Journal of Language Research
ISSN: 2602-4578

*Research Article*

# Cohesion-First Rhetorical Architectures: A Corpus Analysis of Turkish EFL Metadiscourse

Fatih Ünal Bozdağ[1], Osmaniye Korkut Ata University, Department of English Translation and Interpreting, fatihbozdag@osmaniye.edu.tr

**Recommended citation:** Bozdağ, F. Ü. (2025). Cohesion-First Rhetorical Architectures: A Corpus Analysis of Turkish EFL Metadiscourse. *Journal of Language Research (JLR), 9*(2), 66-87.
DOI: https://doi.org/10.51726/jlr.1744953

**Abstract:** This corpus-based study examines how Turkish EFL learners deploy metadiscourse relative to native undergraduates, addressing an under-documented learner population and moving beyond raw counts to network, length-sensitive, and writer-level analyses. Using a hybrid rule-based/ML detector across eight categories, we analyze 285 Turkish essays from TICLE (~202k words) and 171 native essays from LOCNESS (~148k words). Learners prioritize organizational resources: transitions (2.90 vs. 1.95 per 1,000 words; +49%), code glosses (2.27 vs. 1.55; +46%), frame markers (1.36 vs. 0.92; +48%), and especially engagement markers (1.08 vs. 0.33; ≈3.3× native). Native writers invest more in stance: hedges (1.69 vs. 0.92; ≈1.8× learner), boosters (0.50 vs. 0.27; ≈1.9×), and self-mentions (0.57 vs. 0.34; ≈1.7×). Co-occurrence networks reveal a cohesion-centered hub for learners, with stance devices peripheral, versus a stance-centered hub for natives that integrates hedging, emphasis, and authorial presence. Length-sensitive analyses show learners scale cohesion/engagement as texts lengthen, whereas natives scale stance; writer-level entropy is slightly higher in TICLE (M=1.54 bits) with a narrower spread, indicating more uniform reliance on cohesion-first architectures. These results characterize a robust "cohesion-first" profile for Turkish EFL writing and a "stance-integrated" profile for native writing, clarifying where instructional attention can support progression toward expert academic prose.

**Keywords:** *metadiscourse, corpus linguistics, Turkish EFL, academic writing, rhetorical development*

## INTRODUCTION

The ability to guide readers through academic arguments while positioning oneself appropriately as a scholar remains one of the most challenging aspects of second language (L2) academic writing. This challenge centers on metadiscourse—the linguistic apparatus writers use to organize discourse, project stance, and engage readers (Hyland, 2005). Hyland (2017) defines metadiscourse as "self-reflective linguistic expressions referring to the evolving text, to the writer, and to the imagined readers of that text" (p. 17). In practical terms, metadiscourse is language about the discourse itself: rather than contributing new propositional content, it shapes how that content is processed and interpreted. Consider the phrase "However, it seems that these results may suggest a modest improvement," which simultaneously signals contrast (however) while hedging the claim (it seems, may suggest), demonstrating how metadiscourse operates on multiple rhetorical levels.

Metadiscourse serves two complementary but distinct rhetorical functions. Organizationally, it guides readers through the text structure, ensuring coherence and easing comprehension through devices such as transitions (however, therefore) and frame markers (first, in conclusion). Interpersonally, it reveals the author's stance, calibrates commitment, and encourages reader involvement through hedges (perhaps, might), boosters (clearly, indeed), and engagement markers (consider, note that). When deployed effectively, these resources transform dense academic prose into reader-friendly argumentation by clarifying logical links, signaling appropriate caution or confidence,

[1] *ORCID: https://orcid.org/0000-0002-9959-4704*
*Submitted: 17.07.2025*
*Accepted:28.11.2025*

and fostering author-audience rapport (Jomaa & Alia, 2019). For L2 writers, mastering this dual functionality proves particularly challenging because success depends not only on linguistic command but also on internalization of discipline-specific rhetorical conventions.

This challenge is especially acute for Turkish EFL writers, whose L1 rhetorical traditions differ markedly from English academic norms. When these expectations diverge, direct transfer from Turkish discourse patterns can produce miscalibrated stance and compromised textual cohesion. Hedges such as perhaps, might, and it seems often lack exact functional equivalents in Turkish; consequently, learners may underuse them or deploy them in non-target-like ways, projecting unintended certainty or inappropriate timidity (Ardhianti et al., 2023). Such patterns can make academic prose appear either overly categorical or insufficiently confident, potentially compromising not only clarity and persuasiveness but also academic credibility and publication success.

Extensive corpus research has documented these difficulties with increasing precision across diverse L2 populations. Studies consistently show that L2 writers differ systematically from native speakers in both the range and frequency of metadiscourse resources, typically favoring explicit cohesive devices while underusing interpersonal stance markers (Hyland & Milton, 1997; Hinkel, 2005). While this research has primarily examined East Asian learners, Turkish EFL writers present a compelling case for investigation given their distinct linguistic heritage and growing presence in international academia. Furthermore, existing studies have relied heavily on frequency-based analyses, leaving unexplored the functional relationships between metadiscourse categories and their development patterns across different rhetorical contexts.

Building on this foundation, the present study extends metadiscourse research in two important directions. First, we provide the first systematic corpus-based analysis of Turkish EFL learners' metadiscourse patterns, comparing eight functional categories across the Turkish International Corpus of Learner English (TICLE) and the Louvain Corpus of Native English Essays (LOCNESS). Second, we move beyond simple frequency counts to examine co-occurrence patterns, length effects, and writer-level diversity, offering a multidimensional view of rhetorical development. Using automated detection across these benchmark corpora, our analysis aims to: (a) identify systematic differences between Turkish learners and native speakers in metadiscourse deployment, (b) pinpoint specific rhetorical areas where Turkish learners would benefit from targeted pedagogical intervention, and (c) contribute to theoretical understanding of L2 metadiscourse by documenting culture-specific patterns that extend beyond previously studied populations. Our findings reveal that Turkish learners create "cohesion-first" textual architectures that prioritize explicit reader guidance and structural signaling, while systematically underusing the stance markers that characterize mature academic prose—a pattern with clear implications for advanced EFL instruction.

Grounded in prior work on learner stance and cohesion (Hyland & Milton, 1997; Hinkel, 2005; Ädel, 2006), we formulate testable expectations that guide the analysis. Relative to LOCNESS, we expect Turkish learners to show higher normalized rates for interactive categories (transitions, frame markers, code glosses, engagement markers) and lower rates for interactional categories (hedges, boosters, self-mentions); to exhibit cohesion-centric co-occurrence networks in which connective resources anchor the local hub; and to scale metadiscourse with length by adding connective and reader-address devices rather than stance signals. We further anticipate a modestly higher median writer-level entropy in TICLE—reflecting even deployment of a narrower repertoire—paired with a shorter high tail than LOCNESS, where some native writers diversify markedly across categories. By combining frequency profiles, network topology, length-sensitivity, and writer-level diversity on matched argumentative essays, the study provides an integrated account of how Turkish learners assemble metadiscourse in practice and where those assemblies diverge from native usage, extending comparative findings beyond previously examined L2 populations (Akbaş, 2014; Kirişçi & Duruk, 2022; Mur-Dueñas, 2011).

.



## THEORETICAL FRAMEWORK

### Defining and Classifying Metadiscourse Markers

Metadiscourse refers to the linguistic layer through which writers step outside propositional content to comment on, organize, and interpret that content for readers. Hyland and Tse (2004) characterize it as the moment "when the writer intrudes into the discourse to supply a commentary on that discourse" (p. 161). Rather than extending the knowledge base about a topic, metadiscourse signals how surrounding information should be processed—clarifying connections, foregrounding attitude, or managing readership expectations.

The field has converged on Hyland's interpersonal model (2005, 2017), which distinguishes between interactive and interactional resources. Interactive resources manage information flow and textual coherence through five categories: transitions express semantic relations (*however, therefore*), frame markers announce structure (*first, in conclusion*), endophoric markers refer to other text locations (*as noted above*), evidentials attribute propositions to sources (*according to Smith, 2020*), and code glosses rephrase or illustrate concepts (*for example, in other words*). Interactional resources position the author and involve the reader through hedges that soften claims (*might, perhaps*), boosters that express commitment (*clearly, indeed*), attitude markers that encode evaluation (*unfortunately, surprisingly*), engagement markers that address readers directly (*consider, note that*), and self-mentions that foreground authorial presence (*I, we, our study*). Table 1 presents this taxonomy with illustrative examples. This framework has proved stable across languages, genres, and proficiency levels, providing a reliable foundation for cross-linguistic comparison (Hyland et al., 2022).

