

A corpus-based multidimensional computational analysis of episodic variation in Chinese relational clauses

Jiatian Sun¹, 

¹ *Language and Linguistic Science, University of York, Heslington, YO105DD, UK*

ABSTRACT

With the purpose of exploring the mechanism of change in Chinese relational clauses, this paper firstly includes transitive verbs, intransitive verbs and adjectives in the study of relational clauses, and carries out a comparative analysis from the perspectives of syntactic form, semantic expression, and distribution of thesis elements, and finds that relational clauses constituted by transitive verbs are indeed the most typical members of Chinese relational clauses. Then, we examine its performance in the type of relativization, main clause syntactic position of core words, vitality pattern, and structural features, and conclude that the argument elements of the relational clauses present a vitality contrast pattern and have a simpler structure with an average of about 4 syllables, while the distribution of the central words of the Chinese relational clauses conforms to the order of the noun-dominant syntactic position. Finally, ERP technology is used to explore the processing advantages of subject-relative clauses and to regulate the vitality and denotation of the verbal thesis elements of the clauses, and it is found that the difference in processing difficulty between subject and object-relative clauses increases when the subject of the clauses is a vital noun and the object is a non-vital noun.

Keywords: ERP technology, Chinese relational clauses, corpus

1. Introduction

The development and shaping of Modern Chinese began with the literary revolution of the May Fourth period and reached maturity around 1930. It is generally believed that the ephemeral changes of modern Chinese are grammatical changes, which mainly involve both lexical and syntactic levels, including lexical changes, structural changes, and the evolution of grammatical systems [1-2]. Chinese academics have made a great deal of efforts in establishing the theory of grammaticalization reflecting the characteristics of the Chinese language, as well as exploring the mechanism of Chinese grammaticalization and the laws of evolution [3-4]. For a long time, Chinese language evolution studies have mostly focused on observing the changes in the use of specific language items, with little attention

 Corresponding author.

E-mail address: 756542784@qq.com (J. Sun).

Received 20 February 2024; Revised 15 April 2024; Accepted 20 November 2024; Published Online 16 April 2025.

DOI: [10.61091/jcmcc127b-390](https://doi.org/10.61091/jcmcc127b-390)

© 2025 The Author(s). Published by Combinatorial Press. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

paid to the systematic nature of language evolution and a lack of overall description of Chinese language evolutionary features [5]. In addition, studies on the evolution of Chinese over time generally examine narrative texts, which are genre-specific and limited in scope, lacking the support of a stylistically and genre-balanced ephemeral corpus, and failing to reflect the changing landscape of modern Chinese at multiple levels [6-8]. In view of this, this study uses an ephemeral balanced corpus to try to describe the ephemeral changes of modern Chinese relational clauses from multiple perspectives based on syntactic structures.

Chinese relational clauses are also called relational clauses, definite clauses or adjective clauses, which are usually used to modify nouns [9]. Chinese relational clauses can co-occur with “indicative + quantifier”, and the co-occurrence will involve the issue of the order of the two, in which the academic community usually focuses on the order of the subject and object relational clauses [10-11]. Subject-relative clauses extract the subject, with the central noun referring to the same trace left by the subject, while object-relative clauses extract the object of the transitive predicate verb, with the central noun referring to the same trace left by the direct object [12-13]. In order to better observe the tendency of subject clauses and object clauses in prepositioning and postpositioning, corpus statistics are introduced to provide a research path for the temporal evolution of features such as the pattern of positional order in Chinese relational clauses [14-17].

Starting from a real corpus, the study introduces different types of predicates to reveal the different features they exhibit in terms of syntactic-semantic structure and part-of-speech distribution, as well as the tendency of the transitive verb theorist in terms of number and role omission. Subsequently, the distributional dynamics of Chinese postpositive relational clauses is analyzed in terms of the accessibility of the relativization type, the main clause syntactic position of the core word, and the vitality pattern of the relational clauses, among other covariates. ERP techniques are used to examine the neural processing mechanisms of Chinese relational clauses and the role of the influence of the vitality and denotationality of the verbal thesis elements of the clauses.

2. Corpus-based analysis of the change patterns of Chinese relational clauses over time

2.1. Corpus and methodology

2.1.1. Description of the corpus. Lao She's famous novel “The Four Families - Famine” (hereinafter referred to as “Famine”) and Wang Shuo's work “I Am Your Dad” (hereinafter referred to as “Dad”) were used as corpus sources. Both Lao She's and Wang Shuo's novels belong to the Beijing flavor novels and are written in the excellent Beijing language, but Lao She's novels represent the style of Beijing dialect half a century ago, while Wang Shuo's novels represent the Beijing dialect since the 1990s, so both of them also have different linguistic characteristics, which is conducive to the study of the comprehensive examination of the Chinese language's relational clauses.

In selecting the works, attention was paid to consider the unity of the works in terms of word count and narrative style. “Famine” has a word count of 196,000 words and “Dad” has a word count of 203,000 words, and both novels use third-person narration, and both are in narrative style. Only the narrative portion of the novels will be examined, not the dialogic portion of them.

2.1.2. Predicate types. The types of predicates in Chinese relative clauses include: transitive verbs, intransitive verbs and predicative adjectives. Chinese adjectives have both predicative and adjectival properties, with predicative being stronger than adjectival. Among them, predicative adjectives are unmarked as predicates and marked as determiners.

The types of predicates in Chinese relational clauses include the following three categories:

1) Transitive verbs. This type of verb carries at least two arguments, and there is also a high or low level of transitivity within it.

2) Intransitive verbs. These types of verbs have only one argument element.

3) Adjectives. This type of predicates carries one argument element. Adjectives in Chinese have stronger transitive properties than tensed verbs, because adjectives, like typical verbs, can have body markers, can be modified by adverbs, can carry resultant complements, and so on.

