

AUTOMATING THE ANALYSIS OF REPORTING VERBS IN ACADEMIC CITATIONS: AN NLP-ASSISTED CORPUS STUDY

FAIZA ABID¹, AYESHA JUNAID², DR. ZAFAR IQBAL BHATTI³,
KHURRAM SHAHZAD⁴

¹ASSISTANT PROFESSOR, DEPARTMENT OF LINGUISTICS AND COMMUNICATION, UNIVERSITY OF
MANAGEMENT AND TECHNOLOGY, LAHORE, PAKISTAN, faiza.abid@umt.edu.pk

²ASSISTANT PROFESSOR OF LINGUISTICS, DEPARTMENT OF ENGLISH LANGUAGE AND LITERATURE
COLLEGE OF SCIENCES AND HUMANITIES, PRINCE SATTAM BIN ABDUL AZIZ UNIVERSITY, AL-KHARJ
11942, KINGDOM OF SAUDI ARABIA, a.sarfrazkhan@psau.edu.sa

³PROFESSOR OF LINGUISTICS, SCHOOL OF ENGLISH, MINHAJ UNIVERSITY, LAHORE, PAKISTAN,
ravianz902@gmail.com

⁴APPLIED LINGUISTICS RESEARCHER & ASSISTANT LANGUAGE TEACHER (ALT), TOKYO, JAPAN

Abstract

Verbs that are included in reporting like to argue, suggest, claim and demonstrate are the main part of academic writing that can allow authors to command the knowledge, make judgement, and place positions in the conversation concerning disciplinary issues. Although the idea of reporting verbs as resources of stance in citation situations has long been studied through the application linguistics discipline, the prevailing analytic traditions are labor intensive and only scalable to large and heterogeneous corpora in general (Hyland, 1999, 2002; Thompson and Ye, 1991). Recent methods in natural language processing (NLP) provide practical targets to automate part of this analysis, thus being able to make reporting-verb behavior descriptions more systematic, replicable, and large-scale (Jurafsky and Martin, 2023; Teufel and Moens, 2002). This paper is a corpus-based research on demonstrating a corpus approach to reporting verbs in citation sentences and involving both quantitative method of distributional analysis and qualitative methodology of interpretation of rhetorical purpose. It uses a large computational linguistics citation dataset (the ACL Anthology Network resources) by combining it with citation sentiment annotations as well as explore the distribution of verb choice on positive, negative, and neutral citation situations (Athar, 2011; Radev, Muthu Krishnan, Qazvinian, and Abu-Jbara, 2013). A frequency profiling and polarity-conditioned comparison as well as expert validation are both supported by an NLP pipeline consisting of tokenization, part-of-speech tagging, stemming and verb-phrase extraction, and gives insight into the quality of automated extraction. The results, formulated during the present paper show that the automated processing could retrieve effective reporting verb patterns at scale and introduces logical variation attributed to evaluative contexts, and also persists in revealing difficulties facing NLP in discerning polyfunctionality, hedging and context dependent rhetorical strength.

KEYWORDS: natural language processing, traditional verb representing reporting, academic writing, citation analysis, corpus linguistics, rhetorical impact, sentimental of scientific exposition.

1. INTRODUCTION

1.1 Reporting verbs as a site of attribution, stance, and disciplinary identity

Scholarly writing plays a dual role; information as well as interpersonal roles. Besides the content deliverable, it creates credibility, reflects, or provides distance against other existing scholarship, and directs the reader through the contentious or doubtful knowledge. Reporting verbs are some of the most visible leico- grammatical means used by the authors to achieve these functions in the context of referencing. A writer can ensure modulation in the frame of a proposition cited by replacing argue by note, demonstrate by suggest, claim by show, etc. (Hyland, 2002; Thompson and Ye, 1991). In turn, the reporting verbs have been considered one of the key areas where to examine the stance, evaluation, and disciplinary rhetoric in English to Academic Purposes and applied linguistics. Classical research in applied linguistics has shown that the verb of reporting is not an act of reporting information. They encode appraisal orientations and bring about epistemic relations between the citing and cited voices. As it was demonstrated by Thompson and Ye (1991), the verbs of reporting in academic prose may have an evaluative potential thus affecting an interpretation of the quoted text. Hyland (1999, 2002) also argued that there is a systematic difference in reporting practices across fields and are embedded in the norms through which fields

negotiate the nature of evidence and poor presentation of claims. Based on this, reporting verbs are both linguistic decisions and community-indexing cues: they contribute to the formation of a scholarly self, and, at the same time, influence the meaning-making of an already existing research.

1.2 The Methodological Bottleneck: Manual Analysis and the Limits of Scale

Although it is known that reporting verbs are essential features in analyzing a case, manual classification, coding and interpretation has often been relied upon to analyze them. Manual means cannot be done without in the domain of fine-grained rhetorical understanding, but subject to impractical restrictions, which have proven to be problematic:

1. They are also time consuming thus restricting the magnitude of datasets.
 2. They cannot be easily scaled because they require much training and elaborate code manuals to reproduce.
 3. They are prone to inconsistency in case borderline cases are perceived differently by different analysts.
 4. They normally limit cross-disciplinary comparisons to fairly small samples
- These restrictions are significant as reporting verbs are not only frequent but also variable; using the same lemma in different syntactic structures and different rhetorical tasks may take different rhetorical work in different contexts by the influence of the words surrounding them and the section of the genre and disciplinary norm. Patterns at a large scale may therefore be invisible when it is analyzed in small groups of texts. Simultaneously, methodological scaling up also runs the risk of reducing reporting verbs to counts, without having rhetorical content. The overarching NLP claim in this regard is that it does not aim to substitute the linguistic interpretation but to automate the mechanical process, i.e., to find tokens of candidate verbs, normalize morphological variants, extract co-text, that allows the analyst to commit more attention to theoretically consequential interpretation and close validation.

