A Theme-based Text Analysis of an Academic Abstract in English and Chinese

Jiangping Zhou
China West Normal University, China

Abstract—This paper centers on the employment of Theme-related theories to analyze how academic abstracts unfold coherently in English and Chinese versions. By quantitatively exploring the distribution of these theories and ways they are progressed, the findings show that the two versions adopt simple Themes and multiple Themes in a completely opposite way, and their progressive patterns also differentiate with each other significantly. However, they seem to keep a similar trend of developing unmarked and marked Themes through the two versions of the abstract.

Index Terms—simple and multiple themes, unmarked and marked themes, thematic progressive patterns

I. INTRODUCTION

Academic abstract is the most significant and compacted part of an article, which generalizes the research contents, purposes, approaches and findings within, by and large, two or three hundred words or so. In order to achieve this purpose, it is of vital importance for scholars to present these ideas and at the same time keep the articles' readability in a cohesive and coherent way. Mathesius in Prague school propounds the concept of Theme-Rheme structure, which serves to be one of the important means and then M. A. K. Halliday (1985) in his monograph further develops the structure systematically and applies it into text analysis. Scholars both home (Hu, 1980; Cao & Li, 2015; etc.) and abroad (Halliday, 1967; 1968; North, 2005; Leong, 2016; etc.) have employed the theory in analyzing different types of texts. This paper intends to quantitatively analyze an English abstract and its Chinese translation from an MA thesis in psycholinguistics and explores the ways that the Theme-Rheme structures of the abstract are interwoven. Thus, two research questions are raised: 1) how is the theme distributed in the two abstracts? 2) what is the similarity and differences between the two abstracts while the author constructs them?

II. THEORETICAL FRAMEWORK

Theme is glossed as the point of departure of the message and the speaker or writer (hereafter speaker) uses it as a means to facilitate the addressee or reader (hereafter addressee) to decode the interpretation of the message (Halliday, 2014). Theme-Rheme structure is not only a means of interweaving the textual information, but also an important grammatical device to organize the text as a whole. Theme is an element of the clause, expressing what is known or obvious to addressees. It can be realized by circumstantial elements or the utterances occurred in the context. Therefore, Theme is closely associated with the text.

According to the components of forming a theme, it is sub-categorized into simple Theme, multiple Theme and clausal Theme. Being simple Theme does not suggest that it is simply realized by a word or a group; as a matter of fact, the realization of a Theme can be extended to incorporate more than one group, either nominal groups, or adverbial groups, or prepositional phrases, or even the combination of the three. It is being simple Theme because only one of the three meta-functions (ideational, interpersonal or textual) is realized by all these elements in the clause. For instance, the Themes These students and the personal experience narratives of Chinese college students in (1) are realized by nominal group and the combination of nominal group and prepositional phrase respectively in the clause; the Theme in (2) is simply realized by the prepositional phrase you guan xushiyupian yu yuyanchansheng he gongzuojiyi xiang jiehe de yanjiu (the studies associating narrative text with language production and working memory).

(1) These students have revealed that the personal experience narratives of Chinese college students exhibit all of the six categories presented by Labov.

(2) But, have relating narrative text and language production and working memory. However, the studies associating narrative text with language production and working memory are rare at present.
realized by either adverbial groups or nominal groups. The usual linear order for the co-occurrence of the three metafunctional meanings is textual Theme, followed by interpersonal Theme, which precedes the topical Theme. In actual writing, the co-occurrence of the three functional elements is very rare, so there are scant such cases occurring in the two abstracts. Example (3) illustrates the combination of textual Theme and topical Theme. Clausal Theme refers to the cases that the realization of the Theme is achieved by a clause or a nonfinite clause. For example, I must say in (4) is the clausal Theme of the whole clause.

(3) Finally, SPSS17.0 was used to descriptively analyze these experimental numbers.

