Theme Structure and Thematic Progression Patterns in the Introduction Section of Psychology and Biotechnology Research Articles: A Comparative Study

Atheer Abulrahman Hunjur

Lecturer, English Department

Jubail Industrial College

Jubail Industrial City, Saudi Arabia

atheerashash@gmail.com

Hesham Suleiman Alyousef

Professor, English Department

College of Sciences, King Saud University

Riyadh, Saudi Arabia

Abstract

Studies on thematic structure and thematic progression patterns focused on abstract and discussion sections of scientific research articles. Therefore, there is a need for more studies investigating other sections of research articles. In this paper, we investigated thematic structure and thematic progression patterns in Biotechnology and Psychology research article introductions to capture any significant differences between these two fields. These two fields represent, respectively, hard and soft sciences. Halliday's (2014) Systemic Functional Linguistics approach to language and McCabe's (1999) thematic progression patterns were adopted in this study as the frameworks for the analysis. After collecting 30 articles with a total word count of 19,517 words, 15 in Biotechnology and 15 in Psychology, the data was

DOI: 10.24113/ijohmn.v11i3.313 Received 23 February, Accepted 2 June 2025

manually analyzed using Microsoft Excel. Frequency count of Theme aspects was carried out using Excel. After that, a t-test was conducted to evaluate the significance of the differences across the two fields. The results indicate a preference for using unmarked Themes more than marked Themes in both disciplines. There was also a similar tendency in both disciplines to use more textual Themes than interpersonal Themes. Finally, the most employed thematic progression patterns were the constant pattern, followed by the zig-zag. The split Rheme and split Theme patterns were rarely used. Future studies need to investigate thematic structure and TP patterns in the introduction section of research articles across other disciplines in order to capture the differences between fields.

Keywords: Systemic functional linguistics (SFL), thematic structure, thematic progression patterns**,** Biotechnology research articles, Psychology research articles, research article introductions.

1. Introduction

There has been an increased focus and interest in the analysis of the genre of research articles (RAs). This interest is due to the importance of research and the language used in it to persuade readers of its significance and value. Theme structure and thematic progression (TP) patterns are two of the most influential and important aspects used in analyzing RAs. This could be due to their focus on the organization of a text and how that affects coherence and cohesion, which would consequently affect comprehensibility and persuasiveness.

Different sections have been focused on in different studies. Some have focused on the abstract section (e.g., Salager-Meyer, 1992; Martin, 2003; Stotesbury, 2003; Lores, 2004; Pho, 2008; Gillaerts and Van de Velde, 2010; Hu and Cao, 2011; Kanoksilapatham, 2013; Khedri, Chan & Ebrahimi, 2013; Ebrahimi and Chan, 2015; Ebrahimi, 2016; Alyousef, 2021). Others have focused on the method and/or the discussion sections (Ebrahimi, 2017; Ebrahimi & Heng, 2013; Martínez, 2003). However, there remains a need for more research

focusing on other sections of RAs, such as the introduction, the literature review, and the conclusion. The introduction section is one of the most suitable sections in an RA for a thematic analysis because of the authors' attempt to engage readers and provide an organized text with a suitable flow of ideas for easier comprehensibility. Alyousef and Alzahrani (2020) support the need for thematic analysis in the introduction section because "the communicative purposes of an introduction lend themselves readily to long cohesive chains, allowing a broad range of TP patterns" (p.121).

Another interesting point for researchers is the differences in the use of thematic structure and TP patterns between disciplines (e.g., Ghadessy, 1999; Ebrahimi & Khedri, 2011; Ebrahimi, 2016; Leong, 2016; Alotaibi, 2020; Alyousef, 2021) and between authors of varying backgrounds, experiences, and degrees of proficiency in English (Aronsson, 2005; Leong, 2007; Hasselgård, 2009; Hawes & Thomas, 2012; Wei, 2013; Hawes, 2015; Park & Nam, 2015; Alyousef & Alzahrani, 2020). To the best of our knowledge, there is a lack of studies investigating variations and/or similarities in the choice of Theme and TP patterns in the introduction section of Biotechnology and Psychology RAs. To do so, this paper aims to fill in the gap by answering the following questions:

i. Are there any significant variations or similarities in the choice of Theme in the introduction section of Biotechnology and Psychology RAs?
ii. Are there any significant variations or similarities in the choice of TP patterns in the introduction section of Biotechnology and Psychology RAs?

2. Literature Review

2.1 Theoretical Framework: Systemic Functional Linguistics

Halliday (2014) introduced in the 1950s the Systemic Functional Linguistics (SFL) approach to language which views language and linguistic choices as functioning hand in hand with the social context of its use. The systems of language are employed to serve social

functions. For people to make sense of communication and interact in different situations, “the grammar has to interface with what goes on outside language: with the happenings and conditions of the world, and with the social processes we engage in” (Halliday, 2014, p. 25). Language in this view is a social semiotic system in which context provides experience and interpersonal relationships that are then transformed into meaning. The meaning is then transformed into wordings. These wordings are then produced via human organs into sounds to be heard, or written, for communication to take place. These steps are the four strata by which language is organized: context, semantics, lexicogrammar, phonology or graphology.

Starting with context or Register, SFL categorizes context into three variables: field, tenor, and mode (Halliday, 2014). Field refers to the situation in which language takes place. It covers the nature of the situation and the domain, or topic. It provides an answer to the question (what is going on?). Tenor refers to the participants in the situation, their roles, status, relationship, affect, and values. It provides an answer to the question (what are the social roles and relationships between the participants?). Finally, mode refers to the channel of communication. It includes language roles in a situation or its importance, its medium be it written or spoken, its rhetorical mode, its channel be it phonic or graphic, and turn, dialogic or monologic. Field, tenor, and mode are thus related but also contrasting variables. Their combination provides various contexts and different uses of language.

The semantics stratum of SFL covers a very important cornerstone in SFL which is the functionality of language. The three variables of context stated above correlate, respectively, with three language metafunctions: ideational, interpersonal, and textual. The ideational metafunction refers to the way humans’ experiences, perceptions, and consciousness are represented in language. Halliday (2014) describes this metafunction as “language as reflection” (p. 30). It is divided into experiential metafunction and logical metafunction. The former refers to ideas, topics, and content and the latter refers to the

logical relations between ideas. The second metafunction is the interpersonal metafunction. It refers to the roles participants take in communication to express attitudes, opinions, and feelings. Language in the interpersonal metafunction provides the vessel for us to express ourselves in a situation. This metafuntion is more active than the ideational and has been described as "language as action" (Halliday, 2014, p. 30). The interpersonal metafunction is expressed through clauses. The third metafunction is the enabling one by which we organize language to refer to what has been said, the world, and ideas. It relates to the organization of the text, be it written or spoken. Language provides the output of the ideational and interpersonal metafunctions via the textual metafunction.