2.2. Distribution of Chinese relational clauses

2.2.1. Distribution of Relational Clause Predicates in Chinese. The distribution of Chinese relational clauses is examined according to the predicate types of relational clauses, i.e., transitive verbs, intransitive verbs and adjectives, as shown in Tables 1 and 2. Tables 1 and 2 reveal the distribution of predicate types of relational clauses in Chinese narrative corpora, with the proportions of transitive verbs and adjectives being relatively close to each other, with the proportion of adjective predicates being higher than that of transitive verbs in “Famine” and the proportion of transitive verbs being slightly higher than that of adjectives in “Dad”. The proportions of intransitive verbs are closer in “Famine” and “Dad” at 13.7% and 12.3% respectively.

Table 1. The predicate type distribution of small sentences in famine

Predicate type	Total	Proportion
Transitive	379	38.8%
Intransitive verb	134	13.7%
Adjective	459	47.0%
Special structure	5	0.5%
Total	977	100.0%

Table 2. The predicate type distribution of the small sentence in “dad”

Predicate type	Total	Proportion
Transitive	517	42.6%
Intransitive verb	149	12.3%
adjective	523	43.0%
Special structure	26	2.1%
Total	1215	100.0%

Next, the distribution of predicates in Chinese relational clauses will be further examined according to the syntactic morphology of predicates, i.e., single forms and complex forms. The single forms refer to the bare-bones forms, i.e., the types of predicates formed by single transitive verbs, intransitive verbs and adjectives; the complex forms refer to the types of predicates formed by syntactic morphological means such as overlapping, adding arguments and body markers, etc. Some common linguistic facts are also found, as shown in Tables 3 and 4.

It can be noticed that the statistical results in Tables 3 and 4 are very similar, although “Famine” and “Dad” are from different writers and half a century apart in time span, there is a high degree of similarity in terms of linguistic facts (in the distribution of predicates in relational clauses). The distribution of single and complex forms of transitive and intransitive verbs is almost the same. Transitive verbs seldom adopt single forms, i.e., bare-rod forms, and only 2.1% of transitive verbs adopt bare-rod forms to form relational clauses, while 97.9% of transitive verbs adopt complex forms to form relational clauses. Intransitive verbs are less likely to adopt single forms, with 21.8% of single forms and 78.2% of complex forms in “Famine” and 21.2% of single forms and 77.8% of complex forms in “Dad”, which suggests that both Chinese transitive and intransitive verbs tend to adopt complex forms when they actually form relational clauses, especially transitive verbs.

Table 3. A single form and complex form of the small sentence predicate in famine

Predicate type		Quantity	Proportion
Transitive	Single form	8	2.1%
	Complex form	371	97.9%
Intransitive verb	Single form	29	21.8%
	Complex form	104	78.2%
Adjective	Single form	201	42.9%
	Complex form	267	57.1%
Total		980	

Table 4. The single form and complex form of the small sentence predicate in "dad"

Predicate type		Quantity	Proportion
Transitive	Single form	11	2.1%
	Complex form	511	97.9%
Intransitive verb	Single form	35	21.2%
	Complex form	123	77.8%
Adjective	Single form	245	48.8%
	Complex form	257	51.2%
Total		1167	

2.2.2. The Distribution of Thematic Elements in Chinese Relational Clauses:

1) Omitted and complete tenses. Since the relativization operation must extract a must-have thesis element, there is no thesis element omission for both adjectives and intransitive verbs, and only transitive verbs have the possibility to omit one of the thesis elements. The distribution of thesis elements of transitive verbs in the formation of relational clauses will be examined next as shown in Table 5. The results show that "Famine" and "Dad" have very similar distributions, with nearly three-quarters of the transitive verbs in the perfect tense in the relational clauses, with no case of argument element omission, and only one-quarter of the transitive verbs will be found to omit one of the argument elements.

Table 5. The theory of the origin of the verb

Source of corpus	Meta-type	Quantity	Proportion
“Famine”	Complete form	279	75%
	Elliptical form	95	25%
	Tot	374	100%
“Dad”	Complete form	376	74%
	Elliptical form	132	26%
	Tot	508	100%

2) Grammatical roles of ellipsis AvsO. The core thesis elements carried by the transitive verbs, i.e., the giver A and the receiver O, are examined one by one for the omitted forms in “Famine” and “Dad” as shown in Table 6. An interesting phenomenon can be found, in both corpora the case of A as an implicit theorist is very common, while the case of omitting O is rare, i.e., the chances of O obtaining explicit theorists are much greater than that of A.

Table 6. The explicit argument of a and o in the verb

Source of corpus	Meta-type	Quantity	Proportion
“Famine”	A	5	5.1%
	O	93	94.9%
	Tot	98	100%
“Dad”	A	4	3.1%
	O	126	96.9%
	Tot	130	100%

2.3. Analysis and comparison of the change patterns of the corpus over time

By using Wconcord software to analyze the corpus text, this study finally extracted and filtered to 526 relational clauses from the corpus for analysis in this paper. All the relational clauses were automatically and manually annotated by ICTCLAS2002 software according to the following three features: the syntactic position of the central word in the main clause, the type of the extracted clauses, and the vitality of the central word and its cognate nouns within the clauses.

2.3.1. Distributional representation of relationalization types. First of all, let's look at the specific distribution of the extracted types of relational clauses. It mainly involves three types of postpositive relational clauses: subject-relative clauses (SRC), object-relative clauses (ORC) and genitive-relative clauses (GRC). Their distribution is shown in Table 7. The table shows that out of 155 cases of postpositive relational clauses, there are only 4 cases of collocative-genitive relational clauses, whereas subject-relational clauses account for about 61.9% (96/155) and object-relational clauses account for about 35.5% (55/155). Subject-relative clauses are significantly more distributed than object-relative clauses ($\chi^2_{(1, N=155)} = 12, p < 0.005$). This result is in general agreement with the results of previous written corpus examinations of Chinese preterite relational clauses, and is also consistent with the predicted direction of noun phrase accessibility hierarchies. It also shows that different inflectional strategies for Chinese relational clauses are consistent with the accessibility rank based on syntactic position.

