

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: najeebashahbaz00@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

Abstracts of research articles are genre-governed texts, which have to simultaneously respond to a number of communicative tasks-orienting the reader, previewing the purpose of the research, methods, results, and establishment of scholarship-within an approximate word-limit of 150–250 words. The clause-level realization of these communicative purposes, despite this textual reductionism, has not been investigated widely in Pakistani academic writings. Using Halliday & Matthiessen's (2014) account of Theme-Rheme, Dane's (1974) four patterns of thematic progression and Hyland's (2000) five-move genre description of RAAbstract, a qualitative-dominant mixed-method study was designed to investigate 25 RA abstracts written by Pakistani scholars of Applied Linguistics and published in 5 HEC-recognized journals during 2018–2024. Inter-rater reliability for clause-by-clause coding of 387 clauses (388 transitions) were established as = 0.84 (Theme type) and = 0.79 (progression patterns). It was observed that the use of unmarked topical Theme is predominant (56.3%). It was also found that thematic progression patterns were not distributed randomly in terms of moves of the genre. Constant Theme is favored in Method moves (59%), linear progression in Product and Conclusion moves (45% and 61%), respectively; split-rheme organizes parallels of findings in Product moves (24%); and derived theme maps multidimensional problem-situation in Introduction moves (20%). More critically, 48% of the abstracts failed to make a Conclusion move and did not have any interpersonal Theme, nor any progression to claim significance; these are identified as problem areas of the corpus texts written by Pakistani scholars of Applied Linguistics who clearly are very much genre-sensitive but show a definite pattern of weakness in their constructive abilities towards making an evaluative move. Corpus and coding data are available from the researcher upon reasonable request.

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

1. Introduction

Research article abstracts are remarkable communicational feats. Crammed into 150-250 words, they must at once orient readers, signal the study's scholarly relevance, detail methodology, present the findings, and establish the contribution of the work within a specific discipline. This communicational task is not only organizational, but inextricably linguistic; in many cases, it is the way that the information is packed and strung together at the clause level that gives an abstract the feel of cohesive academic argument, rather than simple informational listing. This is where SFL comes in; in the systemic functional framework (Halliday 1994, Halliday & Matthiessen 2014), every clause is said to simultaneously realize three types of meaning: ideational, interpersonal, and textual. The textual metafunction is manifested in the

Theme-Rheme structure of the clause, where the Theme is taken to be the element with reference to which the rest of the clause is organized, and the Rheme is where the proposition is unfolded and elaborated. Across consecutive clauses, the patterns that occur among the Themes of each clause-thematic progression (Dane 1974)-are one of the primary means through which writers build argumentational cohesion and authority in specialized discourse.

This research asks how Pakistani researchers tackle this communicative and linguistic feat within their context. English has become the medium of academic publishing, even though most researchers are not native English speakers. Since their promotion is directly linked to their publishing output (especially in refereed international journals, HEC 2008), the genre competence of Pakistani researchers is an important concern of higher education in Pakistan. However, in the little existing genre analysis on Pakistani abstracts, scholars have merely identified the moves of the genre (e.g. Swales, 1990), without attempting to show whether and how the moves are signaled by writers' clause-level Theme choices. This study, therefore, is designed to answer the following three research questions: (RQ1) What types of Themes dominate each of Hyland's (2000) five moves of research article abstracts? (RQ2) What patterns of thematic progression occur, where and for what rhetorical purpose? (RQ3) How do thematic progression patterns and the genre move structure interact to construct the textuality of Pakistani research article abstracts?

