https://doi.org/10.17323/jle.2025.27389

Stylistic Redundancy and Wordiness in Introductions of Original Empirical Studies: Rhetorical Risks of Academic Writing

Elena Tikhonova [1], **Olga Zavolskaya** [2], **Nataliia Mekeko** [3]

[1] MGIMO University, Moscow, Russia

[2] Sergo Ordzhonikidze Russian State University for Geological Prospecting, Moscow, Russia

[3] RUDN University, Moscow, Russia

ABSTRACT

Background: The introduction of a research article plays a central role in shaping scientific argumentation. However, this section is often especially prone to stylistic overload, which can obscure the clarity of the author's position. While the issues of redundancy and wordiness have been broadly acknowledged in applied linguistics, there is still limited understanding of how these features are distributed in relation to rhetorical structure, particularly within Russian-language academic texts.

Purpose: To identify rhetorically sensitive areas of stylistic overload in the introductions of Russian-language research articles in the field of education.

Method: The analysis is based on a corpus of 40 introductions from empirical articles published in 2024 in leading Russian peer-reviewed journals in education. The rhetorical Move-Step model developed by Swales was used as the framework for annotation. Each fragment was manually coded for two dimensions: the type of deviation (wordiness or redundancy) and its communicative impact (according to the IMPACT scale). Pearson's chi-squared test was used to assess statistical significance.

Results: Stylistic overload was found to cluster in specific rhetorical steps, especially those related to establishing the importance of the topic (M1_S2), identifying gaps in the literature (M2_S1), and stating research objectives (M3_S2). The most frequent features included syntactic overcomplexity, vague abstract nouns, and overused credibility markers. A high level of negative communicative impact (IMPACT = HIGH) was observed in 60 fragments, most of which were located in the mentioned segments. Statistical testing (χ^2, $p < 0.0001$) confirmed a significant relationship between rhetorical function and the type of deviation.

Conclusion: The results confirm that stylistic overload in introductions is not accidental but structurally motivated. This supports the need for rhetorically informed strategies in teaching academic writing. The annotation scheme developed in the study may be applied in future corpus-based analyses of academic Russian.

KEYWORDS

academic writing, wordiness, redundancy, rhetorical analysis, CARS model, Russian academic texts

Citation: Tikhonova, E., Zavolskaya, O., & Mekeko, N. (2025). Stylistic Redundancy and wordiness in introductions of original empirical studies: Rhetorical risks of academic writing. *Journal of Language and Education, 11*(2), 125-136. https://doi.org/10.17323/jle.2025.27389

Correspondence:
Elena Tikhonova,
tikhonova.e.v@inno.mgimo.ru

Received: December 13, 2024
Accepted: June 10, 2025
Published: June 30, 2025

INTRODUCTION

Issues of stylistic clarity and textual economy remain central to current research on academic writing (Swales, 1990; Hyland, 2005; Tikhonova & Mezentseva, 2024). Despite the growing body of methodological guidance on academic style, many authors continue to struggle with expressing their ideas efficiently. This often results in the use of redundant structures and vague formulations that blur meaning (Flowerdew & Forest, 2015; Leufkens, 2023). Among the most common stylistic problems are *wordiness* and *redundancy*, both of which undermine analytical precision and textual clarity while increasing the cognitive load for

readers (Tikhonova, Mezentseva & Kasatkin, 2024; Strunk & White, 2000; Demir, 2019).

Stylistic wordiness refers to the use of low-informational expressions, epistemic modifiers, and nominalizations that reduce semantic precision and clarity (Tikhonova & Mezentseva, 2024; Salager-Meyer, 1994). In contrast, redundancy encompasses lexical tautologies, syntactic repetition, and appositive elaborations that overload the sentence and obscure the core message (Williams & Bizup, 2017; Kravtchenko & Demberg, 2022; Tikhonova et al., 2024). While such stylistic deviations have been examined in numerous international studies (e.g., Hyland, 2005; Biber & Gray, 2010; Fruehwald, 2010; Goonaratna, 2002a; 2002b; Alontseva & Ermoshin, 2019), their specific manifestations and rhetorical effects in Russian-language academic texts remain underexplored. Of particular interest is the structural distribution of wordiness and redundancy across rhetorical segments of academic writing.

Studies indicate that the opening segments of an Introduction, particularly those associated with establishing the research territory (Move 1), are especially vulnerable to stylistic overload. Writers often resort to sweeping generalizations, repetitive thematic statements, and formulaic clichés, which diminish the cognitive density of the text and hinder the reader's focus on a specific research problem (Samraj, 2002; Kanoksilapatham, 2005; Gong & Barlow, 2022). The second rhetorical move (Move 2), aimed at identifying a research gap by highlighting limitations, contradictions, or unresolved issues in previous studies, proves no less susceptible. In an effort to balance critique with scholarly politeness, authors tend to rely on mitigating strategies—explanatory repetitions, vague phrasings, and abstract formulations. These choices often compromise analytical clarity and contribute to verbal redundancy. Corpus-based studies confirm this tendency: in academic texts from the social sciences and education, indirect criticism and cautious evaluative language are particularly prevalent in segments aligned with Move 2 (Wahyuningtyas & Wulandari, 2023; Aoulad-Ouda & Chellaoui, 2023). Finally, the third rhetorical move (Move 3), which presents the aim and scope of the research, is also affected by stylistic standardization. As noted by Hyland (2008), this stage is often marked by template-like declarations and predictable syntactic patterns that frequently echo earlier content, thereby weakening textual originality and cognitive depth. In sum, all three rhetorical moves of the Introduction exhibit varying degrees of vulnerability to stylistic and structural redundancy, with the second move emerging as the most rhetorically strained.