Table 7. Distribution of relational types

Relational type	Quantity	
	N	%
SRC	96	61.9
ORC	55	35.5
GRC	4	2.6
Total	155	100

2.3.2. **Distributional Representation of the Syntactic Position of Core Noun Main Clauses.** Then we look at the distribution pattern of the syntactic position of the core nouns (HN) of these three types of postpositive relational clauses in the main clauses as shown in Table 8. Of the 155 cases of postpositive relational clauses, 12 cases modify the subject of the main clauses, which accounts for about 7.74%; 122 cases modify the object of the main clauses (including 32 cases of verbal objects, 34 cases of phrasal objects, and 56 cases of present objects), which accounts for about 78.71%; and 21 cases modify the other constituents (i.e., prepositions object, topic), totaling 13.55%. Relational clauses are significantly more dominant in modifying the object of the main clause than in modifying the subject of the main clause ($\chi^2_{(1, N=130)} = 87.28, p < 0.001$). This tends to agree with the observations of the corpus and the novel corpora that the difference may be related to differences in stylistic features.

Table 8. The distribution of the main sentence method of the core noun

	Subject		Object		Categorical object		Present object		Preposition object		Topic		Total	
SRC	9	9.4%	23	23.96%	16	16.67%	35	36.46%	8	8.33%	5	5.21%	96	100%
ORC	3	5.45%	7	12.73%	16	29.09%	21	38.18%	5	9.09%	3	5.45%	55	100%
GRC	0	0	2	50%	2	50%	0	0	0	0	0	0	4	100%
Total	12	7.74%	32	20.65%	34	21.94%	56	36.13%	13	8.39%	8	5.16%	155	100%

2.3.3. **Structural features and discourse functions.** This section examines the occurrences of the thesis element component of the transitive verb in postpositive relational clauses, the occurrences of body markers, whether other definite components appear before the core noun, the discourse function of the relational clauses, and the average length of the clauses. The specific results are shown in Table 9 (A1: Whether or not an argumentative element occurs in the subordinate relative verb, A2: Whether or not the subordinate clause contains a body marker, and A3: Whether or not other definite constituents occur before the core noun).

The tendency of thesis element constituents of transitive verbs to appear in postpositive relational clauses. Of the 136 cases of transitive postpositive relational clauses, the thesis element explicitly appears within the clause in 109 cases, which accounts for about 81.55% of the cases. This is consistent with the typical performance of preterite relational clauses in Chinese.

Table 9. The structural characteristics and discourse function of the posterior relation clause

Relative clause	A1				A2				A3				Mean length
	YES		NO		YES		NO		YES		NO		
	N	%	N	%	N	%	N	%	N	%	N	%	
	109	81.55	32	22.45	7	.58	146	95.42	125	83.89	24	116.11	4

2.3.4. Trends in the Distribution of Chinese Relative Clauses and Their Center Phrases. According to the types of Chinese relational clauses and the syntactic positions of the center words in the main clauses, the distribution of all the relational clauses extracted from the corpus is summarized as shown in Table 10. In terms of the distributional trend of syntactic positions of argument elements, the central words of the relationalized argument elements have obvious subject and object tendencies. The subject position carries the most relational clauses (45.22%), followed by direct object (38.84%), and thereafter epithets, prepositional objects, existential objects, genitive case, etc. in order. The highest percentage of subject-relative clauses (50.03%) was extracted from the corpus, followed by object-relative clauses (44.84%), and thereafter genitive and comparative object-relative clauses in that order.

Table 10. The overall distribution of Chinese relation clauses and their central words

Central word	Subject clause	Object clause	Lattice	Comparative object	Total
Subject	145	85	8	3	241(45.22%)
Object	93	103	6	5	207(38.84%)
Watchword	22	22	-	-	44(8.26%)
Preposition object	9	25	-	-	34(6.38%)
Existential object	2	2	-	-	4(0.85%)
Lattice	1	2	-	-	3(0.56%)
Total	272(50.03%)	239(44.84%)	14(2.63%)	8(1.5%)	533(100%)

2.3.5. The Effect of Vitality on the Distributional Dynamics of Chinese Relative Clauses. In this section, the constraining power of the vital factors of the thesis elements in the sentence on the choice of the type of relational clauses will be examined. The data results are shown in Table 11. When the center element refers to life, whether it modifies the subject or object of the main clause, the proportion of subjective relational clauses is higher than 85% or more, which is absolutely dominant and is the main modification means; accordingly, the proportion of objective relational clauses modifying these two types of center elements is extremely low, only around 10%. When the center word refers to inanimate, the situation is diametrically opposite. Regardless of whether they modify the subject or the object center word, the types of clauses are dominated by object-relative clauses, with a high percentage of 82% of object-relative clauses modifying inanimate subjects and 68% of those modifying objects, while subject-relative clauses occur less frequently. In addition, there were twice as many lifeless centralizers in the object position as in the subject position. The most typical center word clause pattern produced at this point is OO, which shows that the lifeless noun always tends to be subject role and take the object in the sentence.

In summary, the vitality of the referent of the center word of Chinese relational clauses has a very high correlation with the choice of the type of the clause and its center word. The two typical skewed distributions of Chinese relative clauses and their center words according to their vitality fully verify that animate objects tend to be doer subjects and inanimate objects tend to be receptive objects.

Table 11. Four typical relational clauses and their central words

Number of central words	Subject clause		Object clause		Total
	Percentage	Number	Percentage	Number	

In life	Subject	140	90%	15	10%	155
	Object	55	87%	8	13%	63
Inanimate	Subject	15	18%	68	82%	83
	Object	45	32%	96	68%	141

3. Analysis of the strengths of episodic variation in Chinese relational clauses

3.1. Research methodology

In this section, we examine the neural mechanisms of Chinese clause processing by using neuroelectrobiological equipment to record subjects' brain waves during the reading of Chinese relational clauses and analyze the event-related potential (ERP) components. That is, by comparing the positive and negative wave components (mainly including N400, P600) evoked during the processing of subject and object relational clauses, we compare the size of working memory resources consumed by the two types of clauses, so as to determine whether subject relational clauses, which are used more frequently, are easier to process than object relational clauses.