**Table 1.** Hyland's (2005) Metadiscourse Taxonomy with *Categories, Functions, and Examples*

| Category | Function | Examples (illustrative usage in a sentence) |
|---|---|---|
| ***Interactive Resources*** | | |
| Transitions | Express logical connections or shifts | *however, therefore, furthermore, in contrast* (e.g., "The sample size was small; **however**, the results were reliable.") |
| Frame markers | Indicate text structure or sequence | *first, next, in conclusion, to summarize* (e.g., "**In conclusion**, we revisit the hypothesis.") |
| Endophoric markers | Refer to information in other parts of the text | *as noted above, see Fig. 2, as discussed later* (e.g., "The results (**see Table 1 above**) confirm the trend.") |
| Evidentials | Refer to source of information/evidence | *according to X (Year), [reference]* (e.g., "**According to Smith (2020)**, academic writing requires explicit cohesion.") |
| Code glosses | Rephrase or exemplify for clarity | *for example, for instance, in other words* (e.g., "Several precautions were taken—**for example**, all data were anonymized.") |
| ***Interactional Resources*** | | |
| Hedges | Express uncertainty or caution | *might, perhaps, possibly, it seems* (e.g., "It **might** be the case that these results are context-specific.") |
| Boosters | Express certainty or emphasis | *clearly, definitely, indeed, it is evident that* (e.g., "This **clearly** demonstrates a significant improvement.") |
| Attitude | Convey writer's attitude or | *unfortunately, happily, surprisingly, important* (e.g., |

| Category | Function | Examples (illustrative usage in a sentence) |
|---|---|---|
| markers | evaluation | "**Surprisingly**, the control group outperformed expectations.") |
| Engagement markers | Address or involve the reader directly | *consider, note that, you can see* (e.g., "**Note that** this analysis excludes outliers.") |
| Self-mentions | Explicit reference to the author(s) | *I, we, my, our* (e.g., "**In our study**, **we** found a strong correlation.") |

*Note. Categories based on Hyland's (2005) metadiscourse framework. Interactive resources help guide readers through the text, while interactional resources involve readers in the discourse.*

Interactive resources form the structural scaffolding of academic prose, guiding readers through logical architecture and signaling how propositions connect. Transitions weave local argument strands together, while frame markers announce major rhetorical moves and endophoric markers direct attention to supporting materials. Evidentials embed claims within existing scholarship, and code glosses ensure terminological clarity. Working in concert, these devices transform dense empirical content into navigable argumentative paths (Hyland & Tse, 2004).

Interactional resources manage the writer-reader relationship by registering authorial presence and calibrating certainty. Hedges present propositions as provisional (it appears that, might, suggest), while boosters amplify commitment (undoubtedly, it is evident that). Attitude markers guide evaluative interpretation (unfortunately, crucially), engagement markers establish conversational tenor (consider the implications), and self-mentions clarify agency and responsibility (we examined, I argue). Together, these resources enable writers to deliver complex content while negotiating stance and fostering audience engagement (Hyland, 2005).

Recent work both reaffirms and refines this framework. A large-scale bibliometric synthesis shows the growing predominance of the interpersonal model across metadiscourse research, while mapping how the field has diversified methodologically and thematically (Hyland & Jiang, 2024). In parallel, a comprehensive systematic review of academic-writing studies takes stock of model choices (including Hyland's), clarifies 'broad' versus 'narrow' metadiscourse traditions, and outlines directions for more context-sensitive, integrative analyses (Pearson & Abdollahzadeh, 2023).

### Cross-Linguistic Patterns in L2 Metadiscourse

Corpus research reveals systematic differences between L2 and native speaker metadiscourse deployment that transcend specific language backgrounds. These patterns reflect the complex interplay of limited linguistic resources, L1 rhetorical transfer, and pedagogical models encountered during instruction. L2 writers consistently demonstrate heavy reliance on explicit cohesive devices, often employing a narrower, more formulaic repertoire than native speakers. Milton and Tsang (1993) found that Hong Kong learners showed pronounced preference for high-frequency transitions while underusing contrastive and adversative alternatives. Similarly, Altenberg and Tapper (1998) documented Swedish learners' overuse of additive connectors (*and, also*) alongside underuse of more sophisticated logical relationships. This pattern suggests that L2 writers prioritize surface-level textual cohesion through familiar devices rather than deploying the varied connective resources that characterize expert prose.

Frame markers exhibit similar imbalances across L2 populations. Ädel (2006) found that Swedish learners relied heavily on sequencing devices (*first, second, finally*) while avoiding more sophisticated structural signals common in native academic writing. Code glosses show parallel restrictions: L2 writers tend to overuse basic exemplification (*for example*) while underemploying reformulation and clarification strategies that native speakers use to manage reader comprehension more subtly. Stance marking represents the most consistent area of L2-L1 divergence across linguistic

backgrounds. Hyland and Milton's (1997) comparison of Hong Kong and British secondary essays revealed that L2 writers used fewer hedge types overall, relying almost exclusively on modal verbs (*may, might*) while native speakers employed diverse lexical hedges (*perhaps, likely, it is possible that*). This restricted repertoire compromised learners' ability to mark subtle gradations of certainty, resulting in claims that appeared either overly categorical or unnecessarily weakened.

Hinkel's (2005) large-scale study extended these findings to university-level writing, showing that L2 writers across multiple language backgrounds commanded markedly restricted stance repertoires. Rather than drawing on formal hedging typical of published scholarship (*generally, tends to, somewhat*), learners relied on conversational modifiers (*really, a bit, sort of*). Intensifiers conveying categorical certainty (*always, completely, absolutely*) were likewise overrepresented, suggesting direct transfer of oral register features into academic contexts. Boosting patterns reveal parallel restrictions. While native speakers calibrate emphasis through varied lexical and phrasal intensifiers, L2 writers often concentrate on a narrow set of high-frequency items or avoid boosting altogether, resulting in prose that lacks appropriate rhetorical force (Hu & Cao, 2011). Cross-linguistic studies consistently show this pattern regardless of L1 background, indicating shared developmental challenges in stance calibration.

Self-mention usage reflects cultural and pedagogical influences alongside linguistic constraints. Studies across diverse L2 populations—including Hong Kong (Hyland, 2002), Spanish (Mur Dueñas, 2011), and Arabic contexts (Al Rubaye, 2015)—document consistent reluctance to employ first-person reference. This avoidance may stem from L1 rhetorical traditions that favor impersonal construction, pedagogical advice against first-person usage, or uncertainty about appropriate levels of authorial presence in English academic discourse. Recent corpus research has moved beyond frequency-based comparisons to examine how L2 writers integrate different metadiscourse functions. Crossley, Kyle, and McNamara (2016) found that stance diversity correlates positively with essay quality ratings among native speakers, suggesting that effective academic writing requires coordinated deployment of multiple interactional resources. However, L2 writers often struggle to achieve this integration, instead concentrating metadiscourse effort on familiar cohesive categories while leaving stance marking underdeveloped.

Co-occurrence analyses reveal systematic differences in how L2 and native writers combine metadiscourse functions. Ädel (2006) documented Swedish learners' tendency to cluster connective devices together while isolating stance markers, contrasting with native speakers' practice of integrating hedging, boosting, and self-reference within similar textual spans. This suggests that L2 metadiscourse development involves not only expanding individual category repertoires but also learning to coordinate different rhetorical functions strategically. While metadiscourse research has documented consistent patterns across diverse L2 populations, Turkish EFL writers remain underrepresented in this literature despite their growing presence in international academic contexts. Limited studies have examined Turkish EFL learner context (Akbaş, 2014; Kirişçi & Duruk, 2022), yet the documented tendency for L2 writers to prioritize explicit cohesion while underusing stance markers provides a foundation for systematic corpus-based investigation of Turkish learners. Language-specific and cultural factors may produce distinctive patterns worthy of comprehensive analysis, while existing research has relied primarily on frequency-based approaches, leaving unexplored the functional relationships and developmental patterns that corpus-based methodologies can now illuminate.

## METHODOLOGY

### Research Design and Questions

This study employs a corpus-based comparative design to investigate systematic differences in metadiscourse deployment between Turkish EFL learners and native English speakers. The research addresses three specific questions: (1) How do Turkish learners and native speakers differ in their use of interactive and interactional metadiscourse categories? (2) What co-occurrence patterns characterize



metadiscourse deployment in each population? (3) How do metadiscourse patterns vary with text length and across individual writers? The analysis applies Hyland's (2005) comprehensive metadiscourse taxonomy, examining eight functional categories through automated extraction and statistical comparison. This approach enables systematic quantification of rhetorical differences while controlling for genre and contextual variables that might otherwise confound cross-population comparisons.

**Corpora Selection and Description**

The study compares metadiscourse patterns across two established benchmark corpora including TICLE - Turkish subcorpus of International Corpus of Learner English ICLE (Granger et al., 2009) and LOCNESS (Granger, 1998) English native speaker corpus. TICLE (Kilimci and Can, 2008) comprises 285 argumentative essays totaling approximately 202,000 running words, written by upper-intermediate Turkish EFL students. The Louvain Corpus of Native English Essays (LOCNESS) provides a native speaker baseline through 171 comparable argumentative essays containing approximately 148,000 running words, produced by British and American undergraduate students.

The TICLE sample comprises 285 argumentative essays (≈ 202,000 running words) written by L1-Turkish EFL undergraduates enrolled at universities in Türkiye. Writers are upper-intermediate learners of English and produced the texts in regular academic contexts (course assignments/exams) using standard argumentative prompts. The LOCNESS comparison set contains 171 argumentative essays (≈ 148,000 running words) authored by L1-English undergraduates at universities in the United Kingdom and the United States. For comparability, the LOCNESS subset used here includes only undergraduate essays (school/A-level texts were excluded). Both corpora consist of standalone essays (one text per writer in the released data), and files are anonymized; no demographic variables beyond L1 and educational level are available.

Both corpora focus exclusively on argumentative essay writing within academic contexts, ensuring that observed differences reflect linguistic and rhetorical preferences rather than genre variation. The Turkish learners represent a relatively homogeneous population in terms of proficiency level and L1 background, while the native speaker corpus provides appropriate developmental comparison through undergraduate-level writing. This controlled foundation enables attribution of systematic differences to language proficiency and cultural rhetorical patterns rather than contextual factors.