However, these observations are typically based on English-language data and have yet to be empirically confirmed on a corpus of texts rooted in the Russian academic tradition. The overwhelming majority of Russian academic journals continue to use Russian as the primary medium of scholarly communication, despite increasing pressure from the global English-dominated publication landscape. This persistence is shaped not only by national science policy but also by the linguistic challenges associated with translating research from Russian into English - especially due to semantic and structural differences between the two languages (Raitskaya & Tikhonova, 2020; Smirnova et al., 2021). As a result, Russian-language academic texts have developed their own stylistic conventions, embedding distinctive rhetorical and genre-based practices (Raitskaya & Tikhonova, 2020). Among these, features such as redundancy and verbal padding have become entrenched, particularly in introductory sections. This points to a significant gap in our understanding of rhetorical and stylistic risks specific to Russian academic writing, particularly in applied disciplines such as education. The present study seeks to address this gap by conducting a quantitative, corpus-based analysis grounded in the CARS rhetorical model (Swales, 1990) and recent classifications of stylistic deviation (Tikhonova et al., 2024; Tikhonova & Mezentseva, 2024).

The purpose of the present study is to identify patterns in the distribution and communicative impact of stylistic deviations, namely wordiness and redundancy. The following research questions are posed:

RQ1: What are the frequency characteristics and rhetorical localization of wordiness and redundancy in the Introduction sections of Russian-language empirical research articles in education?
RQ2: Which rhetorical moves and steps in the Introduction are most prone to stylistic deviations?
RQ3: What impact do these stylistic deviations have on the communicative clarity of the text?

The study tests the following hypotheses:

H1: The highest concentration of wordiness and redundancy occurs in rhetorical segments related to justifying the significance of the research topic (M1_S2) and identifying gaps in existing literature (M2_S2).
H2: While wordiness occurs more frequently than redundancy, the latter more often results in a higher level of negative communicative impact.
H3: Stylistic deviations are structurally motivated and shaped by genre-based expectations.

METHOD

Corpus

The material for analysis consisted of Introduction sections from 40 (forty) original research articles published in 2024 in Russian peer-reviewed journals ranked between positions 1 and 18 in the Russian Science Citation Index (RSCI) for the field of Education and Pedagogical Sciences. All articles were written in Russian by authors affiliated with Russian

universities and research institutions. The choice of a Russian-language corpus reflects the study's objective to examine stylistic practices of academic writing in their authentic, unadapted form, as they occur within the national scientific context.

Inclusion Criteria

The selection of materials was based on the following formal and substantive criteria:

(1) Bibliometric status of the journal: inclusion in the top ten RSCI journals in the field of Education and Pedagogical Sciences;
(2) High categorical ranking: classification of the journal in the first or second category according to the internal Science Index system;
(3) Presence of a standalone Introduction: structurally distinct and separate from other sections such as literature review or methodology;
(4) Genre and rhetorical organization: presence of rhetorical moves that allow segmentation of the Introduction using the Move-Step model;
(5) Topical relevance: alignment of the article's content with the fields of pedagogy, teaching methodology, educational technologies, or educational psychology;
(6) Open access: ensuring legal and ethical transparency of the analysis.

Corpus Profile

Each article was entered into an analytical database that included the following parameters:

1. Journal title;
2. Journal category;
3. Year of publication;
4. Author team;
5. Full article title;
6. Journal name;
7. Open access availability

The entire corpus was organized into a working spreadsheet (Appendix 1), allowing for the tracking of bibliometric and contextual characteristics of each analytical unit. The corpus consists of 40 Introductions from original empirical studies published in journals such as *Voprosy obrazovaniya (Education Issues)*, *Vysshee obrazovanie v Rossii (Higher Education in Russia)*, *Obrazovanie i nauka (Education and Science)*, *Psikhologicheskaya nauka i obrazovanie (Psychological Science and Education)*, *Yazyk i kultura (Language and Culture)*, *Rusistika (Russian Studies)*, *Integratsiya obrazovaniya (Integration of Education)*, *Sovremennye problemy nauki i obrazovaniya (Contemporary Problems of Science and Education)*, *Perspektivy nauki i obrazovaniya (Perspectives of Science and Education)*, *Psikhologo-pedagogicheskie issledovaniya (Psychological and Pedagogical Research)*, and *Vestnik RUDN: Series Psychology and Pedagogy*.

This corpus (Appendix 2) provides a solid basis for a reliable and representative analysis of current academic writing practices in educational science and related disciplines within the context of contemporary Russian scholarly communication.

Rhetorical Structure Annotation

To annotate the rhetorical segments of introductions, the Move-Step model developed by Swales (1990; 2004) was used (Table 1).

The rhetorical annotation based on the Move-Step model was necessary for functionally aligning segments that contain stylistic deviations and for conducting statistical analysis of their distribution in relation to rhetorical function.

Classification Scheme for Wordiness and Redundancy and Annotation Procedure

To annotate stylistically overloaded elements in the corpus of introductions, the typology of wordiness developed by Tikhonova and Mezentseva (2024) and the classification of textual redundancy proposed by Tikhonova, Mezentseva, and Kasatkin (2024) were used. Both models were adapted for operationalization in manual annotation and implemented as a system of tags assigned to each segment of text showing signs of stylistic deviation.

The categories of wordiness and their corresponding tags include:

1. General phrases and low-information introductory expressions (WORDINESS_GENERAL);
2. Excessive use of epistemic modifiers (WORDINESS_HEDGING);
3. Syntactic overload and structural complexity (WORDINESS_COMPLEXITY);
4. Empty or abstract references without specific content (EMPTY_REFERENCE);
5. Nominalization that reduces verbal dynamism (NOMINALIZATION);
6. Formulaic phrases that serve no analytical function (FORMULAIC_PHRASE).

The categories of textual redundancy and their tags include:

1. Lexical tautology and semantic repetition (REDUNDANCY_LEXICAL);
2. Redundant syntactic structures (REDUNDANCY_STRUCTURE);
3. Appositive explanations that duplicate content (APPOSITIVE_PHRASE).

