big data and cognitive computing

MDPI

Article

Hierarchical Discourse-Semantic Modeling for Zero Pronoun Resolution in Chinese

Tingxin Wei [1], Jiabin Li [2], Xiaoling Ye [3,4,]* and Weiguang Qu [5,6,]*

[1] International College for Chinese Studies, Nanjing Normal University, Nanjing 210097, China; weitingxin@njnu.edu.cn
[2] School of Chinese Language and Literature, Nanjing Normal University, Nanjing 210097, China; 210101004@njnu.edu.cn
[3] School of Education Science, Nanjing Normal University, Nanjing 210097, China
[4] Adolescent Education and Intelligence Support Lab of Nanjing Normal University, Laboratory of Philosophy and Social Sciences at Universities in Jiangsu Province, Nanjing 210097, China
[5] School of Computer and Electronic Information/School of Artificial Intelligence, Nanjing Normal University, Nanjing 210023, China
[6] Zhongbei College, Nanjing Normal University, Danyang 212334, China
* Correspondence: 02447@njnu.edu.cn (X.Y.); wgqu@njnu.edu.cn (W.Q.)

Abstract

Understanding discourse context is fundamental to human language comprehension. Despite the remarkable progress achieved by Large Language Models, they still struggle with discourse-level anaphora resolution, particularly in Chinese. One major challenge is zero anaphora, a prevalent linguistic phenomenon in which referential elements are omitted, increasing complexity and ambiguity for computational models. To address this issue, we introduce CDAMR (Chinese Discourse Abstract Meaning Representation), a novel annotated corpus that systematically labels zero pronouns across diverse syntactic positions along with their discourse-level coreference chains. In addition, we present a hierarchical discourse-semantic enhanced model that separately encodes local discourse semantics and global discourse semantics, and models their interactions via structured multi-attention mechanisms. Experiments on both CDAMR and OntoNotes demonstrate the approach's cross-corpus generalizability and effectiveness, achieving F1 scores of 59.86% and 60.54%, respectively. Ablation studies further confirm that discourse-level semantics significantly enhance zero pronoun resolution. These findings highlight the value of cognitively inspired discourse modeling and the importance of comprehensive discourse annotations for languages with limited explicit syntactic cues such as Chinese.

Keywords: discourse semantics; coreference chains; zero pronoun resolution; topic modeling

Academic Editor: Domenico Ursino

Received: 10 August 2025
Revised: 2 September 2025
Accepted: 4 September 2025
Published: 9 September 2025

Citation: Wei, T.; Li, J.; Ye, X.; Qu, W. Hierarchical Discourse-Semantic Modeling for Zero Pronoun Resolution in Chinese. *Big Data Cogn. Comput.* **2025**, *9*, 234. https://doi.org/10.3390/bdcc9090234

1. Introduction

Large language models (LLMs) have demonstrated remarkable performance across a wide range of natural language processing tasks, including question answering, summarization, and generation [1–3]. Despite the advances, studies have shown that LLMs still exhibit limitations in specific areas of information extraction, notably in coreference resolution [4–6]. Recent research has explored the use of prompt engineering for coreference resolution with LLMs [7], while other approaches based on the BERT model family have reported notable improvements [8].

Among the various forms of coreference, zero pronouns represent a particularly challenging phenomenon. Resolving zero pronouns involves identifying their corresponding antecedents within a discourse, making it a fundamental task for many downstream semantic applications in NLP, including question answering [9,10] and text summarization [11]. It further supports machine translation by supplying obligatory pronouns in target languages and strengthens information extraction and dialogue systems by completing argument roles and coreference chains. Hence, improvements in ZPR contribute to stronger and more reliable performance in a wide range of semantic tasks.

A zero pronoun is a phonetically absent but syntactically and semantically meaningful referential form, typically referring back to an antecedent previously mentioned in the discourse. As a grammatical and pragmatic phenomenon, it is prevalent in many pro-drop languages such as Chinese [12], Italian [13] and Japanese [14]. Specifically, Chinese zero pronouns are particularly challenging due to their high frequency, lack of morphological cues, and strong reliance on discourse context; unlike pro-drop languages such as Spanish or Italian, where verb inflections constrain possible antecedents, or Japanese and Korean, where topic-marking particles may guide interpretation, Chinese offers few syntactic hints, thereby increasing ambiguity and making resolution more difficult. Example 1 illustrates a representative instance of a zero pronoun from the CDAMR corpus, which was developed as part of this study.

> Example 1: 遥远的北京城，有一座天安门，广场上升旗仪式非常壮观。我对妈妈说， 我多想去看看 *pro*。(Far away in Beijing, there is a square named Tiananmen where the flag-raising ceremony is spectacular. I told mom that I really want to see *pro*.)

In this example, the zero pronoun *pro* corefers with the mention “天安门” (“Tiananmen Square”) which appears in the preceding sentence with no morphological or syntactic marker signaling the reference. As zero pronouns lack explicit surface forms, they cannot be directly represented by conventional word embeddings. This absence of lexical realization poses a significant challenge for zero pronoun resolution and has attracted considerable attention from researchers in recent years.

To address the issue of zero pronoun representation, Chen and Ng [15] proposed a method that utilizes the leading word and governing verb as surrogate features. However, this approach suffers from a key limitation: the semantic and syntactic properties of these two components often differ substantially, which can lead to inaccurate representations of the zero pronoun. An alternative approach leverages sentence-level context surrounding the zero pronoun [16–19]. For instance, in the sentence “兔姑娘又从小路上走过，ϕ皱起了眉头” (“The girl rabbit walked along the path again, and ϕ frowned”), the zero pronoun ϕ can be represented using the preceding context “兔姑娘又从小路上走过” and the subsequent clause “皱起了眉头.” This context-based approach improves representation by incorporating semantic and distributional cues from the surrounding text, and it has led to significant advances.

Recent work by Sun [20] further enhances contextual representation using convolutional neural networks with self-attention and multi-hop attention mechanisms to better capture informative cues around zero pronouns and their candidate antecedents. Bi et al. [21] proposed an MRC-based framework that leverages semantic dependency structures to capture deep predicate–argument relations, improving zero pronoun resolution without relying on costly syntactic parses. In contrast to these approaches, our model explicitly incorporates discourse-level semantic information, such as topic structure and coreference chains, which enables it to better represent zero pronouns and their antecedents in broader discourse contexts.

However, sentence-level context alone is often insufficient, especially when the sentence is short or lacks informative content. From a linguistic perspective, zero pronouns are a type of anaphora that are typically interpreted within a broader discourse. This insight indicates that effective zero pronoun resolution should draw upon discourse-level semantic information.

In this paper, we first construct a new corpus named CDAMR, which annotates coreference chains from discourse perspective to reveal long-distance semantic relations among entities. Then, inspired by cognitive linguistic perspectives on how humans comprehend discourse through layered semantic cues, we propose a hierarchical modeling framework that integrates topic-level, local contextual, and shared referential semantics for more effective zero pronoun resolution. Experimental results demonstrate that incorporating discourse information significantly improves the performance of zero pronoun resolution.

