

Chinese Causal Relation: Conjunction, Order and Focus-to-Stress Assignment

Liang Zhang¹, Aijun Li², Yingyi Luo²

¹China University of Political Science and Law, China

²Chinese Academy of Social Sciences, China

Liangzhang1025@126.com, liaj@cass.org.cn, yingyi.luo@gmail.com

Abstract

According to the observation and Chinese linguistic theories of minor sentences, there is not much presence of conjunction in Chinese discourse [1], suggesting that Chinese native speakers could construct semantic relation without the conjunctions, and that the use of Chinese conjunction in discourse may play a particular pragmatic role. Thus this study targets at investigating if the presence and absence of causal conjunctions influence the production of causal relations with different narrative orders in terms of focus position and duration in isolation and in discourse. In our spoken dialogue corpus analysis, the causal conjunction was preferably NOT used in the expression of causal relations; In the stress assignment of causal sentences in the production and perception experiments, the conjunction “YinWei” shows a stable and powerful effect of assigning the stress to its succeeding cause clause in isolation and in discourse; While this effect is modulated by the dynamic discourse focus in terms of the CE linear order and the delta reading duration of the two clauses.

Index Terms: Chinese causal relations, conjunction, linear order, isolation focus, discourse focus, focus-to-stress assignment

1. Introduction

Semantic inference is an important aspect of language comprehension. In natural language processing, up to now the discourses are decoded greatly depending on the surface markers of semantics relations such as “because”, “so”, “although”, which occupy 45% in Spanish and 39% in German. The percentage of the surface semantic markers in Chinese is much lower (30%) [2], which brings a hard task in the computational processing of Chinese language to decode the semantic inference. It does not cause trouble to native speakers communicative intention comprehension, not surprisingly, especially in the real-time dialogue. Thus this study targets at investigating whether the presence and absence of causal conjunctions (“YinWei”(because) and “SuoYi”(so)) influence the production of causal relations with different narrative orders with and without the context.

Causality is an important aspect of human cognitive function. In languages, there are multiple strategies in expressing causality, one of which is causal conjunctions. Conjunctions, as significant markers of semantic functions, are assumed to provide processing instructions which specify the enfolding discourse structure, and further to infer the logic relations between words, sentences, even paragraphs [3, 4].

Another factor that influences the comprehension of causal relations is the narrative order. Some evidence from Chinese

suggest that CE (Cause-Effect) is the dominant order in expressing the causal relation [5], while other researchers argue

the other way (Effect-Cause, EC) [6, 7, 8, 9, 10, 11]. Two factors should be taken into account: the target corpus, in which the proportion of EC in written corpus of news is 69%, while the proportion in spoken corpus rises up to 82% [6]. This result is consistent with Li’s investigation with different genres [12]. Another factor is the presence of conjunction, which may have great impact on Chinese corpus analysis. Song and Tao [7] focused on the sentences with conjunction, but neglected causal relations that were led by other connectives besides “YinWei” or even without any surface markers, which is possibly the default expression of semantic relations for the native speakers of Chinese.

Meanwhile, the relationship between narrative order of causal sentences and the focus of the complex sentences is not clear yet. Some researchers claim that the focus is a pragmatic concept, since its function is to emphasize the important message in the sentence, which is further closely related to the communicative intention of the speaker [13, 14]. Therefore, research should be conducted on both the static focus, which investigates the default focus of the causal sentences in Chinese, as well as the dynamic focus, which examines the narrow focus in discourse.

2. E1: The Spoken Dialogue Corpus Analysis

We first examined the spoken discourse, including causal sentences with and without conjunctions “YinWei” and “SuoYi”. Specifically the question is the distribution of causal relations in different linear orders with causal conjunctions absent and present for different communicative intentions in spoken dialogues.

