International Journal of Humanities Social Sciences and Education (IJHSSE)
Volume 10, Issue 10, October 2023, PP 10-16
ISSN 2349-0373 (Print) & ISSN 2349-0381 (Online)
https://doi.org/10.20431/2349-0381.1010002
www.arcjournals.org

Unifying Syntax and Semantics in Cognitive Concept Generation for Natural Language Expression

Li Ping Jiang

School of Foreign Languages & International Business, Guangdong Mechanical and Electrical Polytechnic, Guangzhou city, China

Corresponding Author:** ***Li Ping Jiang, *School of Foreign Languages & International Business, Guangdong Mechanical and Electrical Polytechnic, Guangzhou city, China*

Abstract: *Formalism and functionalism present differing views on the synergy between syntax and meaning in language generation. Formalist grammar emphasizes autonomous syntax, often excluding cognitive and semantic considerations. In contrast, functionalist grammar, exemplified by Halliday, sees grammatical structures as integral representations of multifaceted meaning. Cognitive syntax, influenced by Langacker, and construction grammar, as exemplified by Goldberg, emphasize cognitive functions' role in shaping grammar. While seeming contradictory, formalist and functionalist grammars offer complementary insights into the syntax-meaning relationship. This paper suggests that, in the author's cognitive function generative grammar framework, formalist and functionalist syntax are complementary. The paper explores synthesizing syntax and semantics in cognitive concept generation for natural language expression.*

Keywords: *Syntax; Cognition; Cognitive concept; Natural language expression*

1. INTRODUCTION

Formalist and functionalist grammars offer different perspectives on the relationship between the meaning basis and grammatical autonomy in natural language generation. Formalist grammar emphasizes the implicit syntactic form relationship between syntactic components and does not consider the cognitive, semantic, and pragmatic factors of these components. It views syntax as autonomous and self-governed, and the formal constraints of grammars can ignore the functional meaning of sentences. Chomsky (1981:1995) explains language production through implicit self-governed rules of syntactic form. In contrast, Halliday (1994), the first proponent of functionalist grammar, rejects the one-sided view of language generation presented in formalist grammar. According to Halliday, the grammatical formal structure is the external symbolic representation of the functional meaning of natural language, including conceptual, interpersonal, and textual functions. This view asserts that the form of language symbols and the meaning of language functions are integrated, inseparable, and reflect each other.

Cognitive syntax, headed by Langacker (1991), and construction grammar schools like Goldberg (1995) have inherited and expanded upon the functionalist viewpoint. These schools view grammar as being dominated and constrained by natural language cognitive functions such as construction or sentence cognition, function, semantics, and pragmatics. Although the differences between the formalist grammar school and the functionalist grammar school seem to be incompatible, they actually address the relationship between language syntax form and sentence function meaning from different perspectives. For instance, Xu (2002) argues that although there is no pure semantic component in formalist syntax, it does not deny the close relationship between semantics, pragmatics, and syntax. Therefore, the author of this statement believes that formalist syntax and functionalist syntax are not necessarily contradictory, but rather complementary. According to the author's cognitive function generative grammar school, the language form of natural language and its functional meaning are integrated and inseparable.

2. THE SEMANTIC BASIS OF COGNITIVE CONCEPTS GENERATION

The cognitive function generative grammar school posits that the language nerve in the brain generates cognitive thinking from the objective world, which in turn generates the basic functional meaning of language through constructions or sentences known as cognitive conceptual semantics (Mestre & Periñán, 2016). This includes four major aspects of natural language: conceptual function, pragmatic function, textual function, and stylistic function. Conceptual function pertains to the abstract meaning of language symbols, separate from their context. Meanwhile, pragmatic function encompasses all aspects of language use in context, including the associations between language symbols. Discourse function refers to the organization of language components into texts at the semantic level, including thematic structure, information structure, and organization mode. It also includes cohesion, coherence, information, purpose, situation, acceptability, and intertextuality, among other semantic functional meanings. Finally, stylistic function pertains to stylistic meaning, genre meaning, theme meaning, and rhetorical meaning.