This study makes contributions at three levels. The first contribution is at the theoretical level: it applies research on the textual metafunction of SFL to the Pakistani academic context and shows how it may be used to describe the clause-level information organization of writing within a specific genre. Secondly, at the methodological level, it presents the fruitful interaction between clause-level SFL analysis and genre-level move structure analysis. Finally, it offers findings and implications for pedagogy: the data presented here can serve as an empirical basis for move-aware instruction of writing the abstracts of research articles at the graduate level in Pakistani institutions of higher education. The study only examines empirical research article abstracts written in English by Pakistani academics in the field of Applied Linguistics and related social sciences; an examination of natural sciences, engineering, or medical papers is not attempted since there are differences in information structuring between these disciplines and the social sciences/humanities.

1.1 Operational Definitions

Theme: The first constituent(s) of the clause which gives the point of departure of the clause (Halliday & Matthiessen, 2014, p.89).

Rheme: The rest of the clause, which expands the statement made in the theme, adding new information.

Thematic Progression: The sequential pattern of Theme-Rheme choices in clauses, where new information in one Rheme is chosen as the next Rheme.

Genre move: A textual stretch that fulfills a particular discourse purpose within a genre (Hyland, 2000; Swales, 1990).

1.2 Research Questions

1. What are the categories of Theme and their distributions over the abstract moves of Pakistani Applied Linguistics research paper abstracts according to Hyland (2000)?
2. What thematic progression patterns are found in the Pakistani Applied Linguistics research paper abstracts?
3. What are the effects of the categories of theme and thematic progression patterns on the organization and communication of the abstract moves of the Pakistani Applied Linguistics research paper abstracts?

2. Literature Review

2.1 SFL, the Textual Metafunction, and Thematic Progression

The SFL model, where every clause can express simultaneously ideational, interpersonal, and textual meaning (Halliday & Matthiessen, 2014), uses the textual metafunction (Theme-Rheme) to establish grammatical structure upon which the other two meanings can build a cohesive text. The theory of discourse semantics proposed by Martin (1992) using macro-Theme, hyper-Theme, and periodicity provides a macro-level linkage of grammatical structures to the text level. Eggins (2004) and Thompson (2014) note that the initial position in the clause is rhetorically available in terms of genre expertise—an opening in an expert text is more likely to be marked, an opening in a novice text likely to be unmarked (Eggins 2004, p.180). Authorial identity and institutional pressures, examined in terms of hedging and attitude, are also interpersonal aspects of thematic choice as presented by Hyland (1998, 2000, 2002).

Halliday (1994) classifies Themes into ideational (topical themes), discourse continuity (textual themes through conjunctions and conjunctive adjuncts), and attitude (interpersonal themes through comment adjuncts), noting as well that Themes can be marked or unmarked and simple or multiple. Fries (2013) links thematic development with the organization of text units, bridging grammar and genre: thematic choices establish the "method of development" for segments of texts, and provide the rationale for what constitutes, say, a paragraph, a section or a chapter (Fries, 2013). Dane's (1974) model categorizes the flow of information across clause sequences in four ways: Constant Theme, where a topic recurs; Linear Progression, where the topic of the previous Rheme becomes the topic of the next Theme; Split-Rheme, where the second part of a compound Rheme becomes the topic of the next Rheme; and Derived Theme, where successive clauses' themes are all derived from a common hyper-theme. These patterns, and the predominance of each within particular sections of the same genres, has been supported across languages and contexts (Dane, 1974, Wang, 2007, Babaii & Behnam, 2020, Xiao & Cao, 2021). Regarding second language acquisition and EFL, the prevailing trend appears to be subject-as-Theme, and the tendency towards constant-theme patterning within a text (Guo, 2012), suggesting that L2 acquisition dynamics do play some role.