Our dialogue corpus (Discourse-CASS) [15] includes 43 free speech dialogues from a Chinese instant messenger platform, lasting for 8 hours and 15 minutes. Furthermore, according to the Rhetorical Structure Theory [16], we classified the causal relations into three groups in terms of the perceived communication intentions, based on the context: Cause refers to the emphasis of the content of the cause clause of the two speakers; Result, the result clause; Cause-Result emphasizes on both. The corpus is annotated by two well-trained linguistic students, and double-checked by another linguistic student, in order to diminish the error as much as possible. See Table 1 for the result.

2.1. Result and Discussion

As shown in Table 1, it is quite clear that the causal relations were dominated by expression without conjunctions (91%),

while the causal conjunction only occupies 9% in the data. The Cause and Result relations are marked only with “YinWei” and “SuoYi” respectively. Besides, there was no obvious preference of the linear order, with CE occupying 55%, and EC 45%, which is not significantly different from the chance level ($p = 0.479$). However, speakers did change the linear order because of the communicative intention. They tended to put the emphasized information in the second clause, whether the content was cause or result, accordingly EC was used when cause in the rhetorical relation was emphasized, and CE when the result in the rhetorical relation was emphasized. On the other hand, in the Cause-Result relation, with emphasis on both clauses, the speaker used the CE order only, and with the cause clause heavily marked by causal conjunctions (45 out of 345). What’s more, the causal conjunction “YinWei” seemed to emphasize the expression of the succeeding causal clause, since it only occurred in the Cause relation.

Table 1: *The distribution of conjunction and linear order of causal relations in spoken dialogues.*

Conjunction Order	Presence		Absence		TOTAL
	CE	EC	CE	EC	
Rhetorical Relation					
Cause		70	8	496	574
Result	2		297		299
Cause-Result	45		345		390
Total	47	70	650	496	1263

Two observations could be concluded. First, the default expression of causal relations was sentences without conjunctions. Second, when the conjunction was presented, it was closely related to the expression of the communicative intentions.

3. E2: Causal Relation in Isolation

In this Experiment, we designed a questionnaire and recorded the sentences to investigate the focus position of causal sentences with conjunctions, and the logic relation of the same sentences without conjunctions in isolation.

3.1. Design and Materials

As shown in Table 2, a quite simple sentence structure was used as the parental sentence (I1, the one without conjunction to indicate logic relation): He will listen, she will listen. We use third person singular pronouns as subjects in sentences, since they share the same pronunciation in Chinese, but refer to different genders. Only one conjunction SuoYi or YinWei before the second clause was used in each sentence to indicate the linear order of CE (I2) or EC (I3), which is consistent with the pattern we found in the dialogue corpus. Four sentences were designed for each type, with the verbs in four tones.

Table 2: *The target sentences, with the verb in tone 1.*

Type	Sentence	Conjunction
I1	他要听, 她要听。 (He will listen, she will listen.)	None
I2	他要听, 所以她要听。 (He will listen, so she will listen.)	Post-Effect
I3	她要听, 因为他要听。 (She will listen, because he will listen.)	Post-Cause

3.2. Participants and Procedure

Fifty-six undergraduate and graduate students filled in the on-line questionnaire on the platform of WenjuanXing. They were native speakers of Chinese without a diagnosed reading or learning disability. The first task was to choose the most suitable conjunctions for I1, with three other conjunctions listed, indicating relations such as progressive, conjunction and condition. More than one answer could be chosen, and the final result showed the proportion of participants selecting a particular answer in the total number of participants. The second task was to select the most prominent syllable in ten sentences (two target sentences, and eight filler sentences in other logic relations). The clause carrying the prominent syllable was considered as the prominent clause.