The typical adult language processing system generates cognitive thinking by interpreting objective reality and constructing sentences using the four basic functions of natural language: cognitive, conceptual, semantic, and syntactic (Jiang, 2021). To convey functional meaning, natural language relies on limited, invisible syntactic rules and independent word symbols with both form and meaning. These language forms are generated through sound, form, and meaning, and can be used to create infinitely meaningful oral or written constructions, sentences, or discourses using a finite set of natural language means. This is just as Wang Li mentioned “semantics is the foundation of grammar.” (Hu, 2000).

According to Cheng (2011), the natural language used in everyday communication serves the function of facilitating social communication and thinking. The cognitive function principles of natural language consist of neurocognitive, cognitive operation, lexical relation, connected development, and conceptual semantic principles (Schuller & McTear, 2021). During the process of natural language generation, perception, movement, and somatosensory experiences contribute to the conceptual content, while connected grammar expresses the language semantics. Additionally, cognitive linguistics posits that meaning is derived from conceptualization, which is governed by the cognitive process's characteristics and laws. Conceptualization is achieved through autonomous and dependent linkages in semantic expression. In fact, synthesizing the scientific and reasonable achievements of my country's linguistic tradition, formalist grammar, functionalist grammar, and cognitive grammar, the author believes semantics is the basis of natural language generation. Strictly speaking, lexical system and syntactic system are independent and independent limited natural language generation subsystems and means. Grammar is formed through the cognitive thinking of the brain, and the invisible syntactic rules are conventional, invisible and inaudible.
According to Dai (1990: 12-17), the grammatical structure is derived from the symbols of reality, and the basic means of physical and mental perception and experience of time and space are derived from our physical structure interacting with the outside world. The author believes that constructions or grammatical rules themselves are similar to syntactic adhesives and have no substantial meaning, but only provide invisible combinations for word symbols to be combined into phrases, clauses, constructions or sentences. The natural language word symbol is the smallest form of spoken or written language symbol that is independent, acquired, and has both sound, shape and meaning.

Accordingly, the author argues that natural language constructions or syntactic rules in each language are created over time through daily language use and in response to objective reality in the world (Jiang, 2022), following certain innate cognitive thinking laws. These rules become conventional and fossilized in language as invisible formal rules. Shi (2002:2) asserts that "syntax rules are the projections of the laws of real objects in language" (Sarathy et al., 2018), highlighting the close relationship between language and reality. Over time, these acquired natural language words and constructions become customary and fixed in various national languages, passed down from generation to generation. Langacker (1991) opposes the innate theory of syntactic system formation, arguing that grammar is gradually formed through acquired use, with some pragmatic formats becoming grammatical formats due to their high frequency of use. Similarly, word symbols in natural language, such as Chinese characters, phonetic words in English, or kana text symbols in Japanese, are inventions of their ancestors to express cognitive thinking in accordance with certain scientific

language cognitive thinking rules. These symbols become fixed and passed down from generation to generation for thousands of years. The author likens the relationship between language rules and symbols to that of Chinese chess pieces and board rules, which are independent but must abide by the agreed-upon rules of the chessboard to move or check legally.

The natural language generation is achieved through limited language formal means such as invisible constructions or syntactic rules and word symbols with both sound, form and meaning. However, the real soul, core, and foundation of natural language constructions, sentences or discourses is cognitive concept semantics. The finite word symbols with both sound, form, and meaning are arranged in an orderly manner according to the inherently autonomous and independent finite invisible constructions or syntactic rules to linearly generate infinite natural language expressions that convey cognitive concepts and semantics. This is the basic principle and the secret of truth for the generation of spoken or written natural language.