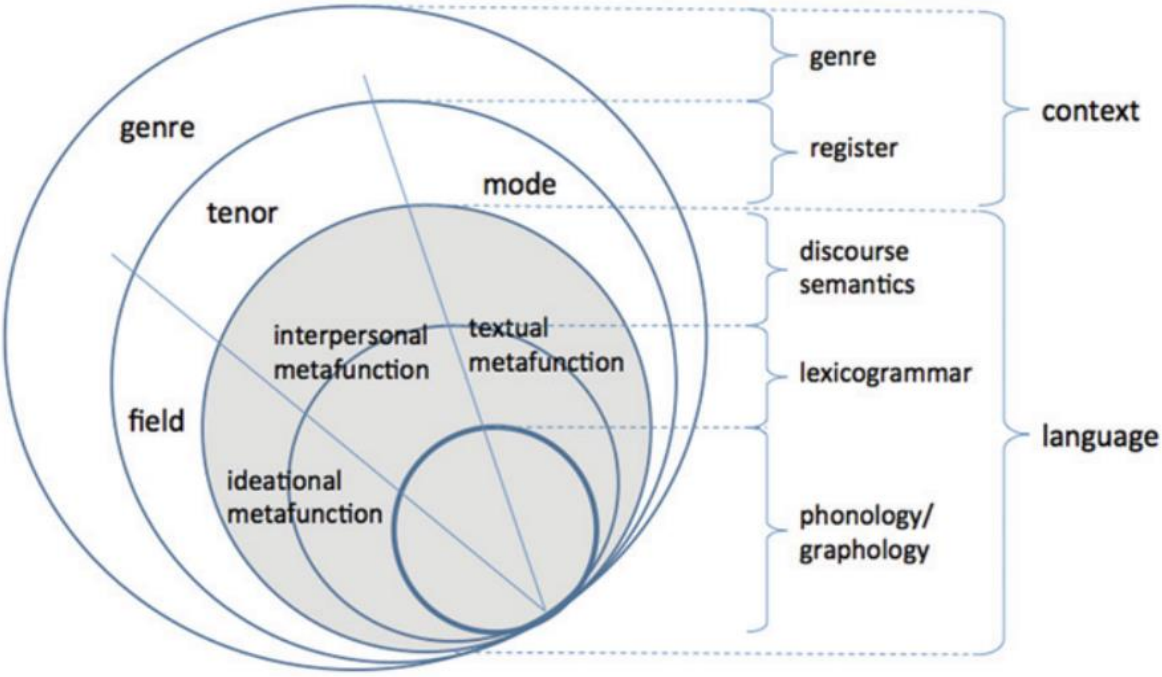
Moving on to the lexicogrammar stratum, this stratum covers the syntactic organization of a text be it written or spoken. It covers both vocabulary and grammar. The focus here is the clause as Halliday (2014) suggests that "meanings of different kinds are mapped into an integrated grammatical structure" (p.10) in the clause. In SFL, the semiotic values resonate with the semantic metafunctions which affect the choice of lexicogrammar resources to serve each metafunction. The main lexicogrammatical resources are Transitivity, Mood, and Theme. For example, tenor-related values could determine that the relation between two speakers is of different status, a professor and a student for instance, this would resonate with the interpersonal metafunction and affect the choice of words, grammar, or Mood when making a request.

The four strata of language are under register or the context of situation. However, there is a higher level that affects all the elements discussed above, i.e., the context of culture or genre (Martin, 1992). In the level of genre, the different ways of achieving a certain goal of communication are affected by different cultures. Martin (1992) defines genre as "staged, goal-oriented social processes" (p. 505). The participants in these social processes are members of their specific culture. Martin and Rose (2003) explain that genre is social

because the interactants are people in the community. These interactants have a goal for their communication and follow certain steps to fulfill their goal, hence goal-oriented. A genre might be a lecture, a wedding, making a purchase, a medical consultation, an RA, and more. The interaction of the different levels in SFL is illustrated in Figure 1.

Figure 1

Systemic Functional Linguistics Framework (adopted from Hao, 2018, p. 4)



2.1.1 Thematic Structure in SFL

One very important aspect of communication and interaction between humans is being able to achieve the goal of communication. This is aided by the organization and flow of ideas, as disorganized ideas in an utterance may hinder effective communication. The textual metafunction of the semantic stratum in SFL studies the ways by which people organize their output and messages in different contexts of culture and situations. Textual meanings are directly related to cohesive and coherent devices used to organize texts, spoken or written. Therefore, an analysis of the textual metafunction is vital to understand the ways

by which organization and flow of ideas in a text is achieved in different genres and registers (Alyousef & Alzahrani, 2020; Alyousef, 2021).

An important structural system within the textual metafunction is the system of Theme in a clause. Thematic structure is important to maintain cohesion in texts. It occurs within clauses. Halliday (2014) defines Theme is "the element that serves as the point of departure of the message; it is that which locates and orients the clause within its context" (p. 89). Speakers or writers choose the point of departure for their message and make it prominent by occurring first in a clause. This then guides the addressee's processing of the message in the intended direction. The remainder of the message is the Rheme. Speakers or writers select what is prominent (typically Given information) and place it first in a clause. They set the scene for the clause. The rest of the clause is the Rheme part of the message, which is typically New information. In other words, the Theme is clause initial and the Rheme is clause final.

Themes are divided into three types: topical, interpersonal, and textual. Topical (or experiential) Themes are obligatory elements of every clause. A topical Theme can either be a participant (or a nominal group), a circumstance (prepositional or adverbial phrase), or a process. Examples 1a-1d illustrate the different types of topical Themes.

1a. The professor + needs the assignment in two weeks.

1b. The assignment, + the professor needs in two weeks.

1c. In two weeks, + the assignment should be submitted.

1d. Submit + the assignment in two weeks.

The topical Theme in 1a is *the professor* which is a participant (the subject). In 1b, the topical theme is also a participant, but this time it is a complement of the verb phrase. In 1c, the topical Theme is *in two weeks* which is a prepositional phrase. 1d has a process as a topical Theme *submit*.

Topical Themes can also be marked or unmarked. Markedness of a topical Theme is based on the usual form of the clause. In a declarative sentence, the usual form is that it starts with a subject, as illustrated in example 1a. Therefore, the topical Theme in 1a is unmarked. 1d also has an unmarked topical Theme as it is an imperative and the usual form of imperatives begins with a verb or a process. 1b and 1c, however, are marked as starting a declarative sentence with a complement or a prepositional phrase is unusual. Marked Themes have an important function of guiding the reader in interpreting messages. They "have a facilitative effect as they render the message unambiguous by setting the scene for the clause carrying that message." (Alyousef & Alzahrani, 2020, p. 117).

There are, however, other elements that may precede the topical theme. These are considered either textual or interpersonal themes. They are optional components of a clause. Textual themes can be either continuative (signaling a move in communication, such as *yes, well, oh, now*), conjunction (words that link or bind the clause, such as *and, or, but, so*, *until, because, assuming that*) and conjunctive adjuncts (adjuncts that relate the text to previous clauses, and they can be moved in the sentence, such as *nevertheless, however*) *(*Halliday, 2014). Interpersonal themes can be vocative (any item used to address someone, such as a personal name or title), modal, comment adjuncts (express attitudes and judgment on the message), and finite verbal operator (auxiliaries such as *is, do, can, could, may*, etc.) (Halliday, 2014). Table 1 provides examples of multiple Themes.