In the next step, thirty undergraduate and graduate students (aged between 18-25, with gender balanced) participated in the recording experiment. The twelve target sentences (3 types * 4 tones) were recorded together with other 96 filler sentences. The obtained audio recordings (in wav format) were annotated and labeled via Praat. The clause carrying the prominent syllable was considered as the prominent clause. The duration of each clause and the most prominent syllable in the target sentences were annotated and extracted, with the duration of the conjunctions excluded. The duration was normalized into Z-score. The delta duration (S1 duration minus S2 duration) is further analyzed. Fundamental frequency F0 and duration are usually taken as parameters indicating prosodic prominence though [17, 18, 19, 20], in this currently research, we mainly focused on the duration representation because the result will be further compared with our parallel cognitive research in reading time from eye-tracking experiment. Chi-square test and Paired T test were adopted in the analyzing of stress assignment and reading duration respectively.

3.3. Result

3.3.1. Logic Analysis

In task one, over 53% of the participants consider causal relation is the most suitable and natural one for I1; While 39% chose progressive relation, and 29% chose the conjunction, and 16% chose conditional relation. Since only 53% of the participants choose causal relations for I1, in the next analysis, we did not treat I1 as a causal relation, but as a baseline to conduct comparative analysis within each type.

3.3.2. Focus Analysis

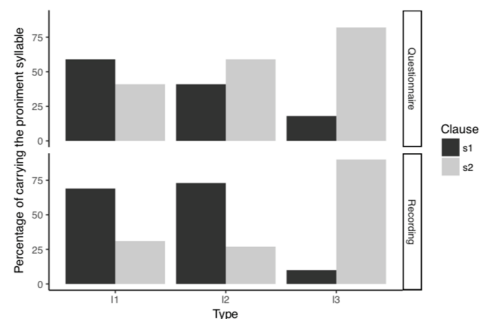


Figure 1: *The percentage of carrying the prominent syllable in isolation (s1 for the first sentence; s2 for the second sentence).*

A similar pattern was found in I1 and I2, that there is no significant difference between the two clauses in questionnaire ($ps>.05$), and that S1 shows a higher possibility of carrying the prominent syllable ($ps<.001$). The cause clause (S2) in I3 shows a significantly higher percentage of carrying the prominent syllable than effect clause (S1), both in questionnaire ($p<.001$), and in recording ($p<.001$), indicating that the subjects emphasize the cause clause with conjunction “YinWei” in production, meanwhile interpret the cause clause as emphasized I3 in perception.

3.3.3. Duration Analysis

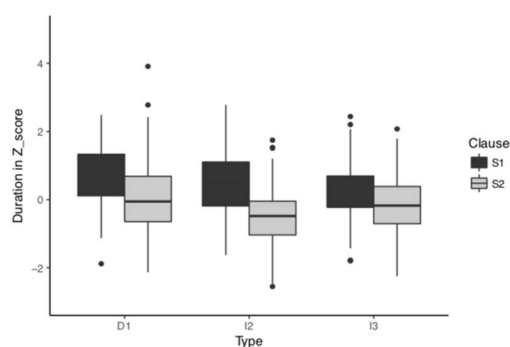


Figure 2: Box-plot of duration of the two clauses

Figure 2 demonstrates the distribution of reading duration of two clauses in three types. Paired T test is employed to examine the differences. The result shows that S1 is significantly longer than S2 in three types ($ps<.001$). S2 (cause clause) in I3 is significantly longer than S2 (effect clause) in I2 ($p=.013$), while the S1s in I2 and I3 are not significantly different from each other. Therefore, the difference in duration of S1 and S2 in I3 is significantly smaller than that in I2 ($p<.001$).

3.4. Discussion

Though less than half of the participants consider the parental sentence as a causal one, the causality still occupies the highest proportion among the options, which may further support the causality-by-default hypothesis [21, 22, 23]. Compared with I2 in CE order with a conjunction “SuoYi”, I3 in EC order with conjunction “YinWei”, shows quite a consistent pattern of the focus position in the production and perception of isolated sentences, in that the cause clause becomes the default focus, and its duration is significantly longer than the second clause in I2, which is marked by “SuoYi”.