(4) Kate, I must say this fish is cooked beautifully. (Halliday, 2014, p.110)

Halliday also demarcates the unmarked Theme and the marked Theme. The former is defined as the phenomenon that the Theme and the subject of the clause are overlapping while the latter refers to the case that the Theme is not the clausal subject, but another element of the clause such as the complement or the circumstance. Marked Theme occurs when the speaker is trying to yield some special effect or to emphasize the complement or the circumstance or some other elements. The Themes are markedly realized by with the development of working memory in (5) and Youguan Shiro tichude fanchou (As to the categorizations proposed by Shiro) in (6) respectively.

(5) With the development of working memory, many scholars are studying language production together with working memory, and received remarkable achievements

(6) 有关 Shiro 提出的 范畴， 我们[...].

‘As to the categorizations proposed by Shiro, we...’

Another important concept pertaining to Theme is Thematic progressing, which is initiated by the Czech scholar Daneš (1970, 1974). He first draws attention to the patterned inter-relationships between Themes and Rhemes in the text and identified five various broad patterns. Subsequently, this concept is further developed by M. A. K. Halliday (1967;1968), and underlying the theoretical framework, many scholars in China (Xu,1982; Zhu, 1990; Liu, 2012) have also explored the theory, and proposed their own sub-classifications. The most frequently occurring patterns are Theme-identical pattern, simple linear pattern and Rheme-identical pattern, which together make the abstracts elected coherent.

Theme-identical pattern refers a progressive relation that Themes of two or more than two clauses are identical, but their Rhemes are different. It can be exemplified as (7) and diagrammed in Figure 1. (T for Theme and R for Rheme)

(7) //The study of personal experience narratives [T1] has received much attention in recent years, [R1] //and [the study of personal experience narratives][T1] has become part of the work together with language production. [R2] //

\[
\begin{array}{c}
\text{T1} \\
\rightarrow \text{R1} \\
\text{T2} \rightarrow \text{R2} \\
\vdots
\end{array}
\]

Figure 1: Theme-identical pattern

The simple linear pattern can be defined as the fact that the Rheme or part of the Rheme of the previous clause is employed as the Theme of the next clause and by so doing, the message is further progressed. Consider the example (8) and Figure 2.

(8) //Sixty students altogether from different universities and different majors [T2] participated in the study. [R3]//

This research [T3=R3] adopted Labov’s (1972) six functional categories of narratives and Shiro’s (2003) evaluative language to analyze these produced narratives made by Chinese college students. [R4] //

\[
\begin{array}{c}
\text{T1} \\
\rightarrow \text{R1} \\
\text{T2} = \text{R1} \rightarrow \text{R2} \\
\vdots
\end{array}
\]

Figure 2: Simple linear pattern

Rheme-identical pattern is contrary to the Theme-identical pattern in that the former refers to the phenomenon that Rhemes of different clauses are the same but their Themes remain different, while the latter’s Themes are similar as
defined above. This progressive pattern is illustrated in (9) and diagrammed in Figure 3.

(9) ///We [T1] propose that narratives and language production are closely related, [R1]/// and so [T1] are narratives and working memory; [R2]// thus, it [T2] is necessary to contrast produced narratives with memorized narratives. [R3=R1]///

\[ \text{T1} \quad \text{T2} \quad \text{R1} \]

\[ \text{...} \]

Figure 3: Rheme-identical pattern

What should be expounded here is that ‘identical’ in Theme-identical pattern or Rheme-identical pattern does not necessarily be the same. It might be derived or related in cohesion or some other means.

III. METHODOLOGY

A. Material Analysis

An English abstract and its Chinese translation from the CNKI are chosen as the materials of my analysis on the distribution of Themes. Abstracts of the thesis are about the relationship between Language Production and Verbal Working Memory in Oral and Written Narratives of Personal Experience (Zhou, 2010). The detailed analysis of English and Chinese versions is presented in the following and Appendix provides with notational conventions of the two versions.