Table 1

Multiple Themes structure

Maybe	*The professor*	*needs the assignment in two weeks*
Interpersonal (adjunct)	**Topical**	**Rheme**

But	*does*	*the professor*	*needs the assignment in two weeks?*
Textual (conjunction)	**Interpersonal (finite)**	**Topical**	**Rheme**
But + nevertheless	*the assignment*	*must be submitted in two weeks.*	
Textual (conjunction + conjunctive)	**Topical**	**Rheme**	

2.1.2 Thematic Progression

Theme-Rheme structures develop cohesion in clauses and beyond clauses. TP patterns play an important role in the cohesion of the whole text. TP refers to "the way in which the Theme of a clause may pick up, or repeat, a meaning from a preceding Theme or Rheme" (Paltrigde, 2006, p. 148). Progression of Theme-Rheme structures is crucial to cohesiveness which in turn is important in increasing or decreasing the comprehensibility of a text (Salmani-Nodoushan, 2007). The theory of TP began with Daneš in 1974 who categorized it into three main patterns: simple linear theme (or the zig-zag pattern), constant Theme (or reiteration), and the multiple Theme (or split Rheme/ fan pattern).

In the linear, or zig-zag, pattern, the Rheme of a clause is the Theme in the following clause. According to Banks (2008), this pattern is the most common in RAs as it helps in developing and building arguments. The constant pattern, on the other hand, is when the Theme of a clause is repeated at the beginning of the next clause, signaling that each clause will introduce further information about the same Theme. This can be done by paraphrasing the Theme, using pronouns, or synonyms. Repeating the same Theme across different clauses in an RA "develops a topic-focused text" (Alyousef & Alzahrani, 2020, p. 117), thereby strengthening the persuasiveness of the text. The final pattern of TP in Daneš's (1974) model

is the split Rheme (or multiple Theme/ fan pattern) which occurs when the Rheme of a clause includes a number of points that are then discussed individually in the following Themes (Paltridge, 2006).

Following Daneš's (1974) model, McCabe (1999) added a fourth pattern of TP, the split Theme. In this pattern, the Theme of a clause includes more than one point that are then discussed one by one in the following Themes. The split Rheme and the split Theme patterns have been described as "expository" since they provide more information or further illustrate and explain a point (Jing, 2015).

McCabe (1999) argues that there are themes that do not belong to any of the TP patterns. In other words, they are the Themes that are not "chained to a previous Theme or Rheme in close proximity" (p. 180). These are called peripheral themes and they are divided into pragmatic, grammatical, extralinguistic, and metatextual. However, since the focus of the paper is on thematic structure and TP patterns, the different types of peripheral themes will not be analyzed.

2.2 Previous Studies

2.2.1 Theme Structure and Thematic Progression in Research Articles

Linguistic variations between disciplines have been proven in numerous studies employing SFL to investigate RAs. This was also advocated by Hyland (2011) who has shown that there are differences in the way authors use words to achieve their goals in different fields. Hyland (2011) states that:

> Writers in the soft fields cannot… report their research with the same confidence of shared assumptions. They must rely far more on focusing readers on the claim-making negotiations of the discourse community, the arguments themselves, rather than relatively unmediated real-world phenomena. This means that arguments have to be expressed more cautiously. (p. 205)

This suggests that variations in thematic structure and careful progression of Themes are even more pertinent in soft discipline for authors to be able to persuade their readers. Therefore, more comparative research is needed to confirm this disciplinary variation.

The classification of disciplines into hard and soft began with Becher (1989). Hard disciplines are those "built on logical deduction and reasoning, knowledge was considered impersonal and value-free, and mathematical modeling used to breakdown complex ideas and search for universal 'truths" (Clifford, 2009, p. 134). These include Math, Physics, Medicine, Biology, Engineering, etc. On the other hand, soft disciplines "offer a more discursive engagement with a body of knowledge, where personal experiences and ideas were valued and knowledge seen to be value-laden, and analysis seen as complex and disputed" (Clifford, 2009, p. 134). Applied Linguistics, Human Behavior, Psychology, and Sociology are examples of soft disciplines. Despite the criticism this classification has been subject to, research has proven that there are linguistics variations between hard and soft sciences and, therefore, we decided to use these labels.

One of the first comparative study that examined thematic structure across disciplines was by Ghadessy (1999). He analyzed 150 abstracts written by native and non-native authors. These abstracts were from 30 different disciplines. He concluded that the corpus provided rich data for the analysis of Themes and that there were variations in the structure of Theme-Rheme between disciplines.

Another study by Ebrahimi and Khedri (2011) has revealed that Applied Linguistics and Chemical Engineering RAs differ in thematic structure. This led them to confirm "the hypotheses that academic research article abstracts are shaped by their disciplinary background" (p. 287). Specifically, they found that the most prevailing type of Theme is the textual Theme in both disciplines. However, it was more used in Chemical Engineering than

in Applied Linguistics. Regarding TP, it was found that the linear chain and constant chain were used more frequently in Chemical Engineering than in Applied Linguistics.

In a similar vein, Ebrahimi (2016) investigated the Thematic structure and TP in 120 RA abstracts from two soft disciplines (Applied Linguistics and Economics) and two hard disciplines (Agriculture and Applied Physics). They found discrepancies between the use of marked and unmarked themes across the two disciplines. Marked Themes were used more in soft sciences than in hard sciences. They have also found that the textual Theme was used more than the interpersonal Theme in all four fields. Ebrahimi (2016) justifies this by arguing that it is due to the "argumentative, factual and less personal tone of the analyzed RA abstracts" (p. 112). Regarding TP, they found a high percentage of miscellaneous patterns, that is a Theme not related to previous Themes or Rhemes.

Leong (2016) also examined the thematic organization in 200 RA abstracts in two disciplines: science and humanities. In both disciplines, the linear pattern was used similarly and more commonly than the other patterns. They, however, differed in the simple-multiple Theme patterns. The science discipline included more simple topical Themes, whereas the soft discipline included more multiple-Themes structure. This could be due to the nature of hard sciences which is reflected in their RAs. Hyland (2011) suggests that hard sciences have facts and evidence to support their arguments whereas authors in the soft discipline use more effort in building their arguments. This extra effort could be reflected in the use of multiple Themes.

A more recent study by Alotaibi (2020) investigated the differences in thematic structure between three hard disciplines (e.g., physics, chemistry, and computer science) and three soft disciplines (e.g., business administration, applied linguistics, accounting). In line with Ebrahimi (2016), textual Themes were used more commonly than interpersonal Themes

in all the examined disciplines. However, when comparing soft and hard disciplines, interpersonal Themes were more common in soft disciplines.

Alyousef (2021) conducted an intradisciplinary investigation of thematic structure and TP patterns in 119 RA abstracts in seven dentistry subdisciplines. He found that the topical Theme was the most frequent type in the 7 dentistry subdisciplines abstracts, followed by the textual Theme, and finally the interpersonal Theme which was “negligible” (p.37). Regarding marked and unmarked Themes, all 7 dentistry subdisciplines used unmarked themes more frequently than marked Themes. It was also found that the linear pattern is the most frequent TP pattern. It was used more frequently in 5 out of the 7 subdisciplines. The constant TP pattern was the second most applied TP pattern in Alyousef’s corpus with no significant differences between the subdisciplines. The least used pattern was the split Rheme which could be “due to the concise nature of an RA abstract, which requires an author to shift from one proposition to another” (p. 41).