3.1.1. Subjects. Twenty (10 males and 10 females) college student volunteers were selected to participate in the experiment, aged between 22-27 years old, with a mean age of 24.51. All subjects were right-handed, with normal or corrected normal vision, volunteered to participate in this experiment and agreed to cooperate with the experiment in its entirety.

3.1.2. Experimental materials. The stimulus sentences used in the experiment consisted of 37 subject-relative clauses (SR) and 37 object-relative clauses (OR), and the selected corpus was strictly controlled in structure. Each stimulus sentence is composed of two parts, the first part is a phrase, which provides a reasonable situation for the sentence, makes the meaning of the sentence more complete and natural, and also avoids the relative clause part as a stimulus appearing in the first word of the sentence; The latter part is a nested clause of the sentence, all of which are of the "noun main clause verb phrase" structure of the clause, where the SR clause is the "clause verb noun" structure and the OR clause is the "noun clause verb" structure. The two types of relative clauses are consistent except for the clause part. Since the influence of vital factors on the processing of clauses is not discussed in this experiment, the arguments of clause verbs are all vital nouns. As:

[a] The robbers of the noblewoman were taken to the police station after the crime. (SR)

[b] After the incident, the robber of the noblewoman was taken to the police station. (OR)

The 74 stimulus sentences were investigated for acceptability and the questionnaire respondents were asked to rate the sentences on a scale of 1-5, with 1 indicating unintelligible and 5 indicating easily understandable. The SR clauses obtained a mean value of 3.91 for acceptability and a mean value of 3.75 for OR, and the results of the repeated measures showed that there was no significant difference between the two types of stimulus sentences in terms of their acceptability, $F(1, 37) = 2.882$, $p = 0.095 > 0.05$. In addition to the 74 stimulus sentences, another 97 sentences of non-relational clauses were selected as filler sentences, which were consistent with the stimulus sentences in terms of structure and length, such as "Experts point out that drinking a glass of light salt water on an empty stomach in the morning is favorable for slimming down."

3.1.3. Experimental procedures. The subjects sat in front of a computer monitor with their eyes horizontally fixed in the center of the screen and their hands flat on the keyboard. All experimental sentences are presented word by word in the center of the screen, with Chinese characters in 48 bold font and white, and a black background. The prompt "" is presented before the first word of each sentence appears, indicating that the stimulus is about to be presented, the prompt presentation time is 1550ms, each word of the sentence is presented at 550ms, the interval between each stimulus word (SOA) is a black screen of 550ms, and the last word of the sentence is marked with a full stop, indicating

the end of the sentence. After each sentence is presented, a declarative sentence related to the semantic content of the sentence will be presented, and the participants are asked to judge as soon as possible whether the declarative sentence is consistent with the meaning of the sentence presented word by word, which is consistent with pressing the key marked "right" on the keyboard, and does not conform to pressing the "wrong" key. The probability of making a "right" and "wrong" judgment is equal, 50%, and it is randomly arranged. The time limit for making a judgment is 6 seconds, and if a judgment is not made within 6 seconds, it will be treated as a judgment error and will go directly to the presentation of the next sentence.

Before the start of the formal experiment, the computer presented an experimental guide explaining the experimental task, the actions to be performed, etc., and a short exercise to familiarize with the experimental procedure. The experiment was divided into 4 vignettes, and the presentation order of all 72 stimulus sentences and 96 filler sentences was pseudo-randomized to ensure that no two stimulus sentences were adjacent to each other. There were short breaks between each vignette, and the entire experimental manipulation lasted approximately 50 min.

3.1.4. Electroencephalographic recordings. Experiments were performed using NeuroScan's 32 conductive electrode caps, with data recorded from 30 main electrodes: *FP1, FP2, F7, F3, FZ, F4, F8, FT7, FC3, FCZ, FC4, FT8, T3, C3, CZ, C4, T4, TP7, CP3, CPZ, CP4, TP8, T5, P3, PZ, P4, T6, O1, OZ, O2*. The sampling rate was 250 Hz, the bandwidth was 0.05-40 Hz, and the electrode-scalp contact resistance dropped to 5 K Ω during the experiment. The recording reference electrode was placed in the left mastoid and converted to take the average of the bilateral mastoids as the reference for offline data analysis.

3.1.5. EEG data processing. Artifact subtraction, an early and most commonly used method for removing eye movement artifacts, is relatively simple and intuitive, and the removal effect is obvious [18].

Since the measured EEG contains an EOG component, we assume that the EEG is a linear combination of the real EEG and EOG, and that the EEG is not correlated with the EOG, and the artifact subtraction method is performed on the basis of this assumption. The calculation formulae, which are mainly based on this assumption, are as follows:

$$y_t(i) = y(i) - kx(i) - b \quad (1)$$

Where $y_t(i) = y(i) - kx(i) - b$, $t(i)$ denotes the denoised EEG; $y(i)$ denotes the measured EEG; $x(i)$ denotes artifacts; and k denotes the proportionality constant, i.e., corrected EEG is the EEG that removes a certain proportion of the EOG from the measured EEG signal.

Where the k value is commonly estimated by least squares linear regression, i.e.:

$$k = \frac{\sum ((x(i) - \overline{x(i)})(y(i) - \overline{y(i)}))}{\sum (x(i) - \overline{x(i)})^2} \quad (2)$$

Or:

$$k = r_{xy} \cdot \frac{SD_y}{SD_x} \quad (3)$$

where $\overline{x(i)}$ denotes the arithmetic mean, SD denotes the standard deviation, and r_{xy} denotes the correlation coefficients for the X , Y channels. In practice, the k value is calculated for each lead signal of each subject and then corrected by applying equation (1). In this case, $b = \overline{y(i) - kx(i)}$ denotes the intercept of the regression equation, and the subtraction of b is the baseline for the removal of the eye movement signal.

