Lire Journal (Journal of Linguistics and Literature)
https://lirejournal.ubb.ac.id/index.php/LRJ/index
P-ISSN: 2598-1803 E-ISSN: 2581-2130
Volume 10 Number 2 2026

RECLASSIFYING CHINESE PHRASE STRUCTURES: FROM FORMAL CONFIGURATION TO SEMANTIC RELATIONS

*Sun Jiaze[1], Mulyadi[2]
[1]*Faculty of Cultural Sciences, Universitas Sumatera Utara, Medan 20155, Indonesia*
[2]*Faculty of Cultural Sciences, Universitas Sumatera Utara, Medan 20155, Indonesia*

***Corresponding Author:** Sun Jiaze**, Email:** 763815563@qq.com

ARTICLE INFO
Received: 12-03-2026
Revised: 07-04-2026
Accepted: 25-04-2026
Published: 01-05-2026
Volume: 10
Issue: 2
DOI:
https://doi.org/10.33019/lire.v10i2.604

KEYWORDS
Chinese phrase structures; semantic relations; formal configuration; phrase classification; Chinese syntax

ABSTRACT
The classification of Chinese phrase structures has traditionally been based on formal configuration, especially categories such as subject-predicate, verb-object, and modifier-head constructions. Although this approach is descriptively useful, it does not always explain why phrases with similar forms show different syntactic behavior, or why different forms may perform similar grammatical functions. This study therefore proposes a reclassification of Chinese phrase structures from formal configuration to semantic relations. Using a qualitative corpus-based descriptive design, the study analyzes 800 modern Chinese phrase tokens drawn from dictionaries, literary texts, news discourse, and linguistic studies. The findings show that semantic classification more effectively explains phenomena such as same form, different function, different form, same function, and form-meaning divergence. The study identifies core semantic relations, including agent-action, patient-action, instrument-action, location-action, object-action, restrictive, coordinative, and complementive relations, and demonstrates that these are closely related to sentence-convertibility, collocational compatibility, nominalization tendency, and ambiguity resolution. Through examples such as 人才交流 (*réncái jiāoliú*), 学者讨论 (*xuézhě tǎolùn*), 晒太阳 (*shài tàiyáng*), and 修房子 (*xiū fángzi*), the study argues that semantic relations provide a more adequate explanatory basis for Chinese phrase classification. It concludes that formal structure remains an important descriptive layer, while semantic relations offer a more explanatory framework for analysing Chinese phrase structures, especially in cases involving ambiguity, form–meaning divergence, and differences in syntactic behaviour.

1. INTRODUCTION

The classification of Chinese phrase structures has long occupied a central position in Chinese grammatical theory because it determines how multi-word units are segmented, interpreted, and linked to sentence formation. In the earliest stage of modern Chinese grammar, Ma (1983) established a strong formal orientation by treating phrase patterns as observable structural units, while Ding et al. (1961) further stabilized this tradition through systematic descriptions of subject-predicate, verb-object, modifier-head, coordinate, complement, and preposition-object constructions. From this perspective, phrase classification was not merely a descriptive tool but a foundational mechanism for organizing the grammar of Chinese.

The structural basis of phrase classification was reinforced by structuralist linguistics, especially by the assumption that language can be described through recurring distributional and configurational relations. Bloomfield (1980) provided the broader theoretical background for such a view, and within Chinese linguistics this tendency was pushed further by Fan (1980a) and Fan (1980b), who helped shape a phrase-centered analytical orientation in which phrase structure and sentence structure were treated as closely aligned. This framework offered a neat and operational typology, but it also encouraged the assumption that similar forms necessarily share similar grammatical properties.

Later descriptive scholarship preserved and refined this formal model. Huang and Liao (2011) maintained the canonical classification of Chinese phrases in a way that remains highly influential in grammar teaching, while Lu (2003) translated these structural distinctions into an accessible pedagogical grammar. Meanwhile, Yang and Li (1987) showed that structurally patterned expressions could also be studied historically, reinforcing the idea that formal relations are essential to Chinese grammatical description. Taken together, these studies examined phrasehood primarily through structural patterning and made formal classification the default framework for analyzing Chinese phrases.

Yet the limits of this tradition had already been noticed in earlier analytical reflection. Lü (1979) argued that phrase-internal relations should not be conflated with sentence-level syntactic functions, because phrases and sentences are not analytically identical units. This warning is important because it exposes a core weakness in the formal approach: previous studies successfully described phrase types, but they did not fully explain why phrases with similar configurations often display different grammatical behavior. What remained insufficiently clarified, therefore, was the deeper principle governing phrasehood beyond surface arrangement.

A typological perspective makes this problem even more visible. Liu (2004) showed that word order in Chinese is intimately connected to semantic organization, and Xu (2014) further emphasized that structural interpretation cannot be separated from discourse-semantic linkage. In the same vein, Cheng and Sybesma (2014) demonstrated that Chinese constituent structure, especially in the nominal domain, is partly shaped by semantic proximity and interpretive dependency rather than by form alone. These studies significantly expanded the field by showing that phrase structure is not exhausted by visible configuration, even though they stopped short of proposing a full semantic reclassification of Chinese phrases.