2.2.2 Introduction Section in Research Articles

Research articles in different disciplines follow almost the same schematic structure (or organization) in which an introduction is an integral part of the article. It is one of the most important parts as it has the challenging task of presenting the topic and persuading the reader of its importance and significance. To do so effectively, Swales (1990, 2004) introduced the Creating a Research Space (CARS) model which provides the moves and steps authors should include in their introductions to RAs. The three moves are: establishing a territory, establishing a niche, and occupying the niche. These moves also include several steps, e.g., highlighting a gap, providing justification, stating the aims, and providing significance.

For authors to be able to cover all these moves and steps and to achieve the required persuasive effect of introductions, coherence and cohesion are very important to aid

comprehensibility and to direct readers' attention to the desired focus. This means that the introduction section provides rich data for the analysis of Theme-Rheme structures and TP patterns. This analysis will greatly help novice researchers and EFL/ESL learners in improving their academic writing skills. It will also help teachers in improving the material they use for teaching writing.

Zein et al. (2023) investigated the thematic structure and TP patterns in 16 RA introductions published in four Indonesian journals focusing on medicine (in relation to cultural values and beliefs), linguistics, intellectual property rights, and culture, language, and literature. They found that simple Theme structure dominated the introductions in their corpus. This also indicates a dominance of topical Themes over textual and interpersonal Themes. Textual Themes came in second place of the most frequent types of Themes, followed by the least frequent type, interpersonal Themes. Most importantly, they found a gap in the distribution of simple and multiple Theme structures which could result in poor organization or flow of ideas. However, these results cannot be compared with similar studies of Theme in international journals because the researchers collected only four RA introductions from each journal specializing in social humanities studies.

Another study by Alyousef and Alzahrani (2020) investigated Thematic structure and TP patterns in 117 Electrical Engineering introduction sections in RAs. Their focus was on the differences between native and non-native authors. They found that textual Themes were more used than interpersonal Themes. They also found more instances of simple Topical Themes than multiple Themes. Unmarked topical Themes were also higher in both sets of data than marked Themes. Regarding TP patterns, the zig-zag pattern was the highest used pattern, followed by the constant and the split Rheme, respectively.

3. Methodology

3.1 Data

The corpus consisted of thirty articles, comprising a total of 19,517 words. Fifteen of these articles were from ISI journals in Biotechnology and the other fifteen were also from ISI journals in Psychology. As a result, characteristics related as to the number of authors, whether they were native or non-native writers, their experience in publication, age, and degree were not considered. All the articles were published between the years 2020-2023. The journals specialized in Biotechnology are *Artificial Cells, Nanomedicine and Biotechnology*, *Biotechnology & Biotechnological Equipment, Electronic Journal of Biotechnology, and Food Technology and Biotechnology*. The Journals specializing in Psychology are *BMC Psychology, Frontiers in Psychology, International Review of Social Psychology, and Australian Journal of Psychology.* Appendix A provides more details regarding each article and the word count of each introduction section.

The selection of Biotechnology as a representative of hard sciences is due to the importance of this field in research. It is a link between medicine and advancements in technology. Therefore, it is pertinent to understand the way language is organized in these texts and how it affects coherence and cohesion. The choice of Psychology as a representative of soft sciences is also due to its importance and influence in research. Almost all disciplines are linked to Psychology. Applied Linguistics, for example, is linked to Psychology through Psycholinguistics or language education and a learner's psychology. This means that researchers from other disciplines might need to read RAs in Psychology. Therefore, its language and organization of ideas should be clear and easy to follow since a lot of unspecialized readers might need to read studies in Psychology.

The corpus for this study was selected randomly. First, ISI-indexed journals with high-impact factors in both disciplines were identified. Then we randomly selected a group of

articles in each field. A process of elimination was applied afterward. For a more varied sample to represent the population better, word count between articles differed. However, both groups of data have a similar total word count (10,083 in Psychology and 9,434 in Biotechnology) to avoid the possible effect of length on the variation of Theme types, markedness, and TP patterns.

3.1 Unit of Analysis

Following the recommendations of Fries (1995) and McCabe (1999), the T-unit was the unit of analysis. T-units have been also called independent conjoined (or coordinated) clause complexes (Fries, 1995). It is theoretically established that the T-unit is the optimal unit for Theme analysis especially if the aim is to analyze TP patterns (Fries, 1995). In a T-unit, the independent clause is treated as a single unit along with any dependent clauses included in it. If a sentence has two independent clauses, it has two T-units. If the dependent clause comes first, it is the Theme, and the rest of the unit is the Rheme. However, "if the independent clause comes first, then the first ideational element of that clause is understood to be the Theme, with the rest of that clause and any other subsequent dependent clauses forming the Rheme" (McCabe, 1999, 70-71).

To divide the data into separate T-units, the whole corpus was copied, each article from its journal website, and pasted into a Word document. Then, the T-units were manually divided in the whole corpus, using double-slant lines. These were then copied and pasted to an Excel sheet for the manual annotation process to begin. Each T-unit was identified by a code that has either P or B for the discipline, C for the clause, and the serial number of the clause. For example, PC005 stands for clause number 5 in Psychology articles. This code was used for excerpts presented in this paper. A revision of the division of the T-units was conducted before beginning the annotation of each article.

3.2 Annotation of Theme Types: Delimitations

When conducting the annotation of Theme types, it is not always a black-or-white situation. It is natural to encounter a number of issues when conducting the analyses. First, since T-units are independent clauses, compound sentences with coordinating conjunctions may not always have an expressed subject in the second clause. Subjects would most of the time be ellipsed, as shown below.

2a. BC019 *In this review, we discuss the mechanism of TTF,*

2b. BC020 *(ellipsed) summarize the results of pre-clinical and clinical trials,*

2c. BC021 *and (ellipsed) discuss future potentials of this new anti-cancer treatment modality.*

Examples 2a, 2b, and 2c are T-units of the same sentence which was divided into three T-units. The last two t-units, 2b and 2c, do not have an expressed subject. However, the subject is ellipsed. It can be inferred from the context as being the main subject of the whole clause complex, *we*. Therefore, ellipsed subjects in our data were treated as unmarked topical Themes.

Another issue that was faced was the case of anticipatory "it." Following Thompson (2014), phrases with anticipatory "it" were treated as Themes. For example, in a sentence like "*Hence, it is difficult to extract acetaldehyde dehydrogenase from L. plantarum to meet the large-scale production [19, 20]"* (BC361), the phrase "*it is difficult*" is the Interpersonal Theme of the T-unit since it is a "thematized comment" (Thompson, 2014, p. 170), "*to extract*" is the Topical Theme, and the rest of the T-unit is the Rheme. This contradicts Halliday's (2014) argument that the empty pronoun "it" is the Theme. However, highlighting only the empty "it" as the Theme obscures development in the text (Thompson, 2014). Therefore, it is recommended to include the comment in the Theme as well (Thompson, 2014; Alyousef, 2021). Finally, cases of existential "there" were also encountered in

analyzing the data. When dealing with cases of existential "there," it has been recommended by many linguists (e.g., Martin, 1995, McCabe, 1999, Halliday, 2014; Thompson, 2014, Alyousef, 2021) to treat existential "there" as the Theme of the unit. This is because even though existential "there" in itself has no representational function, it typically introduces the existence of New information (Rheme). Therefore, existential "there" was considered in the present study as the Theme.