English version

[Para. One]

///The study of personal experience narratives [1s] [unmarked] [T1] has received much attention in recent years. [R1] // and [2] [the study of personal experience narratives] [1c] [unmarked] [T1] has become part of the work together with language production. [R2] /// [Type 1]

//Labov [1s] [unmarked] [T2] (1972) has studied the functional categories of narratives [R3] // and [2] [Labov] [1c] [unmarked] [T2] concludes that every narrative should be composed of six categories, that is, coda, orientation, complicating action, result, evaluation and coda. [R4] /// [Type 1]

// Shiro [1s] [unmarked] [T3] (2003) also analyzed the production of narratives from the aspect of the evaluative language. [R5] // thus [2] [Shiro] [1c] [unmarked] [T3] realized the association between language production and narratives. [R6] /// [Type 1]

// With [2] the development of working memory, [1c] [marked] [T4] many scholars are studying language production together with working memory. [R7] // and [2] [many scholars] [1c] [unmarked] [T5=R7] received remarkable achievements (Baddeley, 1986; 2000; 2001; 2003; Baddeley & Hitch, 1974). [R8] /// [Type 2]

// However, [2] the associative studies [1c] [unmarked] [T6=R7] are rare at present. [R9] /// [Type 2]

We [1s] [unmarked] [T7] propose that narratives and language production are closely related, [R10] // and [2] so [1c] [unmarked] [T7] are narratives and working memory. [R11] /// [Type 1]

//thus [2], if [1c] [unmarked] [T8] is necessary to contrast produced narratives with memorized narratives. [R12=R10] /// [Type 3]

// In order to acquire a deep comprehension of narrative characters in the two experiments [marked] [T9], we contrastively analyzed narrative structures and evaluative language in oral and written narratives as well as in memorized oral and written narratives. [R12] ///

[Para. Two]

///This thesis [1s] [unmarked] [T1] focuses on the most unforgettable experience made by Chinese college students. [R1] // and [2] [This thesis] [1c] [unmarked] [T1] focuses extremely on the features of narrative production and working memory. [R2] /// [Type 1]

//Sixty students altogether [2] from different universities and different majors [1c] [unmarked] [T2] participated in the study. [R3] // This research [1s] [unmarked] [T3=R3] adopted Labov’s (1972) six functional categories of narratives and Shiro’s (2003) evaluative language to analyze these produced narratives made by Chinese college students. [R4] /// [Type 2]

// Finally [2], SPSS17.0 [1c] [unmarked] [T4] was used to descriptively analyze these experimental numbers. [R5] ///

[Para. Three]

/// These studies [1s] [unmarked] [T1] have revealed that [2] the personal experience narratives of Chinese college
students [1c] [unmarked] exhibit all of the six categories presented by Labov.[R1] // That is to say[3], Labovian categories [1c] [unmarked] [T1] seem to exist in a non-western language like Chinese irrespective of the oral and written versions or memorized oral and written versions used in the narrative. [R2] // [Type 1] 
// The only [2] inconsistency with respect to the Labovian model [1c] [unmarked] [T2] was in the use of “abstract” and “coda” categories in written and oral versions and memorized written and oral versions. [R3] // The frequency of these categories [1s] [unmarked] [T3=R3] was significantly higher in the written narratives and memorized narratives. [R4] // [Type 2] 
// Moreover [2], the use of “abstract” and “coda” categories [1c] [unmarked] [T4=R3] seemed optional among Chinese college students, [R5] // whereas [2] the other four categories [1c] [unmarked] [T5] were used obligatorily. [R6] // [Type 1] 
// These findings [1s] [unmarked] [T6] are in line with the findings of other studies [R7] // which [1s] [unmarked] [T7=R7] suggest // that [2] not [3] all stories [1c] [unmarked] have abstracts and codas (Johnstone, 2001). [R8] // [Type 1] 
// The use of evaluative language [1s] [unmarked] [T8] was also analyzed in terms of the categories presented by Shiro. [R9] // The two experiments [1s] [unmarked] [T9] found that [2] the use of evaluative sentences [1c] [unmarked] [T10=T8] is significantly higher in the written version and the memorized written version. [R10] // [Type 1] 
// However, [2] the order of occurrence of the evaluative categories [1c] [unmarked] [T11=R9] was found to be the same in both narrative modes and both memorized narrative modes. [R11] // [Type 2] 
// Thus [2], the “emotion” category [1c] [unmarked] [T12] was most frequent in both versions in two experiments. [R12] // However [2], when [2] the frequency of these categories [1c] [unmarked] [T13=R12] is compared in terms of oral and written versions and memorized oral and written versions, [R13] // the use of the “emotion” category [1s] [unmarked] [T14=R12] is significantly higher in the oral and memorized oral narratives. [R14] // [Type 2] [Type 1] 
// This [1] [unmarked] [T15=R14] indicates the high emotional state of the oral and memorized unforgettable narratives. [R15] // [Type 2] 