In order to effectively remove the eye movement artifacts, the signals from at least two mutually

orthogonal channels need to be measured, one for the horizontal eye movement channel and the other for the vertical eye movement channel, which in turn needs to be used to estimate the ocular artifacts using a multiple regression method. The formula for multiple regression calculation using the eye movement signals of the two channels is:

$$k_{yh,v} = \frac{r_{yh} - r_{yv}r_{vh}}{1 - r_{hv}^2} \cdot \frac{SD_y}{SD_x} \quad (4)$$

where k_{yh} , denotes the proportion of the horizontal eye movement signal $h(i)$ to the EEG signal taking into account the effect of the vertical eye movement $v(i)$ signal, and r denotes the Pearson product moment. Similarly, $k_{yv,h}$ can be calculated. The correction formula is:

$$y_t(i) = y(i) - k_{yh,v}h(i) - k_{yv,h}v(i) - b \quad (5)$$

3.2. Analysis of ERP results

3.2.1. ERP results under vitality conditions:

1) ERP of relational tokens. The ERP waveforms of relational markers are shown in Figure 1. The differences in ERPs induced by the relational marker “yes” for the four types of sentences were first shown in the 250-300 ms time window, with the OR-S raw O raw relational marker inducing a significantly larger negative component (AN) in the anterior brain regions. ANOVA results within this time window showed a borderline significant main effect of type, $F(1,13) = 4.27$, $p = 0.07$, with only the difference in AN evoked by OR-S-born O-born versus SR-S-born O-born reaching significance, whereas there was no significant difference in the negative waves evoked by OR-S-born O-born versus SR-S-born O-born ($p = 0.681$); the main effect of vitality was not significant, $F(1,13) = 0.86$, $p = 0.373$.

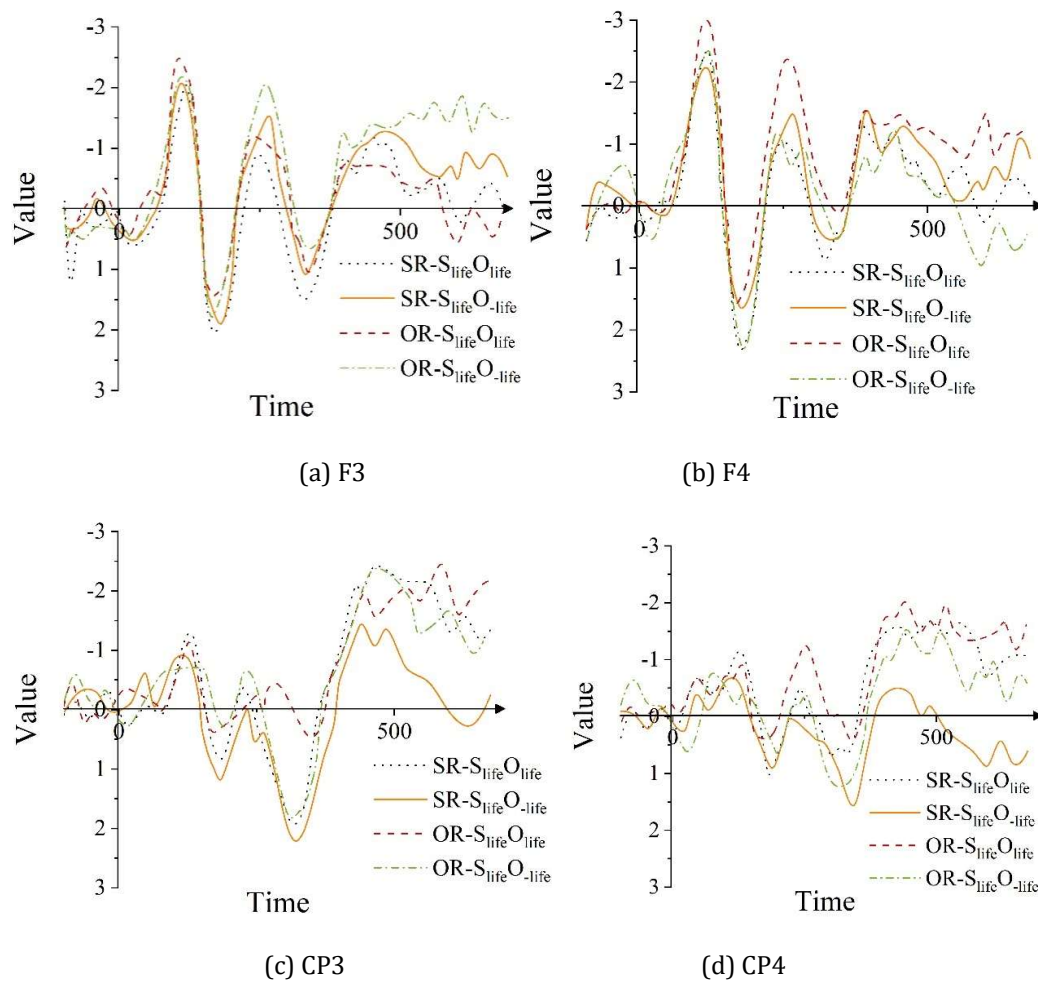


Fig. 1. ERP waveform of relational markup

2) ERP of clause-centered words. The ERP waveforms of the subordinate clause-centered phrases are shown in Fig. 2, and there were differences in the N400 evoked by the subordinate clause-centered phrases of the four types of sentences in the 350-450 ms time window, and the results of the ANOVA within this time window showed a significant main effect of type, $F(1,13) = 8.47$, $p = 0.014$, with OR-S-born O-borns evoking a larger N400 in posterior brain regions than SR-S-born O-born and OR-S-born O-born induced larger N400s in central and posterior brain regions than SR-S-born O-born, with the difference being particularly significant in the S-born O-born condition. The vitality main effect was marginally significant, $F(1,13) = 4.11$, $p = 0.066$, and the vitality $\times$ anterior and posterior brain regions interaction was significant, $F(2,26) = 9.31$, $p = 0.008$, with OR-S-born O-born inducing larger N400s than OR-S-born O-born in central and posterior brain regions, and SR-S-born O-born inducing no significant difference.