The inadequacy of purely formal classification becomes most obvious when one encounters phrases whose surface forms are identical but whose syntactic behavior diverges. This is precisely the type of problem that formal grammars struggle to resolve, especially in cases involving collocational restrictions, nominalization asymmetries, or differences in sentence-conversion potential. More recent psycholinguistic and neurocognitive studies strengthen this point: Wolpert et al. (2024) show that in Mandarin, word order alone is not always a reliable cue to argument structure, while Zhang et al. (2024) demonstrate that semantic factors can actively drive syntactic parsing in Chinese sentence processing. Likewise, Liu et al. (2024) show that semantics interacts with syntax during Chinese relative clause processing.

Recent formalized and computational work also points in the same direction. Peng et al. (2021) argue that Chinese sentence pattern analysis becomes more explanatory when hierarchical

representation is used, and Hao (2024) further shows that complex Chinese relational structures are difficult to capture through formalized clause patterns alone. At the level of semantic modeling, Li et al. (2020) and Chen et al. (2020) show that relation extraction depends heavily on semantic representation rather than on surface structure only. What these studies have examined is the enrichment of structural analysis through semantic information; what remains unaddressed is the systematic use of semantic relations as the basis for phrase classification itself.

The same tendency is visible in recent applied linguistic and NLP research. Wang et al. (2021), Qin et al. (2021), and Wan et al. (2023) all demonstrate that semantic relation modeling improves the interpretation of complex linguistic data, particularly where formal ambiguity is present. Related work by Xu et al. (2022) and Mao et al. (2020) similarly confirms that language-processing performance increases when latent semantic dependencies are made explicit. Although these studies operate largely at the level of event extraction, relation detection, or structured classification, their broader implication is highly relevant for Chinese linguistics: structural form alone is often insufficient for explaining relational meaning.

More recent work continues to confirm the importance of semantic differentiation in Mandarin analysis. Yu et al. (2022) and Chu et al. (2022) illustrate how context-sensitive modeling improves linguistic classification, while Zhu and Grüter (2025) show that both native speakers and learners of Mandarin predict upcoming arguments in dative constructions on the basis of verb constraints rather than form alone. In a broader semantic perspective, Yang and Baayen (2025) demonstrate that Mandarin lexical organization reveals patterned semantic structuring that cannot be reduced to superficial formal similarity. These developments show that semantic relations are not merely supportive interpretive features; they are increasingly central to how Mandarin structure is modeled and understood.

Despite these advances, a major research gap remains. Existing scholarship has described Chinese phrase structures extensively, and recent studies have increasingly acknowledged the importance of semantic interpretation, but there is still limited work that systematically examines how semantic relations can complement formal classification while also explaining syntactic behaviour in a principled way. This gap matters because previous studies have primarily emphasized structural typology or semantically enriched interpretation, whereas less attention has been given to the role of semantic relations as an explanatory dimension in phrase analysis. To address this gap, the present study examines how semantic-relational analysis can refine the interpretation of Chinese phrase structures, particularly in cases where formal configuration alone is insufficient to account for ambiguity, form–meaning divergence, and differences in grammatical behaviour.

This study does not argue that semantic relations categorically replace formal classification. Rather, it argues that semantic-relational analysis improves explanatory adequacy, especially in cases where formal configuration alone cannot sufficiently account for ambiguity, form–meaning divergence, and grammatical behaviour.

2. METHODOLOGY

2.1 Research Design

This study employed a qualitative corpus-based descriptive design to examine whether semantic-relational analysis provides greater explanatory adequacy than formal configuration in the analysis of modern Chinese phrase structures. The study did not aim to measure frequency patterns statistically or to test causal relationships experimentally. Instead, it compared two classificatory perspectives: the conventional form-based classification used in Chinese grammar and a semantic-relational classification focusing on the internal meaning relation between phrase constituents. This design was chosen because the central research problem concerns explanatory adequacy, particularly in cases where formal structure alone does not sufficiently account for ambiguity, form–meaning divergence, or differences in syntactic behaviour.

2.2 Data Source and Unit of Analysis

The dataset comprised 800 tokens of modern Chinese phrase constructions drawn from four source categories: standard Chinese dictionaries, modern literary texts, news discourse, and linguistic studies on Chinese phrase patterns. The corpus was distributed across these four source types in order to balance normative reference material with contextual language use. This source design was intended to ensure both descriptive reliability and contextual diversity, so that the analysis would capture not only canonical phrase forms but also naturally occurring usages across different communicative settings.

The sampling strategy was purposive and theoretically informed rather than random. Tokens were selected on the basis of their relevance to the analytical objectives of the study, namely: representing established phrase patterns in modern Chinese, including cases in which formal structure and semantic relation were aligned, and incorporating cases in which ambiguity, form–meaning divergence, or mismatched grammatical behaviour rendered formal classification alone insufficient. This sampling logic was adopted because the study was intended to evaluate explanatory adequacy rather than to model corpus-wide frequency distribution.