In the research of causal sentences in isolation, one observation could be concluded that the conjunction “YinWei” together with the EC order, emphasizes its succeeding cause clause, by assigning the prominent syllable and extending the total reading duration.

In the next experiment, we aim to further explore the production of causal sentences in discourse, in which we manipulate the presence and absence of the conjunction, the narrative order, and the discourse focus position. Based on the observations in previous experiment, we predict that the pragmatic function of conjunction “YinWei” may result in significant differences in focus assigning and clause duration. Moreover, the dynamic focus assigned by the discourse may be overwhelmed by the static focus cued by “YinWei”, because of its emphasis function.

4. E3: Causal Relation in Discourse

4.1. Participants

Twenty participants recorded isolated causal sentences also participate in this experiment. An experiment assistant acted as the listener through the experimental sessions.

4.2. Design and Materials

The discourse focus position (PRE and POST) and linear order (CE and EC) of the causal sentences were crossingly manipulated, in the condition of conjunction absence and presence, allowing us to test their interplay on the production with the parameter of stress assignment and reading duration. Same sentences in previous experiment were recorded. In order to manipulate the discourse focus position, the participants were prompted by two types of questions. The answer focuses on the result in target sentences after question 1: “A will go to the concert. What about B?”; While the answer focuses on the cause in target sentences after question 2: “Why will B go to the concert?”. The name we choose in the position of A and B are typically gendered names in Chinese. Since the gendered pronouns share the same pronunciation, the differences of the pronoun indicates the differences of linear orders of causal and effect in the sentences without conjunction. In total, 32 target sentences (conjunction * linear order * discourse focus * 4 tones) were recorded. The reading material were divided into two groups according to discourse position, each group was recorded by ten speakers.

4.3. Procedure

Two subjects were recorded simultaneously in order to simulate a spoken dialogue, with one asking the background sentence and one answering the causal sentences. The experiment assistant acting as a listener sat out of the recording room to prevent non-verbal communications such as eye contact and gestures. There was a short oral introduction before the experiment. The listener was asked to fill in an on-line answer sheet about the focus position and reference of the pronouns of each sentence based on the reading of the speakers; while the speakers were told to read the discourse thoroughly, to understand the logic of the sentences before reading, and to try the best to express the particular relation, in order to help the listener answer the questions. In this way, the speakers would be more involved. No communication was allowed between the speakers and the listener throughout the experiment. The experiment began after a familiarization practice with ten sentences with or without semantic relation markers recorded.

4.4. Data Analysis

We use the same method in the previous experiment to annotate focus position and reading duration. The Linear mixed model (LMM) was adopted in analyzing reading duration (fixed effects factors are Linear Order and Focus Position; random effects factors are speaker ID and sentence ID).

4.5. Result

4.5.1. Focus Analysis

Figure 3 illustrate the percentage of carrying the prominent syllable of S1 and S2 in different conditions. Without causal conjunction, a position effect is found that S1 is always focused, regardless of the manipulated conditions. Same effect could be

observed in group with conjunction in CE, while the causal conjunction of “YinWei” overturns the position effect, by emphasizing the cause clause in both Post and Pre focus ($p < .001$). The discourse focus effect significantly influences the effect clause in CE group only ($p = .005$), by assigning more prominent stress on effect in the Post focus condition.

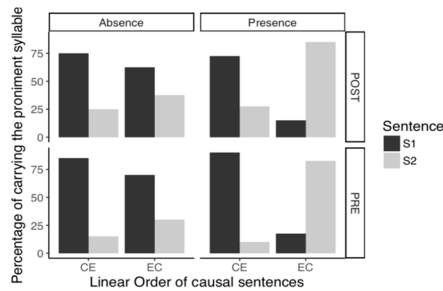


Figure 3: The percentage of the prominent syllable in discourse.