3.2.1 Markedness

As has been discussed earlier, marked themes occur in sentences that do not start with the usual topical theme. An example of this is "*In the present study, four hypotheses were tested*." (PC333). Here, the sentence began with a circumstance "*in the present study*" which does not follow the usual order of a declarative sentence. A usual order begins with a participant, as in "*Four hypotheses were tested in the present study."* Therefore, in any T-unit beginning with a Topical Theme that does not follow the usual order of the Mood of the sentence, the Topical Theme is annotated as marked.

3.2.2 Thematic Progression Criteria

Following Alyousef and Alzahrani (2020, p. 120), TP patterns were analyzed under two conditions to build "a more systematic approach to the analysis of TP patterns." First, a pattern detected should be between a theme of one clause and a preceding or a subsequent Theme or Rheme via one of the following semantic relations: *repetition, synonymy, subordinate, general word*. This would facilitate the identification of different thematic development patterns, especially the constant pattern. The second condition is that the distance between clauses should be specified by three clauses (McCabe, 1999). For this study, we will be specifying three clauses as the maximum number of clauses for detecting TP patterns.

However, there remain a number of issues that were encountered while annotating the data. Since a distance of a maximum of three clauses was specified, sometimes it is difficult to link a Theme to a clause within the preceding three clauses. Moreover, there are also puzzling cases of circumstances of manner, adverbials of time and place, Wh- elements in interrogative sentences, and existential "there." Following McCabe (1999), these instances were annotated as peripheral Themes since they do not belong to any of the TP patterns, and they do not complicate text processing.

At other times, a Topical Theme might be new, i.e., not related to a previous Theme or Rheme. McCabe (1999) labeled these as New Themes under peripheral Themes. However, since they represent processes or participants and have a representational function, they were labeled as New Themes distinct from peripheral themes in the present study.

3.3 Annotation Procedure

After revising the T-units in the Excel file, all Themes in the corpus were annotated as topical, textual, or interpersonal. Then these Themes were identified as either marked or unmarked. After that, combinations of multiple Themes were annotated. A revision of the annotation process was conducted to ensure higher reliability and consistency. After finalizing the annotation of Theme types and markedness, the process of annotating TP patterns, new, and peripheral Themes began by identifying the development and progression of Themes in a maximum distance of three T-units. A revision of this step was conducted followed by a second revision of the whole annotation process a few weeks later.

Using Excel, the frequency and percentage of each Theme aspect was calculated. Finally, a t-test was run using R Programming to see if the differences were significant. R programming tools have been used as the main statistical analysis tool in many fields of linguistics, especially corpus and discourse analysis studies. This is due to its wealth of flexible functions and statistical packages designed by professional scientists and

statisticians. Not only that but it is also free of charge and can be used endlessly (Levshina, 2015).

4 Results and Discussion

This section presents the findings of our investigation. It starts with a summary of the results presented in Table 2, followed by a detailed comparison and discussion of the various thematic aspects found in this study.

Table 2

Frequency and percentage of each thematic aspect in the corpus

		Biotechnology		Psychology		df	t	*P*-Value
		Freq.	%	Freq.	%			
Theme Types	Topical	463	77%	428	77%	27.83	0.60	0.54
	Textual	119	20%	113	20%	26.58	0	0.82
	Interpersonal	16	3%	16	3%	27.82	0	1
Markedness	Unmarked	389	84%	320	77%	27.22	1.32	0.19
	Marked	74	16%	96	23%	27.73	-1.34	0.19
Thematic Progression	Zig-zag	96	21%	95	23%	28	0.06	0.95
	Constant	147	33%	135	32%	27.98	0.43	0.66
	Split Rheme	11	2%	24	6%	23.24	-1.70	0.10
	Split Theme	3	1%	2	0%	27.81	0.33	0.73
	New	163	36%	119	29%	27.96	1.87	0.07
	Peripheral	31	7%	41	10%	27.91	0.93	0.36

It should be noted that the T-unit count differed across the two disciplines. Although the word count in Psychology articles was higher (10,083) than the Biotechnology articles (9,434), the number of T-units found in the latter was slightly higher (463 T-units) than in the Psychology articles (428 T-units). This suggests that there may be a difference in the

complexity of sentences and clauses across different disciplines. Although this difference is not significant, *P*-value = 0.54, it is worth further investigation.

4.1 Theme Types

Our analysis indicates that all three Theme types were used in our corpus. However, a great degree of variance in the use of the three main Theme types was identified. Naturally, since a topical Theme is an obligatory element in a clause, it was the most occurring type in both disciplines, 77% in Biotechnology and 77% in Psychology. The results indicate that the frequency of using topical Themes is almost identical in the two fields. This is supported by the very high *P*-value of the t-test, 0.54.

The percentage of using interpersonal and textual Theme types was identical in the two investigated fields. The second most used Theme type in the data was the textual Theme, with 20% in both disciplines. Interpersonal Theme was the least occurring type in both disciplines, 3%. The higher percentage of using textual Themes as compared to interpersonal Themes was also found in previous research on Theme (e.g., Ghadessy, 1999; Ebrahimi & Khedri, 2011; Ebrahimi, 2016; Alotaibi, 2020; Alyousef, 2021). It has been argued that it is logical to find more instances of Textual Themes than Interpersonal Themes (Ghadessy, 1999). This is due to the number of conjunctions needed for texts to be cohesive and coherent. This in turn improves the flow of ideas, which could also persuade readers of the ideas presented.

However, the variance in the use of simple topical Themes and multiple Themes is striking. This was also found in previous research (Ebrahimi & Khedri, 2011; Ebrahimi, 2016; Leong, 2016; Zein et al., 2023). It has been argued that the use of multiple Themes could increase the continuity of a text and enhance the efforts to persuade readers (Ebrahimi & Khedri, 2011). Therefore, it is an important structure of cohesion in RAs. Some examples of Theme types from our data are presented in 3a-3c.

3a. PC001 ***Mental health disorders*** *cause more work absence and work disability than musculoskeletal problems [1, 2].*

3b. PC010 ***However, the majority of studies*** *had limited controls for confounding, including whether participants had depressive disorders at baseline [14].*

3c. PC224 ***Specifically, 18.3%*** *were limited range polydrug users (alcohol, tobacco and cannabis).*

These examples illustrate how Theme types are used in the data. Extract 3a provides an example of the use of a simple topical Theme. No other type of Theme is used in this example, and the Theme here is the subject of the T-unit. In 3b and 3c, we notice an extension of ideas before the topical Theme. In 3b, there is the use of a cohesive device "*however"* to show how the ideas are connected. In 3c, the author used an adjunct to focus the reader's attention on a specific point. Sentences 3b and 3c contain a combination of Theme types. In 3b, it is textual^topical while 4c interpersonal^topical.

Not only does the data have these two combinations of multiple Themes, but the analysis also revealed that it has five more combinations of multiple Themes. The most used combination of multiple Themes is the textual^topical Themes, followed by the interpersonal^topical Themes, then the textual^textual^topical Themes, and textual^interpersonal^topical. Other combinations of multiple Theme types were found only in the Psychology articles. These are interpersonal^textual^ topical, textual^interpersonal^textual^topical, and finally textual^textual^ interpersonal^topical. Table 3 presents the frequency and percentage of using each combination in the data.