2.2 Genre Analysis of Abstracts and Pakistani Academic Writing

Hyland's (2000) five-move model for research article abstracts (Introduction, Purpose, Method, Product, Conclusion) based on 600 abstracts across eight disciplines has been the most empirically supported and widely applied. In an analysis of 15 non-native English abstracts taken from Spanish linguistics journals, Lors (2004) shows how Product and Conclusion moves are systematically underevaluated; a study of English and Spanish language social science abstracts shows the same pattern, where the Conclusion is the least frequently instantiated move (Martn 2003). These studies show a clear cross-linguistic pattern in the underevaluation of evaluative moves. Studies of Pakistani research article abstracts have similarly shown a near-universality in Purpose and Method moves, while the Conclusion move is often absent, again attributed to institutional settings which promote output over rhetorically complete discourse (see Lors 2004 and Martn 2003 for comparative work). While both clause-level analysis in Pakistani research article abstracts and genre-based research on them have explored respectively SFL thematic progression and move distribution as well as lexico-grammatical aspects, few have attempted to combine SFL coding of thematic progression at the clause-level with genre-analysis of the moves; this study provides an analytical addition in this vein.

3. Methodology

3.1 Research Design and Corpus

This research employs a qualitative-dominant mixed-method design guided by a constructivist-interpretive epistemology; the inquiry is principally a qualitative SFL analysis of thematic and genre-move organization at clause level. This is supplemented by descriptive quantitative methods for generating frequencies and percentages that will be useful in tracing

distributional patterns and in comparative cross-move analysis. The frequencies themselves, however, are rhetorically interpreted within the analytical framework and not as independent quantitative data. Purposive sampling was employed to achieve a corpus representing Pakistani applied linguistics research article abstract writing under present-day HEC publication conditions.

The sampling frame included the abstracts of published articles from five HEC recognized journals: the Journal of Applied Linguistics and Language Research (JALLR), the Pakistan Journal of Language Studies (PJLS), Language in India, the International Journal of English Language and Literature Studies (IJELLS), and the Pakistan Journal of Social Sciences (PJSS). This selection was based on criteria of HEC recognition, consistent application of English Language and applied linguistics research and social sciences reporting, inclusion of at least one author affiliated with an institution in Pakistan, and inclusion of a range of editorial influences. The inclusion criteria for a given research article abstract were that it must (1) report an empirical research study, (2) have at least one author affiliated with an institution in Pakistan, (3) have been published between 2018 and 2024, and (4) be in the field of Applied Linguistics or in an related social science field. Through stratified temporal sampling one abstract from each year-interval in 2018-2024, five abstracts were selected from each journal.

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)
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	387 (M = 15.5; range 11–19)
Clause transitions analyzed	388

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

3.2 Analytical Procedure

The analysis was carried out in four distinct phases. First, each abstract was broken down into Hyland's (2000) five moves based on criteria of function (Introduction, where a problem or theoretical context was presented; Purpose, where a clear intention was stated; Method, where the materials or analytical techniques used were revealed; Product, where the outcomes were stated; Conclusion, where an interpretation of findings or suggestions were given), with move boundary classifications noted prior to the initial clause segmentation in order to minimize potential impact upon thematic interpretation. Second, every finite and non-finite clause with a recoverable subject was established as a thematic unit following the Halliday and Matthiessen (2014) framework with embedded relative clauses as noun group modifiers treated as part of the modification and non-finite initial adjuncts included for their regular thematic function. Third, every clause was coded for Theme type (topical, textual,

interpersonal); marked/unmarked status; and simple/multiple status (where textual Themes were subcategorized into additive, adversative, causal, and concessive). Fourth, the relationships between successive clauses were categorized according to Dane's (1974) schema of Constant Theme, Linear Progression, Split-Rheme, Derived Theme or Mixed/Indeterminate, with data categorized according to the type of clause transition observed between a clause and its subsequent clause rather than based on the themes of each individual clause, resulting in the categorization of 388 transitions across 387 coded clauses including cross move transitions. Finally, summaries were derived through descriptive statistics: column percentages were calculated for each theme type category and for each type of thematic progression pattern across the number of defined move boundaries for each thematic feature, rounded to one decimal place.

