

TEXTUAL METAFUNCTION AND THEMATIC PROGRESSION IN ABSTRACT WRITING: A GENRE-BASED SFL ANALYSIS OF PAKISTANI RESEARCH ARTICLES

Najeeba Shahbaz

M. Phil Scholar, Dept of Applied Linguistics, GC University Faisalabad, Punjab, Pakistan.
Email: najebashahbaz00@gmail.com

Amina Nazar

M. Phil Scholar, Dept of Applied Linguistics, GC University Faisalabad, Punjab, Pakistan.
Email: nazaramna429@gmail.com

Dr. Hafiz Muhammad Qasim (Corresponding author)

Assistant Professor, Dept of Applied Linguistics, GC University Faisalabad, Punjab, Pakistan.

Email: muhammadqasim@gcuf.edu.pk

Abstract

Research article abstracts are genre-governed texts that have to achieve a set of multiple communicative functions, at once orient the reader, preview the research purpose, summarize methods, summarize findings, and emphasize scholarly contribution, all in around 150–250 words” (Swales & Feak, 1994: 221). However, relatively little attention has been paid to the realization of these functions at the level of the clause in the context of Pakistani academic writing. Using Halliday and Matthiessen’s (2014) framework for Theme in SFL, Daneš’s (1974) four patterns of thematic progression, and Hyland’s (2000) five-move structure for research article abstracts, a qualitative corpus-based analysis was conducted on 25 Pakistani research article abstracts from the discipline of Applied Linguistics from five HEC-recognized journals (2018–2024). There were 286 clauses coded one by one, and $\kappa = 0.84$ (Theme type) and $\kappa = 0.79$ (progression patterns) confirmed the reliability of the coding scheme. Topical Theme occurs as the dominant choice throughout all the moves ($n = 227$, 79.4%). As expected, the distribution of thematic progression patterns varies from move to move. Constant Theme is the most common overall ($n = 99$, 37.9% of progress-bearing clauses), followed by Derived Theme ($n = 96$, 36.8%), and Linear Theme ($n = 48$, 18.4%). The Conclusion move (M5) was not evident in 32% of the abstracts (8 out of 25); furthermore, none of the 39 clauses classified in M5 contains an Interpersonal Theme. The findings provide insights into typical tendencies and potential future trends in the genre in the studied corpus. Corpus and coding data are available from the corresponding author upon reasonable request.

Keywords: Systemic Functional Linguistics, Textual Metafunction, Thematic Progression, Theme–Rheme, Genre Analysis, Research Article Abstracts, Pakistani Academic Writing.

Introduction

Abstracts of research articles are excellent communicative achievements. They condense 150–250 words into a small package that needs to be informative and communicative in many ways—it gives readers an idea of the topics, significance, methodology, results and contribution of a given



study to a particular discipline. Doing this requires not only skillful organization of information but also fine-tuned use of language.

The information should be organized at the clause level and be such as to create not a compilation of sentences but rather an argument supported by logically organized facts and reasoning.

This is where Systemic Functional Linguistics (SFL) becomes helpful. In SFL, clause-based analysis reveals how in any clause the writer realizes simultaneously three levels of meaning - ideational, interpersonal and textual. The latter aspect, i.e., textual meaning, is achieved in the form of Themes and Rhemes within a clause, where the Theme is a point of departure of the message, whereas the Rheme unfolds and extends the idea, proposition. By establishing specific regularities among The Themes of the successive clauses in an article-thematic progression (Daneš, 1974), writers build argumentative force and authority.

In this research, I have analyzed how this remarkable communicative and linguistic achievement is reached by researchers from Pakistan.

This context is particularly pertinent because all the higher education Pakistani researchers working at local institutions are English as a Second language speakers, but at the same time, the use of English for publishing academically is institutionally mandated. Since it is not possible for non-native English speakers to fully acquire command of this grammatical skill through immersion in a native language environment, the acquisition and mastery of the English scientific register must be learned. Therefore, genre practices of Pakistani academics have significant pedagogic value.

Up to now, most genre analysis research on the abstracts published by the Pakistani academics identified and accounted for major generic moves without taking into account their relationship with any grammatical elements, such as Themes and Rhemes, or for possible systematic correlations with thematic progression patterns. It is precisely for filling this gap that my study is initiated. Specifically, this research aims to answer three questions:

RQ1: What are the types of Themes and their distributions across the five moves of Pakistani Applied Linguistics research article abstracts (Hyland, 2000)?

RQ2: What patterns of thematic progression occur in Pakistani Applied Linguistics research article abstracts, and how are these patterns distributed across genre moves?

RQ3: What associations exist between thematic progression patterns, Theme types, and the genre move structure of Pakistani Applied Linguistics research article abstracts?

We have formulated three types of contributions in this research study. In terms of its theoretical contribution, this study will demonstrate an integrated analysis between SFL textual meta-function and generic structure; in the Pakistani context, it would be possible to explain in a novel way how Hyland's (2000) framework about the move structure, Halliday and Matthiessen's (2014) work on themes and rhemes and Daneš' (1974) studies of thematic progression cohere to reveal some new grammatical and discoursal phenomena. This study will be of significance from the point of view of methodology because it will offer a new, holistic approach towards the combined structural and genre features of English Academic writing. Pedagogically, this research provides the base to develop a move-conscious approach to abstract writing at the level of postgraduate students at Pakistan's higher education institutions.