The contributions of our work can be summarized as follows:

- We construct CDAMR, a novel annotated Chinese corpus containing both full document-level AMR graphs and explicit discourse-level coreference chains, enabling fine-grained modeling beyond sentence boundaries.
- We propose a hierarchical discourse-semantic enhanced model that integrates local discourse semantics and global discourse semantics, which includes document-level topics and shared coreference chains; multiple attention mechanisms are employed to model the interactions between these semantic levels, effectively capturing long-range cross-sentence dependencies.
- The experiments on CDAMR and OntoNotes validate the cross-corpus generalization of our model and confirm the effectiveness of cognitively inspired discourse semantics in enhancing zero pronoun resolution for Chinese.

2. Related Work

2.1. Linguistic Studies on Zero Anaphora

Over the past decades, anaphora has attracted considerable attention from linguists and psychologists. Most studies agree that the construction and resolution of anaphora is a complex phenomenon involving structural, cognitive and pragmatic factors that interact in intricate ways [21,22]. Scholars of cognitive linguistics and pragmatics argue that anaphora is a primary linguistic device for establishing and maintaining discourse coherence [23,24] throughout the dynamic course of discourse production [25,26]. An increasing number of empirical studies have investigated cognitive factors influencing anaphora resolution, and several influential models have been proposed to explain the process [27], among which the topic continuity model is the most widely accepted [28]. Specifically, topic continuity provides a global semantic framework that guides the interpretation of referents across discourse, making it essential for understanding zero pronouns. As a means of avoiding repeated reference to previously mentioned entities, the zero pronoun is frequently used to maintain the topic continuity during discourse comprehension. The zero pronoun, together with other lexical expressions referring to the same entity, forms a coreference chain within discourse. Therefore, the coreference chain is crucial for zero pronoun resolution throughout the discourse. To capture the full spectrum of cognitive and linguistic cues involved in zero pronoun resolution, our model integrates three complementary levels of information: topic information, which tracks referential continuity at the global discourse level; local discourse context, which captures short-distance syntactic and semantic features; and global discourse semantics, which reflects broader coherence and structural patterns across sentences.

2.2. Studies on Zero Pronoun Resolution

Early approaches to Chinese zero pronoun resolution relied primarily on rule-based systems [29] and conventional machine learning methods [30,31]. With the advancement of deep learning, neural network-based models have increasingly been applied to this task. Chen and Ng [15] proposed a feed-forward neural network that encodes a zero pronoun using its leading word and governing verb. To enhance zero pronoun representation, Yin et al. [16] introduced a memory-based neural network that learns from both the surrounding text and its antecedent mentions. For candidate encoding, Yin et al. [17] designed a hierarchical candidate encoder to capture global information across the candidate span.

Recognizing the rich semantic cues available in context, recent work has focused on incorporating contextual information more effectively. Yin et al. [18] applied a self-attention mechanism to both zero pronoun and candidate encoding, allowing the model to focus on the most informative parts of the associated text. Further extending this idea, Lin et al. [19] observed that not only the context surrounding the zero pronoun, but also that of its candidate antecedents, carries valuable semantic information. They proposed a hierarchical attention network with a pairwise loss function to jointly model the representations of zero pronouns and candidates more effectively.

With the rapid development of large language models, zero pronoun resolution has drawn growing attention, both as a persistent challenge and as a means to enhance LLM-based systems. Wang et al. [32] provide a comprehensive survey highlighting that ZPR aligns with the trend of LLM development, while also stressing challenges such as data scarcity, benchmark overfitting, and the lack of targeted evaluation metrics. Ueyama and Kano [33] further demonstrate the utility of incorporating zero anaphora resolution into dialogue systems, showing that explicitly completing omitted arguments can improve the coherence of LLM-generated responses. However, despite these advances, recent studies have shown that LLMs still exhibit notable limitations in discourse-level information extraction, particularly in coreference resolution [4–6]. To address these challenges, Sun [20] enhances contextual modeling with CNN–attention mechanisms, while Bi et al. [21] propose an MRC-based framework leveraging semantic dependency structures.

Despite the progress made by these models, they primarily rely on sentence-level context, focusing only on the immediate sentences where the zero pronouns and candidates appear. Consequently, they tend to overlook discourse-level semantic information, which is crucial for resolving long-distance anaphora in complex texts.

3. Datasets

In existing studies on zero pronoun resolution, the most widely used dataset is OntoNotes 5.0 [15–19], a large-scale multilingual corpus encompassing five languages, including English, Chinese, and Arabic, which was released by the Linguistic Data Consortium (LDC). It provides rich annotations for various linguistic phenomena, such as syntactic structures, predicate-argument structures, and coreference resolution. In version 5.0, the Chinese portion of the corpus was further enhanced with annotations for zero anaphora in subject positions. However, this corpus exhibits several limitations in its coreference annotation. First, coreferential mentions are annotated primarily based on local context, without adopting a comprehensive discourse-level perspective. Second, zero pronoun annotations are restricted to subject positions only.

In natural discourse, however, zero pronouns can appear in a variety of syntactic roles including subject, object, and even attributive positions, with their antecedents potentially located anywhere within the discourse. Such incomplete annotations hinder the accurate identification of zero pronoun distributions and the construction of coherent coreference chains, ultimately limiting the performance of resolution models trained on this data.

To address these limitations, we introduce CDAMR (Chinese Discourse Abstract Meaning Representation), a new discourse-level semantic resource that provides more comprehensive annotations of zero pronouns and their discourse-spanning coreference chains.

3.1. Construction of CDAMR

CDAMR is derived from the CAMR (Chinese Abstract Meaning Representation) corpus, which pairs Chinese sentences with Abstract Meaning Representation (AMR) graphs [34,35]. CDAMR comprises 333 documents, each annotated with discourse-level semantic structures and inter-sentential coreference relations.

For coreference annotation, all mentions referring to the same entity, including noun phrases, overt pronouns, and zero pronouns in various syntactic positions, are linked into coreference chains representing shared discourse referents. To capture fine-grained semantic variation, the annotations also specify the semantic relation between each mention and its conceptual referent, including relation such as synonymy, hypernymy and homonymy. In annotating, each sentence is annotated using the format C*i*: (relation/instance), where C*i* denotes the chain type and its index. For example, P refers to person chains, M to entity chains, and L to location chains. Within the parentheses, the segment following the "/" specifies the lexical realization of the chain at that position, together with its index in the text. The segment preceding the "/" indicates the semantic relation between the instance and the chain concept, where REF marks the antecedent that introduces the chain; "zero" indicates that the instance is a zero pronoun referring to the chain concept; and "syn" indicates that the instance is a synonym of the chain concept. Additional relations include identical form, pronoun, part-of, member-of, hypernym, attribute-of, spacial reference and metaphorical use.

Additionally, sentence-level AMR graphs from CAMR can be aligned and retrieved as needed to support detailed semantic analysis at both the sentence and discourse levels.

Figure 1 illustrates an example of a coreference chain annotation in the CDAMR corpus, demonstrating how various mention types are linked across sentences to represent a shared discourse referent.