The unit of analysis was the phrase token, rather than the sentence as a whole or the phrase type in the abstract. In the present study, a phrase token refers to one analysable occurrence of a multi-constituent Chinese phrase in a specific source or contextual environment. A token was regarded as valid when it consisted of at least two constituents forming a recognisable phrase unit in modern Chinese and when it could be examined simultaneously in terms of both formal configuration and semantic relation.

2.3 Analytical Procedures

The analysis proceeded in four stages. First, each phrase token was classified according to conventional structural categories in Chinese grammar, including subject-predicate, verb-object, modifier-head, coordinate, complement, and preposition-object constructions. This first stage provided the descriptive baseline against which the semantic-relational analysis could be compared.

Second, each token was coded for its internal semantic relation. The coding categories included agent–action, patient–action, instrument–action, location–action, object–action, restrictive, coordinative, complementive, and other semantically relevant relations identified

Lire Journal (Journal of Linguistics and Literature)
https://lirejournal.ubb.ac.id/index.php/LRJ/index
P-ISSN: 2598-1803 E-ISSN: 2581-2130
Volume 10 Number 2 2026

during corpus examination. At this stage, the analysis focused on the underlying relational meaning between constituents rather than on surface order alone.

Third, the formal classification and the semantic-relational classification were compared systematically in order to identify recurrent mismatch patterns. Three comparison types were used: (1) same form, different semantic relation; (2) different form, same semantic relation; and (3) form–meaning divergence. This stage was central to evaluating whether semantic-relational classification explained phrase behaviour more adequately than formal labelling alone.

Fourth, the study examined how each semantic relation correlated with syntactic behaviour. In this stage, phrase tokens were analysed in relation to collocational compatibility, nominalisation tendency, and sentence-convertibility. These functional properties were used not as separate dependent variables in a statistical model, but as interpretive indicators of whether a given classification framework could account for actual grammatical behaviour more satisfactorily.

2.4 Coding Framework and Reliability Procedures

The study employed a layered coding framework comprising four analytical levels: formal configuration, semantic relation, comparative correspondence, and functional implication. The formal layer recorded observable structural patterns in accordance with established classifications in Chinese grammar. The semantic layer identified the underlying relational meaning between phrase constituents. The comparative layer examined points of convergence and divergence between formal classification and semantic-relational classification. The functional layer considered how these relations were associated with grammatical behaviour, including collocational compatibility, nominalisation tendency, and sentence-convertibility.

To enhance coding consistency, the criteria for classification were established prior to full-scale annotation. A preliminary coding guide was developed to define each formal category and each semantic-relation category, together with representative examples and boundary conditions for difficult or borderline cases. This guide served as the primary analytical reference throughout the coding process and was used to maintain consistency across the dataset.

Table 1. Analytical Framework for Phrase Classification

Coding Layer	Analytical Focus	Main Categories	Function in the Study
Formal configuration	Surface structural pattern	Subject-predicate, verb-object, modifier-head, coordinate, complement, preposition-object	Provides the baseline classification of traditional Chinese grammar
Semantic relation	Internal meaning relation	Agent-action, patient-action, instrument-action, location-action, object-action, restrictive, coordinative, complementive	Provides the proposed semantic-relational classification
Comparativ	Correspondence	Same form-	Identifies

e layer	between form and meaning	different function; different form-same function; form-meaning divergence	the limits of formal classification
Functional layer	Syntactic consequences	Collocational compatibility, nominalization tendency, sentence-convertibility	Evaluates explanatory adequacy of semantic classification

2.6 Methodological Rationale

This methodological design was selected because the study addresses a problem of grammatical explanation rather than one of distributional frequency alone. A purely structural procedure would reproduce the traditional typology of Chinese phrase classification, but it would not adequately explain why structurally similar phrases may differ in semantic interpretation and syntactic behaviour. Conversely, an exclusively intuitive semantic approach would risk weakening classificatory discipline by relying too heavily on interpretive judgment without a stable descriptive framework.

For this reason, the study combined corpus-based description, purposive sampling, layered coding, contextual ambiguity analysis, and expert review. This combination made it possible to preserve the descriptive value of formal classification while also examining whether semantic relations offered greater explanatory adequacy in cases where formal configuration alone proved insufficient. The resulting design therefore provides a more explicit, transparent, and analytically auditable framework for investigating how semantic-relational analysis can complement and refine the classification of Chinese phrase structures.

3. RESULTS AND DISCUSSION

3.1 Results

An effective classification of Chinese phrase structures should improve the explanation of actual syntactic behavior. The results of this study show that a purely formal classification is descriptively useful but explanatorily limited. Based on the 800-phrase corpus, structural labels such as subject-predicate, verb-object, modifier-head, and preposition-object do not always predict how phrases behave when they are nominalized, expanded, embedded, or converted into sentence constituents. By contrast, classification based on semantic relations more directly captures the internal logic of phrase composition and therefore provides a stronger basis for explaining syntactic distribution.