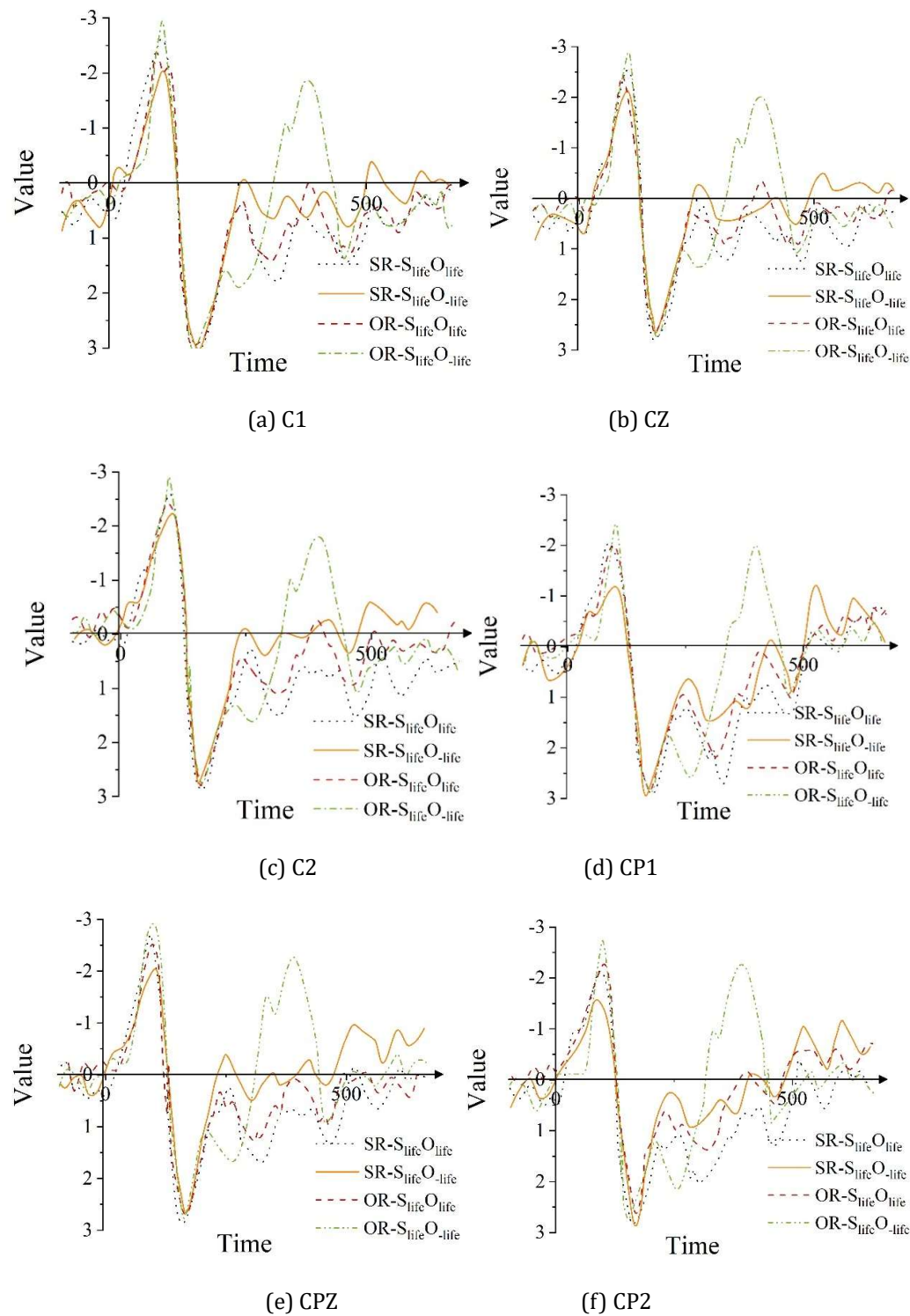


Fig. 2. ERP waveform of clause center language

3) ERP of main clause verbs. The ERP waveforms for main clause verbs are shown in Figure 3, with differences in P600 evoked in the 550-700ms time window for the four types of sentences, where the ANOVA did not reveal a significant type main effect, $F(1,13) = 0.029$, $p = 0.875$, and a non-significant vitality main effect, $F(1,13) = 2.91$, $p = 0.115$, but the type $\times$ vitality $\times$ anterior and posterior brain regions interaction was significant, $F(2,26) = 5.79$, $p = 0.027$. Further simple effects analyses showed that S-born O-born induced larger P600s in anterior brain regions than S-born O-born in the OR condition, and subordinate verbs in OR also induced larger P600s in anterior brain regions than SR in

the S-born O-born condition, but the differences were both only marginally significant ($0.09 > p > 0.05$).