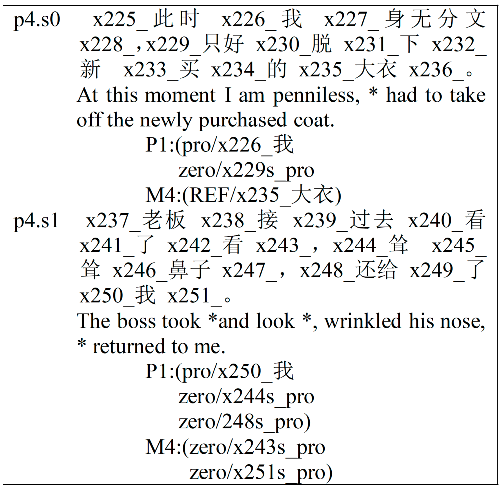

Figure 1. Example of coreference chain annotation in the CDAMR corpus. p4.s0 and p4.s1 denote the paragraph and sentence indices, respectively. P1 represents the first person coreference chain in the discourse, while M4 represents the fourth entity coreference chain. REF marks the coreference chain concept, whereas pro and zero indicate the relation types between mentions and the chain concept. xi denotes the position index of the mention within the discourse. For example, p1(pro/x250_我) indicates that the word "我" at position x250 in the text is annotated as a pro form of the coreference

chain P1, while M4(zero/x243s_pro) indicates that a zero pronoun occurs before position x243, serving as the zero form of the coreference chain M4. The '*' indicates omitted zero pronouns in the English translation.

To construct the corpus, an annotation scheme was first developed and iteratively refined through multiple rounds of pilot annotation. Following the finalization of the guidelines, five postgraduate students majoring in linguistics manually annotated the documents. To ensure consistency and a shared understanding of the annotation principles, a series of team workshops were conducted to clarify ambiguous cases and resolve inter-annotator disagreements.

3.2. Annotation Quality Evaluation

Since the number of annotated coreference chains and mentions varied across annotators, the Inter-Annotator Agreement (*IAA*) score was used to measure internal annotation consistency, which is computed with the following Formula (1):

$$IAA = \frac{A}{N} \tag{1}$$

where N denotes the total number of annotations, and A represents the number of annotations on which all five annotators fully agreed.

After an initial trial annotation and a subsequent adjudication process to resolve disagreements, the second round of annotation yielded an *IAA* score of 0.850 for coreference chain agreement and 0.813 for mention-level agreement. Based on these results, the remaining documents were annotated in accordance with the established guidelines.

3.3. Comparison with OntoNotes

There are several differences between CDAMR and OntoNotes in terms of corpus annotation:

- CDAMR and OntoNotes differ significantly in the scope and types of zero pronoun annotation. OntoNotes annotates primarily zero pronouns in subject positions [36,37], whereas CDAMR includes annotations for zero pronouns in subject, object, and modifier positions. Many studies have shown that in pro-drop languages, zero pronouns tend to refer to subjects [38], making resolution in subject positions relatively easier. In contrast, resolving zero pronouns in object and modifier positions is more challenging and thus needs greater attention.
- The OntoNotes corpus comprises the texts from broadcast news, blogs, phone recordings and interviews while the CDAMR corpus covers a diverse range of genres, including narrative, descriptive, expository, argumentative texts, poetry, and dialogues.
- CDAMR and OntoNotes also differ in the quantity and density of zero pronoun annotations.

Table 1 presents a comparison of the annotation and distribution density of zero pronouns in the two corpora.

As illustrated in Table 1, CDAMR contains an average of 13.64 zero pronouns and 8.31 coreference chains per document, in contrast to 8.84 zero pronouns and 2.99 coreference chains per document in OntoNotes. Moreover, each coreference chain in CDAMR contains an average of 8.72 zero pronouns, and each sentence contains 0.53 zero pronouns on average, whereas the corresponding figures in OntoNotes are only 4.31 and 0.32, respectively. The average sentence length in CDAMR is 18.02 words, slightly shorter than the 20.34 words in OntoNotes. This indicates that the higher density of coreference chains and zero pronouns in CDAMR is not due to longer sentences, but rather reflects the more detailed annotation scheme adopted in our dataset. This further demonstrates that CDAMR provides more comprehensive and enriched zero pronoun annotations. This substantial increase can be attributed to CDAMR's more comprehensive annotation approach. Such dense and detailed annotations make CDAMR a valuable resource for training and evaluating models on zero pronoun resolution and discourse-level coreference tasks.

Table 1. Comparison of zero pronoun annotation between CDAMR and OntoNotes.

	CDAMR	**OntoNotes**
Zero pronoun positions annotated	Subject, Object, Modifier	Subject
Text genres	Narrative, Descriptive, Expository, Argumentative, Poetry, Dialogues	Broadcast news, Blogs, Interviews, Phone recordings
Documents	333	1563
Sentences	8592	45,270
Words/sentence	18.02	20.34
Coreference Chains	2766	4763
Chains/Document	8.31	2.99
Zero Pronouns	4541	13,824
Zero Pronouns/Document	13.64	8.84
Zero Pronouns/Chain	8.72	4.31
Zero Pronouns/Sentence	0.53	0.32

3.4. Distributions of Zero Pronouns in CDAMR

To gain a deeper understanding of zero pronouns in language, we first provide a fine-grained analysis of their semantic and syntactic distribution. Table 2 presents the semantic relation types between mentions and concepts within coreference chains in CDAMR. The results show that identical-form coreference (i.e., repeated mention of the same expression) is the most common strategy, accounting for 37.81% of all mentions. Excluding identical forms, pronouns and zero pronouns make up nearly 70% of the remaining mentions, indicating their central role in maintaining coreference continuity. Additionally, other semantic coreference types such as synonyms, part-of, member-of, and attribute-of relations are also observed, though less frequent, reflecting the richness of coreference phenomena in CDAMR.

Table 2. Semantic relation types between mentions and concepts within coreference chains in CDAMR.

Mentions	**Identical Form**	**Pronoun**	**Zero Pronoun**	**Synonym**	**Part-of**	**Member-of**	**Hypernym**	**Attribute-of**	**Space**	**Metaphor**
Count	8075	4576	4541	1926	603	597	430	244	191	169
Percentage	37.81%	21.42%	21.26%	9.02%	2.82%	2.80%	2.01%	1.14%	0.89%	0.80%

Table 3 further analyzes the syntactic positions of zero pronouns. The majority (75.53%) appear at the beginning of clauses, often functioning as dropped subjects that link back to previous discourse, which is a typical discourse strategy in Chinese. A smaller proportion occur within clauses (17.51%), usually as omitted objects, while only 6.96% appear at the beginning of full sentences.

Table 3. Position distribution of zero pronouns in sentences in CDAMR.

Position	**Beginning of Sentences**	**Beginning of Clauses**	**Within Clauses**	**Total**
Count	316	3430	795	4541
Percentage	6.96%	75.53%	17.51%	100%

Building on this distributional insight, we then examine the distance between zero pronouns and their antecedents to evaluate how far zero pronouns typically refer back in context. As shown in Table 4, we report results under two settings: one that allows overt pronouns as valid antecedents and one that excludes them. This distinction helps reveal the true semantic reach of zero pronouns, given that overt pronouns themselves are often ambiguous and require resolution. The comparison provides insights into how discourse context and antecedent selection strategies influence zero pronoun resolution in real-world texts.

Table 4. Distance distribution of zero pronouns to antecedents in CDAMR.