**Metadiscourse Detection and Validation**

Metadiscourse identification employed a hybrid approach combining rule-based pattern matching with supervised machine learning classification. This dual strategy maximizes both precision and coverage by leveraging the reliability of linguistic patterns alongside the contextual sensitivity of statistical models. The rule-based component utilizes approximately 250 linguistic patterns distributed across Hyland's eight metadiscourse categories, targeting both individual lexical items (perhaps, clearly, however) and multi-word academic phrases (there is no doubt that, it appears that, in my view). These patterns distinguish genuine metadiscourse functions from superficially similar expressions used in conversational or temporal contexts, ensuring that identified markers serve authentic rhetorical purposes within academic argumentation.

The pattern inventory comprises a total of 105 linguistic patterns distributed across the eight categories as follows: transitions (16), evidentials (15), frame markers (14), code glosses (12), engagement markers (12), self-mentions (12), boosters (12), and hedges (12). These patterns include both single-word markers (e.g., however, clearly) and multi-word academic phrases (e.g., on the other hand, according to, in other words). To distinguish genuine academic metadiscourse from superficially similar expressions, the system applies contextual filtering rules that analyze a 50-character window around each detected pattern. For instance, the transition marker however is accepted when it signals a

logical contrast between propositions ("The results supported the hypothesis. However, subsequent analysis revealed limitations.") but rejected when it appears in quantity/degree constructions ("However much I try, I can't understand this problem."). Likewise, first-person pronouns count as self-mentions only when performing academic speech acts ("I argue…", "We propose…") rather than narrating personal actions ("I went…", "We can go…"). The disambiguation algorithm scores each context by counting academic indicators (verbs such as demonstrate, analyze; nouns such as study, findings) versus non-academic indicators (narrative verbs such as went, came; settings such as family, home), and accepts a token as metadiscourse only when academic indicators predominate. This contextual filtering, combined with ML validation, maintains high precision while preserving broad coverage.

**Machine Learning Enhancement**

The machine learning component addresses cases where pattern matching alone proves insufficient by incorporating rich linguistic and contextual features. Advanced natural language processing techniques analyze multiple dimensions of each potential marker:

- Lexical features: word length, capitalization patterns, punctuation context
- Syntactic features: grammatical categories, dependency relationships, phrase structures
- Contextual features: surrounding grammatical context, sentence position
- Academic features: disciplinary vocabulary patterns, formal register indicators
- Semantic features: sentence boundary positions, discourse-level relationships

The classification system was trained on 5,000 balanced synthetic examples (equal positive and negative instances). Training employed an 80/20 stratified train-test split (random_state=42), with the training set (4,000 examples) subject to 5-fold stratified cross-validation for hyperparameter optimization. The 1,000-example holdout set was reserved for internal validation and never exposed during training. Final system performance was evaluated on 86 manually annotated test cases completely independent of the training data, representing diverse metadiscourse contexts across all eight analytical categories. The hybrid system achieved 89.6% true positive detection rates and 92.1% false positive avoidance, yielding an overall composite accuracy of 90.8%. This validation approach ensures no data leakage between training and testing while providing robust estimates of real-world performance.

**System Validation**

Comprehensive validation employed 86 manually annotated test cases representing diverse metadiscourse contexts across all eight analytical categories. The hybrid system achieved 89.6% true positive detection rates and 92.1% false positive avoidance, yielding an overall composite accuracy of 90.8%. Only markers meeting a stringent 0.60 confidence threshold were included in final analyses, ensuring that identified instances represent authentic metadiscourse rather than spurious matches. This conservative approach maintains analytical integrity when comparing corpora with different linguistic characteristics.

**Analysis Framework and Metrics**

Primary analysis computes marker density as occurrences per 1,000 running words. In addition, we report 95% Wilson score intervals for each category's share of all metadiscourse tokens (category count ÷ total metadiscourse tokens). For compact display—and to place these intervals on roughly the same numeric scale as the per-1,000 rates—we rescale the share intervals by a constant factor (×10) in the tables. These intervals quantify uncertainty in category shares, not in the per-word rates; we do not attach CIs to the per-1,000 rates in the present analysis. Category dispersion is the proportion of documents containing at least one token from a category (documents with ≥1 token ÷

total documents), indicating whether a function is broadly distributed across writers or concentrated in a subset.

Co-occurrence analysis tracks joint appearance of metadiscourse categories within individual documents, generating weighted networks that reveal functional associations between different rhetorical strategies. This approach illuminates whether writers integrate multiple metadiscourse functions within their essays or concentrate effort on specific categories. Writer-level entropy quantifies individual metadiscourse diversity using Shannon entropy calculations expressed in bits of information. Higher entropy values indicate writers who demonstrate rhetorical flexibility across multiple categories, while lower values suggest specialization in particular functions—potentially reflecting either strategic focus or limited repertoires.

Keyness analysis employs log-likelihood ratios (G-test statistics) to determine statistical significance of cross-corpus differences while controlling for size variations and expected frequency distributions. Effect sizes use $\log_2$-ratio calculations to indicate practical magnitude of observed differences. Jensen-Shannon distance provides a symmetric metric capturing overall distributional similarity between learner and native speaker metadiscourse profiles. Length sensitivity analysis partitions essays into deciles by word count, calculating normalized deviation from corpus-wide means to reveal how metadiscourse density scales with document length. This approach identifies whether rhetorical strategies change systematically as argumentative complexity increases.

Comprehensive text processing ensured analytical consistency through standardization of formatting, punctuation, and correction of typographical errors common in learner writing. Advanced linguistic analysis enabled accurate sentence identification and grammatical parsing necessary for context-sensitive metadiscourse detection. All analytical procedures were applied uniformly across both corpora to ensure that observed differences reflect genuine rhetorical patterns rather than methodological artifacts. The combination of automated detection with rigorous validation provides a reliable foundation for systematic cross-population comparison while maintaining the scale necessary for robust statistical analysis.[i]

## FINDINGS

The comparative analysis reveals a systematic pattern in how Turkish EFL learners and native English speakers deploy metadiscourse resources. Turkish learners consistently prioritize explicit textual guidance and reader engagement, employing high frequencies of transitions, frame markers, code glosses, and engagement markers to ensure comprehension and maintain reader attention. Native speakers, by contrast, invest more heavily in calibrated stance marking through hedges, boosters, and self-mentions, creating prose that modulates certainty and projects appropriate academic authority. This fundamental difference—cohesion-focused versus stance-focused metadiscourse—emerges across multiple analytical dimensions and has clear implications for L2 academic writing instruction. The following sections examine this pattern through complementary analytical approaches. Section 4.1 documents frequency differences across eight metadiscourse categories, while Section 4.2 explores how these categories co-occur within texts, revealing distinct rhetorical architectures in learner versus native writing. Section 4.3 investigates how metadiscourse patterns change with essay length, and Section 4.4 examines writer-level diversity to understand individual variation within each group.

### Frequency Patterns: Cohesion versus Stance Priorities

The frequency analysis, normalized to occurrences per 1,000 words, reveals contrasting rhetorical priorities between Turkish learners and native speakers. As shown in Table 2, Turkish writers demonstrate markedly higher usage of organizational and clarificatory devices, while native speakers employ more interpersonal stance resources.

**Table 2.** *Frequency Analysis of Metadiscourse Categories Across Corpora*

| Category | TICLE | | | | LOCNESS | | | | | Effect size |
|---|---|---|---|---|---|---|---|---|---|---|
| | Count | /1,000 | CI [low, high] | Disp. | Count | /1,000 | CI [low, high] | Disp. | Sig. | Rate ratio ($log_2$) |
| **Overused by Turkish learners** | | | | | | | | | | |
| Engagement markers | 218 | 1.08 | [0.90, 1.15] | .400 | 49 | 0.33 | [0.29, 0.50] | .158 | *** | 3.27× (+1.71) |
| Transitions | 585 | 2.90 | [2.55, 2.93] | .751 | 288 | 1.95 | [2.02, 2.48] | .649 | ** | 1.49× (+0.57) |
| Code glosses | 457 | 2.27 | [1.97, 2.32] | .695 | 229 | 1.55 | [1.59, 2.00] | .632 | * | 1.46× (+0.55) |
| Frame markers | 274 | 1.36 | [1.15, 1.43] | .554 | 136 | 0.92 | [0.90, 1.24] | .439 | † | 1.48× (+0.56) |
| Evidentials | 296 | 1.47 | [1.24, 1.54] | .530 | 173 | 1.17 | [1.17, 1.55] | .368 | *ns* | 1.26× (+0.33) |
| **Underused by Turkish learners** | | | | | | | | | | |
| Hedges | 186 | 0.92 | [0.76, 1.00] | .351 | 250 | 1.69 | [1.74, 2.17] | .550 | *** | 0.54× (−0.88) |
| Boosters | 55 | 0.27 | [0.20, 0.33] | .161 | 74 | 0.50 | [0.46, 0.72] | .310 | *** | 0.54× (−0.89) |
| Self-mentions | 68 | 0.34 | [0.25, 0.40] | .112 | 84 | 0.57 | [0.53, 0.80] | .181 | *** | 0.60× (−0.75) |

*Note.* Rate /1,000 = normalized frequency per 1,000 running words. Disp. = dispersion (proportion of essays containing ≥1 token from the category). CI values are 95% Wilson intervals for category shares (category count ÷ total metadiscourse tokens), rescaled ×10 for display; they do not represent uncertainty in the per-1,000 rates. Sig. = statistical significance based on log-likelihood keyness analysis. Effect size = rate ratio (TICLE ÷ LOCNESS); $log_2$ ratio shown in parentheses (positive = overuse in TICLE; negative = underuse). ***p < .001, **p < .01, p < .05, †p < .10 (marginal), ns = not significan

Turkish learners substantially outpace native speakers in four categories that privilege textual transparency and reader guidance. Transitions occur at 2.90 per 1,000 words in TICLE versus 1.95 in LOCNESS (rate ratio = 1.49, p < .01), reflecting heavier use of explicit logical links. Code glosses show a comparable uplift (2.27 vs. 1.55; ratio = 1.46, $p < .05$), indicating frequent reformulation and exemplification. Frame markers likewise rise (1.36 vs. 0.92; ratio = 1.48, †p < .10), signalling stronger structural signposting. By contrast, evidentials are close to parity (1.47 vs. 1.17; ratio = 1.26, n.s.), suggesting that citation practices do not differ reliably between the groups. The clearest divergence is in engagement markers: learners use direct reader-addressing devices over three times as often as natives (1.08 vs. 0.33; ratio = 3.27, $p < .001$). Dispersion underscores how widespread these habits are within TICLE: transitions appear in 75% of essays and code glosses in 70%, while engagement markers occur in 40% of learner essays versus only 16% of native essays.