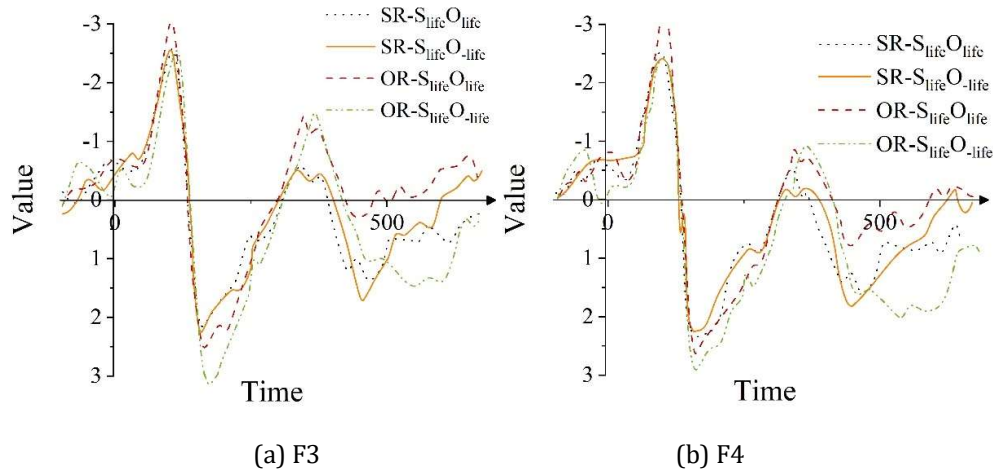


Fig. 3. ERP waveform of main verb

3.2.2. ERP results for the proportionality condition:

1) ERP of relational tokens. The ERP waveforms of relational markers are shown in Figure 4. The differences in ERPs induced by the relational marker “yes” for the four types of sentences were first shown in the 250-300ms time window, with the OR-S raw O raw relational marker inducing a significantly larger negative component (AN) in the anterior brain regions. ANOVA results within this time window showed a nonsignificant main effect of type, $F(1,13)=1.39$, $p=0.28$, a significant main effect of denotative, $F(1,13)=5.27$, $p=0.05$, and a significant type $\times$ denotative interaction, $F(1,13)=7.41$, $p=0.019$. Further simple effects analyses showed that in the OR condition, S-born O born induced a larger AN than S-generation O-born ($p=0.004$), and in the S-generation O-born condition, OR induced a larger AN than SR ($p=0.014$).

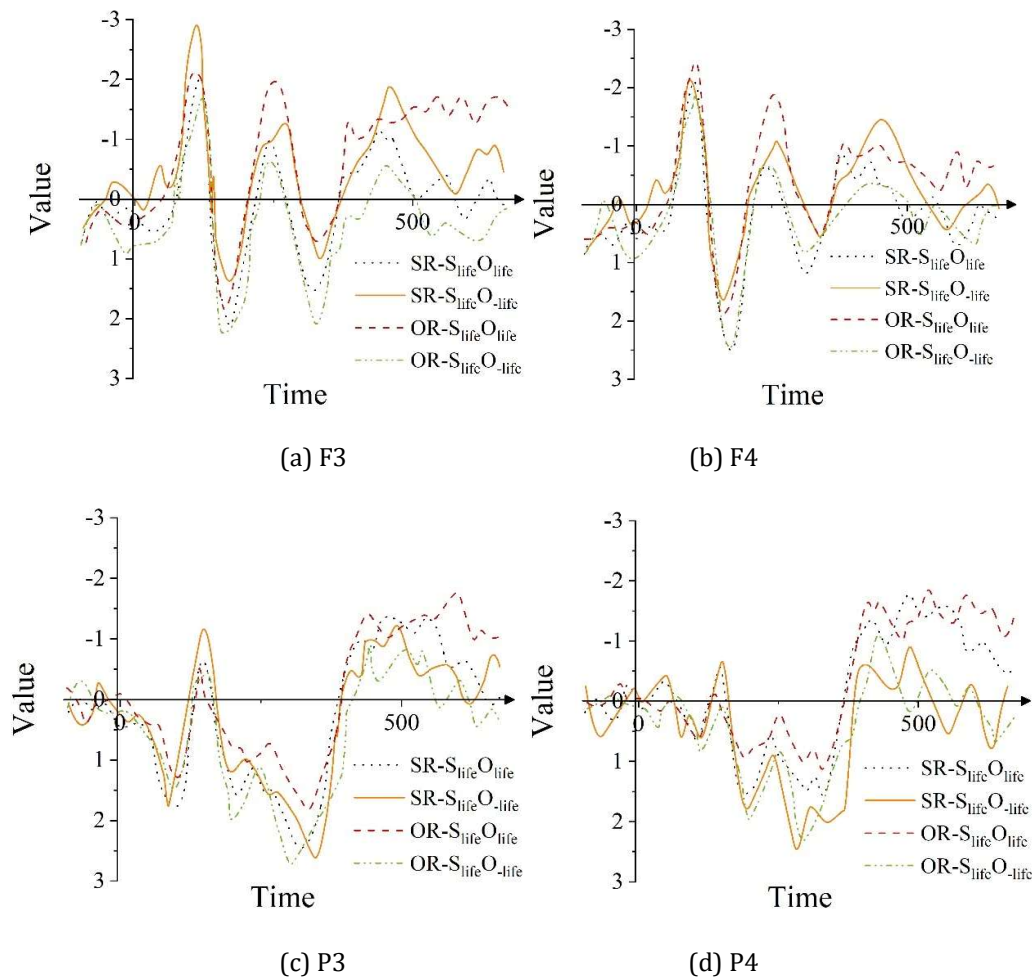


Fig. 4. ERP waveform of relational markup

2) ERP of Clause Center Phrases. The ERP waveforms of clause-centered phrases are shown in Fig. 5. There were some differences in the N400 induced by the clause-centered phrases of the four types of sentences in the 350-450 ms time window, and the ANOVA results showed a nonsignificant main effect of type, $F(1,13)=1.55$, $p=0.241$, and a nonsignificant main effect of referentiality, $F(1,13)=0.62$, $p=0.454$, but the type $\times$ accusative interaction was marginally significant, $F(1,13)=4.51$, $p=0.055$, and further simple effects analyses showed that only OR-S-born O-borns induced larger N400s than SR-S-born O-borns, and that other types did not show any differences in the ERP components induced.