The corpus analysis further shows that one of the main weaknesses of traditional phrase analysis lies in the conflation of three different levels: phrase-internal semantic relation, sentence-level structural relation, and surface syntactic function. At the phrase level, the relation between constituents is fundamentally semantic: a constituent may denote an agent, patient, instrument, location, object, or qualifier in relation to another. At the sentence level, however, the same phrase may function as a subject, object, predicate, or modifier. When these levels are

Lire Journal (Journal of Linguistics and Literature)
https://lirejournal.ubb.ac.id/index.php/LRJ/index
P-ISSN: 2598-1803 E-ISSN: 2581-2130
Volume 10 Number 2 2026

not clearly distinguished, structurally similar phrases may appear to belong to the same type even though they show different grammatical behavior. The present findings therefore support the view that semantic organization should be treated as a crucial explanatory dimension in phrase analysis, particularly where formal configuration alone proves insufficient.

To operationalize this claim, the study proposes a semantic taxonomy that distinguishes Chinese phrases according to their internal relational meaning. The corpus shows that this taxonomy explains more effectively three recurrent phenomena: same form, different function; different form, same function; and form-meaning divergence. Table 2 presents the core semantic categories identified in the present study.

Table 2. Core Semantic Categories for Chinese Phrase Classification

No	Semantic relation	Illustrative example	General interpretation
1	Coordinative	政府与学校 (*zhèngfǔ yǔ xuéxiào*, "government and schools"); 阅读与欣赏 (*yuèdú yǔ xīnshǎng*, "reading and appreciation")	two constituents of equal semantic status
2	Restrictive	美丽风景 (*měilì fēngjǐng*, "beautiful scenery"); 大胆想法 (*dàdǎn xiǎngfǎ*, "bold idea")	one constituent semantically restricts the other
3	Complementive	讲清楚 (*jiǎng qīngchu*, "explain clearly"); 吃饱 (*chī bǎo*, "eat fully")	the second element completes or specifies the first
4	Agent–action	老师讲课 (*lǎoshī jiǎngkè*, "the teacher teaches"); 鸟飞 (*niǎo fēi*, "the bird flies")	the first constituent denotes the doer of the action
5	Patient–action	修房子 (*xiū fángzi*, "repair houses"); 看小说 (*kàn xiǎoshuō*, "read novels")	one constituent denotes the affected entity of the action
6	Judgment / identificational	首都北京 (*shǒudū Běijīng*, "Beijing, the capital"); 班长小李 (*bānzhǎng Xiǎolǐ*, "Xiao Li, the class monitor")	one constituent identifies or judges the other
7	Depictive / descriptive	穿得朴素 (*chuān de pǔsù*, "dressed plainly"); 通红 (*tōnghóng*, "bright red")	one constituent describes the state or manner of another
8	Double-object /	问他问题 (*wèn tā wèntí*, "ask	the relation

Lire Journal (Journal of Linguistics and Literature)
https://lirejournal.ubb.ac.id/index.php/LRJ/index
P-ISSN: 2598-1803 E-ISSN: 2581-2130
Volume 10 Number 2 2026

recipient-related	him a question"); 给我知识 (*gěi wǒ zhīshi*, "give me knowledge")	involves transfer toward a recipient

Table 2 shows that semantic classification does not merely rename traditional structural categories. Rather, it reorganizes phrasehood around relational meaning. The value of this approach lies in its ability to distinguish phrases that may share a structural form but differ in semantic composition. In this sense, semantic classification reflects the fact that Chinese phrases are not miniature sentences, but relatively stable intermediate units whose syntactic potential depends on the semantic relations among their constituents.

A second result concerns the interface between semantic relations and formal structures. The corpus shows that one semantic relation may be realized through more than one formal pattern, and conversely, one formal pattern may encode more than one semantic relation. This confirms that the connection between form and meaning in Chinese phrase structure is systematic but not one-to-one. Table 3 summarizes the major correspondences identified in the corpus.

Table 3. Correspondence Between Core Semantic Relations and Typical Formal Structures

No	Core semantic relation	Typical formal structure	Example
1	Agent–action	subject-predicate (N + V)	老师讲课 (*lǎoshī jiǎngkè*, "the teacher teaches"); 孩子跑 (*háizi pǎo*, "the child runs")
2	Agent–action	verb-object with implicit agentivity	开车 (*kāichē*, "drive a car"); 插秧 (*chāyāng*, "plant rice seedlings")
3	Patient–action	verb-object (V + N)	看小说 (*kàn xiǎoshuō*, "read novels"); 修机器 (*xiū jīqì*, "repair machinery")
4	Patient–action	modifier-head / passive-derived phrase	修好的设备 (*xiūhǎo de shèbèi*, "repaired equipment"); 解决的问题 (*jiějué de wèntí*, "solved problems")
5	Instrument–action	preposition + noun + verb	用笔写字 (*yòng bǐ xiězì*, "write with a pen"); 通过网络学习 (*tōngguò wǎngluò xuéxí*, "learn through the internet")
6	Location–action	preposition + noun + verb	在教室上课 (*zài jiàoshì shàngkè*, "study in the classroom"); 在公园散步 (*zài gōngyuán sànbù*, "walk

			in the park")
7	Object–action	verb-object (V + N)	讨论问题 (*tǎolùn wèntí*, "discuss a problem"); 调查案件 (*diàochá ànjiàn*, "investigate a case")