3.3 Treatment of Overlapping Move Realizations

The abstract register is also more likely to result in sentences that can serve more than one communicative purpose in a given clause. Any clauses which appeared to fulfill two functions were coded to whichever of the two functions appeared to be primary based on three ranked criteria: (a) the core illocutionary force; (b) sentence placement within the abstract; and (c) clear lexicogrammatical markers. All judgments of dual-function sentences were recorded in the coding record and noted to the second analyst during reliability coding. Conflicting dual-function judgments ($n = 6$) were made with recourse to the coding manual and coded conservatively. The Mixed/Indeterminate progression category ($n = 16$; 4% of all transitions) is intended to capture transitions at move boundaries where logical progression moves from one move to another and are principled boundary-transition records rather than uncertain Theme codings.

3.4 Reliability and Validity

A structured coding manual (with operational definitions and examples worked out) was created before the analysis began. A second SFL-trained analyst was used to code a random 20% (5 abstracts, 76 clauses) independently. Kappa Cohen values were $= 0.84$ for type of theme and $= 0.79$ for progression patterns which represent almost perfect and substantial agreement, respectively, according to Landis and Koch (1977). Disagreements were resolved by referring to the coding manual, and borderline examples were documented and conservatively analyzed. Construct validity was insured by ensuring close conformity to the theoretical taxonomies outlined in Halliday and Matthiessen (2014) and Dane (1974). The full range of the publicly available corpus documents was used; there were no human participants involved.

4. Findings

4.1 Genre Move Distribution

The distribution of Hyland's (2000) five moves across the 25 corpus abstracts is illustrated in Table 2. The higher proportion of both Move 2 (Purpose, $n=25$) and Move 3 (Method, $n=24$) compared to other moves indicates a general preference for detailing procedure: indeed, 96% (24/25) and 100% (25/25) of the abstracts contain the second and third moves respectively. A far smaller proportion of abstracts include Move 1 (Introduction, $n=17$) (68% or 17/25), while only half contain the final move of the sequence (Move 5, Conclusion, $n=13$) (52% or 13/25). This systematic imbalance between the amount of information devoted to method and that to argument shows that fully half the abstracts fail to offer an argument. The exclusion of Move 5 eliminates the space where significance is presented, a fact with corresponding clause-level correlations considered below (4.3, 4.4).

Table
Genre Move Distribution Across 25 Corpus Abstracts

2

Move	Communicative Function	Freq. (N = 25)	%
M1: Introduction	Problem statement; theoretical context	17	68%
M2: Purpose	Research aim; objective; question	25	100%
M3: Method	Data; procedure; sampling	24	96%
M4: Product	Findings; results presentation	22	88%
M5: Conclusion	Significance; implication; future direction	13	52%

Note. Move 5 is absent in 12 of 25 abstracts (48%).

4.2 Theme-Rheme Analysis: Coded Extract

Table 3 shows the clause-level Theme-Rheme coding of the Corpus Abstract 07. It is one of the 25 texts making up the dataset, chosen as an illustrative abstract on two grounds: firstly, it is representative of the dataset by encompassing all five Hyland (2000) genre moves, and secondly, by demonstrating the complete array of thematic types and thematic progression used in all the texts in the corpus. The code provided in Table 3 is generated by the same four-step procedure applied to all 25 abstracts, and the main corpus-based evidence for the results presented in sections 4.3 and 4.4 is the complete set of 387 coded clauses.

Table

3

Clause-level Theme–Rheme Coding of Corpus Abstract 07 (authentic corpus extract)

Cl.	Theme	Rheme	Theme Type	Move
1	Academic writing in second language contexts	has been identified as a site of sustained pedagogical concern.	Topical, Unmarked, Simple	M1
2	In Pakistan specifically,	EFL writers face significant challenges in conforming to international genre conventions.	Topical, Marked (Circ. Adj.)	M1
3	This study	investigates the discourse structure of research article abstracts by Pakistani academics.	Topical, Unmarked, Simple	M2
4	Drawing on Halliday and Matthiessen's SFL framework,	the study employs thematic analysis to examine clause-level information organization.	Marked (Non-finite), Multiple	M3
5	The corpus	consists of 25 abstracts from HEC-recognized journals (2018–2024).	Topical, Unmarked, Simple	M3