The material for my study consists of research articles (RA) in the fields of Applied Linguistics and related social science disciplines, originally published by academics at Pakistani universities, whereas those from hard and applied science, such as engineering, medicine, and natural



science disciplines, were avoided, because the structures of information ordering might differ between these disciplines.

1.1 Operational Definitions

Theme: The first constituents of the clause, which make the starting point of the communication (Halliday & Matthiessen, 2014, p. 89).

Rheme: The part that follows the Theme and contains new information.

Thematic progression: The scheme of relationships between Themes and/or Rhematic portions of successive clauses (Daneš, 1974).

Genre move: A unit of text performing a specific function of discourse in a particular genre (Hyland, 2000; Swales, 1990).

2. Literature Review

2.1 SFL, the Textual Metafunction, and Thematic Progression

The SFL model posits that every clause comprises the ideational, interpersonal, and textual meaning simultaneously (Halliday & Matthiessen, 2014). The textual metafunction is enacted thematicity, which serves as the grammatical foundation for other meanings to form an integrated text. In macro terms, Martin's (1992) model of discourse semantics - macro-Theme, hyper-Theme and periodicity - extends from the grammatical into the texture. Eggins (2004) and Thompson (2014) point out that the first position of a clause has a significant role to play with reference to the level of knowledge of genre: a high expertise is predicted to result in a high probability of thematic fronting, while a low expertise is to the contrary (Eggins, 2004, p. 180). Additionally, there are also interpersonal aspects associated with thematic choices related to the authorial voice and institutional pressures, such as the use of hedging and attitude (Hyland, 1998, 2000, 2002). Halliday (1994) proposes that there are three kinds of themes: ideational (topical) themes, textual themes (which make discourse coherent with the use of conjunctions and conjunctives), and interpersonal themes (expressing attitudinal stances with comments). They may be marked or unmarked and simple or multiple. Fries (2013) establishes that the organization of text units depends on thematic development. This work creates a link between the grammar and the genre as they identify "the method of development of segments of text and are the basis of justification for what constitutes paragraph, section and chapter" (Fries, 2013, p. 172). Daneš (1974) establishes four categories of information flow that link together a sequence of clauses. There are constant themes where the topic continues throughout the text, linear progression where the Rheme of one clause becomes the Theme of the next, and the splits-rheme type where subsequent clauses develop one of the components of a complex rheme and, and the derived type where the subsequent clause themes originate from a common macro-theme. These patterns and their frequent use at specific points of the genre are documented across different linguistic contexts and genres (Babai & Behnam, 2020; Daneš, 1974; Wang, 2007; Xiao & Cao, 2021). In second-language contexts and EFL, the prevailing patterns seem to be subject-as-Theme with a predominance of Constant Theme pattern over the course of the text (Guo, 2012), indicating that the nature of L2 acquisition influences such a selection process.

2.2 Genre Analysis of Abstracts and Pakistani Academic Writing

The move sequence Introduction-Purpose-Method-Product-Conclusion described by Hyland (2000) and based on a corpus of 600 abstracts from eight disciplines has been more robustly researched and is probably more commonly used than any other sequence. Lorés (2004), studying 15 non-native abstracts from Spanish linguistics abstracts, found that the Product and Conclusion moves were consistently under-realized. Similar findings regarding the under-realization of the

Conclusion, as the most problematic move, were reported in Martin (2003). Martin had compared the abstract move sequences of English and Spanish abstracts from the social sciences and found a cross-linguistic trend of the less discursal moves being favoured. Investigations of Pakistani research articles have found a similar pattern where most abstracts contained a Purpose and Methodology move, but the Conclusion was absent in several examples (see Lores, 2004; Martin, 2003 for contrastive analysis). It is possible that, while the under-use of moves is not specific to the Pakistani context, few studies have integrated the use of clause-level SFL codes with genre-level analysis.

3. Methodology

3.1 Research Design and Corpus

This research was primarily qualitative SFL analysis on the clause-level of thematic as well as genre-move organization of research article abstract writing with descriptive quantitative statistics. However, descriptive frequencies and percentages were used for tracking the distributional regularities and for a comparative analysis between moves; such data were, however, interpreted rhetorically within the framework of SFL, not as purely quantitative statistics. The corpus, used in this study, was built by means of purposive sampling that was intended to represent the research article abstract writing in Pakistani Applied Linguistics under the current HEC publication standards.

The sampling frame consisted of five HEC-recognized Pakistani Journals, i.e. Pakistan Journal of Language Studies (PJLS), Pakistan Journal of Social Sciences (PJSS), the Journal of Applied Linguistics and Language Research (JALLR), Language in India and the International Journal of English Language and Literature Studies (IJELLS). These criteria of selection were justified in light of their (1) HEC recognition status, (2) focus on the topic of English Language and Applied Linguistics research, (3) having at least one Pakistani researcher (affiliated with a Pakistani university/research institute) on their editorial and or writing board, and (4) the nature of editorial diversity of publications. Besides the above-mentioned criteria, an abstract will qualify as inclusion for this study only if (1) it reports an empirical study or research, (2) it has at least one researcher with affiliation to a Pakistani university/research institute, (3) has been published from 2018 to 2024 inclusive, and (4) belongs to the field of Applied Linguistics or a related social science. The corpus consisted of 25 Pakistani Applied linguistics research abstracts.