Chinese version
[Para. One]

// 近年来[1s] [marked][T1], 个人经历叙事成为了研究关注的焦点, [R1] // [个人经历叙事] [1s] [unmarked][T2=R1] 是语言产生研究常涉及的[2]. [R2] // [Type 2] 
// Labov[1s] [unmarked][T3](1972) 就叙事语篇的功能范畴进行了研究,[R3] // [Labov][1c] [unmarked][T3] 总结出任何叙事语篇都应该由“概括”、“发生”、“发展”、“结果”、“评价”和“结尾”这六大范畴所构成.[R4] // [Type 1] 
// Shiro[1s] [unmarked][T4](2003) 从评价性语言的角度也对叙事语篇的产生进行了分析,[R5], // [Shiro][1c] [unmarked][T4] 实现了语言产生同叙事语篇的相结合.[R6] // [Type 1] 
// 近年[2] 工作记忆研究的发展[1c] [marked][T5] 大量学者将工作记忆与语言产生相结合研究.[R7], // [工作记忆][2]学者[1c][unmarked][T6=R7] 取得了大量的成果(Baddeley, 1986; 2000; 2001; 2003; Baddeley & Hitch, 1974). [R8] // [Type 2] 
// 然而[2], 关于[2] 叙事语篇与语言产生和工作记忆相结合的研究[1c][marked][T7] 还不多见. [R9] // 我们[1s][unmarked][T8] 认为叙事语篇与语言产生[1s][unmarked][T9=R7] 有着必然联系，同样与工作记忆也应有一定的关联.[R10], // [Type 2] 我们[1s][unmarked][T10=T8] 很必要将直接产生的叙事语篇与记忆后所产生的叙事语篇进行对比.[R11] // [Type 1] [Type 1] 
// 这两个实验中[marked][T1], 为了更加深入了解叙事特征[marked][T1], [T11] 我们对比分析了口头叙事与书面以及记忆口述与记忆书写个人经历叙事中的叙事结构与评价性语言.[R12=R11] // [Type 2] [Type 3] 
[Para. Two]

// 本研究[1s][unmarked][T1] 调查了中国大学生在叙述个人最难忘的一段经历时[R1], // [本研究][1s][unmarked][T1] 重点分析了这些叙事语篇的语言产生与工作记忆的特征.[R2] // [Type 1] 
// 本文共有[2] 名来自不同大学、不同专业的大学生[1s][unmarked][T2] 参加了该研究[R3]. // [本研究][1s][unmarked][T3=R3] 采用了 Labov(1972) 的叙事语篇的六大功能范畴和 Shiro(2003) 的叙事评价性语言范畴对大学生所产生和形成的叙事语篇进行了分析.[R4] // [Type 2] 
// 最后[2], 我们[1c][unmarked][T4] 使用了 SPSS17.0 对实验数据进行了描述统计分析.[R] // [Type 2] 
[Para. Three]

// 研究结果[1s][unmarked][T1] 显示国内大学生的个人经历叙事都体现了 Labov 所提出的六个范畴.[R1] // 也就是[R3], 无论[2] 叙事形式[1c][marked][T2=R1], Labov 范畴都可以在中文中体现出来.[R2] // [Type 2] 
// 与 Labov 模型唯一不一致的是[marked][T3] 就是两个实验中叙事类型的“概括”与“结尾”范畴的使用[R3]. // 就[3] 这些范畴的使用频率而言[marked][T4=R3], 书写叙事与记忆书写叙事明显要高[R4]. // [Type 2] 
// 除此之外[2][marked][T5], “概括” 与“结尾”范畴在国内大学生中似乎是选择性的出现[R5=R3], // [Type 2] [其它四个范畴[1c][unmarked][T6=R5][T7] 是] 每一个叙事语篇所必须的.[R6] // [Type 3] [Type 1] 
// 该实验结果[1s][unmarked][T7] 同其它研究结果一致,[R7] // 这[1s][unmarked][T7] 说明不是所有的叙事都有“概括”
B. Data Collection