Distance	0 *	−1	−2	−3	+1	+2	Other	Total
N + overt pronoun%	78.59	12.05	3.35	1.06	1.89	0.70	2.16	100
Cumulative%		90.84	94.19	95.25	97.14	97.84	100	
N − overt pronoun%	58.16	17.42	6.01	3.08	2.64	1.52	11.16	100
Cumulative%		75.58	81.59	84.67	87.31	88.83	100	

* Distance = 0 indicates the zero pronoun and its closest antecedent appear in the same sentence. Distance = −1 means that the antecedent is located in the immediately preceding sentence, −2 in the second sentence before, and so forth. Distance = +1 indicates that the antecedent is found in the immediately following sentence, +2 in the second sentence after, and so forth.

As shown in Table 4, adopting the conventional candidate selection strategy of selecting antecedents from the current sentence and the two preceding ones yields a coverage of 94.19% for zero pronouns when overt pronouns are included. However, excluding overt pronouns leads to a substantial reduction in coverage to 81.59%, suggesting that approximately 18% of zero pronouns fall outside the scope of the conventional window.

In response to this limitation, we expand the candidate sentence window from [−2, 0] to [−3, +1], resulting in an increase in antecedent coverage from 81.59% to 87.31% under the non-pronominal condition.

4. Methodology

The task of zero pronoun resolution involves identifying the coreferential antecedent of a given zero pronoun (zp) by extracting a set of candidate noun phrases, $NP = \{np_1, np_2, \ldots np_n\}$, from the discourse and determining which candidate is coreferent with zp.

Our approach proceeds as follows:

Local discourse information: For each candidate antecedent, we extract the semantic representation of the sentence containing the candidate, along with the representations of its preceding and following sentences. These provide local context around the candidate.

Global discourse information: We first collect all subjects in the discourse to capture topic information that can enrich zero pronoun representations. Next, we obtain the semantic representations of all sentences in the discourse containing the candidate, representing its discourse-level coreference chain. The topic information and coreference chain semantics together form the candidate's global discourse representation.

Finally, these local and global representations are fed into our neural architecture, which integrates them and models their interactions via an attention mechanism, enabling the model to accurately identify antecedents for zero pronouns.

The overall architecture of our model is illustrated in Figure 2.

4.1. Modeling Local Discourse Information

The component characters of candidate antecedents are fundamental to their semantic representation. Following prior work [17,18], we employ a recurrent neural network (RNN) to encode the embeddings of these component characters to produce a hidden state representation for each candidate.

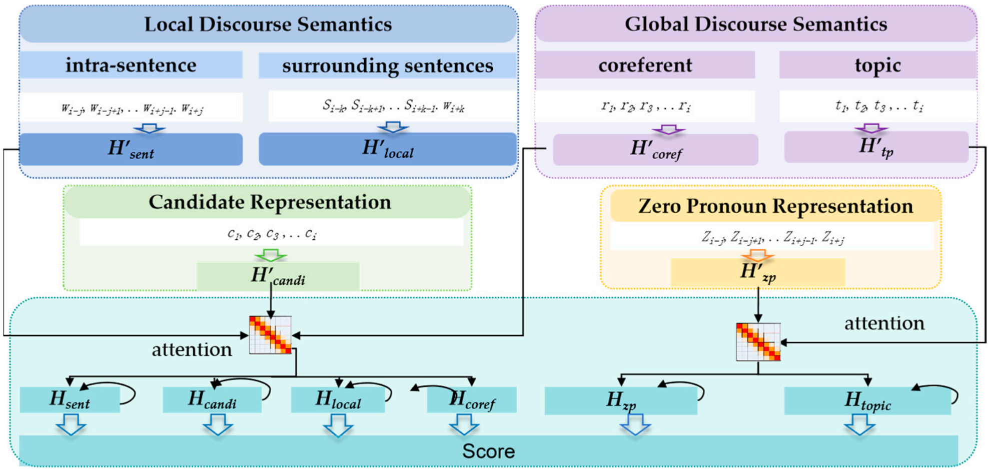

Figure 2. The architecture of the proposed Hierarchical Discourse-semantic Enhanced Model.

Furthermore, since all parts of a discourse are semantically coherent, the sentences adjacent to a candidate can serve as its local discourse context, further enhancing its semantic representation. Specifically, we extract the $2k$ nearest sentences surrounding the candidate, k sentences preceding and k sentences following the candidate's sentence, to form the local discourse context, denoted as $S_{local} = \{s_{-k}, \ldots s_0, \ldots s_{i+k}\}$, where s_0 represents the sentence containing the candidate.

The sentence vectors in S_{local} are then passed through a recurrent neural network to capture sequential dependencies, producing the local discourse representation, as illustrated in the equation below.

$$H'_{local} = RNN(\{s_{-k}, \ldots, s_{+k}\}) \tag{2}$$

The output is the hidden state representing the discourse-level context H'_{local}, which encodes the local discourse information of the candidate.

4.2. Modeling Global Discourse Information

Given that topics generally correspond to sentence subjects, we approximate discourse topics with subjects. For each zero pronoun, we extract all subjects $T = \{t_1, t_2 \ldots t_n\}$ within a window of k sentences in the discourse. These subjects can be easily identified from automatic parse trees or predicate-argument annotations. To model the representation of the i-th topic t_i, we first compute the average word embedding of all the subjects within the defined window. This aggregated representation is then fed into a recurrent neural network to generate the topic representation vector H'_{tp}.

$$v^i_{tp} = \frac{1}{m}\sum\nolimits_{j=1}^{m} v^i_j, \tag{3}$$

$$H'_{tp} = RNN(\{v^1_{tp}, \ldots, v^i_{tp} \ldots v^n_{tp}\}) \tag{4}$$

where v^j_i is the embedding of the j-th subject of the i-th topic. Then we obtain the hidden state of i-th topic information H'_{tp}.

In addition to anaphoric relations, lexical repetition serves as a prevalent mechanism for establishing coreference within discourse. Lexical expressions that refer to the same entity often recur across multiple sentences and at various syntactic positions throughout the discourse. We term these repeated mentions as shared coreferents, which not only share the same lexical form but also corefer with the candidate antecedent. Incorporating such discourse-wide semantic links can provide valuable context for more accurate zero pronoun resolution.

According to the distributional hypothesis [34], a word's meaning can be inferred from its surrounding contexts. Therefore, the sentences containing a candidate's shared coreferents provide valuable semantic evidence to enrich the candidate's representation.

For instance, consider the i-th candidate "村民" (villager). We define the set $R_{coref} = \{r_1, r_2, \ldots\ldots r_m\}$ as all sentences in the discourse containing the lexical form "村民", where m denotes the number of such sentences.

Each sentence in R_{coref} is encoded into a sentence-level vector using the pretrained language model BERT. Since these sentences are dispersed throughout the discourse and not necessarily adjacent, we compute the mean vector of these sentence embeddings. This averaged vector is then passed through a feed-forward neural network to obtain the final shared coreferent representation, as formulated below.

$$v_{coref} = \frac{1}{m}\sum\nolimits_{i=1}^{m} v_i^t, \tag{5}$$

$$H'_{coref} = FNN(wv_{coref} + b) \tag{6}$$

where v_i^t is the i-th sentence vector of candidate t. w and b are the parameters of the feed-forward neural network. This process yields the hidden state of the shared coreferent context H'_{coref} for the candidates.