The sample was collected from the selected journals on the principle of stratified time sampling, one research article abstraction was sampled per annual interval of each Journal in the study from 2018 to 2024 for a total of 25 abstracts.

Table 1 outlines the specification of the Corpus of Study.

Table 1

Corpus Overview

| Parameter | Value |
|-----------------------|--|
| Total abstracts | 25 |
| Source journals | 5 (JALLR, PJLS, Language in India, IJELLS, PJSS) |
| Abstracts per journal | 5 (stratified by year, 2018–2024) |

| Parameter | Value |
|---------------------|--|
| Publication period | 2018–2024 |
| Discipline | Applied Linguistics and cognate social sciences |
| Author affiliation | At least one verified Pakistani institutional affiliation. |
| Article type | Empirical research articles only |
| Total clauses coded | 286 (M = 11.44; range 3–20) |

Note. Corpus inventory and coding documentation are available from the corresponding author upon reasonable request.

3.2 Analytical Procedure

Data were analyzed through four cycles. Firstly, each abstract was broken down into sections based on the five rhetorical functions as defined by Hyland (2000): the Introduction, Purpose, Method, Product, and Conclusion. Move categories were recorded for each abstract before clause segmentation, thus reducing bias in thematic interpretation. Secondly, each finite and non-finite clause with a reconstructable subject was taken as a unit of theme, as per Halliday and Matthiessen's (2014) systemic and functional theory of clause.

Embeddings within the noun group acting as relative clause modifiers of nouns were integrated into the modifier, and, similarly, all initial adjuncts that were non-finite and appeared in a clause's initially positioned thematic unit were included in the analysis due to their consistent functioning as Theme. Thirdly, all clauses were analyzed for their Theme type (Topical, Textual, or Interpersonal), for markedness (marked or unmarked) and simplicity (simple or multiple). Textual themes were then divided into their subdivisions (i.e., additive, adversative, causal and concessive). Fourth, pairs of adjacent clauses within each abstract were analyzed and categorized according to Daneš's (1974) schema of thematic progression (Constant Theme, Linear Progression, Split-Rheme and Derived Theme).

The very first clause in any abstract was assigned a coded value of 'No Progression' by the researcher, since no prior clause exists from which it may proceed; this category has been excluded from proportions calculated for the remaining four patterns.

Lastly, column proportions, column percentages (within each move and the corpus totals) of Theme types and progression patterns were compiled to one decimal place.

3.3 Treatment of Overlapping Move Realizations

Certain clauses seem to perform two communicative functions simultaneously. If so, the clause was designated to the function deemed most prominent, based on the following three ordered preferences: (a) core illocutionary force; (b) position in abstract; (c) the existence of salient lexicogrammatical cues. Any classification of the clause as dual function is in the coding record and is noted to the second analyst on coding of the reliability material. A dual decision, which we conflicted over ($n = 6$), was determined by consultation with the coding manual, and coded conservatively.

3.4 Reliability and Validity

A structured coding manual of operational definitions and worked examples was developed before analysis. All 25 abstracts were subsequently analyzed, and a random 20% sample (five abstracts; 76 clauses) was independently recoded by a second analyst trained in SFL to check

intercoder reliability. Full coding data are available from the corresponding author without restriction. Landis and Koch (1977) suggest Kappa scores of = 0.84 (Theme type) and = 0.79 (progression patterns) are “almost perfect agreement and “substantial agreement”, respectively. Conflicts were settled by referring back to the code book and edge cases are listed and addressed in a conservative manner. Construct validity was guaranteed by a close fit to the relevant taxonomic constructs presented in Halliday and Matthiessen (2014) and Daneš (1974). The dataset was taken from the publicly accessible journal literature; thus, no human participants were employed.

4. Findings

4.1 Move Distribution

Table 2 shows the frequency of Hyland's (2000) 5 moves across the 25 abstracts of the corpus. Purpose (M2) and Results/Product (M4) are the most consistent moves, as they are both present in 24 out of 25 abstracts (96%). The Method (M3) is present in 23 abstracts (92%).

A smaller percentage of the abstract includes the introduction move (M1), with 13 (52%) containing this move.

This evidence suggests that a large percentage of abstract authors move directly from expressing their purpose for conducting the study. Most conspicuously, one third (8 out of 25) of the abstracts do not feature the move Conclusion (M5); they do not include any of the points relating to the implications and significance of the study and/or the directions for further work. Given this disproportion between the relative space dedicated to the method and argument parts of an abstract, then half the corpus fails to give an account of their 'concluding argument'.

Table 2

Genre Move Distribution Across 25 Corpus Abstracts

| Move | Communicative Function | N Clauses | N Abstracts | % Abstracts |
|----------------------|---|------------|-------------|-------------|
| M1: Introduction | Problem statement: theoretical context | 43 | 13 | 52.0% |
| M2: Purpose | Research aim; objective; question | 39 | 24 | 96.0% |
| M3: Method | Data; procedure; sampling | 82 | 23 | 92.0% |
| M4: Product/ Results | Findings; outcomes | 83 | 24 | 96.0% |
| M5: Conclusion | Significance; implication; future direction | 39 | 17 | 68.0% |
| Total | | 286 | 25 | — |

Note. M5 is absent in 8 of 25 abstracts (32.0%). N Clauses = total clauses coded within that move across all 25 abstracts. All percentages are computed from the 286-clause coded dataset.