Native writers, in turn, invest more in interpersonal positioning and commitment calibration. Hedges are nearly twice as frequent in LOCNESS (1.69 vs. 0.92; ratio = 0.54 from the learner perspective, $p < .001$), indicating routine softening of claims and acknowledgement of uncertainty. Boosters (0.50 vs. 0.27; learner ratio = 0.54, $p < .001$) and self-mentions (0.57 vs. 0.34; learner ratio = 0.60, $p < .001$) follow the same pattern, pointing to greater authorial presence and calibrated emphasis among native undergraduates. Dispersion patterns align with this interpretation: hedges occur in 55% of native essays but only 35% of learner essays, and boosters in 31% vs. 16%, respectively. Overall, the profile that emerges is a cohesion-first architecture for Turkish learners (high rates and broad dispersion of transitions, code glosses, frame markers, and engagement) versus a stance-forward

architecture for native writers (higher rates and broader dispersion of hedges, boosters, and self-mentions), with evidentials largely shared across groups.

Figure 1 reveals how these broad frequency patterns manifest through specific lexical choices, showing the most frequent markers within each category for both populations. The data confirm that Turkish learners concentrate their metadiscourse effort on a relatively narrow set of high-visibility items, while native speakers distribute usage more evenly across diverse expressions.

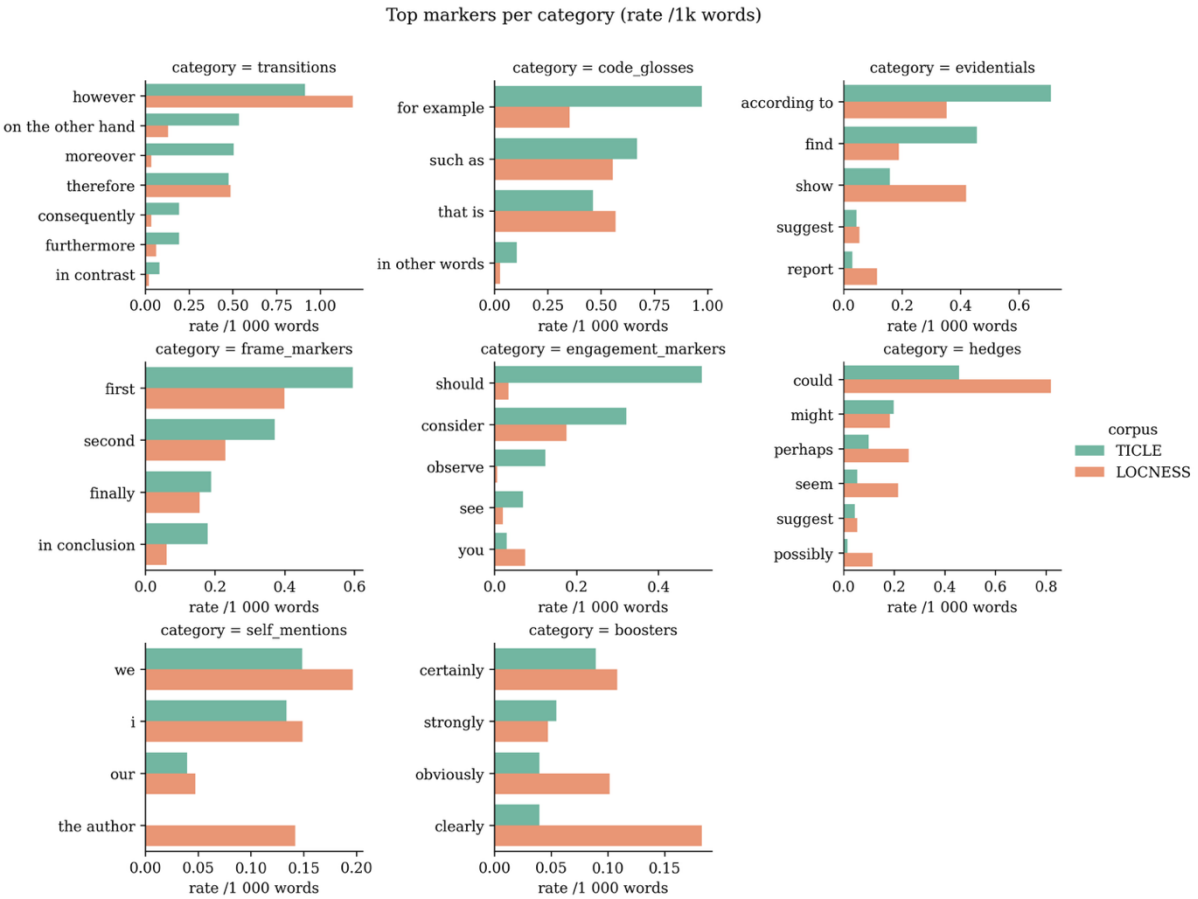



**Figure 1.** The most frequent individual markers (three per category) in TICLE and LOCNESS

In code glosses, Turkish learners show heavy reliance on "for example" at nearly twice the native frequency, while native speakers spread clarification efforts across "such as," "that is," and other alternatives. Transitions reveal similar concentration: both groups favor "however," but TICLE essays far exceed LOCNESS usage of "on the other hand" and "moreover," suggesting classroom emphasis on these particular connectives. Stance vocabulary differences are particularly revealing. For hedging, both populations use "could" frequently, but learners' next preference is "might," while native writers employ "perhaps" and expressions like "seem." This pattern indicates that Turkish learners rely primarily on modal verbs for uncertainty, whereas native speakers command the richer lexical hedging that characterizes published academic prose. Engagement markers underscore the starkest contrast: learners heavily favor "we should" and "consider" at frequencies unmatched in native writing, while native speakers use second-person "you" sparingly, preferring more indirect reader involvement. Self-mentions show the reverse pattern, with native speakers using first-person pronouns more densely and introducing possessives like "our" and expressions like "the author" more frequently.

**Rhetorical Architectures: How Metadiscourse Categories Cluster**

Beyond frequency differences, Turkish learners and native speakers also differ fundamentally in how they combine metadiscourse functions within their texts. This analysis examines which categories tend to co-occur within the same documents, revealing distinct "rhetorical architectures" that characterize each population's approach to academic writing.

To identify these patterns, we tracked which metadiscourse categories appeared together within individual essays, reasoning that categories co-occurring in the same documents reflect writers' preferences for combining different rhetorical functions. Writers who routinely employ both organizational and interpersonal resources within single essays demonstrate integrated metadiscourse strategies, while those who concentrate on particular functional domains reveal more specialized approaches. The resulting co-occurrence networks illuminate how writers mentally organize their metadiscourse toolkit and which functions they view as naturally complementary.

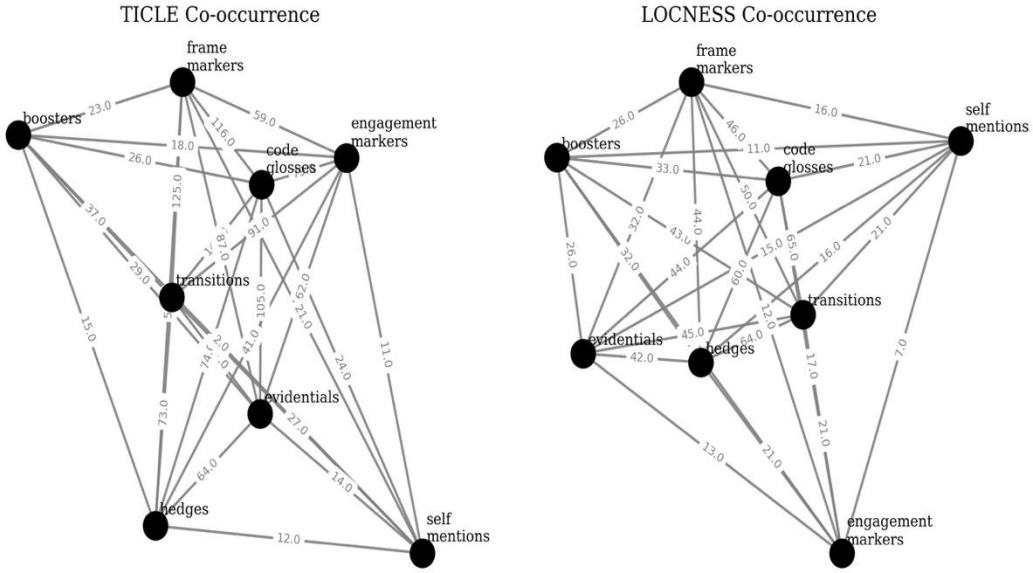



**Figure 2.** Pairwise co-occurrence of metadiscourse categories in TICLE (left) and LOCNESS (right).

Figure 2 (left panel) shows that Turkish learners organize their metadiscourse around a dense "cohesion hub" centered on transitions. In TICLE essays, transitions co-occur most frequently with code glosses, frame markers, and engagement markers, creating thick connecting lines that indicate these categories routinely appear together within the same short text passages. This clustering pattern suggests that Turkish writers think of academic writing primarily as a reader-guidance activity. When they deploy a transition like "however" or "therefore," they frequently follow up with clarifying examples (code glosses), explicit structural signals (frame markers), or direct appeals to readers (engagement markers). The resulting prose provides extensive navigational support, with multiple types of reader-oriented devices appearing in concentrated bursts. Notably, stance categories—hedges, boosters, and self-mentions—remain at the periphery of this network, connected to the cohesion hub only by thin lines. This positioning indicates that Turkish learners treat stance marking as a separate, less integrated aspect of academic writing, rarely combining hedging or boosting with their primary cohesion-building activities.