Cl.	Theme	Rheme	Theme Type	Move
6	Data	were coded for Theme type and progression using Daneš's taxonomy.	Topical, Unmarked, Simple	M3
7	The analysis	reveals that unmarked topical Themes predominate across all moves.	Topical, Unmarked, Simple	M4
8	Furthermore,	Constant Theme characterizes Method moves; Linear Progression marks Product and Conclusion moves.	Multiple (Textual + Topical)	M4
9	These findings	suggest that thematic choices in Pakistani abstracts are genre-sensitive.	Topical, Unmarked, Simple	M5
10	Importantly,	the study contributes to SFL-informed academic writing pedagogy in Pakistan.	Multiple (Interpersonal + Topical)	M5

Note. Corpus Abstract 07 is drawn from the corpus and coded using the procedure described in Section 3.2. Theme identification follows Halliday and Matthiessen (2014). Move assignment follows Hyland (2000). All quantitative findings in Sections 4.3 and 4.4 derive from the complete 387-clause coded dataset. Cl. = Clause; M = Move.

Corpus Abstract 07 contains a number of trends that characterize the larger corpus. Move 1 shows a typical progression from a general thematic subject in the following clause 'In Pakistan specifically' (Clause 2) to a marked circumstantial theme-moving from the more generalized thematic category to the marked circumstantial thematic category to narrow down in two clauses, according to convention at the boundary between the first two moves. The following clause in the Introduction moves (Clause 4) carries a non-finite participial thematic subject ('Drawing on Halliday and Matthiessen's SFL framework') which indicates theory before description, a common feature of marked themes that are used at boundary points, and which, as Fries (2013) points out, is often used to signal transitions in discourse organization. Moves 4-6, comprising the Method move, illustrates a Constant theme pattern through a reliance on co-referential antecedents such as 'the study', 'the corpus', 'data', which continues the focus on description. Move 5 contains what could be called an interpersonal Theme in Clause 10 'Importantly'-a marked theme, the only one of this type, and one that appears only in Move 5 and not at all in the 12 abstracts which lack Move 5.

4.3 Theme Type Frequency

Table

4

Theme Type Distribution Across the Corpus (N = 387 Clauses)

Theme Type	N	%	Primary Move Context
Topical, Unmarked, Simple	218	56.3%	All moves (esp. M2, M3, M4)
Multiple: Textual — additive	54	13.9%	M3, M4
Multiple: Textual — adversative	9	2.3%	M1, M4
Multiple: Textual — causal	7	1.8%	M4, M5
Multiple: Textual — concessive	4	1.0%	M1
Multiple: Textual (all subtypes)	74	19.1%	All moves
Marked: Circumstantial Adjunct	38	9.8%	M1, M3 boundaries
Multiple: Interpersonal + Topical	31	8.0%	M5 exclusively
Marked: Non-finite Participial	19	4.9%	M3 openings
Interpersonal (standalone)	7	1.9%	M5 exclusively
Total	387	100%	—