4.3. Hierarchical Discourse-Semantic Enhanced Model

We first feed the embeddings of the zero pronoun's surrounding words and the candidate's component characters into a recurrent neural network to obtain their initial representations: the pronoun context vector v_{zp} and the candidate representation vector v_{candi} respectively. Additionally, following the work [18], we extract the intra-sentence context vector H'_{sent} by encoding the words surrounding the candidate within the same sentence.

To capture the interactions among these representations, we apply a nonlinear projected attention mechanism. The attention scores are computed between each pair of related vectors to highlight relevant information, using the following formula:

$$att = softmax(\mathrm{ReLU}(W_1^T H_1')^T * \mathrm{ReLU}(W_2^T H_2')) \tag{7}$$

Then we update the vectors using bi-directional attention. The zero pronoun representation H'_{zp} is refined by attending to the topic vector H'_{tp}. The candidate representation H'_{candi} is enhanced through attention over all the three types of discourse information including the intra-sentence context H'_{sent}, the local discourse context H'_{local}, and the shared coreferent context H'_{coref}, as described in the equations below.

$$H_{zp} = Attn(H'_{zp} H'_{tp}) \tag{8}$$

$$H_{candi} = Attn(H'_{candi}, \{H'_{sent}, H'_{local}, H'_{coref}\}) \tag{9}$$

In this way, H_{zp} encodes not only the sentence-level information, but also discourse-level topic semantics. Similarly, H_{candi} integrates both local and global discourse-level information. Meanwhile, H_{tp} is updated by attending to H'_{zp}, and H_{sent}, H_{local}, and H_{coref} are each refined through attention with H'_{candi}, respectively, as formulated below:

$$H_1 = H_1' + att_{12} \cdot H_2' \tag{10}$$

Finally, a self-attention layer is employed to further integrate the enhanced representations and capture higher-order semantics. The equations are as follows:

$$att_{self} = softmax(W_1 tanh(W_2 H')) \tag{11}$$

$$H = att_{self} * H' \tag{12}$$

4.4. Getting Resolution Scores

At the final stage, the outputs from the self-attention layer are fed into a two-layer feed-forward neural network to compute the resolution scores. Unlike prior approaches that only consider pairwise inputs between the zero pronoun H_{zp} and a candidate antecedent H_{candi}, our model leverage a comprehensive set of representation vectors: the zero pronoun H_{zp}, the topic H_{tp}, candidate H_{candi}, the intra-sentence context H_{sent}, the local discourse context H_{local}, and the shared coreferent context H_{coref} to calculate the final resolution score. The resolution score is computed as:

$$s_i = f(W_i s_{i-1} + b_i) \tag{13}$$

where s_i is the output of the i-th layer of the resolution network, W_i and b_i are the corresponding weight and bias parameters, and f is the activation function. The input to the network is defined as $s_0 = \{H_{zp}, H_{tp}, H_{candi}, H_{sent}, H_{local}, H_{coref}, v_{feature}\}$, where $v_{feature}$ denotes encoded additional shallow feature vectors such as positional information and semantic role labels.

The parameter settings are shown in Table 5. The model employs 768-dimensional character embeddings pretrained by BERT, from which word embeddings are obtained through averaging. Sentence embeddings are also derived from BERT, with the same dimensionality of 768.

Table 5. Hyperparameter settings of the proposed model.

Hyperparameters	Value
Learning rate	0.00005
Weight decay	0.0001
Sentence embedding dimension	768
Number of attention layers	2
Hidden layer size	256
Random seeds	0

5. Experimental Results

5.1. Evaluation

The dataset is divided into training and test sets with an 8:2 ratio, as shown in Table 6.

Table 6. Dataset size of CDAMR and OntoNotes in the experiments.

	CDAMR			OntoNotes		
	Train	**Test**	**All**	**Train**	**Test**	**All**
Documents	264	69	333	1391	172	1563
Zero Pronouns	3732	809	4541	12,111	1713	13,824

We use Precision, Recall and F-score to evaluate the performance of our method. Specifically, these metrics are defined as follows:

$$P = \frac{Num_{correct}}{Num_{predict}},\ R = \frac{Num_{correct}}{Num_{gold}}, F_1 = \frac{2 * P * R}{P + R} \tag{14}$$

where $Num_{correct}$ is the number of correctly resolved examples, $Num_{predict}$ is the number of examples predicted by the model, and Num_{gold} is the total number of zero pronouns in the test set.

5.2. Experiments on OntoNotes: Evaluating Discourse Information with Baseline Models

To evaluate the cross-corpus effectiveness and practical applicability of our proposed approach, we further incorporate hierarchical discourse semantic features into baseline models and conduct experiments on the OntoNotes dataset. This allows us to assess how well our method generalizes beyond the CDAMR corpus and verify its effectiveness on a widely used benchmark.

The first baseline is the model proposed in [18]. We enhance it by using BERT-pretrained word embeddings, and refer to this enhanced version as Baseline 1. The second baseline is adapted from the work [19]. Since our method modifies how embeddings are aggregated, specifically by computing the mean vector when incorporating topic information (tp), shared coreferent context (coref), and local discouse information (local), we denote this adapted model as Baseline 2. The experimental results are presented in Figure 3.

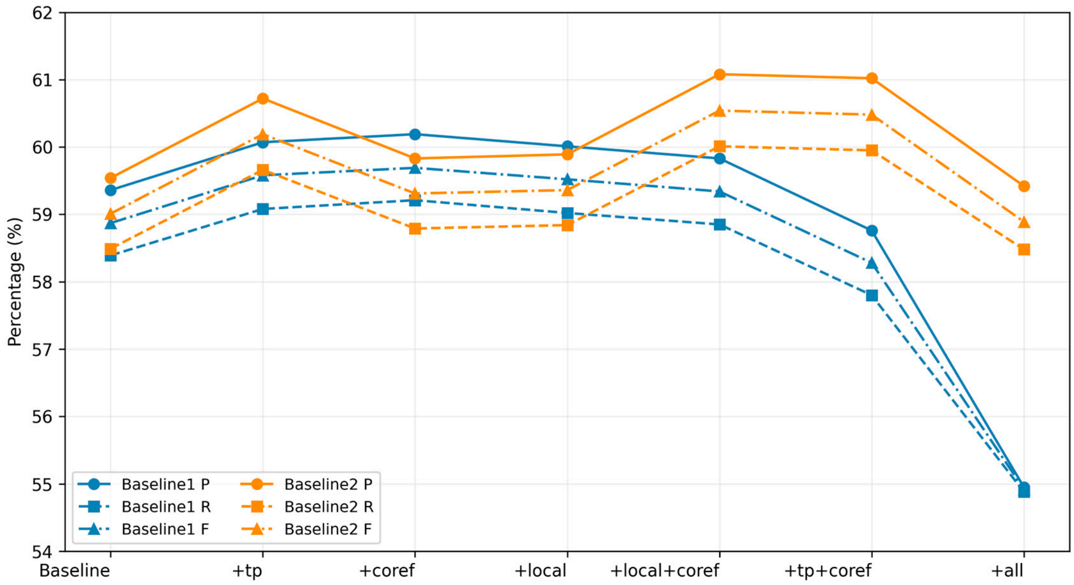

Figure 3. Experimental results with various discourse information on OntoNotes.