4.2 Theme–Rheme Analysis: Coded Extract

In Table 3, the Theme-Rheme classification for abstract 07 is presented at the clause level. This abstract is included as a sample text for two reasons: all five Hyland (2000) rhetorical structure moves (rsms) were present, as were the entire range of Theme-types and thematic-progression

types encountered in the whole corpus. Abstract 07 was coded according to the same 4-stage protocol used for all 25 abstracts.

Results in Sections 4.3–4.6 are based on the whole 286- clause coded database.

The choice of abstract 07 as an illustrative, rather than representative, example was made because it included all five of the established rhetorical structure moves, as well as all the main thematic features found in the overall corpus.

Table 3

Clause-Level Theme–Rheme Coding of Corpus Abstract 07 (Illustrative Corpus Extract)

| l. | Theme | Rheme | Theme Type | Move |
|----|--|---|--|------|
| | Academic writing in second language contexts | has been identified as a site of sustained pedagogical concern. | Topical; Unmarked; Simple | M1 |
| | In Pakistan specifically, | EFL writers face significant challenges in conforming to international genre conventions. | Topical; Marked (Circ. Adj.) | M1 |
| | This study | investigates the discourse structure of research article abstracts by Pakistani academics. | Topical; Unmarked; Simple | M2 |
| | Drawing on Halliday and Matthiessen’s SFL framework, | The study employs thematic analysis to examine clause-level information organization. | Topical; Marked (Non-finite); Multiple | M3 |
| | The corpus | consists of 25 abstracts from HEC-recognized journals (2018–2024). | Topical; Unmarked; Simple | M3 |
| | Data | were coded for Theme type and progression using Daneš’ taxonomy. | Topical; Unmarked; Simple | M3 |
| | The analysis | reveals that unmarked topical Themes predominate across all moves. | Topical; Unmarked; Simple | M4 |
| | Furthermore, | Constant Theme characterizes Method moves; Linear Progression marks Product and Conclusion moves. | Multiple (Textual + Topical) | M4 |
| | These findings | suggest that thematic choices in Pakistani abstracts are genre-sensitive. | Topical; Unmarked; Simple | M5 |
| 0 | Importantly, | The study contributes to SFL-informed academic writing pedagogy in Pakistan. | Multiple (Interpersonal + Topical) | M5 |

Note. Theme identification follows Halliday and Matthiessen (2014). The move assignment follows Hyland (2000). Cl. = Clause; Circ. Adj. = Circumstantial Adjunct. All quantitative findings in Sections 4.3–4.6 derive from the complete 286-clause coded dataset.

4.3 Theme Type Distribution

Theme type is the distribution of Theme types for all 286 clauses that have been coded. It is clear from Table 4 that there is a strong tendency for Topical Themes in this corpus, which made up to 79.4% of the total of Themes ($n = 227$). The choice made by the writers in this corpus shows a pattern where the research entities ‘the study’, ‘the data’, ‘the analysis’ lead the text by presenting the topic first.

This is in line with the trends in Chinese and other EFL writing traditions, where the use of the subject as a Theme in second language writing can be seen in various contexts (Guo, 2012).

Table 5 shows a lower prevalence in the use of Textual Themes, where they occupy 19.9% ($n = 57$) of the text, mainly used as additive connectors and adjuncts such as in “Furthermore” and “In addition”. Interpersonal Themes represent a negligible part of the text, only 2 out of the 286 clauses (0.7%) use Interpersonal Theme, which is located in M1 (1%) and M3 (1%) respectively, and none can be found in Move 5. Textual multiple Themes in this corpus make up to 19.9% of the Themes. Other kinds of textual Themes are used less often than additive connectors and adjuncts, with the occurrence of adversative, causal, and concessive conjunctions each being under 5% of the text occurrence.

Similar to other previous EFL academic writing studies, there is a tendency towards the underuse of logical connectors expressing argumentation (Guo, 2012; Lorés, 2004).

Table 4

Theme Type Distribution Across the Corpus (N = 286 Clauses)

| Theme Type | N | % | Primary Move Context |
|--------------------------------|-----|-------|--|
| Topical (all subtypes) | 227 | 79.4% | All moves (esp. M2, M3, M4) |
| Unmarked, Simple | | | Most frequent subtype; all moves |
| Marked: Circumstantial Adj. | | | M1 and M3 move boundaries |
| Marked: Non-finite Participial | | | M3 openings |
| Textual (all subtypes) | 57 | 19.9% | All moves |
| Additive (e.g., Furthermore) | | | M3, M4 (most frequent textual subtype) |
| Adversative | | | M1, M4 |
| Causal | | | M4, M5 |
| Concessive | | | M1 |
| Interpersonal | 2 | 0.7% | M1 (n=1); M3 (n=1); NONE in M5 |

| Theme Type | N | % | Primary Move Context |
|--------------|------------|-------------------|----------------------|
| Total | 286 | 100.0 % | — |

Note. Topical Theme subtypes share a total of 227; the rows are provided for descriptive context. The two Interpersonal Theme instances are: A16-C04, M1 ('To the best of the author's knowledge', epistemic modal adjunct) and A04-C03, M3 ('we', agent pronoun in participial construction). No Interpersonal Themes appear in M5 clauses.