Note. Interpersonal Themes—standalone (n=7) and as part of Multiple: Interpersonal + Topical (n=31)—appear exclusively in Move 5. Total interpersonal Theme instances: 38. Dominant are simple unmarked thematic Themes (56.3%; n=218). Pakistani writers are tendency to start with the research entity ('the study,' 'the data,' 'the analysis') as the topical starting point of the article. This trend, which seems to be a common one for Pakistani writers in this data set, is similar to the findings in the research written by Chinese and in other EFL writing traditions (Guo, 2012), suggesting a more general tendency toward stable topics across EFL contexts. Textual multiple themes(19.1%) mostly consist of additive connectors and adjuncts('Furthermore,' 'In addition,') (73.0%; n=54 out of 74 textual multiple themes), while other forms are scarce(adversative themes n=9; 2.3%; causal themes n=7; 1.8%; concessive themes n=4; 1.0%), showing clear pattern of underuse of logical connectors that signaling argumentation, which is also evidence from a vast corpus of EFL academic writing studies (Guo, 2012; Lors, 2004). Marked themes (14.7% in total) occur mostly at move boundaries. Non-finite participial themes are mostly foundat the opening of Method moves, and this form demonstrates clear sign of genre-conscious thematic choice. Interpersonal themes are found only in Move 5. None of the abstracts without Move 5 include an interpersonal theme.

4.4 Thematic Progression Analysis

Table 5 shows the distribution of Dane's (1974) four patterns of progression across Hyland's (2000) five moves. This non-random distribution forms the main analytical result of

the investigation: a theoretically readable relationship between genre move and predominant pattern of progression.

Table

5

Thematic Progression Pattern Distribution by Abstract Move (N = 388 Clause Transitions)

Pattern	M1 Intro n = 55	M2 Purpose n = 38	M3 Method n = 104	M4 Product n = 130	M5 Concl. n = 61	Total (%)
Constant Theme	14 (25%)	18 (47%)	61 (59%)	22 (17%)	9 (15%)	124 (32%)
Linear Progression	19 (35%)	11 (29%)	23 (22%)	58 (45%)	37 (61%)	148 (38%)
Split-Rheme	8 (15%)	4 (11%)	9 (9%)	31 (24%)	7 (11%)	59 (15%)
Derived Theme	11 (20%)	3 (8%)	5 (5%)	14 (11%)	8 (13%)	41 (11%)
Mixed/Indet.	3 (5%)	2 (5%)	6 (6%)	5 (4%)	0 (0%)	16 (4%)
Column Total	55	38	104	130	61	388

Note. Column percentages are calculated within each move. Totals may not sum to 100% due to rounding. Indet. = Indeterminate.

4.4.1 Constant Theme in Method Moves: Procedural Cohesion

Constant Theme is accounted for by 59% (n=61) of the Method move transitions, more than any other single pattern in any move category. C4-6 (Table 3) of Corpus Abstract 07 provides a good example:

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-ref.]

C6: [Data]Theme [were coded for Theme type and progression using Dane's taxonomy.] Rheme [M3 - Constant: process anchor]

The thematic chain ('the study' 'the corpus' 'data') remains process-oriented via co-referential anchors while each Rheme adds to a unique step in the methodology. As 'the study' shifts to 'the corpus', and finally to 'data', the research becomes more objectified as researcher-entity yields to data-entity, a display of the social science practice of methodological authority via discursive distance. The ubiquity of this pattern in the Method section of English linguistic articles, identified by Wang (2007), proves to operate in the confined frame of the abstract-evidence of a genre-induced phenomenon functioning at different textual scales.

4.4.2 Linear Progression in Product and Conclusion Moves: Argumentative Momentum

Linear Progression prevails within Conclusion moves (61%; n = 37) and occurs as second most prominent feature in Product moves (45%; n = 58). The following extract (Corpus Abstract 07, Clauses 7-10) demonstrates the Rheme to Theme chaining across the Product-Conclusion interface:

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 move.]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 Theme of C9]

C10: [Importantly,] Theme [the study contributes to SFL-informed academic writing pedagogy in Pakistan.] Rheme [M5 - Interpersonal Theme: significance claim]

The chain from finding (C7-C8) to interpretation (C9) to significance claim (C10) creates the textual representation of scholarly contribution-reader are led through the process of argument rather than asked to determine the relevance of claims for themselves. The total lack of chaining on this interface of abstracts which lack Move 5 directly accounts for the inability to claim significance at the clausal level, adding a grammatical account to Loris's (2004) structural analysis.