To better illustrate and explain the process of natural language generation, Chomsky's classic example of semantically ill-formed sentences can be used as proof. For instance, the sentence "*Colourless green ideas sleep furiously*" may be syntactically correct but is semantically meaningless. This illustrates that the finite word symbols and the finite invisible constructions or syntactic rules must be arranged in a specific manner to convey a coherent and meaningful message that is aligned with cognitive concepts and semantics.

It is evident that the sentence "Colourless green ideas sleep furiously" is grammatically correct but semantically meaningless, as it follows the invisible syntactic rules of the English language. Chomsky used this sentence to demonstrate that formalist syntax is independent of word symbols with both sound and meaning and is an autonomous and independent invisible system of syntactic rules. The author agrees with Chomsky's view that sentence meaning is fundamental to natural language generation, and both independent and autonomous subsystems and means contribute to limited language generation. However, Chomsky overlooks the fact that natural language meaning derives from cognitive thinking of the brain. The grammatically correct but semantically meaningless sentence is a deliberate creation of Chomsky to establish the inherently autonomous and independent nature of formalist syntax grammar.

3. THE GENERATION PROCESS AND MODE OF NATURAL LANGUAGE

The two fundamental elements of language are the limited invisible grammar rules and word symbols that reflect the cognitive thinking of the brain to a certain extent. They represent the conceptualization of grammar and words, allowing humans to freely express their cognitive thinking after acquiring the basic constructions or grammatical rules and words of a language. The generative grammar school, which is based on cognitive function, is the same whether it is for native or foreign language acquisition and generation. Both the oral and written generation of natural language by normal adults rely on the semantics of invisible cognitive concepts and the infinite application of linguistic means, such as finite constructions or syntactic rules and finite word symbols.

All the processes and actions involved in the generation of natural language symbols are under the control of the normal adult brain's language nervous system and cognitive thinking. The organs involved in the generation of natural language in normal adults include the lungs, vocal tracts, vocal cords, oral cavity, nasal cavity, teeth, lips, and other organs, which help in producing either oral or written infinite meaningful linear natural language symbol forms. In the case of written text generation and output, handwriting actions are also required. This process is analogous to various sports such as Chinese table tennis, American basketball, and European football. These sports involve ever-changing movements and superb performances, all of which are executed under the premise of following limited basic sports rules and with the help of certain tools such as a ball. Similarly, the infinite sports art deduced under the control of cognitive thinking is akin to the infinite natural language generation subsystems and means.

Natural language is generated by the cognitive thinking of the brain, which creates cognitive concepts and semantics from experience and objective reality (Jiang, et al., 2023). This process is facilitated by innate autonomy, independence, and the limited invisible form of syntactic rules, and the independent, acquired limited word symbols with both sound and meaning. These elements allow for the infinite

application of linguistic means, which generate spoken or written natural language symbolic entities with both sound and meaning. This principle is similar to playing Chinese chess, where limited means are used to express cognitive concept semantics or sentence meaning formed in the mind.

The generation of natural language is under the control of the normal adult brain language nervous system and cognitive thinking, and involves the participation of organs such as the lungs, vocal tracts, vocal cords, oral cavity, nasal cavity, teeth, lips, and handwriting for written text generation and output. The process is analogous to ever-changing sports movements and superb performances that follow limited basic sports game rules, while also utilizing tools such as a ball. Humboldt once said that language is an infinite use of limited means, which means that the limited words of natural language acquired independently and acquired in accordance with syntactic rules are gradually arranged and combined to generate an infinite natural language oral or written entity with both sound and meaning (Hu, 1999: 274). Zhang (2001: 155-164)argues that cognitive-pragmatic communicative intentions exist in sentence changes such as inverted sentences, split sentences, and end-dropping sentences, and include context, focus, pragmatic-intent pragmatic presupposition cognitive viewpoint, expression center of gravity, salient word order, explicit-inference, and more.