The patterns in Table 3 show that semantic relations provide a more stable interpretive basis than form alone. For instance, both patient–action and object–action phrases may appear in V + N form, yet they behave differently in subsequent syntactic environments. Likewise, agent–action relations may be expressed through canonical N + V structures or through compact phrases in which the agent is only contextually recoverable. These findings support the argument that semantic-relational analysis offers greater explanatory adequacy in explaining why structurally similar phrases may diverge in grammatical behaviour.

A third result concerns the relationship between semantic categories and syntactic function. The corpus shows that phrases with different semantic relations display different degrees of collocational compatibility, nominalization tendency, and sentence-convertibility. Semantic relations therefore do not merely describe internal meaning; they also constrain how phrases behave in broader grammatical environments. Table 4 summarizes these tendencies.

Table 4. Semantic Relations and Their Typical Syntactic Tendencies

No.	Semantic relation	Typical formal variant	Common syntactic tendency
1	Agent–action	N + V	relatively high sentence-convertibility; often predicate-like
2	Patient–action	V + N / passive-derived phrase	stronger nominalization tendency; event-like object compatibility
3	Instrument–action	Prep + N + V	dependent on overt marker; lower independent sentence-convertibility
4	Location–action	Prep + N + V	often adverbial or locative in distribution
5	Object–action	V + N	high compatibility with activity or cognition verbs
6	Restrictive	modifier-head	strong attributive function; limited independent sentence status
7	Coordinative	X + X	flexible, depending on lexical class and discourse context

Lire Journal (Journal of Linguistics and Literature)
https://lirejournal.ubb.ac.id/index.php/LRJ/index
P-ISSN: 2598-1803 E-ISSN: 2581-2130
Volume 10 Number 2 2026

The evidence in Table 4 indicates that semantic relations exert direct influence on grammatical realization. Phrases with semantically complete event structures, especially many agent–action phrases such as 鸟飞 (*niǎo fēi*, "the bird flies") and 老师讲课 (*lǎoshī jiǎngkè*, "the teacher teaches"), *show stronger potential for independent sentence realization. By contrast, patient–action phrases such as 修房子 (xiū fángzi, "repair houses")* or 看小说 (*kàn xiǎoshuō*, "read novels") function more naturally as embedded units, event nominals, or object-like expressions. This pattern explains why some phrases can readily occur after verbs such as 进行 (*jìnxíng*, "conduct"), 推进 (*tuījìn*, "promote"), or 完成 (*wánchéng*, "complete"), whereas others cannot, even when their structural pattern looks similar.

A fourth result concerns ambiguity resolution. One of the clearest advantages of semantic classification is that it provides a principled method for interpreting phrases whose surface forms allow more than one reading. The corpus repeatedly shows that ambiguity in Chinese phrase structure often arises not from uncertainty about form, but from competition between alternative semantic relations. Table 5 presents representative cases.

Table 5. Structural Ambiguity and Semantic Resolution in Chinese Phrases

No.	Potential ambiguity	Example	Contextual interpretation
1	Agent–action / object–action	学习材料 (*xuéxí cáiliào*, "learning materials")	"materials used for learning" / "materials being compiled for learning"
2	Agent–action / patient–action	作业完成 (*zuòyè wánchéng*, "homework completed")	"the student completed the homework" / "the homework has been completed"
3	Restrictive / resultative-descriptive	漂亮衣服 (*piàoliang yīfu*, "beautiful clothes")	"beautiful clothes" / "clothes made beautiful"
4	Agent–action / patient–action	记者采访 (*jìzhě cǎifǎng*, "journalist interview")	"the journalist conducts the interview" / "the journalist is interviewed"

Table 5 demonstrates that ambiguity cannot be resolved through formal structure alone. The same surface phrase may support two or more interpretations depending on whether the noun is construed as an agent, patient, object, or qualifier. Semantic classification addresses this issue by specifying which relational reading is primary in context and by showing how that reading predicts subsequent grammatical behavior. This finding is especially important because it links phrase classification directly to interpretive adequacy rather than to formal labeling alone.

The findings also have implications for Chinese language teaching and register-sensitive usage. In teaching Chinese as a foreign language, learners often experience difficulty not because they fail to memorize phrase forms, but because they do not fully understand the semantic

relations between constituents. A semantic approach helps clarify this progression: relatively transparent patterns such as agent–action (e.g., 我吃 *wǒ chī*, "I eat"; 他走 *tā zǒu*, "he walks") can be introduced at an early stage, while more complex expressions involving instrument, location, cause, or compound semantic layering can be taught gradually as relational extensions. This suggests that semantic classification may provide a more systematic basis for teaching phrase structure than formal labels alone.

