PPLING
International Applied Linguistics Research Conference
https://doi.org/10.29140/9781763711686-1

Reading Beyond Word Counts: Corpus-Assisted Analysis of Stance in Academic Writing

Eyhab Bader Eddin
Dhofar University, Oman

Abstract

In the field of applied linguistics, researchers have increasingly adopted corpus-assisted approaches, particularly in the study of academic writing. They are often, however, associated with frequency lists, word counts and quantitative comparison and therefore risk reducing academic discussion to measurable patterns and neglecting the rhetorical function, meaning in context and interpretation by the researcher. This paper explores the possibilities for rethinking research methods in applied linguistics by considering the use of corpus-assisted methods as a qualitative and interpretive method in the study of academic writing. The study is based on a small, purposively selected corpus of academic abstracts and article introductions in applied linguistics. Corpus tools are not used to measure frequency as an end in itself, but to identify recurrent stance markers, reporting verbs, evaluative expressions and phraseological patterns that contribute to argumentation and the construction of authorial stance. These patterns are then qualitatively analysed in their co-texts to discuss how the writers create research significance, how they interact with previous research and how they situate their research in the field. The paper suggests that the use of corpus evidence can help make linguistic patterns more visible, systematic and traceable while not undermining the role of the researcher as an interpreter in the context of qualitative discourse analysis. Building on established corpus-assisted discourse studies practice, the paper contributes a small-corpus, stance-focused application combining corpus retrieval, functional coding and contextual interpretation in academic writing. The study also has implications for academic writing pedagogy, particu-

E-mail: eabdulrazak@du.edu.om

In Y. Wang, J. Matthews, & B. Zou (Eds.), *Proceedings of the 2nd International Applied Linguistics Research Conference, 2026* (pp. 4–9). Castledown.

larly in helping postgraduate students and novice researchers understand the construction of stance and argumentation in published research.

Keywords: corpus-assisted discourse analysis; academic writing; stance markers; qualitative corpus analysis; small-corpus analysis

1. Introduction

Corpus-assisted approaches have gained increasing importance in applied linguistics, especially in the field of academic discourse and stance. Recent research, however, reveals that the use of corpus analysis is still very much linked to the use of frequency, concordance, collocation and keyword procedures, even when these are applied in the context of discourse-oriented research (Flowerdew, 2024; Gillings et al., 2023). Meanwhile, research on stance has expanded rapidly in various genres, and the focus has shifted from a mere lexico-grammatical perspective to a more general functional interpretation of stance and evaluation (Sun et al., 2024). Within this established corpus-assisted discourse studies (CADS) tradition, the present study examines how a small, purposively selected corpus can support systematic, context-sensitive analysis of stance in academic abstracts and introductions. It focuses on stance markers, reporting verbs, evaluative expressions and phraseological patterns as resources for rhetorical positioning rather than as numerical distributions alone.

2. Literature Review and Theoretical Framework

Corpus-assisted research is a method that enables the researcher to shift between recurring patterns in the language and their rhetorical function through a process of contextual examination of discourse. While corpus analysis has long been characterised by a focus on frequency, keywords, collocation and concordance patterns, there is recent methodological research that highlights the compatibility of corpus analysis with interpretivist qualitative inquiry, especially when the evidence from the corpus is used as a tool for guiding rather than dictating interpretation (Gillings et al., 2024).

Stance is a term that is used in academic writing research as a general category for the linguistic means writers use to demonstrate commitment, evaluation, attitudes and relationships with disciplinary knowledge and readers. More recent corpus-based research has shown that stance varies according to genre, disciplinary context, proficiency and rhetorical purpose. For instance, epistemic positioning through hedges and boosters can be used to show how

writers manage their level of certainty and commitment to their scholarship (Yang & Guo, 2025), and reporting verbs can show previous scholarship as well as allow writers to evaluate and position themselves in relation to cited claims (Deng & Li, 2026). This type of finding indicates that the stance markers cannot be fully understood as lexicalised, decontextualised items.

This study thus takes an interpretive corpus-assisted approach which combines quantitative visibility and qualitative explanation. While concordance patterns reveal potentially important language choices, their significance is established through detailed analysis. This is the way of using evidence from the corpus as a systematic starting point for discourse interpretation, instead of as an end point, mainly based on numerical distribution.

3. Methodology

The corpus comprised 30 English-language empirical research articles published between 2022 and 2025 in peer-reviewed applied linguistics journals. Abstracts and introductions were extracted, producing approximately 55,000 words. Articles were purposively selected for disciplinary relevance, empirical orientation and the availability of both sections for contextual stance analysis; the sample was intended for close qualitative analysis rather than statistical generalisation.

Stance was coded following Hyland's (2005) framework, particularly hedges, boosters, attitude markers and self-mentions; reporting verbs and evaluative phraseology were analysed separately when they contributed to authorial positioning. Categories were non-exclusive: polyfunctional items such as *suggest* and *demonstrate* were coded by contextual function and could receive more than one label when both functions were simultaneously realised. Concordance lines, recurrent lexical and phraseological patterns and relevant co-text were obtained through AntConc (Version 4.3.1; Anthony, 2024), a corpus analysis and concordancing software tool. Instead of focusing on frequency as the main analytical result, the recurrent forms were analysed within their immediate and extended textual contexts.

The analytical process comprised three phases: first, identification of possibly salient stance expressions; second, analysis of the concordance context of these expressions; and third, classification of the rhetorical function of the expressions with respect to the significance of the research, the engagement with previous research and the positioning of the current research. The small-corpus design was selected as it provides a combination of systematic pattern identification and the depth of interpretation needed in discourse-sensitive applied linguistics research. All coding was conducted by the author; no inde-

pendent intercoder reliability check was undertaken, and functional classification therefore remained subject to researcher interpretation.