For instance, in 2012, Chinese writer Mo Yan, who won the Nobel Prize in Literature, was described by the Swedish Nobel Committee as "merging folk tales, history, and the contemporary with hallucinatory realism." Here, the prepositional phrase "with hallucinatory realism" serves as a deliberate and specific qualifier before the collocation verb "merges," indicating the cognitive-pragmatic communicative intention to emphasize the uniqueness of Mo Yan's literary work. Similarly, words function as independent and acquired language symbols with the smallest sound, form, and meaning. However, the choice of positive, neutral, or derogatory words in daily language communication is also influenced by conceptual, pragmatic, and discourse meaning, as well as the influence and restriction of stylistic meaning, with cognitive-pragmatic communicative intentions (Jiang, 2022). For example, in English, Chinese, Russian, French, German, and other languages, most of the words related to the great and just cause of communism are mostly praise words with good cognitive-pragmatic communicative intentions and positive emotions. Thus, we find that cognitive concept semantics first produce constructional meaning or sentence meaning and then gradually borrow limited implicit constructions or syntactic rules (including simple sentence rules, compound sentence rules, and compound sentences) that are inherently autonomous, independent, and acquired rules, as well as independent and acquired words and symbols with both sound, form, and meaning, and other limited language means to generate meaningful and infinite oral or written natural language entities. In other words, meaning is the fundamental purpose of natural language, and the limited invisible constructions or grammatical rules that are inherently autonomous, independent, and acquired, and the limited words and symbols that are independent, acquired, and have both sound, form, and meaning, are merely the basic language means for generating infinite natural language. Indeed, formalist syntax and functionalist syntax are not contradictory, as limited constructions or grammatical rules and limited words are gradually integrated into one (Vishnuprabha et al. 2021). After the combination, similar to the computer package file, infinite natural language generation can occur, leading to a linguistic oral or written communicative symbolic unified entity.

The process of natural language generation involves the objective reality that stimulates cognitive thinking in the normal adult brain, which then generates cognitive concept semantics. In order to express the meaning of sentences formed in the mind, the limited words and symbols of natural language that are acquired independently and through learning follow and comply with the limited words and symbols of innate autonomy, independent autonomy, and acquired learning. This gradual arrangement of language symbols follows the limited constructions or grammatical rules, resulting in the generation of natural language phrases, clauses, and sentences, which are the basic verbal or written language symbols used for communication with both sound, form, and meaning. Furthermore, Pan (2002: 2) and Knowlton & Hunter (2021) also hold the view that language is an externalized symbolic representation of the human mind.

Meaningful and infinite oral or written natural language symbolic forms are the external manifestations and representations of cognitive functions, that is, meaning. On the other hand, inherently autonomous, independent, and acquired finite constructions or syntactic rules are invisible

and meaningless in themselves (Jiang et al., 2022). However, they serve as the invisible syntactic adhesives for the formation of natural language symbolic entities with both sound, shape, and meaning, which are acquired independently and through limited word symbol forms. Furthermore, the grammatical structure of Chomsky's (1995) principle and parameter theory reflects the universal and unique cognitive and thinking laws of different ethnic groups. These laws underlie the basic cognitive and thinking processes that produce cognitive concept semantics, which then guide the selection and arrangement of words and symbols in natural language communication. Thus, the limited syntactic rules and finite vocabulary of a language are the means by which these underlying cognitive and thinking processes give rise to the rich variety of expressive forms found in natural language.