Collection of the data mainly undergoes two steps. The first step is that clauses in the two versions of abstracts are manually annotated. During the process of annotation, I find that embedded clauses and dependent clauses occur frequently in the texts. The embedded clauses are precluded from the analysis, because they are downranked and only function as the modifiers in nominal groups grammatically; while these dependent clauses are included into my analysis because of their being not downranked. The other step is that the frequency of occurrences of each Theme type, markedness and progressive pattern is counted for analysis. The final collected figures are not normalized because the two abstracts are more or less identical in ideation to a great extent.

IV. RESULTS AND DISCUSSION

From the perspective of the composition of Themes in English and Chinese, multiple theme is preferred than simple theme in English, while the Chinese abstract seems to buck the trend as Table 1 demonstrates. Actually, the finding pertinent to the Chinese abstract echoes Cao and Li’s (2015) research on thematic analysis of abstracts taken from the famous international journal of Nature. The simple themes are mainly preferred by these authors, because they are doing their best to make their papers readable and understandable to their readers. Nevertheless, the English version of the abstract is not compatible with previous research findings and some obvious phenomena are found to be present in the English version after scrutinizing the material. 14 out of 21 multiple themes are realized by these common conjunctions and used repeatedly, such as and, thus, and however. This phenomenon implies that the author is not capable of diversifying the multiplicity of the thesis, but he is still trying to employ multiple themes, and hence leads to the high occurrence of multiple themes, which are not appropriate in most cases.

| TABLE 1 |
|---|
| Simple Theme and Multiple Theme in English and Chinese |
|   | Simple Theme (%) | Multiple Theme (%) |
| English | 14 (40) | 21 (60) |
| Chinese | 16 (60) | 11 (40) |

Taking into consideration of the sub-types of Theme, the two versions bear more similarity than the difference. Textual and interpersonal themes are used higher in English version than the Chinese one only in a very slight degree, but the case in topical theme is the other way round. That is, the Chinese version is slightly higher than its counterpart (Figure 4). Figure 4 also tells us that the topical theme is much higher than the other two sub-types of theme, and it is in accordance with M. A. K. Halliday’s idea that the topical Theme is obligatory, and textual and interpersonal themes are optional in each clause. When the markedness of these Themes are counted, they keep a similar trend that the unmarkedness is predominantly higher that the marked cases in both versions (Table 2)), which is in accordance with other related researches. For instance, //This thesis [unmarked] focuses on the most unforgettable experience made by Chinese college students., // and [This thesis] [unmarked] [focuses] extremely on the features of narrative production and working memory. /// The thesis in both clauses is the Theme and at the same time the subject, and thus the thesis is unmarked. Unmarkedness, as a commonly used means, is in line with the way that natural languages presenting the message and caters to our ability of recognition. It is, on the one hand, helpful for the thematic progression, and on the other hand, facilitates the reader’s understanding of the academic papers.
In terms of the thematic progression, two versions of the abstract are mismatched with one another in that the employment of Theme-identical pattern is obviously higher than that of the simple linear pattern and Rheme-identical pattern in English, while it seems that the trend is varied in the Chinese version (Table 3). The distribution of thematic progressive patterns in English version can provide a vindication of previous research findings, no matter the analyzed text is in English (Cao & Li, 2015) or in Chinese (Zhao & Yu, 2000). Academic papers mostly center on a specific topic, which serves as the Theme throughout the text with different Rhemes functioning as directions that the same topic projects from different angles. As to the reasons why simple linear pattern and Theme-identical pattern are both preferred by the author in the Chinese version, a further scrutinization of the material is required. I find that 75% of the simple linear pattern is accompanied by the marked Themes. To be more exact, the real subject of the former clause (part of the Rheme) also functions as the theme of the latter clause and hence forms the simple linear pattern as exemplified in (10).