The native speaker network (Figure 2, right panel) reveals a fundamentally different organization. Here, hedges occupy the central position, forming the thickest connections with boosters, self-mentions, and, to a lesser extent, transitions. This "stance cluster" indicates that native writers routinely combine uncertainty marking, emphasis, and authorial presence within the same essays. Essays by native speakers that include hedging devices also tend to feature strategic boosting and

explicit self-reference, suggesting an integrated approach to interpersonal positioning. Writers who acknowledge uncertainty in some claims (hedging) typically also assert strong confidence in others (boosting) while taking clear authorial responsibility (self-mentions). This pattern reflects sophisticated understanding of how different stance devices work together to create nuanced, persuasive academic voice within individual compositions.

The traditional connective categories—transitions and frame markers—still appear in the native network but with lighter connections, suggesting more selective, strategic deployment rather than the systematic clustering seen in learner writing. Code glosses notably attach to the stance cluster rather than forming a separate organizational network, indicating that native speakers embed clarification within evaluative commentary rather than treating it as pure organizational scaffolding.

**Table 3.** Ranking of Dominant Pairs by Dataset and Strength Level

| Rank | TICLE Dominant Pair | Strength | LOCNESS Dominant Pair | Strength |
|---|---|---|---|---|
| 1 | Transitions + Code glosses | Very high | Hedges + Self-mentions | Very high |
| 2 | Transitions + Engagement | High | Hedges + Boosters | High |
| 3 | Transitions + Frame markers | High | Hedges + Transitions | High |
| 4 | Transitions + Evidentials | Moderate | Boosters + Self-mentions | Moderate |
| 5 | Code glosses + Engagement | Moderate | Transitions + Code glosses | Moderate |

*Note. Rankings are based on frequency analysis of metadiscourse marker pairs in academic writing samples. TICLE = Turkish International Corpus of Learner English; LOCNESS = Louvain Corpus of Native English Essays.*

In TICLE, the strongest pairing combines transitions with code glosses, indicating that Turkish writers routinely follow logical connectors with clarifying examples or reformulations. The second and third strongest links—transitions with engagement markers and transitions with frame markers—further demonstrate how learners cluster multiple reader-guidance functions around their logical connectives. Only at the fourth and fifth positions do evidentials enter the pattern, still linking to the organizational core rather than forming independent stance relationships. LOCNESS presents a fundamentally different ranking. Three of the top five pairings involve hedges: hedges with self-mentions (strongest), hedges with boosters (second), and hedges with transitions (third). This dominance reflects native speakers' tendency to package uncertainty, authorial presence, and occasional emphasis together within compact textual spans. The hedge-self-mention combination is particularly significant, as it represents the sophisticated integration of stance softening with authorial responsibility that characterizes mature academic writing.

Notably, the strongest purely organizational pairing in LOCNESS—transitions with code glosses—ranks only fifth, compared to its first-place position in TICLE. This relative positioning confirms that native speakers treat organizational devices as supporting rather than central elements in their metadiscourse strategy. The co-occurrence differences observed in the networks and pairing patterns achieve statistical significance when tested through keyness analysis. Table 4 presents $\log_2$ ratios comparing TICLE and LOCNESS usage rates, with values above +0.50 indicating substantial over-use by Turkish learners and values below -0.50 indicating substantial under-use compared to native speakers.

**Table 4.** Keyness of Metadiscourse Categories (Rates per 1,000 Words and $\log_2$ Ratios)

| Category | TICLE Rate | LOCNESS Rate | $\log_2$ Ratio | Direction |
|---|---|---|---|---|
| Engagement markers | 1.08 | 0.33 | **+1.71** | over-used |
| Transitions | 2.90 | 1.95 | **+0.57** | over-used |
| Frame markers | 1.36 | 0.92 | **+0.56** | over-used |
| Code glosses | 2.27 | 1.55 | **+0.55** | over-used |
| Evidentials | 1.47 | 1.17 | +0.33 | near parity |
| Hedges | 0.92 | 1.69 | **−0.88** | under-used |
| Boosters | 0.27 | 0.50 | **−0.89** | under-used |
| Self-mentions | 0.34 | 0.57 | **−0.74** | under-used |

*Note. TICLE = Turkish International Corpus of Learner English; LOCNESS = Louvain Corpus of Native English Essays. Bold values indicate statistically significant differences. Positive $\log_2$ ratios indicate over-use in TICLE compared to LOCNESS; negative ratios indicate under-use.*

The keyness analysis confirms the co-occurrence findings with striking clarity. Engagement markers show the largest positive keyness value (+1.71), indicating that Turkish learners use these reader-addressing devices more than three times as frequently as native speakers. This massive difference explains why engagement markers serve as a bridge between cohesion categories in the Turkish network—they represent a distinctly learner-oriented strategy for maintaining reader connection. The remaining over-used categories—transitions (+0.57), frame markers (+0.56), and code glosses (+0.55)—all exceed the significance threshold and correspond precisely to the cohesion hub observed in Figure 2. These positive keyness values provide statistical validation for the organizational clustering that characterizes Turkish learner writing.

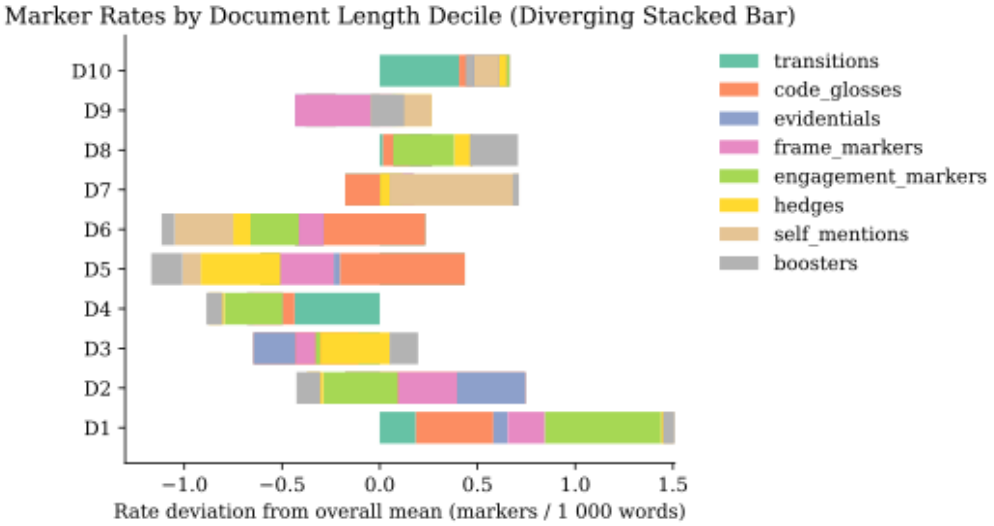
Conversely, the three most under-used categories—hedges (-0.88), boosters (-0.89), and self-mentions (-0.74)—correspond exactly to the stance cluster that dominates native speaker networks. The negative keyness values confirm that Turkish learners systematically avoid the interpersonal positioning strategies that native speakers integrate routinely. Only evidentials approach parity between the groups (+0.33), suggesting that both populations cite sources at comparable rates once academic conventions are established. The magnitude of these differences is pedagogically significant. The largest gap occurs not in any single stance category but in engagement markers, underscoring that direct reader address represents the most distinctive feature of Turkish EFL writing. This finding suggests that classroom instruction has successfully emphasized reader engagement while neglecting the stance calibration that characterizes expert academic discourse.

These contrasting network structures reveal different underlying philosophies of academic writing. Turkish learners appear to conceptualize academic prose as requiring comprehensive organizational support: essays that provide logical signaling also provide structural guidance, clarifying examples, and explicit reader engagement. Their metadiscourse strategy integrates multiple reader-service functions while treating stance marking as a separate, less essential component. Native speakers, by contrast, view academic writing as fundamentally requiring interpersonal negotiation: essays that include stance marking typically integrate multiple forms of positioning and evaluation. Their metadiscourse combines uncertainty acknowledgment, strategic emphasis, and authorial responsibility within individual compositions, treating organizational devices as supporting rather than central elements. The co-occurrence evidence suggests that Turkish learners have successfully internalized a coherent organizational philosophy but have not yet learned to integrate stance functions into this comprehensive approach. This pattern aligns with developmental models suggesting that L2

writers first master systematic textual guidance before developing the integrated interpersonal competencies that characterize expert academic prose.