4.4.3 Split-Rheme in Product Moves and Derived Theme in Introduction Moves

Split-Rheme occurs with 24% of transitions to Product moves ($n = 31$)-the largest percentage of any move type. Split-Rheme distributes a compound Rheme between the Themes of following clauses, thus presenting related information in parallel in a streamlined way rather than with separate clause-introduction words (a helpful resource for an efficient genre such as the abstract). Derived Theme is found with 20% of transitions to Introduction moves ($n = 11$)-the largest percentage of any move type. The multiple, sequential, derived Themes create multi-dimensional representations of the problem; they are nested under a macro-Theme that names the research area. The macro-Theme names the research domain, and the individual themes specify particular problem areas (e.g., gaps in literature, flaws in prior methods, the particular problem analyzed). The shift from macro to micro topic identification is the clause-level articulation of Swales's (1990) CARS (Create a Research Space) model.

5. Discussion

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

The move-conditioned, non-random pattern of development noted above offers an empirical detail to the SFL theoretical claim that genre and lexico-grammar are mutually implicative (Halliday & Matthiessen, 2014; Martin, 1992). Information structure, dictated by the demands of the five-move, argument-oriented abstract structure (Hyland, 2000) must be created via authorial thematic choice in the clause: Method moves must offer process information, and they are developed with co-referential Constant Themes. Conclusion moves must offer argument; they are developed through linear Rheme-to-Theme chaining if they occur. These are the genre-driven, functional equivalences suggested by Fries's (2013) observation that thematic choice is the means of development in a text segment; this is illustrated here not simply across extended discourse, but within the genre constraints of the compressed abstract.

The prominence of unmarked, topical Theme choice (56.3%) among Pakistani writers parallels similar trends noted in Chinese and other EFL writing traditions (Guo, 2012), suggesting a move-level, cross-contextual, SLA-driven trend toward clause organization centered on the subject of the clause. In similar fashion, the prevalence at the move level is telling: the rate of absent Conclusion moves observed here (48%) mimics the observed underrealization of evaluative moves in non-native European EFL abstracts (Lors, 2004; Martn, 2003) and mirrors the overall tendency toward privileging the procedural over the evaluative in Pakistani academic writing contexts. While previous research noted these observations, the present study offers the grammatical evidence for the omission: non-occurrence of Conclusion moves, complete deletion of interpersonal Themes, and the complete lack of Rheme-to-Theme linear meaning-claiming development. Here, we see not just the missing move, but also the complete absence of the clause-level grammatical system-evaluative Themes and R-T argumentative chains-through which the move's rhetorical work would have been accomplished; this is the primary contribution of this study to the literature.

5.2 The Move 5 Deficit: Analytical and Pedagogical Dimensions

The omission of Move 5 from 48% of abstracts is structural, grammatical, and contextual all at once. It erases, in terms of structure, the communicative space required for significance-claiming. In terms of grammar, it eradicates the inter-personal Theme chains and lineal chains that perform the actual significance-claiming. It performs this in terms of structure by creating a pattern, found in Spanish linguistic and social science abstracts (Lors, 2004; Martn, 2003) as well, of omitted evaluative moves in favor of a procedural focus in non-native EFL abstracts. This thesis provides the clause-level component in this genre record-in absence of Move 5 its thematic structures disappear also. The exclusive presence of interpersonal Themes in abstracts containing Move 5 indicates genre competence in this writing genre. Writers include inter-personal Themes when constructing the Move 5, but are prevented by structural features when the Move is omitted. These writers do not lack entire genre understanding but the actual construction of one component. All EFL academic writing traditions share, however, the primary coherence mechanism of topical coherence and constant themes (Guo, 2012) which can become globally unconstructive in absence of logical structure at higher levels. The writer's implicit knowledge of the genre is augmented if the clause level relationalities of Move 5-the inter-personal and logical-relational themes of the actual significance-claiming-are demonstrated. These are conventional; thus, the student writer is trained not in an entirely novel genre but in one within their repertoire by means of genre-pedagogy (Martin & Rose, 2008).