4.5.2. Duration Analysis

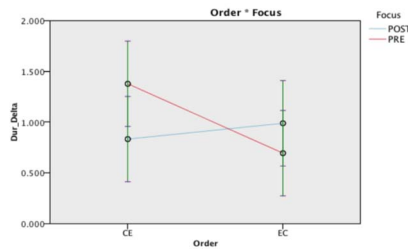


Figure 4: Interaction between linear order and discourse focus position for delta duration with conjunction absent.

LMM analysis of the delta duration with conjunction absent shows: 1) A significant interactive effect of Order and Focus ($F = 11.305, p = .002$). 2) Linear order shows a significant difference in the delta duration ($F = 4.463, p = .036$). To be specific, the duration difference of the two clauses in CE is significantly larger than that in EC. 3) Discourse focus does not show a significant difference in the delta duration ($F = .209, p = .648$).

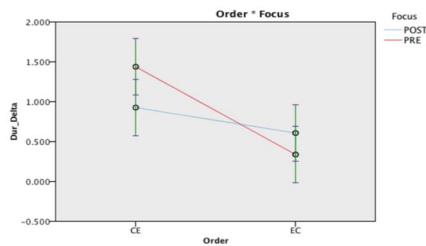


Figure 5: Interaction between linear order and discourse focus position for delta duration with conjunction present.

LMM analysis of the delta duration with conjunction present shows: 1) A significant interactive effect of Order and Focus ($F = 9.286, p = .003$). 2) Linear order shows a significant difference in the delta duration ($F = 30.628, p < .001$). To be specific, the duration difference of the two clauses in CE is significantly larger than that in EC. 3) Discourse focus does not show a significant difference in the delta duration ($F = .308, p = .579$).

4.6. Discussion

In this production experiment of causal sentences in discourse, the stress assignment in causal sentences without conjunction is influenced by position effect, while with conjunction, the assignment is significantly influenced by the emphasis on the cause clause. In both conditions, either position effect or conjunction effect influenced the stress assignment significantly, while the discourse focus only modulate the stress assignment of the Effect clause in CE order. The differences in the clause duration also support this conclusion that linear order by itself impact the delta duration significantly, while the discourse focus modulate the order effect by presenting an interactive effect.

5. Discussion and Conclusion

The three experiments provide from different aspects of production and interpretation yet convergent evidence of the function conjunction “YinWei” to understand Chinese language facts that the causal conjunctions are pretty rare in daily communication. The results in Experiment 2 and 3 are quite consistent with our prediction that the emphasis effect of causal conjunction “YinWei” on the cause clause stays quite stable and powerful in static isolation and in the dynamic discourse. To be specific, the conjunction “YinWei” always assigned the stress on the cause clause in the order of EC regardless of the discourse focus position. On the other hand, the dynamic discourse focus do modulate this effect by assigning stress on the effect clause in CE order with conjunction “SuoYi”. This result of the reading duration also support this conclusion in that the linear order by itself impact the delta duration significantly, while the discourse focus modulate the order effect by presenting an interactive effect.

In annotation the prominent syllable in the reading material, we did not take the prosodic relation of the two clauses into consideration. That is to say only the prominent syllable was labeled, even if the target sentence might consist two intonation phrases. In measuring the clause duration, the conjunctions were excluded, since the stress is assigned on the succeeding clause with a dominant proportion. However, there were examples that speakers assigned the stress on the conjunction, in isolation and in discourse.

To conclude, in the stress assignment of causal sentences in Chinese, the Conjunction “YinWei” shows a stable and powerful effect of assigning the stress to its succeeding cause clause in isolation and in discourse; While this effect is modulated by the dynamic discourse focus in terms of the CE linear order and the delta reading duration of the two clauses.

Future work mainly involve three parts: 1) Fundamental Frequency F0 will be involved as a parameter to examine the focus realization. 2) The hint effect of the punctuation between two clauses on the relation perception [1] will be examined. 3) A comparative study will be conducted with the results of the eye-tracking experiment of Chinese causal relations.