4.4 Theme Type Across Moves (RQ1)

Table 5 shows a cross-tabulation of Theme type by genre move, thus answering RQ1. Thematicity dominates across the five moves (from 75% in M4 to 95% in M2). The highest incidence of Textuality was found in M3 (23%) and M4 (25%), in contrast to the very low occurrence in M2 (5%), thereby indicating a higher probability that additive or connective textual material will be found in moves within the Method or Results sections. Textuality was relatively infrequent in all other moves

Table 5

Theme Type × Move Cross-Tabulation (N = 286 Clauses)

| Theme Type | M1 n = 43 | M2 n = 39 | M3 n = 82 | M4 n = 83 | M5 n = 39 | Total |
|---------------|-----------|-----------|-----------|-----------|-----------|------------|
| Topical | 35 (81%) | 37 (95%) | 62 (76%) | 62 (75%) | 31 (79%) | 227 |
| Textual | 7 (16%) | 2 (5%) | 19 (23%) | 21 (25%) | 8 (21%) | 57 |
| Interpersonal | 1 (2%) | 0 (0%) | 1 (1%) | 0 (0%) | 0 (0%) | 2 |
| Total | 43 | 39 | 82 | 83 | 39 | 286 |

Note. Cell values: n (column % within move). Percentages are calculated within each move column and may not sum to 100% due to rounding. ZERO Interpersonal Themes in M5.

4.5 Thematic Progression Pattern Distribution (RQ2)

The classification of progression types is shown in Table 6, and for the five moves is displayed by frequency in each of the four Danes' patterns. Leaving aside the Researcher-defined category No progression (n=25; opening clause only), a total of 261 progression bearing Clause transition are divided into Constant Theme (n=99; 37.9% of progression-bearing clauses), Derived theme (n=96; 36.8%), Linear theme (n=48; 18.4%), and Split Rheme (n=18; 6.9%). A second important study finding is that this distribution across moves is non-random and this is considered in Section 5.

Table 6

Thematic Progression Pattern Distribution by Abstract Move (N = 286 Coded Clauses)

| Pattern | M1 n=43 | M2 n=39 | M3 n=82 | M4 n=83 | M5 n=39 | Total | % (of 261) |
|----------------|---------|---------|---------|---------|---------|-------|------------|
| Constant Theme | 16 | 12 | 41 | 15 | 15 | 99 | 37.9% |
| Linear Theme | 9 | 5 | 6 | 22 | 6 | 48 | 18.4% |
| Derived Theme | 9 | 6 | 30 | 33 | 18 | 96 | 36.8% |
| Split Rheme | 0 | 0 | 5 | 13 | 0 | 18 | 6.9% |

| Pattern | M1 n=43 | M2 n=39 | M3 n=82 | M4 n=83 | M5 n=39 | Total | % (of 261) |
|-----------------------------|-----------|-----------|-----------|-----------|-----------|------------|---------------|
| No Progression ^a | 9 | 16 | 0 | 0 | 0 | 25 | — |
| Total | 43 | 39 | 82 | 83 | 39 | 286 | 100.0% |

Note. ^aNo Progression is a researcher-assigned category for the opening clause of each abstract (one per abstract; n = 25). It is not part of Danes’s (1974) original taxonomy. The percentage column (% of 261) excludes No Progression clauses. All figures are computed from the 286-clause coding sheet.

4.6 Interpersonal Theme Statistics (RQ3)

Table 7 shows the breakdown of Theme types for M5 clauses and the summary statistics for the Interpersonal Theme in the 17 abstracts that contained at least one Move 5 element. Out of the 39 M5 clauses, 31(79.5%) contained Topical Themes and 8 (20.5%) contained a Textual Theme. None of the M5 clauses contains an Interpersonal Theme.

Across all 17 abstracts that contained M5, the minimum and maximum count for an instance of Interpersonal Theme in Move 5 were 0.

The mean score was 0.00 and the median was 0.0. Two of the 14 available Interpersonal Themes occurred in Move 1 (in A16-C04 and A04-C03).

Table 7

Theme Type Distribution in Move 5 and Interpersonal Theme Statistics (N = 39 M5 Clauses)

| Metric | N | % | Notes |
|--------------------------------------|------|--------|--------------------------------------|
| M5 Topical Themes | 31 | 79.5% | All 17 abstracts with M5 |
| M5 Textual Themes | 8 | 20.5% | All 17 abstracts with M5 |
| M5 Interpersonal Themes | 0 | 0.0% | None observed in any M5 clause |
| M5 Total Clauses | 39 | 100.0% | Across 17 abstracts |
| Abstracts with M5 | 17 | 68.0% | 8 abstracts lack M5 |
| Mean M5 clauses/abstract | 2.29 | — | Among 17 abstracts with M5 |
| Interpersonal Themes in M5: range | 0–0 | — | Across all 17 abstracts with M5 |
| Interpersonal Themes in M5: mean | 0.00 | — | |
| Interpersonal Themes in M5: median | 0.0 | — | |
| Total Interpersonal Themes in corpus | 2 | 0.7% | 1 in M1 (A16-C04); 1 in M3 (A04-C03) |

Note. All statistics were computed directly from the 286-clause coding sheet. Interpersonal Theme instances outside M5: A16-C04, M1 (‘To the best of the author’s knowledge’); A04-C03, M3 (‘we’).

