Identification Methods for Information Structures in Different Types of Scientific Discourse

Jiayang Liu

Harbin University of Commerce, Harbin, 150028, China

Abstract: Currently, with the accelerated development of science and technology, there is an increasing number of scientific and technological exchanges between China and the international community. The exchange of scientific and technological achievements is crucial for the sharing of technological experience, as well as for the development and progress of science and technology itself. Therefore, there is a need for scientific and technological papers to express and explain the scientific and technological achievements of different countries in writing to promote technological information exchange and cooperation. Taking English as an example, as one of the internationally recognized languages, about 85% of international scientific and technological literature is in English, and some key information indexes in the international scientific and technological papers are mostly in English. Therefore, it is necessary to grasp the forms of scientific and technological achievements exchange, master the information structure and recognition methods of different types of scientific and technological discourse, and write English scientific and technological papers well to promote scientific and technological exchange. This article introduces the methods for identifying information structures in different types of scientific discourse, in order to better interpret the methods for identifying information structures in scientific discourse and provide some ideas and references for better writing English scientific papers.

Keywords: Different Types; Technical Discourse; Information Structure; Recognition Methods

1. Introduction

To write an English scientific paper well, it is necessary to grasp its writing method. Unlike other types of problems, English scientific papers belong to written problems, and writing needs to be rigorous and thorough, with a certain degree of logic. The text should be concise, highlighting the key points, and the concepts should be precise. Based on the differences in research disciplines, content, and methods of English scientific papers, the corresponding scientific discourse structure is also different. Grasping the scientific discourse structure of English scientific papers and explaining the relevant content can provide direction and ideas for readers' scientific paper creation.

2. Overview of information structure in scientific discourse

Halliday pointed out that language has three major functions, namely conceptual function, interpersonal function, and technological discourse function. The function of technological discourse not only includes thematic structure and cohesive system, but also information structure. Based on the analysis of the information theory, language is a complex information system unique to human beings. People communicate through language.[1] This process is the process of organizing and transmitting information. The information in human society has a certain concept form, and the concept form information is a unique form of human interaction and communication. Classify these conceptual forms into individual information units, which can be classified into known and unknown information based on their properties, namely old and new information [2]. Chafe (1976) pointed out in his research that known information refers to the information that the speaker believes is already in the recipient's consciousness when speaking, which is old information, while information structure is actually the structure of information units formed by the interaction between known information and new information. From this perspective, information structure refers to the structure of information units formed by the interaction between known information and new information. In general, in the process of discourse communication, the information structure is mainly related to tone, tone group, and step, while in scientific discourse communication, the information structure can be arranged through similar

content components, and can be reflected through selection and limitation. In this regard, the child in this article likes to introduce different types of technological discourse topics to classify and recognize information, with the aim of grasping the distribution patterns and characteristics of information.

3. Methods for identifying information structures in scientific discourse

In scientific discourse, the author's intention is mainly to convey information. Through abstract thinking, the author organizes and summarizes objective phenomena, factual principles, work processes, etc., describes and explains the information, and then transmits it to the reader.[3] Based on this goal, the author screens and processes the network system composed of various propositions that exist in their own minds. Through the brain language mechanism and specific program encoding, the observed surface information is transformed into concepts, which are fixed and presented through word forms [4]. As long as words are formed, the information activity of judgment and reasoning will be carried out from the language. The author has the right to make decisions about the content structure of a sentence and which word items to choose based on a series of different levels, selecting the word or phrase that best meets this purpose, and determining how to best organize the relevant components within the sentence, as well as the best order to unfold the paragraph. In order for readers to effectively extract more useful information, authors need to effectively grasp the structure of scientific discourse, which is crucial for analyzing the information structure of scientific discourse.

Based on the characteristics of the text, there are significant differences between different themes of technological discourse. Literary technological discourse is mainly narrative, while technological discourse is mainly informational. The information structure of narrative technological discourse is generally compact, and its unfolding mode also has certain commonalities, mainly based on chronological order. Its structure is mainly based on the starting, unfolding, climax, and ending modes [5]. Information technology discourse contains various forms, and the types of discourse that technology readers come into contact with are generally diverse. For example, introducing a theoretical discourse is completely different from describing an institution. Different types of discourse can be represented by different information components, and different information components can be represented by information units of different sizes and arrangements. In the writing of technological articles, the information structure can be judged based on the presence or absence of relevant information components. Generally, written articles are easily limited by the characteristics of the article being written.[6] This type of information structure analysis is not based on the surface features of technological discourse, but more on the information or meaning components based on the text, in order to indicate what components or how related components work. Analyzing the information structures of the two different types of technological discourse mentioned above can be carried out from three aspects:

Firstly, the typical patterns of information structure composition in scientific discourse;

Secondly, the evening meeting chain related to certain information structures;

Thirdly, the choice of tense and verb forms in two types of discourse.

Generally speaking, the form of English scientific discourse changes more than the conventional form of English scientific discourse. On the one hand, there are many questions in scientific English, and on the other hand, there is a set of strictly limited questions. A type of question represents a question with the same information component. In a scientific article, the main function of a certain article is to introduce the form, characteristics, and function of objects. If these main components appear in a type of scientific discourse, it indicates that this type of scientific discourse is a type of physical structure topic, and there is another type of discourse topic that involves the process of forming something. For example, the formation of rock belongs to a process type of scientific discourse. In A.J Herbert's article "Structure of Science and Technology English", by analyzing the types of diverse discourse topics, the information structures of science and technology discourse with the same or similar components are classified. The specific classification method is shown in the table 1:

Table 1: Classification methods for information institutions of different types of scientific discourse

Entity described	Describe the process
Steam boiler	Steel making
The strunture system in human body	welding
Bridge structure	Oil refining
Centrifigual goverrnors	The formation of rock