4. Findings and Discussion

The corpus-assisted analysis identified three common rhetorical uses of stance: creating research significance, positioning prior research and setting up the contribution of the current research. These functions were achieved not by means of single lexical items, but by means of combinations of hedges, boosters, evaluative expressions and reporting verbs. This helps to validate the idea that stance should not only be examined based on frequency distributions, but also in relation to the rhetorical context.

A frequency check showed that hedging expressions occurred more frequently than boosters across the corpus, with *may*, *suggest* and *appear* among the most recurrent stance items. Reporting verbs such as *suggest*, *argue* and *demonstrate* also appeared repeatedly across abstracts and introductions.

One of the most salient features was the use of hedging devices like *may*, *suggest*, *appear* and *to some extent* when writers assessed previous findings or put forward interpretations. These forms often moderated the writers' level of commitment, but did not necessarily prevent authors from making claims that are not universally valid. In contrast, the use of boosters like *clearly*, *demonstrate* and *show* was found to occur when writers set up the significance of a problem and/or the relevance of their own study. The contrast suggests that epistemic strength was not consistently used throughout texts, but was rather strategically adjusted with respect to the rhetorical purpose.

Reporting verbs also were a significant factor in the positioning of writers in relation to previous research. Verbs like *argue, suggest, report* and *demonstrate* not only attributed information, but also implied varying levels of endorsement and evaluative distance. For instance, *suggest* allowed writers to acknowledge previous research cautiously, while *demonstrate* provided a more authoritative stance when citing results. This is especially apparent when concordance lines are examined alongside their wider co-text.

The third pattern was in the evaluative and phraseological expressions that were used to formulate research gaps and significance, such as *little attention has been paid to*, *remains underexplored* and *the present study addresses*. These expressions often occurred in rhetorical sequences whereby the previous studies were first recognised, a limitation was identified, and the current study was then introduced as a response. Thus their significance was not just one of repetition but one contributing to argumentative progression. This sequence also parallels Swales's (1990) Create a Research Space (CARS) model, partic-

ularly the movement from establishing a research niche to occupying that niche.

The results here demonstrate the methodological benefits of using a combination of corpus retrieval and close discourse analysis. The frequency and concordance searches reveal recurrent patterns and the contextual interpretation elucidates how these patterns work rhetorically. The analysis thus backs up the idea that quantitative evidence should be used to guide attention and not to determine meaning in an interpretive corpus-assisted approach.

5. Methodological and Pedagogical Implications

Methodologically, the study illustrates the utility of small-corpus analysis in strengthening qualitative research by making recurrent patterns of discourse more visible, traceable and subject to systematic analysis. This approach does not treat corpus evidence as a replacement for interpretation, but rather as a complementary step in the process of analysis that can be done through computational retrieval and close reading. This can make the methodology more transparent in applied linguistics, especially in research related to rhetorical meaning, the positioning of the writer and context-sensitive language use.

The results also have implications for the teaching of academic writing. Stance is often introduced to postgraduate students or novice researchers by means of lists of hedges, boosters and reporting verbs that are decontextualised. Instead, corpus-assisted activities can provide learners with authentic patterns of use and prompt them to analyse how published writers calibrate claims, evaluate previous research and construct research significance. This can contribute to increased rhetorical awareness, as learners will not only be able to identify the stance resources used in academic writing, but also the reasons behind the choice of forms in certain argumentative contexts.

6. Conclusion

This study has demonstrated that corpus-assisted analysis can be used to support qualitative applied linguistics research to make recurrent stance patterns visible without reducing their significance to frequency alone. Its small, purposively selected corpus limits statistical generalisability, and the interpretive coding is still subject to researcher judgement. These limitations are, however, in line with the goal of the study, which is to focus on the contextual depth and methodological transparency. The paper illustrates the use of the small-corpus approach in combination with close discourse analysis, which can sup-

port more systematic interpretations of academic stance while offering practical value for research methodology and academic writing pedagogy.

References

Anthony, L. (2024). *AntConc* (Version 4.3.1) [Computer software]. Waseda University. https://www.laurenceanthony.net/software/antconc/

Deng, L., & Li, X. (2026). Reporting verbs across genre sets and research paradigms: A corpus-based analysis of denotation and evaluation in applied linguistics research articles. *Frontiers in Psychology, 17*, 1799996. https://doi.org/10.3389/fpsyg.2026.1799996

Flowerdew, L. (2024). Corpus-based ESP genre analysis and corpus-assisted discourse studies (CADS): Exploring connections. *Journal of Corpora and Discourse Studies, 7*(1), 22–42. https://doi.org/10.18573/jcads.124

Gillings, M., Learmonth, M., & Mautner, G. (2024). Taking the road less travelled: How corpus-assisted discourse studies can enrich qualitative explorations of large textual datasets. *British Journal of Management, 35*(4), 1667–1679. https://doi.org/10.1111/1467-8551.12816

Gillings, M., Mautner, G., & Baker, P. (2023). *Corpus-assisted discourse studies*. Cambridge University Press. https://doi.org/10.1017/9781009168144

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173–192. https://doi.org/10.1177/1461445605050365

Sun, Y., Yang, K., & Chen, G. (2024). The research trends of corpus-assisted stance research (2004–2023): A systematic literature review. *Current Psychology, 43*(48), 37167–37182. https://doi.org/10.1007/s12144-024-07130-x

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

Yang, Y., & Guo, X. (2025). "We find that..." changing patterns of epistemic positioning in research writing. *Frontiers in Psychology, 16*, 1634848. https://doi.org/10.3389/fpsyg.2025.1634848