The grammatical rules of natural languages exhibit both commonalities and individualities. According to Mei (2008), all grammars share basic components such as subject, predicate, object, attribute, adverbial, and complement. Similarly, modern Chinese and modern English have basic subject-predicate-object grammatical structures as well as theories such as X-bar theory, control and jurisdiction theory, constraint theory, and case theory. However, there are also individual differences, such as the unique subject-object-predicate structure in modern Japanese and ancient Chinese, and the common parataxis structure in modern Chinese and modern Japanese versus the hypotaxis structure in modern English. Additionally, modern Chinese has specific sentence patterns such as subject-predicate-predicate sentence, "ba" sentence, "bei" sentence, linked sentence, concurrent sentence, double-object sentence, and existential sentence that are different from most Indo-European language families like modern English. Overall, the grammatical rules of natural languages exhibit both universal and specific features. The natural language generation process can be described as follows: It begins with the creation of cognitive conceptual meaning, rooted in objective reality, through cognitive thought. This meaning is then transformed into linguistic symbolic entities that encompass sound, form, and meaning, all guided by the principles of constructional or syntactic rules. The outcome is a natural language communicative symbolic entity, shaped by these constructional or syntactic rules.

4. CONCLUSION

He generative grammar school of cognitive function posits that the process of natural language generation for normal adults of each ethnic group can be visualized in Figure 1. Natural language symbolic entities, including words, phrases, clauses, constructions, and sentences, with both sound, form, and meaning are the oral or written symbolic representation and manifestation of the semantics of cognitive concepts. Both are used to express the cognitive concept semantics generated by cognitive thinking in the brain through objective reality, and rely on innate autonomous and acquired limited natural language words and symbols with both sound, shape, and meaning, as well as limited construction or syntactic rules acquired independently to produce unlimited spoken or written natural language with both sound, form, and meaning. These rules are essential for natural language generation, providing implicit cohesion or combination methods for the combination of limited word symbols into phrases, clauses, constructions, or sentences. Additionally, the grammatical rules formed by all natural languages have both commonalities and individualities.

Although the construction and syntactic rules of each natural language are inherently autonomous, independent, and acquired, the specific macro- or micro-syntax, sentence patterns, and grammatical choices for daily communication are influenced and restricted by cognitive-pragmatic communicative intentions and motivations, as well as the semantics of cognitive concepts (Jiang, 2021). For instance, the choice of positive, neutral, or derogatory words in daily communication is also affected by the pragmatic functions and cognitive pragmatic intentions, which originate from people's cognitive thinking through the objective reality. Therefore, the process and mode of natural language generation follow the idea of "the heart is born and the words stand." This view is supported by Knowlton et al. (2021) and Zhu & Luo (2023).

ACKNOWLEDGEMENT

This research was supported by Guangdong Provincial Philosophy and Social Science Planning Fund, 2023 (Grant No.:GD23CWY04).

REFERENCES

Chomsky, Noam.(1981).*Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam.(1995).*The Minimalist Program.* Cambridge, Mass,: MIT Press, 1995.

Goldberg, A. E. (1995).*Constructions A Construction Grammar Approach to Argument Structure*. Chicago and London.

Halliday, M. A.K.(1994).*An Introduction to Functional Grammar(2nd edition)*. London: Arnold. Langacker, R. W.(1991).*Foundations of Cognitive Grammar: Descriptive Applications(II)*. Stanford: Stanford University Press.

Cheng, Q. (2011).*A New Perspective of Conceptual Semantics Research*. Shanghai: Shanghai Foreign Language Education Press, 40-41.

Knowlton, T., Hunter, T., Odic, D., Wellwood, A., Halberda, J., Pietroski, P., & Lidz, J. (2021). Linguistic meanings as cognitive instructions. *Annals of the New York Academy of Sciences*, *1500*(1), 134-144.

Jiang L, Al-Shaibani GKS, Yang F, Cheng M and Huang M (2022) The metonymic mechanism of English translation of Chinese intangible cultural heritage terms from the perspective of cognitive psychology. *Front. Psychol.* 13:957485. doi: 10.3389/fpsyg.2022.957485.

Jiang, L., Lv, M., Wen, Y., Zhang, P., & Huang, Q. (2023). Bilingual Conversion in the Translation of ICH Terms: A Study on the Psychological Processes of Translators. *Journal of Psycholinguistic Research*, 1-18. doi.org/10.1007/s10936-023-09989-6.