Table 1

Rhetorical Structure of the Introduction Section According to Swales (1990; 2004)

Move	Step	Rhetorical Function
Move 1. Establishing a Research Territory	Step 1. Topic Presentation	Introducing the subject area and situating the topic
	Step 2. Justifying the Importance of the Topic	Arguing for the relevance and significance of the research problem
	Step 3. Reviewing Key Studies	Referring to prior research and describing the current state of the field
Move 2. Establishing a Niche	Step 1. Indicating a Gap in Knowledge	Highlighting areas that remain underexplored
	Step 2. Identifying Conflicts or Contradictions	Emphasizing inconsistent findings or theoretical disagreements
	Step 3. Pointing to Methodological or Practical Limitations	Critiquing existing approaches or identifying applied challenges
Move 3. Presenting the Present Work	Step 1. Stating the Aim and Objectives of the Study	Defining the research goal, object, and focus
	Step 2. Formulating Research Questions or Hypotheses	Outlining key research questions or testable propositions
	Step 3. Providing a Brief Methodological Description	Summarizing the approach, methods, and dataset

In cases where a single fragment demonstrated multiple types of stylistic deviation, each tag was recorded as an individual unit of observation. This approach ensured comprehensive typological coverage.

Although redundancy is often considered a subset of wordiness, the two were distinguished analytically in this study. Wordiness is defined as stylistic elements that complicate textual processing due to formal heaviness, such as clichés, nominalizations, or epistemic modifiers. Redundancy, by contrast, is understood as structural or semantic repetition of information already stated.

This distinction allowed for a more precise evaluation of the nature of stylistic deviations and their distribution across the rhetorical segments of the introduction. It also increased the sensitivity of interpreting the level of communicative impact (tagged as IMPACT), which reflects the degree to which a deviation affects the clarity and analytical transparency of the text (Table 2).

This scale served as the main evaluative matrix for each annotated fragment in the corpus. The impact level was assigned through expert interpretation, considering the rhetorical Move and Step as well as the combination of formal and rhetorical markers observed during annotation.

The annotation was conducted manually using Microsoft Excel. For each fragment, the following information was recorded: a unique ID; article number, Move and Step; the text of the fragment; the category of wordiness or redundancy; IMPACT level; and a brief explanatory comment (Appendix 2).

Data Analysis

The quantitative analysis of the annotated corpus aimed to identify stable patterns in the distribution of stylistically overloaded fragments (instances of wordiness and redundancy) depending on the rhetorical structure of the introduction section. The analysis focused on three main relationships: (1) the frequency and type of stylistic deviations were examined across rhetorical Moves and Steps, (2) dominant categories of wordiness and redundancy were identified within specific functional segments, (3) the level of negative communicative impact (IMPACT) was assessed in relation to the rhetorical role of each fragment.

To verify the observed patterns, contingency tables were constructed, which confirmed the presence of statistically significant associations between the rhetorical function of a fragment (Move or Step) and the type of stylistic deviation it contained. These findings support the conclusion that stylistic overload in academic writing is not a random occurrence. Rather, it is structurally conditioned and closely aligned with rhetorical tension and the communicative demands of specific segments within the introduction.

Statistical Data Processing

To confirm the observed patterns in the distribution of stylistic deviations across the rhetorical structure of introductions, a statistical analysis was conducted based on the association of categorical variables. The goal was to determine whether there were significant dependencies between three variables: the rhetorical function of the fragment (Move and

Table 2

Criteria for Evaluating the Impact of a Fragment on Textual Clarity and Analytical Precision (IMPACT)

Level	Label	Description	Basis for Classification
Low	LOW	The stylistic element has minimal effect on text clarity. It is acceptable within genre norms and does not require revision.	A stylistic deviation is present (such as a cliché or tautology), but it does not disrupt logical flow, hinder interpretation, or require rephrasing.
Medium	MEDIUM	Slows reading and comprehension and makes it harder to understand the author's position or the logic of transitions.	The fragment increases cognitive load, reduces precision of expression, introduces unnecessary links in the argument, or creates semantic ambiguity.
High	HIGH	Significantly distorts or impedes understanding of the research aim, central claim, or analytic logic.	The deviation breaks the coherence between parts of the text, forces rereading, may lead to misinterpretation, and undermines the academic credibility of the statement.

Step), the type of deviation (CATEGORY), and its level of communicative impact (IMPACT).

At the first stage, contingency tables were constructed for the following pairs of variables:

(1) Move and Step by CATEGORY (rhetorical position by type of deviation);
(2) Move and Step by IMPACT (rhetorical position by level of impact);
(3) CATEGORY by IMPACT (type of deviation by level of impact).

In all cases, the Pearson chi-square test was used to check the hypothesis of independence between variables. The calculations were carried out using standard statistical tools in Python (SciPy) and Microsoft Excel. The significance level was set at $p < 0.05$.

Additionally, to clarify the nature of the impact within rhetorical segments, the proportion of fragments marked with IMPACT = HIGH was calculated for each step of the Move–Step model. This helped to identify parts of the text that are especially vulnerable to communicative overload or distortion and to justify these segments as priority targets for revision. Manual verification of the annotated tables ensured accurate alignment of each fragment with its rhetorical function.

RESULTS

General Description of the Corpus

A total of 487 text fragments were identified and analyzed, each containing features that deviated from stylistic norms. These fragments were classified either as wordiness or as textual redundancy. The annotated corpus therefore consisted of 487 analytical units, with each unit labeled according to the type of deviation, its rhetorical position based on Move and Step, and its level of communicative impact as measured by the IMPACT scale.

Out of the total number of fragments, 342 were categorized as wordiness. This refers to cases where scientific writing becomes less precise and less informative due to the use of vague or redundant language. These include common rhetorical clichés, "empty" introductory phrases, excessive modifiers, abstract or poorly specified formulations, and the replacement of verbal structures with nominalizations. Nominalizations slow down meaning processing and obscure the main point. These elements do not contribute meaningful content but increase the length of the text and reduce its clarity.