Table 3

Frequency and Percentage of Combinations of Theme Types

	Biotechnology		Psychology	
	Freq.	%	Freq.	%
textual^topical	109	85%	101	83%
interpersonal^topical	13	10%	9	7%
textual^textual^topical	3	2%	4	3%
textual^ interpersonal^topical	3	2%	1	1%
interpersonal^textual^topical	0	0%	4	3%
textual^interpersonal^textual^topical	0	0%	1	1%
textual^textual^interpersonal^topical.	0	0%	1	1%

The occurrence of more combinations of multiple Themes in Psychology articles aligns with Leong's (2016) findings. This suggests that authors in this discipline exert more effort to convince their readers of their proposals. It also goes in line with Hyland's (2011) proposal that authors in soft discipline use more varied Thematic structures and TP patterns for the purpose of building a stronger argument.

4.2 Theme Markedness

The findings revealed that the unmarked Theme was the most frequently used type in Biotechnology and Psychology RA introductions. In Biotechnology articles, 84% of the topical Themes were unmarked. The use of marked Themes was significantly lower than the use of unmarked Themes. Only 16% of topical Themes in Biotechnology articles were marked. The variation in the use of marked and unmarked Themes in Psychology articles was very similar to their variation in Biotechnology articles. 77% of Themes in Psychology articles were unmarked, whereas 23% were marked. This result is expected "since the

unmarked Theme is the typical way of starting the message of a clause" (Alyousef, 2021, p. 21), as writers prefer to place subjects in Theme position (Ebrahimi & Khedri, 2016), or to allow for the continuity of focus (Mirahayuni, 2002). This finding was also reported in previous studies (e.g., Ghadessy, 1999; Ebrahimi & Khedri, 2011; Ebrahimi, 2016; Alotaibi, 2020; Alyousef, 2021). Examples of the use of marked and unmarked Themes in our data are presented in examples 4a and 4b.

4a. P096 ***Occupational accidents*** *pose a significant threat to human health, the economy, society, and the environment [1].*

4b. BC341 ***Currently,*** *more than 500 ALDHs have been isolated from organisms.*

As can be seen in 4a, the subject of the sentence is the Topical Theme. This is the usual order of sentences in the English language. Therefore, it is an unmarked Theme. Example 4b, on the other hand, begins with a circumstantial adverb of manner (or time) "*currently*" which is not the grammatical subject of the sentence, but it serves the function of specifying the time when an action occurred and putting the focus on the concept mentioned. Since it is not the grammatical subject and it does not follow a typical order of T-units and sentences, it is a marked Theme.

Across the two disciplines, it was found that Biotechnology articles included more unmarked Themes than Psychology articles, but it also included less marked Themes than Psychology articles. However, these differences are not significant as the *P*-value is 0.19 for both types. The higher usage of marked Themes in soft disciplines was also captured by Ebrahimi (2016) who argues that the use of marked Themes indicates a more argumentative nature of discourse. He also argues that the use of marked Themes is one of the ways by which authors in soft discipline create cohesion. Moreover, it has been argued that authors in hard disciplines do not exert as much effort as exerted by authors in soft disciplines to convince their readers or build a stronger argument. They simply present their arguments and

scientific evidence (Hyland, 2011; Ebrahimi, 2016). This could explain why we find more marked themes and more effort in soft disciplines than in hard disciplines.

4.3 Thematic Progression

Our analysis of TP patterns revealed that almost all TP patterns are used in Biotechnology and Psychology articles. It was found that the constant pattern is the most used TP pattern, 33% in Biotechnology articles and 32% in Psychology articles. The second most used pattern in the data was the zig-zag pattern, 21% in Biotechnology articles and in 23% in Psychology articles. This finding is in contrast with previous studies which found a higher frequency of using the zig-zag pattern than the constant pattern. Alyousef (2021) and Ebrahimi (2016) found that the zig-zag pattern was the most used in their data. However, their analysis focused on RA abstracts. The nature of an abstract is that it is more concise and limited in length than any other sections of RAs. Therefore, the use of zig-zag patterns would ensure higher levels of coherence, cohesion, and clarity. Introduction sections, on the other hand, introduce the topic along with all other ideas related to it. There are more new terms, more definitions, and more new ideas than in the RA abstract. The authors' elaboration on explaining new concepts in the RA introduction justifies the need for more constant TP patterns.

However, the introduction sections were also studied by Alyousef and Alzahrani (2020) and their result corresponds with previous research. They found a higher frequency of using the zig-zag TP pattern in electrical engineering RA introductions than the constant TP pattern. This discrepancy between the results of this study and theirs may be due to the corpus size. Their corpus consisted of more than 84,000 words while the corpus in this study consisted of only 19,673 words. This could have affected the frequency count by providing more instances in the larger corpus. Another possible reason is that they focused on the differences between native and non-native authors. This effect was not examined in this

study. Therefore, it might be the reason for the difference in our results. Some of the data in which the zig-zag and constant patterns were used are presented in Table 4.

Table 4.

Examples of the Zig-zag and Constant TP Patterns

Constant Pattern	5a. BC023 ***The TTF***	*is transmitted via ceramic gel pads (arrays) on the skull, which requires shaving the patient's scalp to allow direct skin contact.*
	5b. BC024 ***This therapy***	*should be used by the patients for 18 h a day.*
Zig-zag Pattern	5c. PC002 *In a given year,*	*about **19 million Americans** have a major depressive episode [3].*
	5d. PC003 *Nearly half of **them***	*also have **anxiety disorders** [4, 5].*
	5e. PC004 ***These disorders***	*cost employers nearly $200 billion in 2018.*

In the constant pattern example, we find that the unmarked Topical Theme of the first sentence "*The TTF*" is repeated via the use of a synonym "*This therapy.*" It is reiterated, and, therefore, it follows a constant pattern. The second set of examples represents the zig-zag pattern. The phrase "*19 million Americans*" in the Rheme position of the first sentence is repeated via the use of pronouns in the Theme of the second sentence "*Nearly have of them.*" This also occurs in the second and third examples where the Rheme of sentence 6d "*anxiety disorders*" is repeated in the Theme of sentence 5e using demonstrative and reiteration "*These disorders.*" This movement resembles a zig-zag, hence its name.

The split Rheme and the split Theme patterns come in third place and they were rarely used, constituting only 2% of TP patterns in Biotechnology and 6% in Psychology. The two patterns were used only 3 times in Biotechnology articles and 2 times in psychology articles.

The scarcity in the use of these patterns was also captured in previous studies (e.g., Ebrahimi, 2016; Ebrahimi & Khedri, 2011; Alyousef & Alzahrani, 2020; Alyousef, 2021). This could be due to the nature of discourse.

Table 5

Examples of the Split Rheme and Split Theme in the Corpus

Split Rheme	6a. PC429 *In developing this model,*	*we have relied on **two theoretical frameworks** …*
	6b. PC430 ***One** framework*	*is Schwartz's (1992) theory of basic personal values.*
	6c. PC431 *The **other one***	*is the polarity theory.*
Split Theme	6d. PC424 ***Two interconnected lines** of research …*	*have emerged as central to the study of political orientation as a psychological construct.*
	6e. PC425 ***One** line*	*focuses on the very structure and conceptualization of political orientation.*
	6f. PC426 *The **other line** of research*	*looks at how various psychological constructs… fit in the structure of political orientation*

It is not always the case that we split our ideas in utterances or writing unless there is a need for it. Referring to a previously mentioned idea through the zig-zag and constant pattern is naturally more common than splitting our ideas in a subsequent utterance. Some examples of the split Rheme and the split Theme in the corpus are presented in Table 5.