A similar implication emerges in stylistic variation. Phrases that share the same semantic relation may take different formal realizations in spoken and written Chinese while preserving the same relational core. For instance, spoken forms such as 大家讨论 (*dàjiā tǎolùn*, "everyone discusses") may correspond to more formal written expressions such as 集体讨论 (*jítǐ tǎolùn*, "collective discussion"), while colloquial causal constructions such as 因为堵车迟到了 may be reformulated in writing as 由于交通拥堵而迟到. This shows that semantic classification is useful not only for explaining syntax, but also for accounting for pedagogical progression and register-sensitive adaptation.

3.2 Discussion

3.2.1 Semantic classification as a more explanatory framework

This study does not argue that semantic relations categorically replace formal classification. Rather, it argues that semantic-relational analysis improves explanatory adequacy, especially in cases where formal configuration alone cannot sufficiently account for ambiguity, form–meaning divergence, and grammatical behaviour.

The findings suggest that semantic-relational analysis offers greater explanatory adequacy than formal classification in analysing Chinese phrase structures, particularly in cases where structurally similar phrases display different grammatical behaviour. Traditional labels such as subject-predicate, verb-object, and modifier-head remain useful for describing surface configuration, but they do not consistently explain why phrases that look similar behave differently in syntax. In this respect, semantic relations provide an important interpretive dimension that refines, rather than replaces, formal classification.

This point can be seen in the contrast between 人才交流 (*réncái jiāoliú*, "talent exchange") and 学者讨论 (*xuézhě tǎolùn*, "scholar discussion"). Although these expressions may appear structurally comparable, only the former combines naturally with 进行 (*jìnxíng*, "to conduct"), as in 进行人才交流. This difference indicates that what governs acceptability is not surface form alone, but whether the phrase is semantically construed as an event-like unit. Thus, semantic relations provide a more precise explanation of phrase behavior than formal labels alone.

3.2.2 Practical validation through agent–action and patient–action phrases

A particularly revealing contrast appears between agent–action phrases and patient–action phrases. Agent–action phrases such as 鸟飞 (*niǎo fēi*, "the bird flies") or 老师讲课 (*lǎoshī*

jiǎngkè, "the teacher gives a lesson") are relatively self-sufficient because the semantic role of the actor is explicitly encoded. By contrast, patient–action phrases such as 修房子 (*xiū fángzi*, "repair houses") or 看小说 (*kàn xiǎoshuō*, "read novels") normally presuppose an external actor and therefore function more naturally as embedded or expanded units rather than as independent sentences.

The same difference appears in collocational behavior. Patient-oriented phrases more readily combine with event-selecting verbs, as in 推进教育改革 (*tuījìn jiàoyù gǎigé*, "advance educational reform"), because they are semantically compatible with event nominalization. Agent–action phrases, by contrast, are more clause-like and actor-centered. This confirms that semantic classification predicts both sentence-forming potential and distributional constraints more effectively than formal structure alone.

3.2.3 Form–meaning divergence and ambiguity resolution

The results also show that semantic classification is particularly useful for resolving ambiguity. Many Chinese phrases have stable surface forms but allow different interpretations depending on context. For example, 记者采访 (*jìzhě cǎifǎng*) may refer either to "the journalist interviews [someone]" or, in a compressed nominal context, to "the interview with the journalist." Likewise, 作业完成 (*zuòyè wánchéng*, "homework completed") can be interpreted differently depending on whether the focus falls on the actor or on the completed result. These cases demonstrate that ambiguity arises not from form alone, but from competing semantic relations.

A similar divergence appears in lexicalized phrases such as 晒太阳 (*shài tàiyáng*, "sunbathe"), 食堂吃饭 (*shítáng chīfàn*, "eat in the cafeteria"), and 住旅馆 (*zhù lǚguǎn*, "stay in a hotel"). Formally, these resemble simple V + N constructions, yet semantically the nouns do not function as ordinary patients. Instead, they encode locative, environmental, or activity-setting relations. Such evidence indicates that semantic-relational analysis is especially valuable in cases where formal V + N patterns do not adequately capture the actual relational meaning of the phrase. Rather than treating formal pattern as irrelevant, these findings show that semantic interpretation is necessary for explaining why superficially similar constructions may encode different types of constituent relations.

3.2.4 Implications for teaching and stylistic variation

The findings also have pedagogical value. In teaching Chinese as a foreign language, learners often misinterpret phrases not because they fail to memorize forms, but because they do not understand the semantic relation between constituents. A semantic approach makes this clearer: 我吃 (*wǒ chī*, "I eat") and 他走 (*tā zǒu*, "he walks") can be introduced as basic agent–action patterns, while more complex structures such as 用笔写字 (*yòng bǐ xiězì*, "write with a pen") and 在教室上课 (*zài jiàoshì shàngkè*, "have class in the classroom") can be taught as

extensions involving instrument and location relations. This makes phrase learning more systematic and cognitively transparent.