A total of 145 fragments were classified as redundancy. These are cases of structural or semantic repetition, where information is duplicated either lexically or syntactically. As a result, wordiness appears to dominate not only in quantity but also in variety. This reflects a widespread tendency toward vague expression and verbal expansion in academic writing, particularly in introductions. The detailed distributions are presented in Tables 3 and 4.

Table 3

Categories of Wordiness Identified in the Corpus (n = 342)

Category of Wordiness	Frequency	Percentage (%)
WORDINESS_GENERAL	84	24.6
NOMINALIZATION	81	23.7
WORDINESS_COMPLEXITY	62	18.1
FORMULAIC_PHRASE	46	13.4
EMPTY_REFERENCE	44	12.9
WORDINESS_HEDGING	25	7.3

Table 4

Categories of Redundancy Identified in the Corpus (n = 145)

Category of Redundancy	Frequency	Percentage (%)
REDUNDANCY_STRUCTURE	73	50.3
REDUNDANCY_LEXICAL	62	42.8
APPOSITIVE_PHRASE	10	6.9

The distribution of communicative impact levels revealed the following pattern. A total of 220 fragments were classified as having a low level of impact (LOW). These do not distort the perception of scientific content, although they diverge from the expected genre conventions. Another 207 fragments were assessed as having a medium level of impact (MEDIUM), suggesting that they hinder the interpretation of the author's reasoning or the logic of transitions. The most critical group includes 60 fragments with a high level of impact (HIGH), the vast majority of which are associated with wordiness. These are primarily linked to empty nominal expressions, syntactic overload, and unsupported abstract statements (Table 5).

Thus, even at the initial stage of analysis, a high structural density of stylistic deviations can be observed, affecting more than 90 percent of all introductions in the corpus. These findings indicate that stylistic overload in scientific writing is not random but reflects a systemic pattern rooted in the genre conventions of academic discourse.

Dominant Forms of Stylistic Overload

Unlike the previous section, which described the overall set of fragments containing stylistic deviations, this part focuses on the qualitative distribution of deviation categories within the two broader groups: wordiness and redundancy. The purpose of the analysis is to identify the most frequent and representative types of stylistic overload and to assess their contribution to the overall picture of academic writing distortion.

The results show that within the wordiness group, three categories have the greatest weight: general discourse clichés and introductory phrases (24.6 % of all wordiness cases), nominalization (23.7 %), and syntactic complexity (18.1 %). These three categories form the core of wordiness in the corpus, together accounting for nearly 67 % of all relevant fragments. Their predominance indicates a tendency among authors toward generalized and overly formalized statements as well as syntactically complex constructions that reduce textual clarity. Less frequent but rhetorically significant categories include formulaic expressions (13.5 %) and vague abstract references (12.9 %), which confirms the multilayered nature of wordiness as a stylistic phenomenon.

In the group of textual redundancy, structural redundancy emerges as the most prominent category, accounting for more than half of all identified fragments (50.3 %). This type of deviation involves the inclusion of unnecessary syntactic elements that repeat information already expressed. Nearly as frequent is lexical tautology (42.8 %), which points to a common tendency to restate the same idea through repetitive wording and pleonastic constructions. Appositive phrases that duplicate content are less common (6.9 %), but they still contribute significantly to cognitive noise by undermining the structural compactness of the sentence.

Taken together, the quantitative distribution of categories reveals a clear internal hierarchy among the forms of stylistic deviation. The dominant violations are those that directly obstruct access to the semantic structure of the introduction. These include formulaic expressions that add no informational value and constructions that overload the grammatical framework of the text. These findings provide a foundation for further analysis of which rhetorical segments of the introduction are most prone to such forms of stylistic overload and how they may distort the interpretability of the text.

Rhetorical Localization of Stylistic Deviations

One of the central objectives of this study was to identify how stylistic deviations are distributed across the rhetorical structure of the introduction. To achieve this, annotated fragments were mapped onto the steps of the Move–Step

Table 5

Distribution of Wordiness and Redundancy Cases by Communicative Impact Level (IMPACT)

CATEGORY	HIGH	LOW	MEDIUM	Total	Share of HIGH (%)
WORDINESS_COMPLEXITY	23	2	37	62	37.0
REDUNDANCY_STRUCTURE	13	28	32	73	17.8
NOMINALIZATION	11	25	45	81	13.5
REDUNDANCY_LEXICAL	6	22	34	62	9.6
EMPTY_REFERENCE	5	21	18	44	11.3
APPOSITIVE_PHRASE	1	4	5	10	10.0
WORDINESS_GENERAL	1	59	24	84	1.1
FORMULAIC_PHRASE	0	43	3	46	0.0
WORDINESS_HEDGING	0	16	9	25	0.0

model proposed by Swales (1990, 2004), which segments the introduction into functionally distinct parts ranging from the presentation of the research territory to the statement of goals and objectives.

The results of the analysis (Table 6) support the hypothesis that stylistic overload is structurally conditioned. The highest concentration of deviations occurs in those segments of the introduction where authors experience the greatest rhetorical pressure: where they are expected either to demonstrate subject-matter awareness or to formulate the novelty of the research. These include, above all, Step M1_S2 (justifying the relevance of the topic), Step M1_S3 (reviewing previous research), and Step M3_S1 (stating the aim of the study).

In Step M1_S2, a total of 125 stylistically deviant fragments were identified, making it the most overloaded rhetorical segment in the corpus. This part of the introduction is particularly marked by lexical redundancy, generic statements of significance, and nominalizations, all of which reflect an attempt to strengthen the argument through intensified language. Such overuse often results in verbosity and repetitive evaluative expressions that fail to provide any substantive clarification of the problem. Step M1_S3 also exhibits a high density of stylistic issues, with 106 annotated fragments. This segment frequently includes vague generalizations, formulaic phrases, and empty references that are not supported by analytical interpretation. These features suggest a lack of critical engagement with the literature and a focus on formally fulfilling the genre's requirements. Although Step M3_S1 contains fewer annotated fragments (61), it shows a disproportionately high share of deviations rated as having medium or high communicative impact. In this segment, nominalizations are particularly prevalent, leading to overly complex and imprecise formulations of the research aim. Additionally, syntactic overloading is common, which hampers the reader's ability to follow the logic of the study.