As shown in Figure 3, incorporating each type of discourse information notably enhances model performance. Specifically, for Baseline 1, integrating shared coreferent context information (coref) yields the largest improvement among the three discourse types. In contrast, for Baseline 2, topic information (tp) provides the highest performance boost, outperforming the other two types of discourse semantics. Furthermore, when topic information (tp) and shared coreferent context (coref) are combined in Baseline 2, the F1 score reaches 60.48%. When local discourse context (local) is combined with shared coreferent context (coref), the F1 score further increases, marking the highest performance in zero pronoun resolution under these conditions. These results demonstrate that discourse-level information, especially global-level semantics, effectively enriches the representations of both zero pronouns and candidate antecedents, ultimately boosting model performance, thereby confirming the importance of coreferent context and topic information for zero pronoun resolution.

However, not all feature combinations yield positive outcomes. Incorporating all three types simultaneously leads to a performance decline, with the F1 score dropping to 54.92% in Baseline 1 and 58.89% in Baseline 2, which are both lower than their respective baselines without discourse-level enhancements. One reason for the performance decline is that different discourse features may overlap in the information they encode, leading to redundancy. For example, topic information (tp) and local discourse context (local) both capture semantic relevance at different granularities, and when they are combined, the model may overfit to repeated signals. Another reason could be that the model lacks an effective mechanism to disentangle useful interactions from noisy ones. This indicates that simple feature aggregation may introduce noise or redundancy, undermining overall performance.

These observations underscore the importance of carefully modeling discourse features individually and capturing their interactions explicitly. Motivated by this, we proceed to evaluate our proposed discourse-semantic enhanced model on the CDAMR dataset.

5.3. Cross-Corpus Feature Ablation: Evaluating the Richness of CDAMR Annotations Compared to OntoNotes

To better understand the contribution of discourse-level semantic information, we conduct comparative experiments on both the OntoNotes and CDAMR corpora using two widely adopted baseline models [18,19]. Since OntoNotes features are mainly extracted from syntactic parse trees, while CDAMR relies on Abstract Meaning Representation (AMR) graphs, we use feature-ablated versions of these models to ensure consistent input and a fair comparison. The experimental results are summarized in Table 7.

Table 7. Comparative experiments of feature ablation on CDAMR and OntoNotes.

	Dataset	**Feature**	**P%**	**R%**	**F%**
Yin et al. [18]	OntoNotes	+			57.30
		–	43.97	43.25	43.61
	CDAMR	+	51.91	48.05	49.98
		–	51.70	47.98	49.77
Lin & Yang [19]	OntoNotes	+			60.20
		–	44.80	44.02	44.41
	CDAMR	+	51.02	47.35	49.11
		–	50.48	46.84	48.59

As shown, both models achieve higher F1 scores on CDAMR than on OntoNotes when hand-crafted features are removed. Specifically, the F1 scores on CDAMR reach 49.77% and 48.59%, respectively. Moreover, the performance degradation caused by feature ablation is much more significant on OntoNotes than on CDAMR.

This difference mainly arises because the syntactic parse trees of OntoNotes provide over 60 detailed grammatical and positional features that greatly enhance model performance. In contrast, CDAMR offers richer and more accurate discourse-level semantic information through comprehensive coreference annotations, which effectively compensate for the absence of hand-crafted syntactic features. This rich semantic information enables models to learn generalized patterns of discourse and coreference relationships, rather than relying on corpus-specific syntactic cues. Consequently, it supports stronger cross-corpus performance. As a result, CDAMR provides higher-quality discourse semantic representations, leading to better performance in zero pronoun resolution tasks.

These findings justify the motivation for designing our hierarchical discourse-semantic enhanced model, which fully leverages the rich annotations of CDAMR to improve resolution accuracy, as demonstrated in the next section.

5.4. Experiments on CDAMR: Performance of the Proposed Discourse-Semantic Enhanced Model

Following the cross-corpus analysis in Section 5.3, we now evaluate the effectiveness of our proposed discourse-aware model on the newly constructed CDAMR corpus. As CDAMR is the first Chinese zero pronoun dataset annotated with full AMR and coreference information, no prior models have been tested on it. Therefore, we first establish a strong baseline for comparison. Specifically, we extract representative vectors for the sentence-level context of zero pronouns and the component characters of candidate antecedents. These representations are used to compute attention scores, after which the updated vectors H_{zp} and H_{np} are fed into the model. On this baseline, we progressively integrate topic information and discourse-level candidate information into H_{zp} and H_{np}, respectively.

Moreover, based on the candidate window distribution analysis presented in Section 3.2, we expand the candidate sentence window from the conventional range of [−2, +0] to [−3, +1], which is annotated as +candidate in the results. The experimental outcomes are presented in Table 8.

Table 8. Experimental results with various discourse information on CDAMR.

	P%	R%	F%
Baseline	56.51	56.37	56.44
+tp	58.74	58.59	58.66
+coref	57.13	56.98	57.05
+local	55.14	55.01	55.07
+sent	57.87	57.73	57.80
+feature	57.62	57.47	57.54
+candidate	57.18	57.11	57.14
+all	54.95	54.88	54.92
Our model	59.90	59.83	59.86

As shown in Table 8, and consistent with the findings on OntoNotes, incorporating topic information and shared coreferent context yields notable performance improvements. Additionally, extending the candidate window and adding structural features further enhance model performance, highlighting the complementary nature of syntactic and discourse cues.

However, when all types of discourse information are naively concatenated in the baseline model, performance drops significantly, mirroring the decline observed in OntoNotes experiments. This degradation likely stems from the model's inability to effectively disentangle and leverage the unique contributions of each type of discourse information when they simply combined. Due to the implicit nature of zero pronouns, it is difficult for the model to determine in advance which discourse features are most relevant. Simple feature stacking may introduce noise or irrelevant signals, thereby harming performance. This is precisely why we adopt an attention mechanism to allow the model to learn which features are most useful in different contexts through training.

Our proposed discourse-semantic enhanced model treats each type of discourse information independently and explicitly models their interactions via a structured attention mechanism. Consequently, it achieves an F1 score of 59.86%, outperforming all single-feature baselines and validating the effectiveness of our integration strategy. These results further confirm the importance of discourse semantics in enriching the representation of both zero pronouns and their candidate antecedents.

5.5. Error Analysis

To further understand the behavior of our model, we examine specific examples that illustrate both its strengths and limitations. The effectiveness of the model can be seen in the following case:

> “在一座陡峭的山峰上，有一只猴子。它两只胳膊抱着腿，*pro* 一动不动地蹲在山头。” (On a steep mountain peak, there was a monkey. It held its legs with both arms, and *pro* squatted motionlessly at the top).

Possible intra-sentential antecedents for pro include “它 (it)”, “胳膊 (arms)”, and “腿 (legs)”. A model that only considers the current sentence is likely to select “it” as the antecedent. However, “it” is a pronoun itself and lacks referential clarity, which means that resolving *pro* to “it” fails to establish a meaningful reference. In contrast, our model leverages discourse-level information and includes the noun phrase “monkey” from the previous sentence as a valid candidate via a shared coreference context mechanism. By integrating hierarchical discourse features and modeling their interactions through an attention mechanism, our model correctly resolves *pro* to “monkey”.