4.5.1 Constant Theme in Method Moves: Procedural Cohesion

53 of 82 M3 clauses (47.3%). The pattern is shown in Table 3 from clauses 4 to 6 of the following Abstract:4: Constant Theme. 47.08 of 82 M3 clauses (57.3%). The pattern can be shown in the M3 clauses 4 to 6 of Abstract 03, below: Constant Theme. This is the dominant pattern for the M3 clause with 50.0%.

C4: [Drawing on Halliday and Matthiessen's SFL framework,] Theme // [the study employs thematic analysis to examine clause-level information organization.] Rheme [M3]

C5: [The corpus] Theme // [consists of 25 abstracts from HEC-recognized journals (2018–2024).] Rheme [M3—Constant: co-referential]

C6: [Data] Theme // [were coded for Theme type and progression using Daneš' taxonomy.] Rheme [M3—Constant: process anchor]

The thematic group ('the study', 'the corpus', 'data') is continued through coreferentiality, while the individual Rhemes indicate different procedural moves in addition to the move types previously identified. The movement in the Theme from researcher entity to data entity is arguably correlated with a move away from the researcher toward data and procedure - a feature also found by Wang (2007) in the reporting of social science methodology. Wang (2007) also found that the Constant Theme pattern is pervasive in the method sections of English Linguistics Journal articles and our finding that it also occurs in the denser context of an abstract implies it is a genre - motivated feature that can span different levels of text.

4.5.2 Derived Theme in Method and Product Moves: Multi-Dimensional Development

Derived Theme This is the second most frequent clause pattern found in all the data ($n = 96$; 36.8% of progression-bearing clauses) and features quite highly in Moves M3 ($n = 30$; 36.6%) and M4 ($n = 33$; 39.8%). Here, consecutive clause Themes are formed from a major hyper-Theme setting up at the start of the move or throughout the entire abstract, thus creating multi-dimensional development of a methodology or a body of results, without actually having to repeat the Theme itself. The use of Derived Theme is also to be seen within Introduction moves ($n = 9$), where a succession of Derived Themes lists sub-problems associated with a given research field, denoted by a single hyper-Theme. This pattern, therefore, clearly resonates with the CARS move structure in Swales (1990) in that the macro-Theme defines the research space to which the Derived Themes belong.

4.5.3 Linear Theme in Product Moves: Argumentative Momentum

The 'linear' theme type occurs mainly in Product (M4) (22/83 M4 clauses are linear, and 6/83 M5 are linear). In 'linear' progression, the 'rheme' of sentence N becomes the 'theme' of sentence N+1 (rheme-to-theme progression related to argumentation in the corpus). The Rheme-to-theme chain for clauses 7-9 of Abstract 07:

C7: [The analysis] Theme // [reveals that unmarked topical Themes predominate across all moves.] Rheme [M4]

C8: [Furthermore,] Theme // [Constant Theme characterizes Method moves; Linear Progression marks Product and Conclusion moves.] Rheme [M4—Multiple Textual + Topical]

C9: [These findings] Theme // [suggest that thematic choices in Pakistani abstracts are genre-sensitive.] Rheme [M5—Linear: Rheme of C8 becomes Theme of C9]

It's argued that the chain from identifying findings to the final interpretation could guide the reader through the argument instead of leaving inferential effort to them. This was found to be



consistent in the argument, which falls between the Product and Conclusion move interface, where the findings and their interpretive function were connected by using thematic chaining.

4.5.4 Split Rheme in Product Moves: Parallel Findings Presentation

Split Rheme is almost exclusively confined to Product moves (n=13 of 18; 72.2%) and this is where a compound Rheme has distributed its elements among the Themes of successive clauses. These findings tend to occur in tandem, and this device elegantly and efficiently presents several pieces of evidence originating from a single experimental process in what amounts to a 'two-for-the-price-of-one deal' that suits the abstract style's space constraints very well. We find no examples of Split Rheme in M1, M2, or M5, so this can be clearly attributed to the presentational needs of the Product.

5. Discussion

5.1 Genre-Sensitive Thematic Organization: SFL Theory and EFL Comparison

The move-conditioned, non-random sequence of development above supports SFL's claim about the interdependent nature of genre and lexico-grammar (Halliday & Matthiessen, 2014; Martin, 1992). The information structure of each move is correspondingly realized as a thematic pattern at the clause level: Method moves correspond to co-referential Constant Themes; Product and Conclusion moves are associated with Linear Progression if present; the Introduction move, with Derived Themes to map multi-dimensional research problems. The above correlations are well accounted for in Fries's (2013) assertion that thematic choice is one type of thematic pattern that defines the method of development of a text segment, the abstract in this case, within the case here, with the genre constraint of a short genre. The frequent occurrence of Topical Themes in writers of this corpus (79.4%) mirrors that of Chinese and other EFL writing (Guo, 2012), suggesting a cross-linguistic pattern towards Subject-as-Theme.

The conspicuous absence of the Conclusion moves in many of the abstracts (32% abstracts don't have M5) corresponds to the under-realization of evaluation moves in the non-native abstract found in Lorés (2004) and Martín (2003).

