The Features, Classification, and Functions of Event Anaphora in Chinese News Discourse

Jing Wan [1, a], Han Ren [2, b, *]

[1] Center for Lexicographical Studies, Guangdong University of Foreign Studies, China;

[2] Laboratory of Language Engineering and Computing, Guangdong University of Foreign Studies, China.

[a] jingwan@gdufs.edu.cn, [b] hanren@gdufs.edu.cn

Abstract. In the corpus of modern Chinese news discourse, there exists a phenomenon of referentiality involving verbal elements, referred to as event anaphora. Its connotations exhibit distinct differences from similar concepts such as event nouns, cohesive connections, and deixis. This paper investigates the inherent distinctions of event anaphora from three perspectives: structural types, sentence components, and the frequency of reference. Additionally, incorporating discourse and stylistic considerations, the study offers three explanations for biases observed in statistics: the stylistic inclination of "this," the referential target and objective mechanism of the object, and the information structural impact of the frequency of reference.

Keywords: event anaphora; structure type; sentence components; anaphora frequency.

1. Introduction

The study of anaphora in discourse has always been a significant topic in language teaching and language information processing. In recent years, the research scope of anaphora has been extended from nominal elements to verbal ones, and the research of anaphora for verbal elements has received much attention[1,2,3,4,5,6]. The following sentence shows an example of the phenomenon of event anaphora:

(1) 美国近来变本加厉挑起贸易战，宣称要对中国2000亿美元产品加征关税，**这一做法**招致国际舆论广泛谴责。

Recently, the United States has escalated its trade war, declaring its intention to impose tariffs on $200 billion worth of Chinese products, **this approach** that has elicited widespread condemnation in the international public opinion.

In example (1), *this approach* refers to the preceding clause about the event *the United States has escalated the trade war, claiming to impose tariffs on $200 billion worth of Chinese products* In other words, it does not refer to a nominal element but to a verbal one, which always represents an event. Such phenomenon of coreference that a referring mention appears after an event it refers to in news discourse is known as event anaphora.

Event anaphora is concerned by the research community of linguistics and natural language processing area[7,8,9,10,11]. Martin[12] proposed the notion of discourse metaphor, suggesting that within a discourse, the same content may exist in multiple forms. Discourse metaphors often involve relationships between different forms, where components representing events and components referring back to events constitute discourse metaphors. Charolles[13] introduced the concept of conceptual anaphora, where the mention of such an anaphora is an event from the previous text. Since the preceding information is often event-related, conceptual anaphora can also be seen as a form of event anaphora. Wang[2] explored the nature and format of conceptual anaphora in the context of Chinese. Fang[5] focused on the anaphora of abstract nouns in the category of nominal subordination, observing that the referents of abstract nouns often involve event components. Zhou and Yang[3] approached conceptual anaphora from a cognitive discourse perspective, investigating its connotations, pragmatic features, cognitive characteristics, and functional attributes.

However, there are significant differences between event anaphora and conceptual anaphora. Firstly, the anaphoric elements in conceptual anaphora do not mention or evaluate the preceding text, and adjectives are absent. Both situations deviate from linguistic reality. Secondly, the anaphoric elements in conceptual anaphora are mostly nominal, overlooking the role of demonstratives in anaphorically referring to events. Thirdly, the referents of conceptual anaphora are mostly clauses, whereas we argue that they can encompass not only clauses but also sentences and even groups of sentences.

In this paper, we explore the event anaphora in Chinese discourse texts, trying to find the characteristics and types of event anaphora based on discourse semantics. In a discourse text, a coreferencial mention which appears after its referring event, is an event anaphora. The declarative passages are referred to as event antecedents, and the referring text units are termed event anaphors. This approach is based on the linguistic fact that events in discourse not only serve a declarative function but also a referential function. Additionally, in Modern Chinese, due to the absence of strict morphological changes and part-of-speech markers, the transition of events from declarative to referential functions is more practical[14]. Furthermore, considering the sparse exploration in the scholarly community regarding anaphora in verbal elements, and the prevalent phenomenon of antecedents with explicit predicate characteristics, similar to conceptual anaphora[2,10], it is necessary to conceptualize and thoroughly examine this phenomenon. Event anaphora not only serves as a crucial supplement to the acquisition of referential elements in language learning but also represents a key task in the field of co-reference resolution in language information processing. We also explore the phenomenon of event anaphora by using a taxonomy to demonstrate different types of event anaphora.

2. Classification of Event Anaphora

From a formal perspective, certain combinations of anaphoric expressions are conducive to identifying event anaphora. These include bare demonstrative "这" (this), noun phrases containing "这" (this), and noun phrases without "这" (this). We also carry out a statistical analysis by leveraging on a small Chinese news corpus including 100 articles.