6. Limitations

The corpus of 25 abstracts taken from 5 different journals is comparable in size to those used in similar qualitative-interpretive SFL and genre-analytic investigations (Lors, 2004; Martn, 2003; Wang, 2007), though it is too small for reliable generalizations in a statistical sense. The observed patterns should be considered the general characteristics of the corpus of Pakistani Applied Linguistics abstracts examined in this research, not those of Pakistani Applied Linguistics writing more broadly. Before making claims of generalizability and interpreting the frequency of distributions as representative, a larger and more stratified corpus that spans more institutional contexts and time would need to be studied. The research is also focused narrowly on Applied Linguistics and closely related social sciences – thematic organizations in natural sciences, engineering, or the humanities may be entirely different (Hyland, 2000) – and only on the abstract genre, so claims regarding organization patterns in other research article section can not be made based on the results presented here. Even with high inter-rater reliability ($\kappa = 0.84$; $\kappa = 0.79$), the nature of thematic analysis requires subjective decisions, especially regarding move boundaries and situations in which the material may belong to two different Dane categories. While the Mixed/Indeterminate category ($n=16$; 4%) accounts for some subjectivity, subtle, boundary case decisions elsewhere may be the source of unrecorded interpretive subjectivity. Lastly, this research is solely on product and not on writers' intentions, drafting history, or readers' reactions, and the question of whether these patterns reflect conscious choice or implicit knowledge of the genre cannot be answered solely through analysis of the product.

7. Conclusion

This article has shown that it is not by chance that thematic progression patterns occur in Pakistani applied linguistics research article abstracts but that they are genre-constituted. It also offers, in an under-researched domain of Pakistani abstract-level SFL analysis, an empirical explanation of the clause-level thematic correlates of move-level genre patterns that have previously been described in Pakistani academic writing. The procedural coherence of Method moves is foregrounded by Constant Theme; the argumentative progression of Product

and Conclusion moves is underpinned by Linear Progression; product statements are organized by Split-Rheme; and multiparameter problem-situations are depicted in Introductions by Derived Theme. The statistical validation across 387 clauses and 25 authors and 5 journals of these correlations' non-random character supports them as systemic genre features, not idiolectal traits.

The strongest empirical finding is the complete absence (48% of the corpus) of the move as well as its constituent interpersonal Themes and linear, significance-claiming progression, which confirms at the clause level a structural deficit that has previously been identified at the move level (Lors, 2004; Martn, 2003), and suggests a grammatical consequence thereof: by skipping the Conclusion, writers abandon the clause-level thematic mechanisms by which academic significance is constructed. The pedagogical implication of these findings is a need for explicit instruction of the thematic resources, especially logical-relational and interpersonal themes, available for signaling significance in the Conclusion move, in graduate programs in Pakistan, which writers have already begun to utilize at the clause level. Further research should be extended to other sections of research articles, a comparison be made with other cultures, process data be used, and pedagogical interventions be evaluated.

Data Transparency Statement

This research used a corpus of 25 abstracts of research articles written by Pakistani academics working in the field of Applied Linguistics; they were published in five HEC recognised journals during the years 2004-2008. All the 387 clauses were coded by the principal researcher according to the operational definitions outlined in a systematic coding book. The moves were determined and recorded before dividing the clauses up. Five randomly selected abstracts (76 clauses; 20% of the corpus) were independently coded by a second SFL trained analyst to determine inter-rater reliability: = 0.84 for the Type of theme and = 0.79 for the progression patterns. All the coding worksheets, disagreement records, and reliability calculations are stored by the principal researcher. An inventory of the corpus and coding documents can be obtained on reasonable request.

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.
- Eggs, 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.
- 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 organisation. *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.
- 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>