Steps M1_S1 (thematic introduction) and M2_S2–M2_S3 (niche description) were less saturated with stylistic deviations, although they still contain problematic fragments. These are primarily related to formulaic expressions and structural redundancy, particularly when articulating methodological limitations or discrepancies in previous approaches.

The distribution of stylistic violations across the Move–Step model clearly reveals a concentration in the functionally significant sections of the introduction. This indicates that linguistic overload should be viewed not only as a stylistic issue but also as a rhetorical phenomenon shaped by the communicative challenges authors face when addressing the core demands of the genre. Further analysis of the communicative impact levels of these deviations will help identify which types most severely hinder interpretation and therefore require prioritized revision.

Levels of Communicative Impact

The overall structure of the corpus reveals that the vast majority of fragments with stylistic deviations fall into either the low (45.2 percent) or medium (42.5 percent) impact categories. Only 12.3 percent of the annotated fragments, or 60 units, were classified as having a high level of communicative impact (IMPACT = HIGH) as shown in Table 5. However, the significance of this 12 percent should not be underestimated. These fragments represent critical areas of the text where the logic of the argument or the meaning of key statements is distorted.

Particular attention should be given to the category of syntactic complexity (WORDINESS_COMPLEXITY), which displays the highest proportion of critically disruptive cases. Within this category, 37.1 % of the fragments were rated as having a high impact. This finding suggests that overly long and structurally overloaded sentences are especially harm-

Table 6

Rhetorical Distribution of Stylistic Deviations in Introduction Sections

Rhetorical Step	APPOSITIVE_PHRASE	EMPTY_REFERENCE	FORMULAIC_PHRASE	NOMINALIZATION	REDUNDANCY_LEXICAL	REDUNDANCY_STRUCTURE	WORDINESS_COMPLEXITY	WORDINESS_GENERAL	WORDINESS_HEDGING	Total
M1_S1	2	2	3	4	9	4	6	26	0	56
M1_S2	0	11	10	16	19	15	23	27	4	125
M1_S3	4	15	15	21	11	17	11	9	3	106
M2_S1	0	5	6	8	5	6	9	4	4	47
M2_S2	2	1	1	2	3	14	1	0	6	30
M2_S3	1	2	0	7	6	5	2	7	4	34
M3_S1	1	4	9	17	7	10	7	6	0	61
M3_S2	0	0	0	1	1	0	2	2	0	6
M3_S3	0	4	2	5	1	2	1	3	4	22

ful to the clarity of academic writing, particularly in rhetorically dense segments of the introduction.

The next most prominent categories are structural redundancy (17.8 %) and nominalization (13.6 %. In the case of structural redundancy, the primary communicative distortion arises from repeated content, which blurs the intended emphasis and compromises the conciseness of the exposition. Nominalization, by contrast, leads to the absence of verbal dynamism, which causes sentences to lose their specificity and precision, making the logical structure of the text less distinct.

A less pronounced yet still noteworthy effect is observed in the categories of empty references (11.4 % and appositive phrases (10 %). These types of deviations suggest a partial loss of conceptual clarity or the inclusion of excessive explanatory detail that lacks analytical purpose.

Taken together, these findings indicate that a high level of communicative disruption is usually associated not with superficial stylistic variety, but with deeper structural flaws in the text. These include disrupted syntactic flow, misalignment between grammatical form and logical function, and a lack of referential precision. Such observations make it possible to identify high-priority areas for editorial revision, which will be discussed in more detail in the following subsection.

Statistical Validation of Rhetorically Motivated Redundancy

A statistical analysis was conducted to examine the relationship between the rhetorical structure of the introduction, the types of stylistic deviations, and the degree of their communicative impact. The aim was to confirm that the patterns identified earlier are not random but demonstrate a consistent and structurally driven nature. For this purpose, contingency tables were used along with Pearson's chi-squared test, which makes it possible to assess the statistical significance of associations between categorical variables (Table 7).

Table 7

Chi-Squared Analysis of Contingency Tables for Three Pairs of Variables

Variable Pair	χ^2	df	p-value
MOVE_STEP × CATEGORY	159.45	64	< 0.0001
MOVE_STEP × IMPACT	36.84	16	0.0022
CATEGORY × IMPACT	145.84	16	< 0.0001

The analysis of the first contingency table (MOVE_STEP × CATEGORY) revealed a statistically significant relationship between rhetorical positioning and the type of stylistic overload ($\chi^2 = 159.45$; df = 64; $p < 0.0001$). This means that deviations from stylistic norms are not evenly distributed throughout the Introduction. Certain types of overload are typical of specific rhetorical steps. For example, nominalization and syntactic complexity are most frequently found in the formulation of the study's purpose (M3_S1) and hypotheses (M3_S2). In contrast, formulaic expressions and tautologies tend to cluster in steps M1_S2 and M1_S3, where the author needs to justify the topic's importance and provide a literature review.

The second table (MOVE_STEP × IMPACT) also demonstrated a statistically significant association ($\chi^2 = 36.84$; df = 16; $p = 0.0022$), confirming that not only the type of stylistic deviation but also its degree of impact on the text depends on its rhetorical position. In other words, not all segments of the Introduction are equally vulnerable to communicative distortions. For instance, the formulation of research questions and the description of the research niche (M3_S2, M2_S1) more often contain fragments with a high level of impact, suggesting that these rhetorical tasks present consistent challenges for authors.

The strongest association was found between the type of deviation and its communicative impact (CATEGORY × IMPACT), with the chi-square test yielding $\chi^2 = 145.84$ at 16 degrees of freedom ($p < 0.0001$). This result confirms that not all stylistic deviations have the same effect on how the text is perceived. For example, syntactic overload accounts for the highest share of critically impactful fragments, whereas formal clichés and idiomatic phrases more often fall into the category of stylistically undesirable but permissible recurrences.