2.1 Bare Demonstrative "这"

The demonstrative pronoun "这" (this) serves as the anaphoric expression directly referring to the antecedent. We refer to this anaphoric form as the bare demonstrative "这." Based on corpus statistics, instances of anaphoric expressions as bare demonstrative "这" are not very common, accounting for 2.7% of the total occurrences, with 29 instances. For example:

(2) 彭博社3月调查了21位经济学家，根据他们估计的中位数，中国的债务总额在2018年底将达国内生产总值 （GDP）的260%，与一年前持平。在名义上，这意味着债务的增长放缓至与经济大致同步。

Bloomberg surveyed 21 economists in March. According to their median estimate, China's total debt is expected to reach 260% of the Gross Domestic Product (GDP) by the end of 2018, unchanged from a year ago. In nominal terms, this means that the growth of debt has slowed to roughly match the pace of the economy.

In the context of event anaphora, "这" refers to the preceding clause "China's total debt is expected to reach 260% of the Gross Domestic Product (GDP) by the end of 2018, unchanged from a year ago." It can be observed that the anaphoric expression "这" not only functions to denote the preceding clause but also guides the direction of anaphora, serving as a means to maintain the continuity of the discourse topic and demonstrating the "relevance" between antecedent information and current information[15].

From a syntactic perspective, topic elements often appear in the subject position, and the lighter the anaphoric form, the greater the possibility of maintaining a unified topic. "这" serves as the

second lighter form after zero-form anaphora, contributing to the continuity of the topic. On the other hand, the demonstrative "那" (that), which refers to the preceding statement or the event described in the previous context, does not appear as an anaphoric expression in event anaphora. This is partly due to the high frequency of "这" as a discourse anaphoric expression, reaching 93% in corpus-based statistical studies. Additionally, "那" tends to be used when referring to the content narrated in the immediately preceding clause rather than a specific object.

2.2 Noun Phrase with "This"

A noun phrase containing "这" always refers to a noun phrase with "这" functioning as an anaphoric expression, referring back to the antecedent in the previous discourse. This type has the highest frequency in event anaphora, with 907 instances, accounting for 84.7% of the total occurrences. Within this category, noteworthy formats include "这" + quantifier + noun phrase, "这一" + noun phrase, and "这" + noun phrase. The following discusses these three formats separately.

2.2.1　"This" + Quantifier + Noun Phrase

The combination of "这" (this) + quantifier + noun phrase represents the most typical form of anaphoric expression. In this structure, the noun phrase often consists of abstract nouns such as "状况" (situation), "消息" (news), "行为" (behavior), "调整" (adjustment), "说法" (statement), "道理" (reason), "想法" (idea), "建议" (suggestion), etc. Most of these abstract nouns are bisyllabic, and the quantifiers chosen are closely related to abstract nouns, such as "类" (kind), "种" (type), "个" (individual), "次" (occasion). There are also a few instances with monosyllabic noun phrases, such as "这种事" (this kind of thing), "这个话" (this statement), "这个事" (this matter).

2.2.2　"This One" + Noun Phrase

The combination of "这一" (this one) + noun phrase represents another format of anaphoric expression. Through corpus analysis, it is observed that "这一" + noun phrase and "这个" + noun phrase are similar in terms of the accompanying noun phrases. Both can be followed by abstract noun phrases, such as "问题" (issue), "说法" (statement), "事件" (event), "事情" (thing), "事实" (fact), "政策" (policy), "理论" (theory), "道理" (reason), "决定" (decision), "想法" (idea), "措施" (measure), "规定" (regulation), "条件" (condition), "意义" (meaning), "观念" (concept), "动作" (action), "原则" (principle), etc., mostly bisyllabic abstract nouns. Additionally, there are cases with monosyllabic nouns like "话" (speech).

2.2.3　"This" + Noun Phrase

The direct combination of "这" (this) and "名词短语" (noun phrase) is not highly frequent, with only one occurrence in the form of "这方面" (this aspect). However, related instances include "这样" (this way), "这些" (these), and "这个" (this), as well as "这些" + noun phrase. The usage frequency of "这" + noun phrase is lower than that of "这" + quantifier + noun phrase and "这一" + noun phrase. This discrepancy is due to the less stable structure of "这" + noun phrase, forming a "1+2" structure. The "2+2" structure is more in line with the restrictions on this structure in news language. Sometimes, for rhythmic reasons, an additional component may be required for syllable completion.

Analyzing the corpus also reveals a progressive sequence in the use of "这," "这样," and noun phrases containing "这." "这样" seems more like a manner, for example, "Fish should be cut like this before cooking." "这样" is roughly equivalent to "这种方式" (this way), indicating a manner. In contrast, "这一" + noun phrase, as in "这一事实" (this fact), is entirely different; it does not represent a single event or a specific behavior but an overview of events, directly referring to the

preceding complete event. "这一切" (all of this) is similar to "这样" but represents a slightly narrower range of ways, expressing a generalization rather than a specific reference to a part of the preceding text. This highlights the specificity of noun phrases containing "这" in referring to event summaries.