Therefore, data show not only missing move but also no presence of corresponding clause-level resources of Interpersonal Themes and linear significance-claiming chains by which the move's pragmatic functions are fulfilled, the crucial contribution of the paper is to establish clause-level grammatical correlates for this previously structurally described genre-level feature by incorporating Hyland's (2000) move model, Halliday & Matthiessen's (2014) Theme–Rheme structure and Daneš' (1974) thematic progression schema by coding of each Pakistani Applied Linguistics abstracts.

5.2 The Move 5 Deficit: Analytical and Pedagogical Dimensions

This absence of Move 5 from 32% (8/25) of abstracts is simultaneously structurally, grammatically and contextually realized. Structurally, Move 5 occupies the communicative space that makes the significance available. Grammatically, its absence correlates with the complete absence of Interpersonal Themes in the M5 clauses of the dataset, which is itself consistent with the argument that interpersonal stance marking in the genre-constructed academic registers is conventional.

Writers use an Interpersonal Theme when they make Move 5, but Interpersonal Themes are absent when Move 5 is missing.

This research shows the clause-level correlate of this genre-based finding: the thematic structures that realize Move 5 are, in the absence of that move, non-existent. It is found that this structure is a recurring feature of native EFL abstract writing (Lorés, 2004; Martín, 2003), and therefore



these data do not necessarily reflect a general absence of awareness of the genre's structure, but rather an underrealisation of a particular aspect of its evaluative structure. Constant Theme is the main structural component of any academic writing genre in English and Topical coherence is one of the key cohesion systems that operates across the genre-type Guo(2012; these could work within procedural clauses but seem less central to the evaluation, which is an important feature of the evaluative move). The analysis of the role of Move 5 as the clause-level slot for genre-specific interpersonal and linear-argumentative thematic constructions implies that explicit instruction focused on such resources would be beneficial, as these are conventional and learnable, so writers already possessing an understanding of the genre structure would be able to simply pick up these specific linguistic forms, and would not need to construct a wholly new genre (Martin & Rose, 2008).

6. Limitations

The 25 abstract corpus size is comparable to prior qualitative SFL studies (Lorés, 2004; Martín, 2003) and is interpreted as indicative, not generalizable. Also, what has been described is true of Pakistani Applied Linguistics abstracts but not of Pakistani writing in the discipline more generally. A wider and stratified corpus with multiple institutional and time contexts will be needed before generalizing or inferring frequency patterns. The study is specific in that it explores patterns of Pakistani Applied Linguistics and related social sciences; such patterns may not hold for other academic disciplines like the natural sciences, engineering, or the humanities (Hyland, 2000).

Moreover, despite high inter-rater reliability ($\kappa = 0.84$; $\kappa = 0.79$), thematic analysis is an interpretive process, particularly as far as classifying boundary and differentiating move boundary from non-move boundary and coding cases as Derived Themes. Though a separate category, No Progression ($n=25$) was introduced for an account of the introductory statements, there are some unrecorded cases which could pose further interpretive subjectivity and ambiguity for move boundary establishment. Last, it can be pointed out that the focus of this research is solely on textual products, so it is beyond the purview of this investigation to make claims about writers' intentions or the development of the draft text. This study is, therefore, descriptive, with no inferential statistics being run to ascertain the underlying factors, or indeed to posit causal connections. One might wonder whether the identified structural choices signify writers' explicit generic awareness or intuitive conventionality, but a text product-based analysis per se cannot answer this query and limits the generalizability and explanatory power of the findings.

7. Conclusion

We propose, based on corpus-based analysis, that Thematic progression patterns in the abstracts analysed are not distributed randomly but correlate with the genre move structure of the abstract.1 Concerning the question "Are specific themes preferred to build certain move structures in Pakistani Applied Linguistics abstract", our answer is as follows: Topical Themes outnumber others across all 5 moves in the corpus (79.4%), with the next most frequent genre the Textual Theme as the topical subject of moves 3 and 4 (7.9%, and 6.1% respectively). Interpersonal Themes are barely utilized ($n=2$, 0.7%), with not a single example being found at the level of Move 5 of the text genre.2 With regard to "Are any certain thematic progression patterns used more frequently, or pattern-move co-occurrences found in Pakistani Applied Linguistics Abstracts, we observed that the following themes are overall most prominent: Constant Theme (37.9% of progression-bearing clauses), followed by Derived Theme (36.8%), these two progression types accounting for 74.7% of the total. There is a clear relationship between some of these types and specific moves.



As can be expected, Constant Theme is very often seen to be most concentrated in Method moves, and Derived Theme to be well represented in Method and Product moves.

Linear Theme occurs predominantly in the move, which is usually found at the borderline between Method and Product (almost 10% of these two moves), while the relatively rare Split Rheme exclusively pertains to move Product (almost 7% of Move4). There are no Interpersonal Themes or most common progression patterns observed at level 5.3. The most striking finding to emerge is that M5 clauses are totally devoid of Interpersonal themes, while move 5 does not appear in 32% of abstracts. This clause-level finding adds to what has been reported regarding a genre-level pattern in both movement and lexical terms from Pakistani Applied Linguistics texts of this genre (Lorés, 2004; Martín, 2003) by indicating how it would manifest at the grammatical level in terms of thematic development: The most significant finding of this study is, in conclusion, that the grammatical mechanisms by which academic significance is most clearly formulated in the genre-in the form of thematic processes that relate academic work to a prior context of significance in a way that highlights the novelty and the importance of academic findings in relation to that prior work-seem to be mostly suspended when the Move5 “Conclusion “ is not explicit. Our integrated approach, utilizing Hyland’s (2000) move classification along with Halliday and Matthiessen’s (2014) thematic concept of the clause and Daneš’s (1974) thematic progression taxonomy, revealed many characteristics that had previously been unexplored at this level in the specific corpus examined.