6. Acknowledgements

The work is supported by major program of National Social Science Foundation of China (No.15ZDB103); National Key R&D Program of China (No. 2017YFE0111900) , the Research Foundation of Beijing Municipal Science & Technology Commission (Grant No. Z181100008918002).

7. Reference

- [1] Y. Chao, *A grammar of spoken Chinese*. The Commercial Press,, 2011.
- [2] M. Yue, "Annotation and analysis of Chinese financial news commentaries in terms of rhetorical structure," Ph.D. dissertation, Communication University of China, 2006.
- [3] M. A. Gernsbacher, *Coherence cues mapping during comprehension*, 1996.
- [4] T. J. M. Sanders and W. Spooren, "Discourse and text structure," 2010.
- [5] A. Kirkpatrick, "Information sequencing in modern standard Chinese in a genre of extended spoken discourse," *Text - Interdisciplinary Journal for the Study of Discourse*, vol. 13, no. 3, pp. 423–454, 2009.
- [6] Y. O. Biq, "Chinese causal sequencing and yinwei in conversation and press reportage," vol. 21, no. 2, 2014.
- [7] Z. Song and H. Tao, "A comparative study of Chinese and English causal clause sequences in discourse," *Chinese Linguistics*, vol. 2008, no. 4, pp. 61–71, 2008.
- [8] Y. Su, "A cognitive exploration of Chinese connectives," *Current Allerg and Sasthma Reports*, vol. 1, no. 6, pp. 578–578, 2001.
- [9] M. C. Tsai, "A discourse approach to causal sentences in mandarin Chinese," Kyung Hee University, 1996.
- [10] WANG and YU-FANG, "The preferred information sequences of adverbial linking in mandarin Chinese discourse," *Text - Interdisciplinary Journal for the Study of Discourse*, vol. 22, no. 1, pp. 141–172, 2002.
- [11] Y. F. Wang, "The information structure of adverbial clauses," *Taiwan Journal of Linguistics*, vol. 4, no. 1, 2006.
- [12] H. Li, "The research on the varied order of Chinese modificatory complex sentences," Ph.D. dissertation, Anhui University, 2012.
- [13] X. Fan, "On the problem of Chinese constituent order," *Chinese Language Learning*, no. 5, pp. 18–28, 2001.
- [14] Y. Zhang, "关于汉语句子焦点问题的两点思考," *Academic Journal of Zhongzhou*, no. 2, pp. 245–247, 2006.
- [15] Y. Jia and A. Li, "A linguistic annotation scheme of Chinese discourse structures and study of prosodic interactions," in *International Symposium on Chinese Spoken Language Processing*, 2017.
- [16] W. C. Mann and S. A. Thompson, "Rhetorical structure theory: A theory of text organization," *Text - Interdisciplinary Journal for the Study of Discourse*, vol. 8, no. 3, pp. 243–281, 1987.
- [17] D. R. Ladd, *The structure of intonational meaning*. University Microfilms International, 1978.
- [18] Carlos and Gussenhoven, *The Phonology of Tone and Intonation*. Cambridge University Press,, 2004.
- [19] Y. Xu and C. X. Xu, "Phonetic realization of focus in English declarative intonation," *Laboratory Phonology*, vol. 33, no. 3-4, pp. 159–197, 2015.
- [20] C. Gussenhoven, "'focus, mode and nucleus'," *Circulation*, vol. 106, no. 19, pp. 2434–41, 2002.
- [21] T. J. M. Sanders, "Causality in connectives, coherence relations and cognition."
- [22] G. M. Zunino, V. Abusamra, A. Raiter, G. M. Zunino, V. Abusamra, and A. Raiter, "Causality, iconicity and continuity: the effects of prior world knowledge on the understanding of causal relations," *Alfa Rev. linguist*, vol. 60, no. 2, pp. 263–287, 2016.
- [23] M. H. Oudega, "How default is causality-by-default?" *Open Access*, 2011.

[This paper was published at ISCSLP 2018]