The statistical analysis not only supports the earlier observations but also reinforces them with analytical precision. The results indicate that stylistic overload in academic writing follows a predictable and rhetorically driven pattern. These patterns of linguistic redundancy can be taken into account both in the editorial assessment of manuscripts and in the development of academic writing instruction.

High-Risk Zones: Priority Segments for Editing

One of the key outcomes of this study was the identification of rhetorically vulnerable segments within the introduction: parts of the text where linguistic overload is most pronounced and results in communicative distortion. Based on the proportion of fragments with a high impact level (IMPACT = HIGH) across different rhetorical steps of the Swales (1990) model, it is possible to pinpoint stable high-risk zones that require targeted editing.

The highest proportion of high-impact fragments was observed in step M3_S2 (the formulation of research questions or hypotheses), where 33.3% of all fragments were classified as critically overloaded. Although this step was represented by a relatively small number of units in the corpus, it

showed the greatest concentration of rhetorical strain. This suggests that formulating research assumptions within the constraints of academic genre norms presents a significant challenge. Authors often attempt to enhance the perceived significance of their hypotheses through convoluted formulations, relying on nominalization, vague definitions, and unnecessary elaborations.

A high proportion of fragments with IMPACT = HIGH was also found in step M2_S1 (identifying the research gap, 23.4 %) and in step M3_S3 (describing the methodology, 18.2 %). This indicates a systematic difficulty in transitioning from the literature review to the formulation of one's own research position. In step M2_S1, authors often attempt to emphasize novelty through rhetorical intensification, which may involve tautological repetition, syntactic overcomplication, and excessive explanatory phrases. In step M3_S3, difficulties arise from the challenge of articulating the methodology briefly and clearly; when linguistic precision is lacking, this part often becomes overloaded and poorly structured.

These patterns are further supported by the results of the statistical analysis. The contingency table MOVE_STEP × IMPACT revealed a statistically significant association between rhetorical function and impact level ($\chi^2 = 36.84$; $p = 0.0022$). This provides empirical confirmation for the idea that stylistic overload is not distributed randomly across the introduction but is concentrated in segments where rhetorical tension is especially high. In practical terms, this means that stylistic issues are most likely to emerge in the moments when the author must either establish a research niche, justify methodological choices, or create a smooth transition into the main body of the article. Consequently, steps M3_S2, M2_S1, M3_S3, and M1_S3 can be described as rhetorically charged zones, where the risk of communicative distortion is highest. These parts of the text require special attention during editing and can serve as focal points in academic writing instruction aimed at teaching strategic reduction and structural refinement.

DISCUSSION

The results of this study demonstrate that stylistic overload in academic introductions is neither a random nor evenly distributed phenomenon. On the contrary, the observed forms of wordiness and textual redundancy show clear rhetorical localization, corresponding to functionally demanding segments of the introduction's structure. Specifically, the highest concentration of stylistic pressure occurs in parts of the text where the author is expected to simultaneously establish the relevance of the topic, legitimize the research position, and outline the methodological framework.

These findings provide partial confirmation of Hypothesis 1, which proposed that the highest concentration of stylistic overload would occur in segments related to justifying the topic's significance (M1_S2) and identifying research gaps (M2_S2). The data support the first part of this assumption: Step M1_S2 indeed demonstrated the greatest rhetorical load, with a high frequency of both general formulations and nominalizations aimed at reinforcing argumentative weight. However, the expected prominence of Step M2_S2 was not confirmed. Although it contained several stylistic deviations, its overall density and communicative impact were notably lower than in other segments, suggesting that authors may allocate less rhetorical effort to articulating research gaps than to legitimizing the importance of their work.

The concentration of stylistic overload in these key rhetorical steps suggests that redundancy in academic writing may serve a defensive purpose. As Gong and Barlow point out, when presenting the novelty of their research, authors often use repetitive or softening expressions and rely on nominalizations. These choices help them avoid directly challenging previous studies while adhering to conventions of academic politeness. Such strategies are especially common when researchers work under pressure to publish. In those situations, texts often become saturated with surface markers of scientific credibility, such as complex syntax, cautious statements, and standardized rhetorical phrases. Although these features are intended to meet formal expectations, they can reduce the transparency of the writing and make it harder for readers to follow the line of reasoning. This has been noted by several authors, including Biesta and colleagues, who argue that the pursuit of academic recognition can come at the cost of clarity and accessibility.

The concentration of stylistic overload in key rhetorical steps of the Introduction suggests that redundancy in academic writing often serves a protective function. As Gong and Barlow (2022) point out, when formulating the novelty of their research, authors frequently resort to repetitive or mitigating constructions as well as to nominalizations in an effort to avoid direct confrontation with existing findings and to conform to the norms of academic politeness. Such strategies are particularly common in situations where authors operate under pressure to increase their publication output (Çakir et al., 2024). In these cases, the density of a text with formal markers of scientific discourse, such as complex syntax, epistemic caution, and standard rhetorical formulas, is perceived as a necessary condition for academic recognition, even when it reduces cognitive transparency and makes the argumentation harder to follow (Biesta et al., 2024).

The identified predominance of high-impact deviations in step M3_S2, which is responsible for formulating hypotheses, research questions, and objectives, requires special attention. This segment, as shown by the corpus analysis, is marked by the highest density of syntactically overloaded and declaratively redundant constructions that hinder the clear expression of the research aim. Such a concentration of overload points to rhetorical pressure that emerges when the author must simultaneously present the novelty

of the study and comply with established genre and institutional expectations. These observations are consistent with the findings of Alramadan (2020), who demonstrated in a corpus-based study of introductions in applied linguistics articles that the stage of presenting research contribution carries significant rhetorical weight and is often realized through formulaic or nominalized constructions, which reduce the cognitive clarity of the text. Therefore, step M3_S2 can be regarded as a key zone of communicative tension, where linguistic redundancy functions both as a means of institutional positioning and as a potential threat to the clarity of scientific argumentation.