A similar advantage appears in stylistic analysis. Phrases with the same semantic relation may take different formal shapes in spoken and written Chinese while preserving the same relational core. For example, 大家讨论 (*dàjiā tǎolùn*, "everyone discusses") in speech may correspond to 集体讨论 (*jítǐ tǎolùn*, "collective discussion") in formal writing. Likewise, 因为堵车迟到了 may be reformulated more formally as 由于交通拥堵而迟到. These shifts show that semantic classification is useful not only for syntax, but also for explaining how phrases adapt across registers without losing their core meaning.

3.2.5 Overall interpretive significance

Overall, the discussion shows that semantic relations provide a more explanatory perspective on Chinese phrase structures in cases where formal configuration alone is insufficient. Formal classification remains useful for identifying observable structural patterns, but it does not always account for syntactic behaviour, collocational restrictions, ambiguity resolution, or stylistic adaptation. In such cases, semantic-relational analysis offers a more precise account of how constituents are related and why phrases behave differently in actual usage.

Therefore, the present study does not reject traditional structural grammar, nor does it argue that semantic classification universally replaces formal classification. Instead, it proposes that semantic relations should be treated as a crucial explanatory dimension in phrase analysis, especially when formal labels do not adequately capture relational meaning or grammatical behaviour.

4. CONCLUSION AND RECOMMENDATIONS

4.1 Conclusion

This study does not argue that semantic relations categorically replace formal classification. Rather, it argues that semantic-relational analysis improves explanatory adequacy, especially in cases where formal configuration alone cannot sufficiently account for ambiguity, form–meaning divergence, and grammatical behaviour.

The analysis demonstrates that semantic relations such as agent–action, patient–action, instrument–action, location–action, and object–action are closely linked to grammatical properties including sentence-convertibility, collocational compatibility, nominalization tendency, and ambiguity resolution. Through examples such as 人才交流, 学者讨论, 晒太阳, and 修房子, the study shows that semantic-relational analysis is particularly useful in cases where formal configuration alone does not adequately explain phrase behaviour.

The study contributes to Chinese linguistics by showing that formal structure and semantic relation should not be treated as competing absolutes, but as complementary dimensions of phrase analysis. In this sense, the proposed framework offers a more precise and

linguistically coherent way of interpreting Chinese phrase structures, especially in cases involving ambiguity, form–meaning divergence, and mismatched structural behaviour.

4.2 Recommendations

Based on these findings, several recommendations may be proposed. First, future studies should expand the corpus to include a larger range of spoken, written, literary, and digital Chinese data so that the semantic classification framework can be tested across more diverse usage contexts. A broader corpus would also allow finer analysis of marginal and overlapping phrase types.

Second, future research may refine the semantic categories proposed in this study by examining borderline cases, prototype effects, and multi-layered phrase relations in greater detail. This is especially important for phrases whose interpretations shift across discourse contexts or whose formal structures allow more than one semantic reading.

Third, the semantic-relational framework developed here should be further explored in Chinese language pedagogy, particularly in teaching phrase expansion, sentence formation, and ambiguity resolution for non-native learners. A teaching model based on semantic relations may prove more effective than one based solely on formal structural labels.

Finally, future work may investigate the applicability of this framework in computational linguistics, especially in semantic role labeling, phrase disambiguation, and Chinese syntactic parsing. Such developments would help connect theoretical grammar with practical language technologies and further demonstrate the broader value of semantic-based phrase classification

REFERENCES

Bloomfield. (1980). *语言论* [*Language*] (Yuan Jiahua et al., Trans.). 商务印书馆.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Chen, L., Xu, S., Zhu, L., Zhang, J., Lei, X., & Yang, G. (2020). A deep learning based method for extracting semantic information from patent documents. *Scientometrics, 125*(1), 289–312. https://doi.org/10.1007/s11192-020-03634-y

Cheng, L. L.-S., & Sybesma, R. (2014). The syntactic structure of noun phrases. In *The handbook of Chinese linguistics* (pp. 248–274). Wiley. https://doi.org/10.1002/9781118584552.ch10

Chu, D., Wan, B., Li, H., Dong, S., Fu, J., Liu, Y., … Liu, H. (2022). A machine learning approach to spatio-temporal linguistic pattern analysis. *International Journal of Geographical Information Science, 36*(11), 2169–2193. https://doi.org/10.1080/13658816.2022.2087224

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.

Ding, S., et al. (1961). *现代汉语语法讲话* [*Lectures on modern Chinese grammar*]. 商务印书馆.

Fan, X. (1980a). 关于结构和短语问题 [On structure and phrase issues]. *中国语文* [*Chinese Language*], (3).

Fan, X. (1980b). 说句子成分 [On sentence components]. *上海师范学院学报* [*Journal of Shanghai Normal University*], (1).