Jiang, L. (2021). A Study on the Efficiency and Countermeasures of Network-Based Autonomous Learning Platform for Vocational College English Learners. *Frontiers in Artificial Intelligence and Applications*, *345*, 484-492.doi:10.3233/FAIA210437

Jiang, L. (2021). A Case Study of the Diachronic Development of Second Language Grammatical Competence Based on Computer-mediated Negotiated Interaction. *CONVERTER*, 496-504. https://doi.org/10.17762/converter.

Jiang, L. (2022). Effective Utilization of Computer-Mediated Communication Technology in Network-based Foreign Language Teaching, Wireless communications and mobile computing,1-7. https://doi.org/10.1155/2022/1048311

Jiang, L. (2022): Factors influencing EFL teachers' implementation of SPOC-based blended learning in higher vocational colleges in China: A study based on grounded theory, *Interactive Learning Environments*, DOI: 10.1080/10494820.2022.2100428.

Hu, M. (1999). *Selected Readings of Western Linguistics (Second Edition)*. Beijing: Renmin University of China Press, 274.

Hu, Z. (2000).*On Functionalism*. Beijing: Foreign Language Teaching and Research Press.

McShane, M., & Leon, I. (2022). Language generation for broad-coverage, explainable cognitive systems. *arXiv preprint arXiv:2201.10422*.

Mei, D.(2008).*Modern Syntax*. Shanghai: Shanghai Foreign Language Education Press, 94-212.

Mestre, E. M., & Periñán, C. (Eds.). (2016). *Understanding Meaning and Knowledge Representation: From Theoretical and Cognitive Linguistics to Natural Language Processing*. Cambridge Scholars Publishing.

Pan, W.(2002).*Character Standard and Chinese Character Research*. Shanghai: East China Normal University Press, 2002.

Sarathy, V., EDU, T., Oosterveld, B., Krause, E., & Scheutz, M. (2018). Learning cognitive affordances for objects from natural language instruction. In *Proceedings of the Sixth Annual Conference on Advances in Cognitive Systems*.

Schuller, B. W., & McTear, M. F. (2021). Socio-Cognitive Language Processing for Special User Groups. In *Multimodal Agents for Ageing and Multicultural Societies: Communications of NII Shonan Meetings* (pp. 87-95). Springer Singapore.

Shi, Y. (2002).*The Cognitive Semantic Basis of Grammar*. Nanchang: Jiangxi Education Press.

Wei, H.(1990).*Cognition-based Chinese Grammatical Function*. Foreign Linguistics, (2): 12-17.

Vishnuprabha, V., Murali, L., & Viswanathan, D. M. (2021). Natural Language Processing. In *Cognitive Computing Systems* (pp. 349-400). Apple Academic Press.

Xu, L.(2002).*Functionalism and Formalism*. Foreign Language,(2):8-14.

Zhang K.(2001).*The Cognitive Pragmatic Motivation of Sentence Change*[A].Zhang, H.,&Hu Z.,

China Foreign Language Botu Forum [C]. Beijing: Foreign Language Teaching and Research Press, 155-164.

Zhu, Q., & Luo, J. (2023). Generative design ideation: a natural language generation approach. In *Design Computing and Cognition'22* (pp. 39-50). Cham: Springer International Publishing.

AUTHOR'S BIOGRAPHY

Liping Jiang, is currently an associate professor and a Doctor in education; His research fields include Business English research, Cognitive linguistics, He published more than 90 academic papers, some of them indexed by SSCI Q1 and CSSCI. 1531447813@qq.com

Citation: *Li Ping Jiang. "Unifying Syntax and Semantics in Cognitive Concept Generation for Natural Language Expression" International Journal of Humanities Social Sciences and Education (IJHSSE), vol 10, no. 10, 2023, pp. 10-16. DOI:* https://doi.org/*10.20431/2349- 0381.1010002.*