**Metadiscourse Scaling: How Rhetorical Strategies Change with Essay Length**

The cohesion-versus-stance divide observed in frequency and co-occurrence patterns becomes even more pronounced as essays increase in length. This analysis examines how Turkish learners and native speakers adjust their metadiscourse strategies when given more space to develop their arguments, revealing fundamental differences in how each group conceptualizes the relationship between text length and rhetorical sophistication. To investigate length effects, essays in both corpora were ranked by word count and divided into ten equal groups, from shortest (bottom 10%) to longest (top 10%). For each length band, we calculated how much each metadiscourse category deviates from its corpus-wide average, revealing whether writers use more or fewer markers as their essays expand.




** D1 represents the shortest 10 % of essays, D10 the longest 10 %.*

**Figure 3.** Deviation from corpus-wide mean (markers per 1 000 words) across length deciles

Figure 3 (left panels) reveals that Turkish learners respond to increased essay length by systematically amplifying their cohesion-focused strategy. As essays grow longer, transitions, code glosses, and engagement markers all increase substantially above their corpus averages, while stance categories remain consistently below average regardless of text length. In the longest Turkish essays (top 10%), transitions rise approximately 0.5 markers per 1,000 words above the learner average, accompanied by similar increases in code glosses (+0.4) and engagement markers (+0.3). This pattern suggests that Turkish writers equate essay length with the need for more explicit reader guidance—more logical connectors, more clarifying examples, and more direct reader address.

Critically, hedges remain below the corpus mean throughout the entire length scale, indicating that even extended argumentative space does not prompt Turkish learners to engage in the uncertainty marking that characterizes mature academic discourse. This persistence suggests that the stance deficit in Turkish writing reflects conceptual rather than simply linguistic limitations. Native speakers demonstrate the opposite scaling pattern, using additional essay length to develop interpersonal nuance rather than organizational scaffolding. In the longest native essays (top 10%), hedges climb approximately 0.6 markers above the LOCNESS average, accompanied by increases in self-mentions (+0.2) and boosters.

This pattern indicates that native speakers view extended argumentative space as an opportunity for more sophisticated stance calibration—more careful hedging of uncertain claims, more strategic self-positioning, and more nuanced emphasis. The longest native essays feature the richest interpersonal metadiscourse, suggesting that rhetorical maturity involves learning to exploit additional

space for stance development rather than simply adding more organizational signals. Notably, transitions and code glosses actually decrease below corpus averages in mid-length native essays (bands 5-7), indicating that native speakers may reduce explicit cohesion markers when they have sufficient space for more sophisticated rhetorical strategies. Engagement markers remain near baseline throughout the length scale, confirming their peripheral role in native undergraduate writing.

The contrasting length-scaling patterns reveal different theories of what makes academic writing more sophisticated. Turkish learners appear to believe that longer, better essays require more explicit guidance—more transitions to signal logical relationships, more examples to ensure comprehension, and more direct reader engagement to maintain attention. This view treats academic writing as fundamentally a reader-service activity that benefits from increased explicitness.

Native speakers, by contrast, use additional space to develop what might be called "interpersonal complexity"—more nuanced stance marking, more careful claim calibration, and more strategic authorial positioning. This approach treats longer essays as opportunities for more sophisticated intellectual negotiation rather than simply more thorough explanation. These patterns suggest that Turkish learners have mastered one dimension of academic writing development—the organizational scaffolding that supports reader comprehension—but have not yet learned to exploit extended argumentative space for the interpersonal positioning that distinguishes expert from novice academic prose. The length analysis thus reinforces the frequency and co-occurrence findings while adding a developmental dimension: as Turkish learners progress toward longer, more complex arguments, they need explicit instruction in how to use additional space for stance development rather than simply organizational amplification.

**Individual Variation: Diversity in Metadiscourse Strategies**

Beyond group-level patterns, Turkish learners and native speakers also differ in how much individual variation exists within each population. This analysis examines writer-level diversity to understand whether the cohesion-focused and stance-focused strategies represent uniform group tendencies or reflect different degrees of individual flexibility within each writing community.

To measure individual metadiscourse diversity, we calculated how evenly each writer distributed their metadiscourse effort across the eight functional categories. Writers who concentrate on just one or two categories receive low diversity scores, while those who employ a balanced mix across multiple categories receive high scores. This approach reveals whether writers demonstrate rhetorical flexibility or specialize in particular metadiscourse functions.

**Table 5 .**Descriptive Statistics for Writer-Level Entropy

| Corpus | N | M | SD | Median | IQR | Range (min–max) |
|---|---|---|---|---|---|---|
| TICLE | 285 | 1.539 | 0.623 | 1.585 | 0.699 | 0.10–2.70 |
| LOCNESS | 171 | 1.358 | 0.744 | 1.500 | 0.940 | 0.05–2.85 |

***Note**. Entropy values are in bits (Shannon entropy). Higher values indicate a more balanced distribution across metadiscourse categories. Group comparison: ΔM = 0.181 bits, 95% CI [0.048, 0.314], Welch's t(310.26) = 2.67., p = .008, Cohen's d = 0.27 (small). M = mean; SD = standard deviation; IQR = interquartile range (75th − 25th percentile).*

Table 5 reveals uniform metadiscourse profiles among Turkish learners, with most writers clustering near the group median (1.585 bits). The narrow interquartile range (0.699) indicates similar rhetorical strategies across individuals, reflecting shared conventions rather than individualized approaches. This uniformity stems from the cohesion-first approach documented previously. Turkish learners deploy comparable combinations of transitions, code glosses, and engagement markers while

systematically underusing stance categories. The resulting diversity scores reflect statistical balance across categories rather than genuine rhetorical versatility—a pattern of consistent deployment of familiar cohesive functions.

The clustering suggests successful classroom instruction in specific metadiscourse strategies. Turkish learners have internalized a coherent approach emphasizing reader guidance and textual transparency, producing essays with recognizably similar rhetorical approaches despite varying content. Native speakers demonstrate greater individual variation (IQR = 0.940). Some writers exhibit highly specialized profiles, focusing intensively on particular functions—perhaps emphasizing stance marking in arguments or organizational scaffolding in explanations. Others achieve diversity scores above 2.7 bits by integrating multiple metadiscourse functions within single essays.

This variation indicates that rhetorical maturity involves flexible deployment strategies rather than uniform conventions. Native speakers have developed intuitive understanding of when different metadiscourse combinations serve their purposes, producing varied individual profiles despite shared competencies. The comparison also reveals a paradox: Turkish learners achieve higher median diversity (1.59 vs. 1.50) yet show less individual variation than native speakers. This pattern reflects the difference between statistical balance and rhetorical sophistication. Turkish writers distribute markers relatively evenly across categories, creating numerical diversity, but concentrate heavily on cohesive functions while avoiding stance complexity.

Native speakers show lower median diversity because many specialize strategically in particular metadiscourse functions rather than employing all categories equally. Their greater individual variation indicates that this specialization reflects conscious rhetorical choices rather than limited repertoires. Some native writers concentrate on stance marking when argumentative force is needed; others emphasize organizational clarity when explaining complex concepts. The diversity analysis suggests Turkish learners have acquired a coherent metadiscourse system but lack the strategic flexibility characterizing advanced academic writing. The uniform cohesion-focused approach represents genuine competence—Turkish learners reliably provide textual guidance and maintain reader engagement—but this competence has not expanded to include stance calibration that enables rhetorical adaptation.

Advanced academic writing requires knowing when to emphasize different metadiscourse functions. This develops through extensive exposure to diverse academic contexts and internalization of how rhetorical situations require different combinations. For Turkish learners, progression toward native-like flexibility involves moving beyond their mastered cohesion strategy toward context-sensitive integration of stance and organizational functions—not abandoning current strengths but combining them with the interpersonal positioning characterizing expert discourse.

Statistical comparison confirms these differences. Turkish learners' mean entropy ($M = 1.539$ bits, $SD = 0.623$) exceeded native speakers' ($M = 1.358$ bits, $SD = 0.744$), Welch's $t(310.26) = 2.67$, $p = .008$, mean difference = 0.181 bits, 95% CI [0.048, 0.314], Cohen's $d = 0.27$. Native speakers' greater variability ($SD = 0.744$ vs. 0.623) supports the interpretation that they employ more strategically adapted approaches, whereas Turkish learners show uniform distribution across categories.

## DISCUSSION

The systematic preference for organizational over interpersonal metadiscourse among Turkish EFL learners reflects the complex interplay of linguistic, cultural, and pedagogical factors that shape L2 academic writing development. This section examines why the cohesion-first pattern emerges so consistently and what it reveals about the cognitive and social processes underlying metadiscourse acquisition.



**Developmental Sequencing in L2 Metadiscourse**

The Turkish data support developmental models suggesting that L2 writers master surface-level textual features before acquiring the more complex interpersonal competencies that characterize expert academic prose (Crossley et al., 2016). Interactive metadiscourse—transitions, frame markers, code glosses—serves primarily organizational functions that map relatively directly onto universal logical relationships: addition, contrast, sequence, exemplification. These relationships exist across languages and cultures, making them accessible targets for explicit instruction and conscious application. Interactional metadiscourse, by contrast, requires sophisticated understanding of disciplinary epistemology, audience expectations, and cultural norms surrounding knowledge claims. Effective hedging depends on recognizing when claims warrant uncertainty, how much uncertainty to express, and which linguistic forms appropriately calibrate that uncertainty within specific academic communities. Recent work underscores that this interpersonal repertoire is acquired through genre-specific enculturation: as writers are socialized into disciplinary genres, they learn the stance options those genres sanction and prefer (Kurt & Kafes, 2025; Yee, Afzaal, & Aldayel, 2024). Such competencies develop gradually through extensive exposure to disciplinary discourse and implicit socialization into academic culture—processes that formal instruction can support but not easily accelerate. The Turkish pattern suggests that learners naturally prioritize the more cognitively accessible organizational dimension while struggling to develop the culturally embedded interpersonal dimension. This sequencing reflects rational adaptation to learning constraints rather than deficient instruction or limited aptitude.