Gao, L., Li, H., Liu, Z., Liu, Z., Wan, L., & Feng, W. (2021). RNN-transducer based Chinese sign language recognition. *Neurocomputing, 434*, 45–54. https://doi.org/10.1016/j.neucom.2020.12.006

Hao, W. (2024). Extraction of complex sentence relationships based on formalized Chinese clause structures. In *2024 7th International Conference on Machine Learning and Natural Language Processing (MLNLP)* (pp. 1–8). IEEE. https://doi.org/10.1109/MLNLP63328.2024.10800396

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Huang, B., & Liao, X. (2011). *现代汉语（增订六版）* [*Modern Chinese*] (6th rev. ed.). 高等教育出版社.

Li, Q., Li, L., Wang, W., Li, Q., & Zhong, J. (2020). A comprehensive exploration of semantic relation extraction via pre-trained CNNs. *Knowledge-Based Systems, 194*, 105488. https://doi.org/10.1016/j.knosys.2020.105488

Liu, D. (2004). 先秦汉语语序特点的类型学观照 [Typological view on pre-Qin Chinese word order]. *语言研究* [*Language Studies*], (1), 37–46.

Liu, Y., et al. (2024). How semantics works in Chinese relative clause processing: Insights from eye tracking. *Frontiers in Psychology, 15*, Article 1294132. https://doi.org/10.3389/fpsyg.2024.1294132

Lu, J. (2003). *现代汉语语法教程* [*Tutorial on modern Chinese grammar*]. 北京大学出版社.

Lü, S. (1979). *汉语语法分析问题* [*Issues in Chinese grammar analysis*]. 商务印书馆.

Lyu, S., & Li, Z. (2020). A comparative analysis of Chinese and English animal idioms. *Theory and Practice in Language Studies, 10*(6), 708–712. https://doi.org/10.17507/tpls.1006.12

Ma, J. (1983). *马氏文通* [*Ma's grammar*]. 商务印书馆.

Mao, X., Huang, S., Li, R., & Shen, L. (2020). Automatic keywords extraction based on semantic features. *IEEE Access, 8*, 117528–117538. https://doi.org/10.1109/ACCESS.2020.3004628

Peng, W., Wei, Z., Song, J., Yu, S., & Sui, Z. (2021). Formalized Chinese sentence pattern structure and its hierarchical analysis. In *Workshop on Chinese Lexical Semantics* (pp. 286–298). Springer. https://doi.org/10.1007/978-3-031-06703-7_22

Qin, Y., Yang, W., Wang, K., Huang, R., Tian, F., Ao, S., & Chen, Y. (2021). Entity relation extraction based on semantic enhancement. *Symmetry, 13*(4), 539. https://doi.org/10.3390/sym13040539

Wan, Q., Wan, C., Xiao, K., Hu, R., Liu, D., & Liu, X. (2023). CFERE: Multi-type Chinese financial event relation extraction. *Information Sciences, 630*, 119–134. https://doi.org/10.1016/j.ins.2023.01.143

Wang, Y., Wang, L., Yang, Y., & Lian, T. (2021). SemSeq4FD: Integrating global semantic relationship for fault diagnosis text modeling. *Expert Systems with Applications, 166*, 114090. https://doi.org/10.1016/j.eswa.2020.114090

Wolpert, M., Ao, J., Zhang, H., Baum, S., & Steinhauer, K. (2024). The child the apple eats: Processing of argument structure in Mandarin verb-final sentences. *Scientific Reports, 14*, Article 20459. https://doi.org/10.1038/s41598-024-70318-5

Xu, J. (2014). *现代汉语篇章语言学* [*Modern Chinese discourse linguistics*]. 商务印书馆.

Xu, N., Chang, H., Xiao, B., Zhang, B., Li, J., & Gu, T. (2022). Relation extraction based on semantic dependency features. *Buildings, 12*(10), 1633. https://doi.org/10.3390/buildings12101633

Yang, H., & Li, Y. (1987). 《诗经》"名·是·动"式新考 [New study on "name-is-action" in *Shijing*]. *武汉大学学报（社会科学版）* [*Wuhan University Journal (Social Sciences)*], (4), 80–86.

Yang, Y., & Baayen, R. H. (2025). Comparing the semantic structures of lexicon of Mandarin and English. *Language and Cognition, 17*, e10. https://doi.org/10.1017/langcog.2024.47

Yu, B., Deng, C., & Bu, L. (2022). Policy text classification algorithm based on BERT. In *2022 11th International Conference on Information Technology in Medicine and Education (ITME)* (pp. 488–491). IEEE. https://doi.org/10.1109/ICTech55460.2022.00103

Zhang, S., Hu, Z., Zhu, G., Jin, M., & Li, K. C. (2021). Sentiment classification model based on semantic enhancement. *Soft Computing, 25*(1), 463–476. https://doi.org/10.1007/s00500-020-05160-8

Zhang, Y., Taft, M., Tang, J., & Li, L. (2024). Neural correlates of semantic-driven syntactic parsing in sentence comprehension. *NeuroImage, 291*, 120593. https://doi.org/10.1016/j.neuroimage.2024.120593

Zhu, Y. A., & Grüter, T. (2025). Native speakers and learners of Mandarin predict upcoming arguments in dative constructions based on categorical and gradient verb constraints. *Bilingualism: Language and Cognition, 28*(3), 601–612