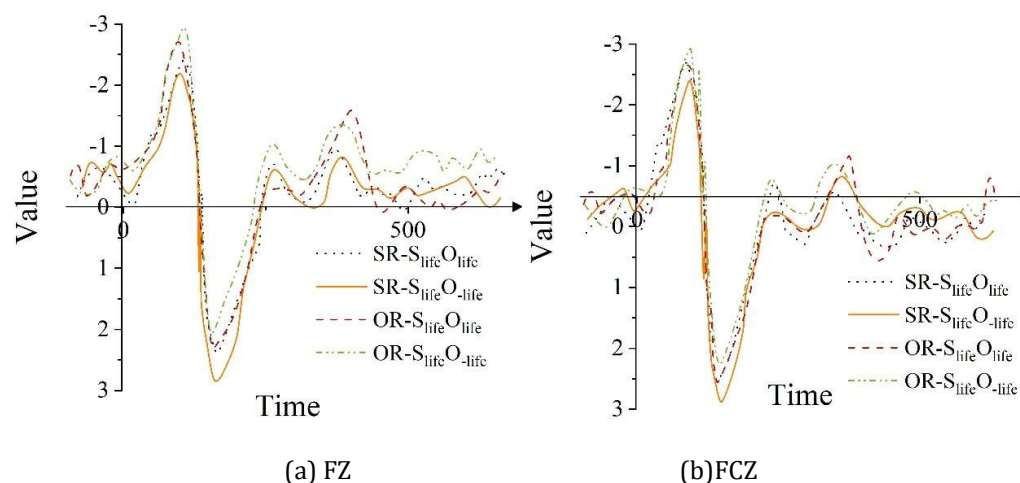


Fig. 5. ERP waveform of clause center language

4. Conclusion

This paper analyzes the utterance characteristics of Chinese relational clauses. In terms of semantic expression, it is considered that adjectives are attributive predicates, while transitive and intransitive verbs are characterized by both attributive predicates and event predicates; in terms of the distribution of thesis elements, transitive verbs are inclined to be perfective. Secondly, the distributional dynamics of Chinese postpositive relational clauses were examined, and it was found that the relativization type, vitality pattern, and clause length behaved consistently, both of which were more dominated by subject-relative clauses and contrastive vitality patterns. Finally, the neural processing mechanism of Chinese relational clauses was examined by controlling the syntactic structure and modulating the vitality and denotation of the verbal thesis elements of the clauses using the ERP technique, and it was found that Chinese subject-relative clauses were easier to process than object-relative clauses, and that the processing dominance of subject-relative clauses was linguistically typologically generalized.

References

- [1] Iurato, A. (2022). Investigating Chinese learner corpus research and learner corpora: Main features, critical issues and future pathways. *Kervan. International Journal of African and Asian Studies*, 26(1).
- [2] Lu, X., & Chen, B. (Eds.). (2019). *Computational and corpus approaches to Chinese language learning*. Singapore: Springer.
- [3] Yang, G., & Jia, M. (2025). Diachronic Research on Chinese Discourse Markers in China: Looking Back and Moving Forward. *SAGE Open*, 15(1), 21582440251321218.
- [4] Arcodia, G. (2022). On structural particles in Sinitic languages: typology and diachrony. *ANNALI DI CA'FOSCARI. SERIE ORIENTALE*, 58(1), 601-648.
- [5] He, X. (2021). Diachronic change and synchronic variation of the Chinese discourse marker shishishang. *East Asian Pragmatics*, 6(3), 283-301.
- [6] Chen, J., Chersoni, E., & Huang, C. R. (2022, May). Lexicon of changes: Towards the evaluation of diachronic semantic shift in Chinese. In *Proceedings of the 3rd Workshop on Computational Approaches to Historical Language Change* (pp. 113-118).
- [7] Li, Y. (2018). The diachronic development of the Chinese passive: From the wei... suo passive to the long passive. *Language*, 94(2), e74-e98.

- [8] Tantucci, V., & Wang, A. (2020). Diachronic change of rapport orientation and sentence-periphery in Mandarin. *Discourse Studies*, 22(2), 146-173.
- [9] Zhou, P., Yao, Y., & Christianson, K. (2018). When structure competes with semantics: Reading Chinese relative clauses. *Collabra: Psychology*, 4(1).
- [10] Hu, S., Gavarró, A., Vernice, M., & Guasti, M. T. (2016). The acquisition of Chinese relative clauses: contrasting two theoretical approaches. *Journal of child language*, 43(1), 1-21.
- [11] Hsiao, Y., & MacDonald, M. C. (2016). Production predicts comprehension: Animacy effects in Mandarin relative clause processing. *Journal of Memory and Language*, 89, 87-109.
- [12] Liu, Y., & Ni, C. (2024). How semantics works in Chinese relative clause processing: insights from eye tracking. *Frontiers in Psychology*, 15, 1294132.
- [13] Wu, F., Kaiser, E., & Vasisht, S. (2018). Effects of early cues on the processing of Chinese relative clauses: Evidence for experience - based theories. *Cognitive science*, 42, 1101-1133.
- [14] Chang, L. P. (2023). The development of relative clauses in L2 Chinese: A corpus-based study. In *Learner corpora: Construction and explorations in Chinese and related languages* (pp. 135-154). Singapore: Springer Nature Singapore.
- [15] Wang, F., & Wu, F. (2020). Postnominal relative clauses in Chinese. *Linguistics*, 58(6), 1501-1542.
- [16] Lin, C. J. C. (2018). Subject prominence and processing dependencies in prenominal relative clauses: The comprehension of possessive relative clauses and adjunct relative clauses in Mandarin Chinese. *Language*, 94(4), 758-797.
- [17] Morbiato, A. (2020). Chinese Sentence-Initial Indefinites: What Corpora Reveal. In *Corpus-Based Research on Chinese Language and Linguistics* (Vol. 6, pp. 59-92). Edizioni Ca'Foscari.
- [18] Chernyy Nick & Gluckman Bruce J. . (2010). Online Artifact Subtraction for Concurrent Neural Recording During Ongoing Electrical Stimulation. *Journal of Medical Devices*, 4(2),