In the two categories mentioned above, there are eight questions. The left question can clearly describe the physical structure, performance, form, etc. These features described by the relevant questions are similar, but there are some differences between them and the questions in the right column in some aspects. The questions in the right column all have words in the form of 'ing', which is a hint for not analyzing information structure. This is the difference between the two parts of the questions caused by generative semantics. Overall, the arrangement of the information structure, information components, and information selection involved in all questions is crucial. All the questions on the left contain this or that type of entity, while the questions on the right are related to process characteristics [7]. Based on process technology discourse, its information components include the following: firstly, the form or state of an object; secondly, object orientation; thirdly, change activities; fourthly, the process of change. The formation of this structure is not accidental, but based on scientific foundations. The objects and events in it are closely related, and the structure is more compact. The interdependent relationships between the information elements formed internally exist. In fact, the topic itself can also express the idea of the topic, and it is feasible to discover how information is formed in scientific discourse through the idea. To gain a deeper understanding of how to discover the formation of information in technological discourse through ideas, one can identify the differences between the two types of questions and grasp the differentiated information provided by the entity structure and process chapters.

In the description of entity structure, information type and information are related, and the components of object structure, object orientation, component performance, component function, etc. are key classification elements and reference indicators.

4. Vocabulary chain and tense and verb type selection in different technological texts

4.1 Vocabulary chain

One major difference in language between physical discourse and process technology discourse can be presented through lexical chains. In structural technology discourse, lexical chains represent components, while in process technology discourse, they represent transformation relationships. Based on the foundation of studying different lexical items using lexical chains, it can be concluded through comparison that all lexical terms in structural technology discourse are connected in a special way through semantic whole and partial relationships, which is the primary characteristic of lexical chains. All items in a word chain are limited to noun referents and pronouns, which is the second characteristic of a word chain. The process of representing a word chain is also one of the specific ways of linking, which means that the various technical words in this word chain are not connected through a defined relationship or do not need to be connected through a defined relationship, and their conversion relationship is easily limited by the technological discourse itself. The two types of technological discourse lexical chains provide a wealth of evidence for analyzing information structure through language reference and information component arrangement. The semantic relationships between the two lexical chains are mainly overall and partial relationships, as well as transformation relationships [8].

4.2 Selection of tense and verb types

By analyzing the verb types and tenses of two different types of scientific discourse, it can be concluded that the verbs in scientific discourse express state or event verb chains. In the verb chain of structural technology discourse, the main goal is to describe objects, representing the process of relationships, which can be described through simple present tenses. State verb words are used as the subject, usually using has, include, etc. The relational process describes various relationships in the objective world, such as attributive relational processes, which endow the subject with a relationship. The process of relationship is mainly a description and judgment of objectively existing objects, with related relational verbs such as pure verbs: be, come, see, etc; Sensory verbs such as' look ',' feel ',' smart ',' take ', etc., can reflect the relationship between the whole and parts in structured technical discourse. Process technology discourse generally takes the simple present tense as the main tense, and there is a widely used passive voice in this type of technology discourse, as well as a phrase with a way or purpose after it. The main verbs in process technology discourse represent dynamic verbs of activities, such as run, jump, move, touch, stamp, wave, walk, sit, open, close, smile, speak, etc. Related verbs can accelerate verb conversion, represent material processes of physical systems, i.e. concrete and tangible actions or events, and can also describe abstract behaviors or processes.

5. Conclusion

Overall, there are various methods for analyzing the information structure of scientific and technological discourse. Combining semantic segmentation theory with traditional analysis of discourse subject and rheme is not comprehensive. Therefore, this article divides the information structure of scientific and technological discourse into different topics, combines the components that make up the discourse, and uses arrangement and tense selection to divide the information structure of scientific and technological discourse, in order to study effective methods for identifying the information structure of scientific and technological discourse. These studies are also helpful for the subsequent creation and analysis of scientific discourse, helpreaders better interpret and grasp the key information content of scientific discourse, in order to find the discourse they need from a large number of scientific discourse.

Acknowledgment

References

[1] Luo J. Analysis of Grammatical Metaphor in Chinese and English Scientific Discourse and Its Conversion [D]. Sichuan: Xihua University. 2012.;DOI:10.7666/d.y2113193.

[2] Cheng X. A Study on the Function of the Application of the Grammtical Metaphor in English for Science and Technology [D]. Heilongjiang: Northeast Agricultural University. 2012; DOI: 10. 7666/ d. Y2234963.

[3] Li N, Fang L, Zhang Y. Multidisciplinary comparative analysis of deep learning recognition methods for academic text structure and function [J]. Modern Intelligence. 2019; 39(12):55-63, 87.

[4] Li Y, Sun J, Zhou G. Recognition and classification of discourse connectors in Chinese scientific and technological language [J]. Journal of Peking University (Natural Science Edition) 2015; 51(2): 307-314.

[5] Zou Y, Zhang Z. Research on the Method of Identifying Important Objects in Network Technology Information Based on Object Grid [J]. Journal of Information Technology. 2014, 33(3):250-258.

[6] Xu L, Qin C, Li J. A Study on the Path of Scientific and Technological Literature Digitalization and Organizational Presentation [J]. Journal of Chinese Library. 2022, 52(3):25-42.

[7] Liu Y, Zhang Y, Zhang Y, Yi L. Constructing a Text Knowledge Network Integrating Text Structure [J]. Library and Information Work. 2021; 65(21):118-130.

[8] Wu T, Li M, Kong F. Construction of a corpus of textual level entity hierarchical relationships based on synonymous reasoning [J]. Chinese Journal of Information Technology. 2020; 34(4):38-46.