In example 6a, the Rheme contained two ideas "*two theoretical frameworks*" that were later split into two Themes, "*One framework*" in 6b and "*The other one*" in 6c. The same occurred in examples 6d-6f but to the Theme of 6d instead of the Rheme. The Theme in 6d

contained two ideas expressed in the clause "*Two interconnected lines*" which was then divided into two Themes, "*One line*" in 6e and "*The other line of research*" in 6f.

The variation of TP patterns across the two disciplines is not significant as they were very similar and almost identical. This is supported by the *P*-value for each of the comparison sets, as seen in Table 2. The major differences in the TP patterns between the two fields were in the use of the constant pattern and the split Rheme pattern. Biotechnology articles included more constant TP pattern than articles in Psychology, respectively, 33% and 32%. This difference, however, is negligible. On the other hand, the split Rheme was manifested more in Psychology articles (6%) than in Biotechnology articles (2%). Although the differences are not significant, this finding goes in line with previous research (Ebrahimi & Khedri, 2011; Ebrahimi, 2016).

Themes that fall outside TP chains were also identified and compared. These are New Themes and peripheral Themes. New Themes were used higher than Peripheral Themes and almost as common as the constant TP pattern. In Biotechnology articles, New Themes were found in 36% of Thematic moves. It also occupied a high percentage of the moves in Psychology articles, 29%. The high percentage of using New Theme could be attributed to one of two reasons. It might be due to situations in which the referent is mentioned in a distance of more than three clauses, thus creating a gap, or it can be due to the nature of introductions. As discussed earlier, introductions are used to present new ideas and definitions. This may result in a higher percentage of New Themes in the introduction section than in any other sections of an RA.

On the other hand, Peripheral Themes were used 31 times in Biotechnology articles (7%) and 41 times in Psychology articles (10%). Although we can spot differences in the use of new and peripheral Themes across the two disciplines, these differences are not significant as indicated by the *P*-value for each (Table 2). No previous study, to the best of my

knowledge, investigated the use of peripheral Themes in a cross-disciplinary comparative study. Therefore, it is difficult to say if this percentage of peripheral Themes is the norm or not. However, since most marked topical Themes are peripheral (McCabe, 1999), the difference in the use of peripheral Themes across the two fields could be attributed to the higher percentage of marked Themes in Psychology articles. Examples of New and peripheral Themes are presented in 7a-7d.

> *(7a) Incremental creativity refers to small changes and modifications to existing practices in the organization, focusing on the exploitation of ideas… (7b) Recently, the literature has acknowledged the distinction between radical and incremental creativity… (7c) Both types of creativity are key drivers of organizational performance… (7d) Nonetheless, previous studies did not investigate or further distinguish the mechanism by which leader humor influences these two types of creativity.* (PC014)

Sentence 7a begins with a Theme that is connected to the previous T-unit. 7b, on the other hand, begins with a marked topical Them, an adverb of time, which we labeled peripheral since it does not relate to a previous Theme, and it does not introduce a new idea. Example 7d has a New Theme "*previous study*." This Theme was not mentioned in the preceding three T-units. This makes it a New Theme.

5 Conclusion

Although Thematic structure and TP patterns play a very important role in creating cohesion and coherence in writing different sections of RAs, there is a lack of analysis of their use in the RA introduction. Moreover, many studies on Thematic structure and TP patterns have concluded that their use differs across disciplines (e.g., Ebrahimi, 2016; Ebrahimi & Khedri, 2011; Alyousef & Alzahrani, 2020; Alyousef, 2021). Therefore, the aim of this study was to compare the frequency of Theme types and TP patterns in Biotechnology and Psychology RA introductions. These two disciplines represent, respectively, hard and

soft sciences. Thirty RAs were extracted from ISI-indexed Journals in both disciplines. Adopting Halliday's (2014) SFL theory and McCabe's (1999) TP patterns, the RA introductions were annotated, and the frequency of Theme types and TP patterns was calculated using Microsoft Excel. After that, a t-test was carried out to test the significance of the differences between the two fields.

The results of this study revealed several interesting findings. Starting with Theme types and markedness, the analysis indicated a preference for using more simple topical Themes than multiple themes. Another finding was the authors' preference for using unmarked Themes more than marked ones. It was also found that textual Themes were used more than interpersonal Themes. Although the two disciplines varied in the frequency of Theme types used and markedness, this variation was not significant. However, more investigation on a larger corpus needs to be carried out.

Regarding TP patterns, the analysis revealed a preference for the use of constant pattern in RA introductions, followed by the zig-zag pattern. The split Rheme and split Theme patterns were rarely used in the data. New Themes were used extensively across the two disciplines as they are important for the introduction of new ideas and concepts. Peripheral Themes were also used by authors in both fields. The analysis also revealed that there were differences in the TP patterns employed across the two disciplines. However, these differences were not substantial. More research on the thematic structure and TP patterns in the RA introduction across other fields is still very important in order to capture a more reliable and clearer picture of the differences between disciplines.

6 Limitations and Recommendations

A number of limitations to this study should be acknowledged. First, the size of the corpus was not representative. Although this moderately sized corpus allowed for a more detailed analysis of the data, expanding the corpus was not possible due to time constraints.

Second, characteristics related to the authors of the analyzed text were not considered. These include native and non-native comparison, number of authors, experience in publication, age, and degree. Exploring these characteristics would have enriched the analysis. However, it was not possible due to the constraints on the scope of this paper and the limited time in which it was conducted.

Although the abovementioned limitations did not affect the quality of this paper, it should be considered in future research. Future researchers need also to carry out more studies on more disciplines. Moreover, this paper found a great variance in the number of T-units between the two fields examined. This calls for the need to examine the complexity of T-units across varied fields. One last point to consider in future research is the analysis of peripheral and New Themes, which are understudied although their use in RAs is evident and effective.

Conflict of Interest: The corresponding author, on behalf of second author, confirms that there are no conflicts of interest to disclose.

References

Alotaibi, H. S. (2020). The thematic structure in research article abstracts: Variations across disciplines. *Cogent Arts & Humanities*, *7*(1), 1756146. https://doi.org/10.1080/23311983.2020.1756146

Alyousef, H. S., & Alzahrani, A. A. (2020). A functional analysis of the thematic organization in electrical engineering research article introductions written in English by native and Saudi scholars: A comparative study. *Arab World English Journal*, *11*(2) 114-141. https://doi.org/10.24093/awej/vol11no2.9

Alyousef, H. S. (2021). The status of theme in research article abstracts in seven dentistry subdisciplines: A text-based study of intradisciplinary variations and similarities in thematic choices and thematic progression patterns. *Journal of Language and Education, 7*(1), 28-45. https://doi. org/10.17323/jle.2021.10574

Aronsson, M. B. (2005). *Themes in Swedish advanced learners' written English.* (Unpublished doctoral dissertation). Göteborg University.

Banks, D. (2008). The significance of thematic structure in the scientific journal article, 1700- 1980. *Odense Working Papers in Language and Communication, 29*, 1-29. https://doi.org/10.1080/23311983.2020.1756146

Becher, T. (1989). *Academic tribes and territories: Intellectual inquiry and the cultures of disciplines.* The Society for Research into Higher Education and Open University Press.