2.3 Noun Phrases without "This"

Noun phrases without the inclusion of "这" (this) refer to anaphoric expressions that do not contain the term "这." Examples include pronouns such as "其" (its), "它" (it), "该" (this), and constructions like "此" + noun phrase and "上述" + noun phrase. This format appears in 134 instances, accounting for 12.5% of the total occurrences. The forms of anaphoric expressions in this category are relatively complex, and they will be discussed separately.

Firstly, pronouns represented by "其" are used as anaphoric expressions. For instance:

(3) 从战略态势看，两国由以往的相对隔离、间接过招演变成现在的直接对立、正面博弈，其源头始于美国战略重心从欧洲、中东突然全面转向亚太，来到中国的家门口。

From the strategic situation, the two countries have transformed from past relative isolation and indirect conflicts to direct opposition and frontal confrontation. Its origin lies in the sudden comprehensive shift of the U.S. strategic focus from Europe and the Middle East to the Asia-Pacific, coming to China's doorstep.

"其"(that) serves as the subject of the clause, summarizing the antecedent "the two countries have transformed." "其" not only represents the two countries but also signifies the strategic situation of the changes, which cannot be expressed by repeating "the changes in the strategic situation of the two countries" due to verbosity. The use of "它们的" (their) would not be suitable, as it leads to both awkwardness and a loss of solemnity. Therefore, "其" as an anaphoric expression contributes to the expression of the event. Table 1 shows the statistical analysis results of the types of event anaphora.

Table 1. The Forms and Examples in Event Anaphora

Form	Construction	Examples	Numbers
Bare 'this'	'this'	this（这）	29（2.7%）
NPs Containing 'this'	'this'+ quantifier +NPs	this type/ this kind	907(84.7%)
Noun phrases without 'this'	Pronoun/'that/those'+ quantifier +NPs	that news/ those problems	134(12.5%)

3. Conclusion

This study on event anaphora contributes to a better understanding of discourse and automatic language comprehension. Due to the lack of morphological markers in Chinese, the maximum model of event anaphora with the same reference can be used for interpretation. Event anaphora is not only a comprehensive reference to the entire preceding passage but also sometimes refers to a single clause or several clauses, with multiple antecedents that can be referenced. Further research into this area is significant for the automatic processing of language information. This paper only touches the surface of the observational perspective on anaphoric expressions, leaving ample room for exploration, awaiting further insights from researchers. Additionally, while the focus has been on news discourse due to its broad applicability, event anaphoric expressions exist in other genres as well, warranting continued investigation.

Acknowledgments

References

[1] Martin, J R. English Text System and structure by Talmy，Leonard.Toward a Cognitive Semantics，Vol2.Cambridge，Massachusetts: MIT Press, 2000.

[2] Wang Xiuli. Encapsulating Anaphora in text analysis. Contemporary Linguistics, 2012, 14(03).

[3] Zhou, Xiaotao, Wang Jun. Encapsulating Anaphora Analysis from the Perspective of Cognitive Pragmatics,Foreign Language Research, 2014.

[4] Yang Chunmei, Wen Xu. Interpretation of Résomptive Anaphora Based on Conceptual Blending Theory, Foreign Language and Literature,2016,32(05)

[5] Fang Qingming. The Study on Abstract Noun' Resomptif Anaphora and Resomptif Cataphora in Modern Chinese,Chinese Language Learning,2016(04)

[6] Yang Jiahui, Han Jingquan. Anaphoric Reference to Event and Temporal Interpretation of Bare Eventive Sentenc, Chinese Teaching in the World. 2021, 35(04).

[7] Li Man and Bai Ruifang. Recognition of English Information and Semantic Features Based on SVM and Machine Learning. Journal of Intelligent & Fuzzy Systems, 2021, 40(2).

[8] Xie, Dongdong, Li Fei, Li Bobo, Teng Chong, Ji,Donghong and Zhang Meishan. Chinese Event Discourse Deixis Recognition: Design of the Dataset and Model. ACM Transactions on Asian and Low-Resource Language Information Processing, 2023, 22(11).

[9] Vandan Mujadia, Palash Gupta and Dipti Misra Sharma. Pronominal Reference Type Identification and Event Anaphora Resolution for Hindi. International Journal of Computer Linguistics Application, 2016, 7(2).

[10] Wampler J. Do Thus: An Investigation into Anaphoric Event Reference. Glossa: a journal of general linguistics, 2021, 6(1).

[11] Li Shengnan, Qu Weiguang, Wei Tingxin, Zhou Junsheng Zhou, Gu Yanhui and Li Bin. A Survey of Chinese Anaphora Resolution. Proceedings of International Conference on Artificial Intelligence and Security, 2021.

[12] Martin J.R. English text: System and structure. Amsterdam: John Benjamins, 1992.

[13] Charolles M. La référence et les expressions référentielles en français, Ophrys, 2002.

[14] Shen Jiaxuan. Nouns and Verbs, The mercial Press, 2016.

[15] Fang Mei. On the grammaticalization of Zhe in Beijing Mandarin: From Demonstrative to Difinite Article, Chinese Language,2002.