With regard to the second hypothesis, which suggested that redundancy, although less frequent than wordiness, tends to result in more serious communicative distortions, the data call for a more nuanced interpretation. Wordiness was indeed more widespread in the corpus, particularly in the form of vague generalizations, hedging expressions and nominalizations. However, redundancy, which is also a manifestation of textual wordiness, was disproportionately associated with high-impact fragments. For example, structural redundancy accounted for 17.8 percent of all high-impact cases, which is noticeably higher than the share observed for most subcategories of wordiness. This indicates that redundancy presents a greater risk to the clarity and coherence of scientific argumentation, thus providing partial confirmation of the second hypothesis.

Linguistic overload in the analyzed corpus reflects not only stylistic but also cognitive factors. According to the cognitive model of writing developed by Flower and Hayes (1981), redundant structures interfere with the reader's ability to construct a stable mental representation of the text and increase cognitive load. This effect is particularly evident in the case of nominalizations and complex syntactic constructions, which make it more difficult to access the core meaning of the text (Graesser et al., 2003; Tikhonova et al., 2024b). The reader experiences greater cognitive strain because more informational links are required to maintain coherence and must be held in working memory.

It is also important to highlight the institutional dimension of the observed deviations. As van Dijk (2008) notes, stereotyped language patterns, formulaic phrases, and repetitive constructions serve not only to connect or emphasize ideas but also to signal the writer's affiliation with the academic community. From this perspective, linguistic overload can be seen as a reflection of genre-based socialization. By reproducing formal patterns, the writer internalizes the conventions of academic writing, even when such choices reduce the informational density of the text.

These findings also provide empirical support for the third hypothesis (H3), which posited that stylistic deviations are structurally motivated and shaped by genre-based expectations. The distribution of wordiness and redundancy across distinct rhetorical steps, along with the significantly higher impact observed in functionally loaded zones such as M3_S2 and M1_S2, confirms that linguistic overload arises not simply from stylistic carelessness but as a response to rhetorical demands embedded in academic writing conventions. Rather than being distributed evenly, deviations cluster around communicatively sensitive parts of the introduction, where authors face the task of justifying their research, identifying a niche, and articulating novelty. This suggests that redundancy operates as a genre-induced strategy for managing rhetorical pressure.

Thus, the discussion of results points to the dual nature of linguistic overload in academic introductions. On the one hand, it reflects a response to the functional and rhetorical challenges of the text. On the other hand, it results from institutionalized genre expectations. Taken together, these factors call for a shift among researchers and educators from prescriptive stylistics to a functional-rhetorical diagnostic approach. Within this framework, redundancy should be understood not simply as a deviation from an ideal linguistic norm, but as a symptom of rhetorical instability.

CONCLUSION

This study aimed to identify the structural patterns of linguistic overload in the introductions of academic articles in the field of education, with a focus on rhetorical positioning, the nature of stylistic deviations, and their communicative impact. The findings demonstrated that textual redundancy in academic writing is not a random deviation from the norm. Rather, it arises in response to rhetorical pressure that emerges in functionally important parts of the text, such as justifying relevance, framing the research gap, and formulating research objectives.

An analysis of frequency distributions, together with the results of chi-square tests, confirmed statistically significant relationships between the type of stylistic deviation, its rhetorical placement, and the level of communicative impact. This supports the conclusion that wordiness and redundancy are not merely stylistic excesses, but systematic rhetorical strategies used by authors in segments of high cognitive and communicative tension. Particularly vulnerable were Step M1_S2 (justification of relevance), M2_S1 (statement of the research gap), and M3_S2 (formulation of hypotheses), where a high concentration of fragments was found to potentially distort the clarity of meaning.

The practical value of this study lies in the development of a functionally oriented approach to diagnosing and editing academic texts. By identifying rhetorically vulnerable segments, it becomes possible to design targeted instructional strategies for teaching academic writing. These strategies are not aimed at eliminating redundancy as such but rather at optimizing the form of expressing scientific meaning in

alignment with the communicative goals of each section of the introduction.

At the same time, the interpretation of the results must take into account several limitations. First, the study corpus is restricted to a single discipline (education) and includes only Russian-language texts. This limits the generalizability of the findings to other academic cultures. Second, although the rhetorical annotation procedure was systematic, the assessment of communicative impact (IMPACT) involved a degree of expert interpretation and therefore requires further inter-rater validation.

Despite these limitations, the study provides empirical support for the concept of rhetorically motivated linguistic overload and lays the groundwork for further exploration of the relationship between the structure of scholarly discourse and communicative clarity. Promising directions for future research include (1) cross-disciplinary comparisons of rhetorical redundancy patterns, (2) analysis of linguistic overload in English-language publications by Russian authors, and (3) the development of algorithmic tools for automatically identifying rhetorically induced redundancy in support of scholarly writing and editorial workflows.

DECLARATION OF COMPETING INTEREST

None declared.

AUTHORS' CONTRIBUTION

Elena Tikhonova: conceptualization; data curation; formal analysis; methodology; project administration; visualization; writing – original draft; writing – review & editing.

Olga Zavolskaya: formal analysis; investigation; writing – original draft.

Nataliia Mekeko: data curation; visualization; writing – original draft.