Danes, F. (1974). Functional sentence perspective and the organization of the text. In F. Daneš (Ed.), *Papers on functional sentence perspective*, 106-128. Academia.

Ebrahimi, S. (2016). Theme types and patterns in research article abstracts: A cross-disciplinary study. *International Journal of English Language & Translation Studies, 4*(3), 104-115.

Ebrahimi, S. (2017). “In order to…….” A functional study of marked theme in method sections from three disciplines. *Journal of Teaching English for Specific and Academic Purposes, 4*(3), 689-699. https://doi.org/10.22190/JTESAP1603689E

Ebrahimi, S., & Heng, C. (2013). Manifestation of theme as a point of departure in the result and discussion section of academic research articles. *Pertanika Journal of Social Sciences & Humanities, 21*, 29-40.

Ebrahimi, S., & Heng, Chan. (2015). Research article abstracts in applied linguistics and economics: Functional analysis of the grammatical subject. *Australian Journal of Linguistics, 35*(4), 381-397. https://doi.org/10.1080/07268602.2015.1070660

Ebrahimi, S. & Khedri, M. (2011). Thematicity in research article abstracts: A cross-disciplinary study. Educational Quest: *An International Journal of Education and Applied Social Sciences, 2*(3), 287-292.

Ghadessy, M. (1999). Thematic organization in academic article abstracts. *Estudios Ingleses de la Universidad Complutense, 7,* 141-161.

Gillaerts, P. & Van de Velde, F. (2010). Interactional metadiscourse in research article abstracts. *Journal of English for Academic Purposes, 9*, 128–139. https://doi.org/10.1016/j.jeap.2010.02.004

Halliday, M. A. K. (2014). *Introduction to functional grammar*. Revised by Christian M. I. M. Matthiessen (4th ed.). Taylor & Francis.

Hao, J. (2018). Reconsidering “cause inside the clause” in scientific discourse – from a discourse semantic perspective in systemic functional linguistics. *Text & Talk*, *38*(5), 525-550. https://doi.org/10.1515/text-2018-0013

Hasselgard, H. (2009). Thematic choice and expressions of stance in English argumentative texts by Norwegian learners. In K. Aijmer (Ed.), *Corpora and Language Teaching, 33*, 121-139). John Beniamins.

Hawes, T. (2015). Thematic progression in the writing of students and professionals. *Ampersand, 2*, 93-100. http://dx.doi.org/10.1016/j.amper.2015.06.002

Hawes, T., & Thomas, S. (2012). Theme choice in EAP and media language. *Journal of English for Academic Purposes, 11*, 175-183. https://doi.org/10.1016/j.jeap.2012.04.005

Hu, G. & Cao, F. (2011). Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English-and Chinese-medium journals. *Journal of Pragmatics, 43* (11), 2795- 2809. https://doi.org/10.1016/j.pragma.2011.04.007

Hyland, K., (2011) *Disciplines and discourses: Social interactions in the construction of knowledge.* In Starke-Meyerring, Doreen and Paré, Anthony and Artemeva, Natasha and Horne, Miriam and Yousoubova, Larissa, (eds.) *Writing in Knowledge Societies*, 193-214. Parlor Press and the WAC Clearinghouse. http://doi.org/10.37514/PER-B.2011.2379.2.10

Kanoksilapatham, B. (2013). Generic characterisation of civil engineering research article abstracts. *3L: Language, Linguistics, Literature, 19*(3), 1-10.

Khedri, M., Chan, S. H., & Ebrahimi, S. F. (2013). An exploration of interactive metadiscourse markers in academic research article abstracts in two disciplines. *Discourse Studies, 15*(3), 319-331. https://doi.org/10.1177/1461445613480588

Leong, A. P. (2007). Developing the message: thematic progression and student writing. *The Journal of Asia TEFL, 4*(3), 93-127. https://doi.org/10.51601/ijersc.v2i3.86

Leong, A. P. (2016). Thematic density of research-article abstracts: A systemic-functional account. *WORD, 62*(4), 209-227. https://doi.org/10.1080/00437956.2016.1248668

Lores, R. (2004). On RA abstracts: From rhetorical structure to Thematic organization. *Journal of English for Specific Purposes, 23*, 280-302. https://doi.org/10.1016/j.esp.2003.06.001

Martin J. R. (1992). *English text: System and structure*. John Benjamins Pub. Retrieved April 2 2023 from http://site.ebrary.com/id/10534220.

Martin J. R. & Rose D. (2007). *Working with discourse: Meaning beyond the clause* (2nd ed.). Continuum.

Martin, M. P. (2003). A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English for Specific Purpose, 22*, 25-43. https://doi.org/10.1016/S0889-4906(01)00033-3

Martinez, I. A. (2003). Aspects of Theme in the method and discussion sections of biology journal article in English. *Journal of English for Academic Purpose, 2*, 103-123. https://doi.org/10.1016/S1475-1585(03)00003-1

McCabe, A. M. (1999). *Theme and thematic patterns in Spanish and English history texts.* (Unpublished doctoral dissertation). University of Aston.

Mirahayuni, N. K. (2002). *Investigating textual structure in native and non-native English research articles: Strategy differences between English and Indonesian writers.* (Unpublished doctoral dissertation). University of New South Wales.

Paltridge, B. (2006). *Discourse analysis*: *An introduction*. Continuum.

Park, K., & Nam, D. (2015). Analysis of thematic structure in L2 writing: A systemic functional perspective. *SNU Journal of Education Research, 24*, 65–88. Retrieved from: https://s-space.snu.ac.kr/bitstream/10371/95224/1/03%20Park%20Kwanghyun,%20Nam%20Daehyeon.pdf

Pho, P. D. (2008). Research article abstracts in applied linguistics and educational technology: a study of linguistic realizations of rhetorical structure and authorial stance. *Discourse Studies, 10*(2), 231-251. https://doi.org/10.1177/1461445607087010

Salager-Meyer, F. (1992). A text-type and move analysis study of verb tense and modality distribution in medical English abstracts. *English for Specific Purposes, 11*(2), 93-113. https://doi.org/10.1016/S0889-4906(05)80002-X

Salmani-Nodoushan, M. A. (2007). Is text cohesion a precursor to reading success? *Journal of Educational Technology, 3*(4), 87-91. Retrieved from: https://files.eric.ed.gov/fulltext/EJ1069214.pdf

Stotesbury, H. (2003). Evaluation in research article abstracts in the narrative and hard sciences. *Journal of English for Academic Purposes, 2*(4), 327-341. https://doi.org/10.1016/S1475-1585(03)00049-3

Swales, J. (1990). *Genre analysis: English in academic and research settings.* Cambridge University Press.

Swales, J. M. (2004). *Research genres.* Cambridge University Press.

Thompson, G. (2004). *Introducing functional grammar* (2nd ed.). Arnold.

Wei, J. (2013). Corpus-based research on topical theme choices in Chinese and Swedish English learner writings. *Theory & Practice in Language Studies, 3*(12), 2202–2208. http://dx.doi.org/10.4304/tpls.3.12.2202-2208

Zein, T. T., Sinar, T. S., Nurlela, Lubis, T., & Rosa, R. N. (2023). Theme structure in an introduction section of articles published in Indonesian national journals. *Studies in English Language and Education, 10*(1), 251-265. https://doi.org/10.24815/siele.v10i1.25181