., and Kotz, S. A. (2010). Event-related potential responses to metric violations: rules versus meaning. *Neuroreport* 21, 580–584. doi: 10.1097/WNR.0b013e32833a7da7

Sanford, A., and Sturt, P. (2002). Depth of processing in language comprehension: not noticing the evidence. *Trends Cogn. Sci.* 6, 382–386. doi: 10.1016/S1364-6613(02)01958-7

Scherer, K. R., Ladd, D. R., and Silverman, K. E. A. (1984). Vocal cues to speaker affect: testing two models. *J. Acoust. Soc. Am.* 76, 1346–1356. doi: 10.1121/1.391450

Schmidt-Kassow, M., and Kotz, S. A. (2009). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *J. Cogn. Neurosci.* 21, 1693–1708. doi: 10.1162/jocn.2008.21153

Schmidt-Kassow, M., Rothermich, K., and Kotz, S. A. (2015). "The role of default stress patterns in German monolingual and L2 sentence processing," in *Rhythm in Cognition and Grammar: A Germanic Perspective*, eds R. Vogel and R. Vijver (Berlin: Walter de Gruyter GmbH and Co KG), 83–110.

Selkirk, E. (2011). "The syntax-phonology interface," in *The Handbook of Phonological Theory, 2nd Edn*, eds J. Goldsmith, J. Riggle, and A. Yu (Oxford: Blackwell), 435–484.

Shi, D., Shu, H., and Zhang, H. (2000). Study of on-line processing of Chinese sentence. (in Chinese). *Acta Psychol. Sin.* 33, 7–12. Available online at: http://118.145.16.229:81/Jweb_xlxb/CN/Y2001/V33/I01/7

Spivey, M. J., and Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: modeling the e ects of referential context and lexical frequency. *J. Exp. Psychol. Learn.* 24, 1521–1543. doi: 10.1037/0278-7393.24.6.1521

Stolterfoht, B., Friederici, A. D., Alter, K., and Steube, A. (2007). Processing focus structure and implicit prosody during reading: di erential ERP e ects. *Cognition* 104, 565–590. doi: 10.1016/j.cognition.2006.08.001

Sturt, P. (2007). Semantic re-interpretation and garden path recovery. *Cognition* 105, 477–488. doi: 10.1016/j.cognition.2006.10.009

Swets, B., Desmet, T., Clifton, C. Jr., and Ferreira, F. (2008). Underspecification of syntactic ambiguities: evidence from self-paced reading. *Mem. Cogn.* 36, 201–216. doi: 10.3758/MC.36.1.201

Van Gompel, R. P. G., Pickering, M. J., and Traxler, M. J. (2001). Reanalysis in sentence processing: evidence against current constraint-based and two-stage models. *J. Mem. Lang.* 45, 225–258. doi: 10.1006/jmla.2001.2773

Venables, W. N., and Ripley, B. D. (2002). *Modern Applied Statistics with S, 4th Edn*. New York, NY: Springer.

Wang, H. (2008). *Han yu fei xian xing yin xi xue (The Nonlinear Phonology in Chinese)*. *(in Chinese)*. Beijing: The Peking University Press.

Wang, J., Chen, H. C., Radach, R., and Inho , A. (1999). *Reading Chinese Script: A Cognitive Analysis*. Mahwah, NJ: Erlbaum: Psychology Press.