\[(10) // \text{近年来} [\text{marked}][T1], \text{个人} \text{经历}  \text{叙事} \text{成为}  \text{了} [...], [R1] //  \\
\text{jin nian lai, geren jingli xushi chengwei le [...],}  \\
\text{Recent years person experience narrative becomes ASP [...],}  \\
\text{‘The study of personal experience narratives has received much attention in recent years,}  \\
\text{[个人 经历 叙事][T2=R1] 是 语言 生 [...][R2].}  \\
\text{[geren jingli xushi] is language production [...].}  \\
\text{person experience narrative is language production [...].}  \\
\text{and [the study of personal experience narratives] has become part of the work together with}  \\
\text{language production.’}  \\
\]

V. CONCLUSION

It is of great help for scholars to arrange the clauses coherently while constructing the abstract of a paper by getting to know how others do. Comparing the deployment of Theme-based theories in an English abstract with that in its Chinese version, I find that the two versions employ a very different approach to develop their simple and multiple Themes. Besides, they also progress the Thematic patterns in a quite dissimilar means. Finally, a notable similar approach that the two versions of the same abstract adopt is that they both prefer the unmarked Themes to the marked ones for the ease of readers to decode the extremely compacted abstract. However, this paper only confines to a single abstract in English and Chinese, so the tentative conclusion needs to be enhanced by more samples in broader fields.
APPENDIX

Notational conventions:
- single underlining: Theme
- plain: topical Theme
- bold: interpersonal Theme
- italic: textual Theme
- Type 1: Theme-identical pattern
- Type 2: simple linear pattern
- Type 3: Rheme-identical pattern

Arabic numerals are used for the convenience of counting the total occurrences.

REFERENCES

[1] Cao, Z. H. & Li, L. (2015). Analysis of the Thematic Progression in the Abstracts of Nature's Articles. *Journal of Xidian University* (Social Science Edition), 25(06): 104-111.

[2] Cheung, M. (2011). Sales promotion communication in Chinese and English: a thematic analysis. *Journal of Pragmatics*, 43 (4): 1061–1079.

[3] Daneš, F. (1970). One instance of Prague school methodology: Functional analysis of utterance and text. In Paul L. Garvin (ed.), *Method and theory in linguistics*, 132–140. The Hague: Mouton.

[4] Daneš, F. (1974). Functional sentence perspective and the organization of the text. In Daneš, F. (ed.), *Papers on functional sentence perspective*, 106–128. The Hague: Mouton.

[5] Halliday, M. A. K. (1967). Notes on transitivity and theme in English (part 2). *Journal of Linguistics*, 3 (2): 199–244.

[6] Halliday, M. A. K. (1968). Notes on transitivity and theme in English (part 3). *Journal of Linguistics*, 4 (2): 179–215.

[7] Hu, Zhuanglin. (1980). Linguistic Features of ABC News Broadcasting. Department of Linguistics, The University of Sydney.

[8] Leong, P. A. (2016). Thematic density of research-article abstracts: a systemic-functional account. *Word*, 62 (4): 209–227.

[9] North, S. (2005). Disciplinary variation in the use of theme in undergraduate essays. *Applied Linguistics*, 26 (3): 431–452.

[10] Zhao, H. & Yu, W. S. (2000). Theme Analysis of Chinese Scientific Texts. *Journal of Chongqing University*, 6 (1): 60-62.

[11] Zhou, J. P. (2010). Language Production and Verbal Working Memory in Oral and Written Narratives of Personal Experience. MA Thesis, Southwest University.

Jiangping Zhou is a PhD candidate in Linguistics at Peking University, Beijing, China. He is also a lecturer of English and Linguistics at China West Normal University. His research interests include Systemic Functional Linguistics, discourse analysis, and corpus linguistics.