Furthermore, our model also demonstrates strong capability in handling longer-range, cross-sentence references. Consider the following example:

> 老人说：”你问的那只骆驼，是不是左脚有点跛？” “是的。” “*pro*是不是左边驮着蜜，右边驮着米？” “不错。” “*pro*是不是缺了一颗牙齿？”
>
> (An old man said, “The camel you asked about, does it have a slight limp on its left leg?” “Yes.” “Is *pro* carrying honey on the left side and rice on the right side?” “That’s right.” “Does *pro* have a missing tooth?”)

In this case, the antecedent of the last *pro* is “骆驼 (camel)” mentioned in the very first sentence. Traditional sentence-level models, which limit candidate antecedents to the current sentence, cannot locate the correct referent because it appears far from the immediate context. By incorporating topic-based global discourse semantics, our model is able to retrieve “camel” as a valid antecedent across multiple preceding sentences. This demonstrates that our approach not only resolves intra-sentential pronouns effectively, but also significantly improves long-distance coreference resolution.

However, it is worth noting that certain types of zero pronouns remain particularly challenging for our model to resolve. One challenging type arises when a zero pronoun appears at the beginning of a sentence. For example, in the sentence “*pro*吃过午饭, 奶奶要睡午觉, 妈妈收了棉被铺到床上 (After *pro* having lunch, grandma went to take a nap, and mom put away the quilt and spread it on the bed.)”, the model incorrectly resolves *pro* to the subject of the previous sentence instead of recognizing that the correct antecedent is “奶奶 (grandma)” within the current sentence. This suggests that the model tends to prefer candidates appearing immediately after the zero pronoun, making antecedents located earlier in the context more difficult to capture.

Another challenging type involves zero pronouns in sentences that use metaphors or descriptive figurative expressions. For example, consider the sentence: “就说仙桃石吧，它好像从天上飞下来的大桃子，*pro*落到了山顶的石盘上。 (Take the Xiantao Stone for instance. It resembles a giant peach falling from the sky, and pro landed on the stone plate at the mountain top.) ” Here, the potential antecedents for the zero pronoun are “仙桃石 (Xiantao Stone)” and “大桃子 (giant peach)”. While human readers can easily infer that pro refers to “仙桃石”, the model struggles to distinguish between the literal referent and the metaphorical entity, and it incorrectly resolves the zero pronoun to “大桃子”, likely due to its closer syntactic position and the vividness of the metaphor.

6. Discussion and Future Work

In this work, we demonstrate that incorporating discourse-level semantic information can substantially improve Chinese zero pronoun resolution, revealing the critical role of broader discourse relationships in implicit language comprehension.

Moreover, our approach remains applicable to typical natural language understanding tasks and is adaptable to other pro-drop languages such as Japanese and Italian, since it primarily relies on using the subjects within a discourse to approximate the discourse topic, which can be straightforwardly obtained through syntactic analysis. While language-specific adaptations may be required to handle unique discourse structures, the central mechanism which utilizes hierarchical discourse-semantics to resolve zero pronouns, remains broadly relevant. In addition, our approach has direct practical value in downstream applications such as dialogue systems, where it can improve coherence and context awareness; machine translation, where it helps preserve meaning across languages with different pronoun-drop patterns; and summarization, where it helps produce more accurate and fluent summaries.

Our findings reveal several important insights:

First, discourse-level semantic information plays a critical role in zero pronoun resolution. While local context provides immediate semantic cues surrounding the candidate antecedents, shared coreferent contexts gather information from across the entire discourse, enriching the semantic representation of each mention or candidate. Modeling broader discourse relationships is essential not only for zero pronoun resolution but also for other forms of implicit language comprehension, in which meaning is inferred beyond what is explicitly stated at the sentence level.

Second, each type of discourse-level semantic information possesses distinct characteristics and intrinsic properties. Therefore, it is crucial to represent each type of discourse semantics separately, and model their interactions.

Furthermore, the comparison between OntoNotes and CDAMR suggests that providing more comprehensive and in-depth discourse-level annotations enables models to better capture discourse relationships, which is especially important for languages like Chinese where explicit syntactic cues are often limited.

The limitations of this work include: First, our experiments are based on BERT-based models to demonstrate the importance of discourse-level semantics; Second, we focus primarily on local discourse semantics and shared coreferent contexts, while other valuable discourse information, such as speaker intentions, and pragmatic cues, remain unexplored and underutilized.

In future work, we aim to develop more accurate methods to model deeper discourse semantics and improve the model's capacity for implicit language comprehension. In particular, exploring alternative feature combination strategies, such as tensor fusion, or hierarchical modeling, may further enhance integration effectiveness. We also plan to incorporate external knowledge to further enhance the model's generalizability and robustness. In addition, we plan to explore the use of CDAMR in downstream tasks such as text summarization and discourse parsing, where annotated discourse-level coreference and zero pronouns can provide valuable context.

Author Contributions: Conceptualization, W.Q., X.Y. and T.W.; methodology, W.Q., T.W. and J.L.; formal analysis, W.Q. and T.W.; resources, J.L. and X.Y.; writing—original draft preparation, T.W.; writing—review and editing, J.L. and X.Y.; supervision, W.Q.; funding acquisition, W.Q. and X.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Major Project of the National Social Science Foundation of China, Grant Reference Number: 21&ZD88; The Humanities Project of the Ministry of Education of China, Grant Reference Number: 24YJAZH198.