**Transfer and Interference Effects**

Cross-linguistic differences in metadiscourse norms likely contribute to the observed patterns, though direct transfer effects prove difficult to isolate without detailed contrastive analysis of Turkish and English academic discourse. Turkish academic writing traditions may emphasize different rhetorical values—perhaps favoring implicit stance marking over explicit hedging, or collective voice over individual authorial presence—creating misalignment with Anglophone academic expectations. The engagement marker findings (3x native frequency) suggest possible positive transfer from Turkish pedagogical culture, which may emphasize direct reader address and explicit guidance more strongly than Anglo-American academic traditions. While this transfer produces "overuse" relative to native norms, it demonstrates sophisticated audience awareness and communicative intentionality that instruction should channel rather than discourage. Similarly, the systematic avoidance of self-mention may reflect cultural values emphasizing modesty and collective knowledge construction over individual scholarly authority. Turkish learners' reluctance to employ first-person positioning could stem from L1 rhetorical conventions that view explicit self-reference as inappropriate self-promotion rather than necessary academic accountability.

**Pedagogical Shaping of Rhetorical Preferences**

The remarkable uniformity in Turkish learners' metadiscourse profiles (narrow entropy distributions, consistent co-occurrence patterns) suggests successful classroom transmission of specific rhetorical strategies. Turkish EFL instruction appears to emphasize explicit cohesion and reader engagement while providing limited exposure to the stance calibration that characterizes published academic discourse. This pedagogical focus likely reflects practical constraints: organizational metadiscourse is easier to teach explicitly, assess reliably, and master demonstrably than the nuanced interpersonal positioning required for advanced academic voice. Textbooks and curricula naturally gravitate toward the more systematic, rule-governed aspects of academic writing while struggling to address the context-sensitive, culturally embedded dimensions of stance marking. The engagement marker pattern particularly suggests direct pedagogical influence, as the specific phrases identified ("we should consider," "note that") reflect classroom-friendly formulas that can be taught and practiced systematically. The resulting prose sounds distinctly "instructed" in ways that reflect successful but incomplete pedagogical intervention.



**Cross-Linguistic Patterns in L2 Metadiscourse Development**

The comparative evidence from TICLE and LOCNESS confirms a pattern repeatedly documented in cross-linguistic research: advanced EFL writers, while competent at signaling surface cohesion, underuse the stance resources that native writers routinely mobilize to calibrate claims. Similar imbalances have been reported for Hong Kong secondary essays (Hyland & Milton, 1997), Swedish undergraduate writing (Ädel, 2006) and Chinese learner corpora (Hyland et al., 2022). Each study shows that learners privilege overt textual guidance at the expense of fine-grained stance modulation—a configuration now replicated in Turkish data. The stance deficit is most visible in hedging. Hyland & Milton (1997) found that Hong Kong learners relied heavily on modal verbs (will, may, would) while native speakers employed more diverse lexical hedges such as perhaps. Turkish learners employ may and might but seldom draw on the richer lexical set—perhaps, apparently, somewhat—that native writers use to signal graded certainty. Akbaş (2014) documented strategic deployment of certainty markers like certainly while simultaneously using hedges to maintain scientific caution. Kirişçi and Duruk (2022) found similar patterns in Turkish research article abstracts, showing that Turkish writers used boosters significantly more frequently (89% vs 33%) while native English speakers used hedges more often (64% vs 19%), reflecting different cultural conventions for expressing certainty in academic writing.

Hinkel (2005) attributes this profile to negative transfer from informal registers, where certainty markers dominate everyday conversation. The present keyness analysis echoes that explanation: boosters appear almost twice as often as hedges in TICLE, whereas LOCNESS reverses the ratio. Hinkel's large-scale corpus study shows that even university-level L2 writers command a markedly restricted repertoire of stance devices. Rather than drawing on the formal hedging typical of published scholarship—expressions such as generally, tends to or somewhat—the learners relied on conversational modifiers like really, a bit and sort of. Underuse of self-mention cuts across many EFL contexts. Hyland (2002) showed that Hong Kong postgraduate students avoided I and we; Mur Dueñas (2011) reported comparable reluctance among Spanish undergraduates. TICLE fits this profile: self-mentions appear at roughly 60% of the native rate and skew toward collective we rather than singular I. Hyland (2005) argues that Anglophone academic culture expects writers to "claim both expertise and responsibility" (p. 180) through first-person reference—a move still resisted by many advanced EFL learners. Network analysis strengthens these interpretations. In TICLE, the densest edges connect transitions to code glosses, frame markers and engagement markers; stance nodes hover at the periphery. Ädel (2006) found the same connective-centric hub in Swedish learner essays. LOCNESS, by contrast, mirrors the stance-centred pattern Hyland (2004) observed in expert research articles: hedges, boosters and self-mentions form tight triangles, indicating that native writers routinely bundle caution, emphasis and authorial presence within the same clause windows. Hyland et al. (2022) report a comparable stance core in native doctoral writing but a fractured one in Chinese PhD texts, again matching the Turkish learner profile. Length sensitivity further differentiates the cohorts. As Turkish essays expand, writers add more transitions, code glosses and engagement markers yet leave stance density virtually unchanged. LOCNESS shows the opposite trajectory: hedges and self-mentions increase with essay length, echoing Crossley, Kyle and McNamara's (2016) finding that stance diversity correlates positively with native essay quality. Evidentials form an exception. Turkish and native writers cite sources at similar frequencies, aligning with findings that citation practices level out once disciplinary conventions are taught (Al Rubaye, 2015). Citation mechanics may therefore be less sensitive to L1 influence than stance or engagement.

**Intra-group variation among Turkish learners**

The Turkish learner cohort shows a distinctive pattern of within-group uniformity that is nonetheless punctuated by a small set of outliers. Writer-level entropy centres around a median of 1.585 bits (M = 1.539, SD = 0.623, IQR = 0.699, range = 0.10–2.70), indicating that most essays distribute metadiscourse across several categories rather than concentrating exclusively on one. Crucially, this apparent "balance" is built largely from organizational resources rather than an even mix of organizational and stance devices.



Dispersion statistics clarify what "balanced" means in practice. Organizational cues are widely shared conventions: transitions occur in 75% of learner essays, code glosses in 69.5%, frame markers in 55.4%, and evidentials in 53.0%. By contrast, the stance palette is sparsely deployed across writers: hedges appear in 35.1% of essays, boosters in 16.1%, and self-mentions in only 11.2%. Engagement markers sit between these poles at 40.0%, showing that direct reader address is a salient but far from universal habit. In short, most Turkish writers converge on the same cohesion-first toolkit, while relatively few bring stance resources into play. Length-sensitive analyses show that intra-group differences scale the intensity of this shared toolkit rather than shifting its composition. As essays lengthen, learners increase transitions (≈ +0.5 per 1,000 words in D10 relative to the TICLE mean), code glosses (≈ +0.4), and engagement markers (≈ +0.3), whereas hedges remain below the corpus average across all deciles. The longest essays therefore amplify cohesion and reader guidance rather than adding interpersonal nuance, suggesting that variation within the group is driven more by *how much* organizational scaffolding is used than by *which* functions are combined.

Item-level evidence points to lexical concentration rather than divergent strategies. Writers who use these resources most heavily tend to recycle a narrow set of high-visibility forms—*however*, *on the other hand*, *moreover* for transitions; *for example* for code glosses; *we should* and *consider* for engagement; *could* and *might* for hedging—while comparatively few essays venture into broader lexical repertoires (e.g., adverbial hedges such as *perhaps*). Thus, individual differences mainly reflect frequency of familiar formulas, not alternative rhetorical mixes. There is, however, a small upper-tail of essays with entropy above 2.3 bits (up to 2.70) that integrate a wider range of categories, occasionally coupling stance with cohesion in the same local spans. These cases demonstrate that broader repertoires are possible within the cohort, but they are atypical. For the majority of Turkish learners, intra-group variation takes the form of denser or sparser application of a common, cohesion-oriented template rather than qualitatively different metadiscursive strategies.

Synthesizing these strands, Turkish learner prose joins a wider constellation of EFL writing that foregrounds cohesion and explicit reader guidance while downplaying calibrated stance. The magnitude varies across studies—Chinese texts tend to over-boost even more strongly (Hyland et al., 2022), whereas Arabic writers may keep both boosters and engagement cues low (Al-Rubaye, 2015)—but the direction is consistent: hedges and self-mentions remain scarce, overt transitions abundant. The TICLE data add one nuance: Turkish learners deploy engagement markers at three times the native rate, a surge not reported for Chinese, Spanish or Arabic corpora. This spike may trace back to local classroom advice that favours inclusive we and directive verbs such as consider or note, even as it discourages singular I. This distinctiveness may reflect Turkish pedagogical culture's emphasis on reader orientation and communicative clarity, values that translate into systematic deployment of inclusive pronouns and directive verbs.

The cross-study concordance suggests that the Turkish pattern is neither isolated nor idiosyncratic. EFL writers from diverse L1 backgrounds share a common tendency to secure textual fit and reader orientation through explicit cohesive signals, while struggling to match native norms for stance differentiation. The Turkish data align with that global profile yet highlight an unusually strong commitment to reader engagement. Future pedagogical interventions will therefore need to preserve the strengths of this cohesive scaffolding while systematically broadening learners' repertoire of stance markers.