REFERENCES

Alontseva, N. V., & Ermoshin, Y. A. (2019). The problem of language redundancy on the example of a scientific text. *RUDN Journal of Language Studies, Semiotics and Semantics, 10*(1), 129-140. https://doi.org/10.22363/2313-2299-2019-10-1-129-140

Alramadan, M. (2020). Stance and engagement: A corpus-based comparison of university students' and published writers' research article introductions in applied linguistics. *Journal of English for Academic Purposes, 44*, 100837. https://doi.org/10.1016/j.jeap.2020.100837

Aoulad-Ouda, M., & Chellaoui, S. (2023). *Hedging and politeness in Moroccan university students' academic writing: A case study of literature reviews*. *Arab World English Journal, 14*(2), 183–197. https://doi.org/10.24093/awej/vol14no2.13

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes, 9*(1), 2–20. https://doi.org/10.1016/j.jeap.2010.01.001

Biesta, G., Takayama, K., Kettle, M., & Heimans, S. (2024). How 'academic' should academic writing be? Or: why form should follow function. Asia-Pacific Journal of Teacher Education, 52(2), 121–125. https://doi.org/10.1080/1359866X.2024.2324582

Çakir, A., Kuyurtar, D., & Balyer, A. (2024). The effects of the publish or perish culture on publications in the field of educational administration in Türkiye. *Social Sciences and Humanities Open, 9*, 100817. https://doi.org/10.1016/j.ssaho.2024.100817

Demir, C. (2019). Writing Intelligible English Prose: Conciseness vs. Verbosity. *Söylem Filoloji Dergisi, 4*(2), 482-505. https://doi.org/10.29110/soylemdergi.617184

Flower, L., & Hayes, J. R. (1981). *A cognitive process theory of writing*. *College Composition and Communication, 32*(4), 365–387. https://doi.org/10.2307/356600

Flowerdew, J., & Forest, R. W. (2015). *Signalling nouns in English: A corpus-based discourse approach.* Cambridge University Press. https://doi.org/10.1017/CBO9781139135405

Fruehwald, E. S. (2010). Exercises for legal writers II: Editing for wordiness. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.1704045

Gong, N., & Barlow, J. (2022). Rhetorical practices and disciplinary identity in novice academic writers: Evidence from corpus-based analysis. *Journal of English for Academic Purposes, 59*, 101148. https://doi.org/10.1016/j.jeap.2022.101148

Gong, X., & Barlow, L. (2022). A corpus-based analysis of research article macrostructure patterns in high-impact journals. *Theory and Practice in Language Studies, 12*(6), 1067–1075. https://doi.org/10.17507/tpls.1206.12

Goonaratna, C. (2002a). Writing well (6) Wordiness, alias verbosity. *Ceylon Medical Journal, 47*(1), 1-3. https://doi.org/10.4038/cmj.v47i1.6393

Goonaratna, C. (2002b). Writing well (7) Wordiness alias verbosity, continued. *Ceylon Medical Journal, 47*(3), 79-80. https://doi.org/10.4038/cmj.v47i3.3432

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text. In A. P. Sweet & C. E. Snow (Eds.), *Rethinking Reading Comprehension* (pp. 82–98). Guilford Publications.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing.* Continuum.

Hyland, K. (2008). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics, 18*(1), 41–62. https://doi.org/10.1111/j.1473-4192.2008.00178.x

Hyland, K. (2008). Genre and academic writing in the disciplines. *Language Teaching, 41*, 543-562. https://doi.org/10.1017/S0261444808005235

Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes, 24*(3), 269–292. https://doi.org/10.1016/j.esp.2004.08.003

Kravtchenko, E., & Demberg, V. (2022). Informationally redundant utterances elicit pragmatic inferences. *Cognition, 225*, 105159. https://doi.org/10.1016/j.cognition.2022.105159

Leufkens, S. (2023). Measuring redundancy: The relation between concord and complexity. *Linguistics Vanguard, 9*(s1), 95-106. https://doi.org/10.1515/lingvan-2020-0143

Raitskaya L.K., & Tikhonova E.V. (2019). Multilingualism in Russian journals: A controversy of approaches. *European Science Editing, 45*(2), 41. https://doi.org/10.20316/ESE.2019.45.18024

Raitskaya, L., & Tikhonova, E. (2020). Overcoming cultural barriers to scholarly communication in international peer-reviewed journals. *Journal of Language and Education*, *6*(2), 4-8. https://doi.org/10.17323/jle.2020.11043

Salager-Meyer, F. (1994). Hedges and textual communicative function in medical English written discourse. *English for Specific Purposes, 13*(2), 149–170. https://doi.org/10.1016/0889-4906(94)90013-2

Samraj, B. (2002). Introductions in research articles: Variations across disciplines. *English for Specific Purposes, 21*(1), 1–17. https://doi.org/10.1016/S0889-4906(00)00023-5

Smirnova, N. V., Lillis, T., & Hultgren, A. K. (2021). English and/or Russian medium publications? A case study exploring academic research writing in contemporary Russian academia. *Journal of English for Academic Purposes, 53*, Article. 101015. https://doi.org/10.1016/j.jeap.2021.101015

Strunk, W., & White, E. B. (2000). *The elements of style* (4th ed.). Pearson Education.

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

Swales, J. M. (2004). *Research genres: Explorations and applications.* Cambridge University Press.

Tikhonova E.V., Kosycheva M.A., & Mezentseva D.A. (2024b). Ineffective strategies in scientific communication: textual wordiness vs. clarity of thought in thesis Conclusion section. *Integration of Education, 28*(2), 249-265. https://doi.org/10.15507/1991-9468.115.028.202402.249-265

Tikhonova, E. V., & Mezentseva, D. A. (2024). Wordiness in academic writing: A systematic scoping review. *Research Result. Theoretical and Applied Linguistics, 10*(1), 133–157. https://doi.org/10.18413/2313-8912-2024-10-1-0-8

Tikhonova, E., Mezentseva, D., & Kasatkin, P. (2024a). Text redundancy in academic writing: A scoping review. *Journal of Language and Education, 10*(3), 128–160. https://doi.org/10.17323/jle.2024.23747

van Dijk, T. A. (2008). *Discourse and Power*. Palgrave Macmillan. https://doi.org/10.1007/978-1-137-07299-3

Wahyuningtyas, L., & Wulandari, F. (2023). A comparative study of rhetorical structures in research article introductions: Social sciences versus education disciplines. *Journal of Language and Linguistic Studies, 19*(1), 333–348.

Williams, J. M., & Bizup, J. (2017). *Style: Lessons in clarity and grace* (12th ed.). Pearson Education.