Data Availability Statement: Database can be downloaded via https://github.com/NJNU-CLW/ILGDS (accessed on 9 August 2025).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chowdhery, A.; Narang, S.; Devlin, J.; Bosma, M.; Mishra, G.; Roberts, A.; Barham, P.; Chung, H.W.; Sutton, C.; Gehrmann, S.; et al. PALM: Scaling Language Modeling with Pathways. *J. Mach. Learn. Res.* **2023**, *24*, 11324–11436.
2. Dasgupta, I.; Lampinen, A.K.; Chan, S.C.; Creswell, A.; Kumaran, D.; McClelland, J.L.; Hill, F. Language models show human-like content effects on reasoning. *arXiv* **2022**, arXiv:2207.07051.
3. Wang, Y.; Yu, Z.; Yao, W.; Zeng, Z.; Yang, L.; Wang, C.; Chen, H.; Jiang, C.; Xie, R.; Wang, J.; et al. PandaLM: An automatic evaluation benchmark for LLM instruction tuning optimization. In Proceedings of the Twelfth International Conference on Learning Representations, Vienna, Austria, 7–11 May 2024; pp. 1–21.
4. Isik, B.; Ponomareva, N.; Hazimeh, H.; Paparas, D.; Vassilvitskii, S.; Koyejo, S. Scaling laws for downstream task performance of large language models. In Proceedings of the ICLR 2024 Workshop on Mathematical and Empirical Understanding of Foundation Models, Vienna, Austria, 7–11 May 2024; pp. 1–18.
5. Kumar, P. Large language models (LLMs): Survey, technical frameworks, and future challenges. *Artif. Intell. Rev.* **2024**, *57*, 260. [CrossRef]
6. Zhu, Y.; Moniz, J.R.A.; Bhargava, S.; Lu, J.; Piraviperumal, D.; Zhang, Y.; Yu, H.; Tseng, B. Can Large Language Models Understand Context? *arXiv* **2024**, arXiv:2402.00858. [CrossRef]
7. Xu, S.; Li, P.; Zhu, Q. Corefprompt: Prompt-based event coreference resolution by measuring event type and argument compatibilities. *arXiv* **2023**, arXiv:2310.14512.
8. Martinelli, G.; Barba, E.; Navigli, R. Maverick: Efficient and accurate coreference resolution defying recent trends. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics, Bangkok, Thailand, 11–16 August 2024; Volume 1, pp. 13380–13394. Available online: https://aclanthology.org/2024.acl-long.722 (accessed on 2 July 2025).
9. Chen, J.; Durrett, G. Robust question answering through sub-part alignment. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Online, 6–11 June 2021; pp. 1251–1263.
10. Wu, W.; Wang, F.; Yuan, A.; Wu, F.; Li, J. CorefQA: Coreference resolution as query-based span prediction. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, 5–10 July 2020; pp. 6953–6963.
11. Liu, Z.; Shi, K.; Chen, N. Coreference-aware dialogue summarization. In Proceedings of the 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue, Singapore and Online, 29–31 July 2021; pp. 509–519.
12. Li, C.N.; Thompson, S.A. Third-person pronouns and zero-anaphora in Chinese discourse. *Synth. Semant.* **1979**, *12*, 311–335.
13. Di Eugenio, B. Centering theory and the Italian pronominal system. In Proceedings of the 13th COLING, Helsinki, Finland, 20–25 August 1990; pp. 1–6.
14. Kameyama, M. Zero Anaphora: The Case of Japanese. Ph.D. Dissertation, Stanford University, Stanford, CA, USA, August 1985; pp. 100–110.
15. Chen, C.; Ng, V. Chinese zero pronoun resolution with deep neural networks. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 7–12 August 2016; pp. 778–788.
16. Yin, Q.; Zhang, Y.; Zhang, W.; Liu, T. Chinese zero pronoun resolution with deep memory network. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 9–11 September 2017; pp. 1309–1318.
17. Yin, Q.; Zhang, W.; Zhang, Y.; Liu, T. A deep neural network for Chinese zero pronoun resolution. In Proceedings of the 26th International Joint Conference on Artificial Intelligence, Melbourne, Australia, 19–25 August 2017; pp. 3322–3328.
18. Yin, Q.; Zhang, Y.; Zhang, W.; Liu, T.; Wang, W. Zero pronoun resolution with attention-based neural network. In Proceedings of the 27th International Conference on Computational Linguistics, Santa Fe, NM, USA, 20–26 August 2018; pp. 13–23.
19. Lin, P.; Yang, M. Hierarchical attention network with pairwise loss for Chinese zero pronoun resolution. In Proceedings of the Association for the Advancement of Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 8352–8359.
20. Sun, S. Self-attention enhanced CNNs with average margin loss for Chinese zero pronoun resolution. *Appl. Intell.* **2022**, *52*, 5739–5750. [CrossRef]
21. Bi, M.; Liu, X.; Zhang, Q.; Yang, Z. Machine reading comprehension combined with semantic dependency for Chinese zero pronoun resolution. *Artif. Intell. Rev.* **2023**, *56*, 7597–7612. [CrossRef]

22. Longacre, R.E. The Paragraph as a Grammatical Unit. In *Discourse and Syntax*; Academic Press: Cambridge, MA, USA, 1979; Volume 12, pp. 115–134.
23. Gundel, J.K.; Hedberg, N.; Zacharski, R. Cognitive status and the form of referring expressions in discourse. *Language* **1993**, *69*, 274–307. [CrossRef]
24. Yang, N.; Zhang, J.; Ma, L.; Lu, Z. A study of zero pronoun anaphora resolution in Chinese discourse: From the perspective of psycholinguistics. *Front. Psychol.* **2021**, *12*, 663168.
25. Hirst, G. Discourse-oriented anaphora resolution in natural language understanding: A review. *Am. J. Comput. Linguist.* **1981**, *7*, 85–98.
26. Huang, Y. Discourse anaphora: Four theoretical models. *J. Pragmat.* **2000**, *32*, 151–176. [CrossRef]
27. Kibrik, A.A. Anaphora in Russian narrative prose: A cognitive calculative account. *Typol. Stud. Lang.* **1996**, *33*, 255–304.
28. Givon, T. *Topic Continuity in Discourse: An Introduction*; John Benjamins: Amsterdam, The Netherlands, 1983; pp. 1–42.
29. Yeh, C.L.; Chen, Y.C. Zero anaphora resolution in Chinese with shallow parsing. *J. Chin. Lang. Comput.* **2007**, *17*, 41–56.
30. Zhao, S.; Ng, H.T. Identification and resolution of Chinese zero pronouns: A machine learning approach. In Proceedings of the 2007 Conference on Empirical Methods in Natural Language Processing, Prague, Czech Republic, 28–30 June 2007; pp. 541–550.
31. Kong, F.; Zhou, G. A tree kernel-based unified framework for Chinese zero anaphora resolution. In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, Cambridge, MA, USA, 9–11 October 2010; pp. 882–891.
32. Wang, L.; Liu, S.; Xu, M.; Song, L.; Shi, S.; Tu, Z. A survey on zero pronoun translation. *arXiv* **2023**, arXiv:2305.10196. [CrossRef]
33. Ueyama, A.; Kano, Y. Dialogue Response Generation Using Completion of Omitted Predicate Arguments Based on Zero Anaphora Resolution. In Proceedings of the 24th Annual Meeting of the Special Interest Group on Discourse and Dialogue, Prague, Czech Republic, 11–12 September 2023; pp. 282–296.
34. Li, B.; Wen, Y.; Qu, W.G.; Bu, L.; Xue, N. Annotating the Little Prince with Chinese AMRs. In Proceedings of the Linguistic Annotation Workshop Held in Conjunction with ACL, Berlin, Germany, 11 August 2016; pp. 7–15.
35. Banarescu, L.; Bonial, C.; Cai, S.; Georgescu, M.; Griffitt, K.; Hermjakob, U.; Knight, K.; Koehn, P.; Palmer, M.; Schneider, N. Abstract meaning representation for sembanking. In Proceedings of the 7th Annotation Workshop & Interoperability with Discourse, Sofia, Bulgaria, 8–9 August 2013; pp. 178–186.
36. Xi, X.; Sun, Q.; Zhou, G. Research and Prospect of Discourse Topic Structure Analysis for Discourse Intentionality. *Chin. J. Comput.* **2019**, *42*, 2770–2794.
37. Sheng, C.; Kong, F.; Zhou, G. Building Chinese Zero Corpus form Discourse Perspective. *Acta Sci. Nat. Univ. Pekin.* **2019**, *55*, 15–21.
38. Wang, Q.; Liang, J. The Division of Labor Between Null and Overt Pronouns in Discourse Anaphora. *J. Foreign Lang.* **2020**, *43*, 2–12.