This study's combination of rule-based pattern matching with supervised machine learning offers a methodological advance for corpus-based metadiscourse research. Traditional approaches relying purely on keyword searches often suffer from low precision, while manual annotation proves too labor-intensive for large-scale analysis. The hybrid system achieves 90.8% accuracy while maintaining the scale necessary for robust statistical comparison. The feature engineering approach—incorporating lexical, syntactic, contextual, and academic indicators—enables discrimination between genuine metadiscourse functions and superficially similar expressions that occur in conversational or temporal contexts. This precision proves particularly crucial when comparing L2 and native corpora,



where different error patterns might otherwise confound analytical results. Future metadiscourse research could benefit from similar hybrid approaches, particularly when investigating less-studied languages or specialized academic genres where existing detection tools may prove inadequate.

The systematic examination of metadiscourse co-occurrence patterns represents a significant advance beyond frequency-based comparisons that have dominated the field. By revealing how writers combine different metadiscourse functions within local textual contexts, network analysis illuminates the rhetorical architectures that characterize different writing populations. The discovery of contrasting organizational patterns—cohesion hubs versus stance clusters—provides insight into the cognitive and strategic processes underlying metadiscourse deployment. These patterns suggest that effective academic writing involves not only mastering individual categories but learning to coordinate multiple functions strategically within compact textual spans. Co-occurrence analysis could prove particularly valuable for developmental research, potentially revealing how metadiscourse integration patterns change as L2 writers gain experience and expertise within specific disciplinary communities.

The integration of frequency, co-occurrence, length, and diversity measures provides a more comprehensive view of metadiscourse competence than any single analytical approach could achieve. Each dimension reveals different aspects of rhetorical development: frequency shows category preferences, co-occurrence reveals functional integration, length effects illuminate strategic scaling, and diversity measures capture individual variation. This multi-dimensional approach proves particularly important for pedagogical applications, as different analytical perspectives suggest different instructional priorities. Frequency analysis identifies which categories need attention, co-occurrence patterns reveal how integration might proceed, and diversity measures indicate whether uniform or varied approaches serve learners' developmental needs.

Several limitations constrain the generalizability of these findings. The TICLE corpus represents upper-intermediate Turkish learners from relatively homogeneous educational backgrounds, potentially limiting insights into how metadiscourse patterns vary across proficiency levels or institutional contexts. Longitudinal studies tracking individual learners' metadiscourse development over extended periods could provide valuable insights into developmental trajectories and intervention effects. The automated detection system, while achieving strong validation results, may still misclassify some ambiguous cases or fail to capture subtle contextual variations that human analysts would recognize. Future research might benefit from hybrid approaches combining automated detection with targeted manual verification of challenging cases. The focus on argumentative essays within academic contexts provides controlled comparison but limits insights into how metadiscourse patterns vary across genres, disciplines, or professional writing contexts. Expanding analysis to include other academic genres—literature reviews, research proposals, conference presentations—could reveal whether the cohesion-stance divide persists across different rhetorical situations.

The study cannot definitively separate L1 transfer effects from universal developmental constraints or pedagogical influences. Detailed contrastive analysis of Turkish and English academic discourse, combined with investigation of Turkish EFL curricula and teaching materials, could explain the relative contributions of linguistic, cultural, and instructional factors to observed patterns. Cross-sectional comparison with other Turkic language speakers or Mediterranean academic cultures might reveal whether the Turkish patterns reflect broader regional rhetorical traditions or more specific national educational practices.

## CONCLUSION

This corpus-based investigation reveals a fundamental pattern in Turkish EFL academic writing: learners prioritize explicit textual organization and reader engagement while systematically underusing the interpersonal positioning that characterizes native academic discourse. This cohesion-versus-stance divide emerges consistently across frequency, co-occurrence, length, and diversity analyses, representing a coherent rhetorical philosophy rather than random developmental gaps. The findings challenge deficit-based approaches to L2 metadiscourse instruction by demonstrating that



Turkish learners possess sophisticated organizational competencies that provide genuine foundation for academic writing development. Their systematic deployment of cohesive scaffolding creates texts with clear logical progression and explicit reader guidance—achievements that merit enhancement rather than replacement.

For L2 writing theory, the study extends our understanding of developmental sequences in academic discourse acquisition while contributing the first systematic analysis of Turkish learner metadiscourse patterns. The distinctive emphasis on reader engagement distinguishes Turkish writing from other L2 populations studied, suggesting that cultural and pedagogical factors shape rhetorical development in population-specific ways. The broader implication extends beyond Turkish EFL contexts to L2 academic writing instruction generally. Effective pedagogy should identify learners' existing competencies, understand their rhetorical strengths, and design interventions that integrate rather than replace successful strategies. This approach respects cultural diversity while building the flexible metadiscourse competence necessary for international academic participation. Future research investigating how these patterns develop longitudinally and respond to targeted instruction will determine whether the cohesion-stance divide represents a stable developmental stage or a persistent characteristic requiring sustained pedagogical attention throughout academic careers.

## REFERENCES

Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. John Benjamins.

Akbaş, E. (2014). *Commitment-detachment and authorial presence in postgraduate academic writing: A comparative study of Turkish native speakers, Turkish speakers of English and English native speakers* (Unpublished doctoral dissertation). University of York.

Al-Rubaye, M. H. K. (2015). *Metadiscourse in the academic writing of EFL and ESL Arabic-speaking Iraqi graduate students* (Unpublished master's thesis). Missouri State University.

Altenberg, B., & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on computer* (pp. 80–93). Longman.

Ardhianti, M., Susilo, J., Nurjamin, A., & Prawoto, E. C. (2023). Hedges and boosters in student scientific articles within a pragmatic metadiscourse framework. *Journal of Languages and Language Teaching, 11*(4), 626–640.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing, 32*, 1–16.

Granger, S., Dagneaux, E., & Meunier, F. (2009). *The International Corpus of Learner English: Version 2*. Presses universitaires de Louvain.

Hinkel, E. (2005). Hedging, inflating, and persuading in L2 academic writing. *Applied Language Learning, 15*(1–2), 29–53.

Hu, G., & Cao, F. (2011). Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English- and Chinese-medium journals. *Journal of Pragmatics, 43*(11), 2795–2809.

Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics, 34*(8), 1091–1112.

Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing, 13*(2), 133–151.

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

Hyland, K. (2017). Metadiscourse: What is it and where is it going? *Journal of Pragmatics, 113*, 16–29.

Hyland, K., & Jiang, F. (2024). Metadiscourse: The evolution of an approach to texts. *Text & Talk, 44*(3), 411–433. https://doi.org/10.1515/text-2021-0156

Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing, 6*(2), 183–205.

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics, 25*(2), 156–177.

Hyland, K., Wang, W., & Jiang, F. K. (2022). Metadiscourse across languages and genres: An overview. *Lingua, 265*, Article 103205.

Jomaa, N. J., & Alia, M. M. (2019). Functional analyses of metadiscourse markers in L2 students' academic writing. *Arab World English Journal, 10*(1), 361–381.

Kilimci, A., & Can, C. (2008). TICLE: Uluslararası Türk Öğrenici İngilizcesi Derlemi. In M. Sarıca, N. Sarıca, & A. Karaca (Eds.), *XXII. Ulusal Dilbilim Kurultayı Bildirileri* (pp. 1–11). Yüzüncü Yıl Üniversitesi.

Kirişçi, D. I., & Duruk, E. (2022). A comparative study of metadiscourse markers in the abstract sections of research articles written by Turkish and English researchers. *Shanlax International Journal of Education, 10*(4), 101–114.

Kurt, B., & Kafes, H. (2025). The role of disciplinary enculturation in stance-taking in L2 academic writing. *Acta Psychologica, 260*, 105720. https://doi.org/10.1016/j.actpsy.2025.105720

Milton, J., & Tsang, E. S. C. (1993). A corpus-based study of logical connectors in EFL students' writing: Directions for future research. In R. Pemberton & E. S. C. Tsang (Eds.), *Studies in lexis* (pp. 215–246). The Hong Kong University of Science and Technology.

Mur-Dueñas, P. (2011). An intercultural analysis of metadiscourse features in research articles written in English and in Spanish. *Journal of Pragmatics, 43*(12), 3068–3079.

Pearson, J., & Abdollahzadeh, E. (2023). Metadiscourse in academic writing: A systematic review. *Lingua, 293*, 103561. https://doi.org/10.1016/j.lingua.2023.103561

Yee, S. W., Afzaal, M., & Aldayel, H. S. (2024). A corpus-based comparison of linguistic markers of stance and genre in the academic writing of novice and advanced engineering learners. *Humanities and Social Sciences Communications, 11*, 284. https://doi.org/10.1057/s41599-024-02757-4

[i] **Data and Code Availability**

The full analysis pipeline (pattern lists, extraction scripts, ML validator, and table/figure generators) is openly available at https://github.com/fatihbozdag/metadiscourse_analysis. The repository includes step-by-step reproduction instructions, pinned dependencies (requirements.txt/environment.yml), and fixed random seeds. We report 95% Wilson score intervals for each category's share of all metadiscourse tokens (category count ÷ total metadiscourse tokens); for compact display, these intervals are scaled ×1,000 to match the magnitude of rate columns. Per-1,000 word rates are computed separately and are not accompanied by CIs in the present analysis.

Due to licensing restrictions, the raw corpora (TICLE/ICLE-TR and LOCNESS) cannot be redistributed. Researchers may obtain them directly from the corpus providers and run our pipeline without modification. To facilitate replication, we release all derived, non-copyrighted artifacts: document-level category counts, per-1,000 rates, dispersion values, keyness outputs, co-occurrence matrices, and entropy values. Additional intermediate outputs are available from the authors upon reasonable request.