Theoretical Implications

The evidence presented supports the SFL hypothesis of the reciprocal influence between genre and lexico-grammar (Halliday & Matthiessen, 2014; Martin, 1992). The move-constrained distribution of Theme types and thematic progression shows genre sensitivity in the textual meta-function of the clause, which, to the extent of this abstract genre, complements Fries’s (2013) method of development framework.

The findings suggest a need for explicit instruction in the thematic resources associated with the

Pedagogical Implications

Conclusion move, particularly Interpersonal Themes (comment and modal adjuncts and the linear Rheme-to-Theme chains by which the value is realized are Conventional and learnable resources among writers’ already known repertoire of genres (Martin & Rose, 2008). Graduate writing programs in Pakistan can be improved by adopting a move-oriented pedagogy which explicates the clause-level thematisation involved at each move, more significantly at Move 5.

Recommendations for Future Research

Future work could investigate a wider range of research article sections to discover whether the same move-thematic progression relationships found in abstracts occur at wider textual levels. Interdisciplinary comparisons to be made to natural sciences, engineering and humanities would allow a clearer idea of the extent to which genre in individual disciplines influences clause-level thematic choices. Analysis of process data (e.g., think-aloud studies, histories of writing process) could help resolve whether thematic patterning in English research abstracts relies on explicit genre knowledge or conventional wisdom. Moves aimed to exploit these move 5 resources need to be developed and evaluated using an experimental design.

Data Transparency Statement

A total of twenty-five research article abstracts produced by Pakistani L1 English authors of Applied Linguistics published in five HEC-recognized journals between 2018 and 2024 form the



present research corpus. 286 clauses contained in these abstracts were systematically coded by the main researcher on the basis of operational definitions developed for the coding guide. Boundaries of moves within the abstracts were identified and noted before clause delineation. The five abstracts selected randomly (26% by abstract count), i.e., 76 clauses, were independently coded by another expert in the relevant field familiar to SFL based system of coding for calculating inter-rater reliability with: Theme type = 0.84 and Theme progressions = 0.79. Coding sheets of disagreements of the two coders and calculation of reliability are available with the corresponding author on reasonable request in the form of a corpus inventory and coding documentation.

References

- Babaii, E., & Behnam, B. (2020). Thematic progression and argumentative structure in academic journal editorials. *Journal of English for Academic Purposes*, 44, Article 100832. <https://doi.org/10.1016/j.jeap.2020.100832>
- Daneš, F. (Ed.). (1974). *Papers on functional sentence perspective*. Academia.
- Eggins, S. (2004). *An introduction to systemic functional linguistics* (2nd ed.). Continuum.
- Fries, P. H. (2013). Theme, method of development, and texts. In L. Fontaine, T. Bartlett, & G. O'Grady (Eds.), *Systemic functional linguistics: Exploring choice* (pp. 333–358). Cambridge University Press.
- Guo, L. (2012). Tracing the history of thematic progression research and its applications in second language writing instruction. *Linguistics and Education*, 23(1), 43–55. <https://doi.org/10.1016/j.linged.2011.09.001>
- Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). Edward Arnold.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.
- Higher Education Commission. (2008). *Criteria for promotion of faculty at HEC recognized degree awarding institutions*. HEC.
- Hyland, K. (1998). *Hedging in scientific research articles*. John Benjamins.
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. Longman.
- Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 34(8), 1091–1112. [https://doi.org/10.1016/S0378-2166\(02\)00035-8](https://doi.org/10.1016/S0378-2166(02)00035-8)
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Lorés, R. (2004). On RA abstracts: From rhetorical structure to thematic organization. *English for Specific Purposes*, 23(3), 280–302. <https://doi.org/10.1016/j.esp.2003.06.001>
- Martin, J. R. (1992). *English text: System and structure*. John Benjamins.
- Martin, J. R., & Rose, D. (2008). *Genre relations: Mapping culture*. Equinox.
- Martín, P. (2003). A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English for Specific Purposes*, 22(1), 25–43. [https://doi.org/10.1016/S0889-4906\(01\)00033-3](https://doi.org/10.1016/S0889-4906(01)00033-3)
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.



ISSN E: 2709-8273
ISSN P: 2709-8265

JOURNAL OF APPLIED
LINGUISTICS AND
TESOL

JOURNAL OF APPLIED LINGUISTICS AND TESOL (JALT)
Vol.9.No.2 2026

Thompson, G. (2014). *Introducing functional grammar* (3rd ed.). Routledge.

Wang, L. (2007). Theme and rheme in the thematic organization of text: Implications for teaching academic writing. *Asian EFL Journal*, 9(1), 164–176.

Xiao, Y., & Cao, Y. (2021). Thematic progression patterns in academic discourse: Cross-disciplinary and cross-cultural perspectives. *English for Specific Purposes*, 62, 1–15. <https://doi.org/10.1016/j.esp.2020.09.001>