1.3 Why NLP is Relevant: Automation with Interpretability

The natural language processing provides a rich set of linguistic analysis tools that are both wide and reproducible. Components such as tokenization, part-of-speech tagging, lemmatization, and dependency parsing are used to aid automatic recognition of verbs within a citation environment to allow comparative analytical studies of a single sentence at a time to compare thousands of citation sentences. Furthermore, machine-learning and neural-network methods do offer workable classification mechanisms that cannot be done reliably manually in large scale, e.g. the sentiment of labeling a citation or the sentiment of citation-function (Athar, 2011; Teufel, Siddharthan, and Tidhar, 2006). Most of them can be interpreted with respect to linguistically significant units, lemmas, reporting frames, syntactic relations, instead of opaque feature sets, which is significantly important to applied linguistics. This approach is taken in the current research since NLP is used to facilitate corpus-based description and provide a certain coverage level that would be impossible to reach without manual analysis. And at the same time the research sets the reporting verbs as rhetorical tools the role of which cannot be explained using individual tokens; the interpretation will require consideration of situational, co-contextual, and disciplinary resources.

1.4 The Role of Evaluative Context: Citation Sentiment as an Analytic Lens

Another interface between natural language processing and applied linguistics that is especially useful is provided by the studies of scientific citation sentiment and citation purpose. Evaluative cues in the field of scientific composition are often subtle, diffused and weakened; nonetheless, the citations can still either be positive, negative or neutral in relation to the work cited (Athar, 2011). Sentiment annotations are useful in the event that there are sentiment verbs that can be used to examine reporting verbs as possible lexical predictors of assessment and attitude. To illustrate, such verbs as demonstrate and show can be expected to occur more often in supportive context, where the citing author presents the cited work as empirically strong, whereas such verbs as claim can be found in the context of detachment or opposition, although such generalizations require rigorously corroborated contexts, as verbs often are governed by adverbials or by modals and negation or by acknowledging the opposing view (Hyland, 2002).

Citation sentiment is therefore used in this study as a structured contextual variable: not only are the frequencies of reporting verbs considered at the global level but also in positively constructed, negatively constructed, and neutrally constructed citation environments. Such a methodological strategy, therefore, makes it possible to analyze beyond mere observations of what verbs are most frequently used but asks the question of how the selection of reporting-verbs and where they are placed in assessments interact.

1.5 Research Objectives and Research Questions

The study retains the following research objectives:

- To analyze the frequency of different reporting verbs used in citation texts and identify patterns in their usage.
- To utilize Natural Language Processing techniques to identify the type of sources being cited based on the language patterns associated with different reporting verbs.
- To study the context in which reporting verbs are used in citations to understand the relationships between the cited material and the author's argument or research.

The study retains the following research questions:

- What are the most frequently used reporting verbs in citation texts across different academic disciplines?
- How do reporting verb usage patterns differ between different types of academic sources (e.g., journal articles, books, conference papers)?
- What are the predominant reporting verbs used in citations to indicate support for the cited material, and how do they vary across disciplines?

Because the empirical dataset used in this article is drawn from computational linguistics, disciplinary comparison is addressed in two complementary ways: (a) by treating the computational linguistics corpus as a demonstrator of method and a detailed case study of one disciplinary domain, and (b) by situating observed patterns against established applied-linguistics findings on disciplinary variation in reporting practices (Hyland, 1999, 2002; Swales, 1990). This two-pronged approach will retain the initial questions, but claims made empirically will be adequately supported by evidence presented in the dataset.

2. REVIEW OF THE LITERATURE

2.1 Reporting Verbs and the Rhetorical Architecture of Academic Discourse

Reporting verbs are as though at a gray area between grammatical and rhetorical action. The functional perspective sees them as achieving the attribution task - of making it explicit, directorate and content - and, at the same time, encode interpersonal nuances by foregrounding the speaker attitude and by assessing this attitude. The applied-linguistic literature has in fact revealed all along that reporting verbs are a vital element within more general systems of metadiscourse, as through which writers reckon the reader, forestall opposition and offer negotiable space to fit with institutional norms (Hyland, 1999; 2002). The genre-analytic model of Swales (1990) is no different; in that it also foregrounds the social interests which organize scholarly writing, it explains how an author provides a research space by placing his/her contribution against the already existing academic work.

A classic study by Thompson and Ye (1991) of evaluative meaning in reporting clauses is an example that demonstrates that the evaluative verbs may be differentiated based on what is the degree of intensity of the assessment they represent, and what reporting frame they license. This theoretical implication has rather important pedagogical consequences: in a new writer, the evaluative range of the reporting verbs is often exploited insufficiently, or the verb forms that are used imply unwanted positions by accident. In turn, the application of linguistics is turning towards reporting verbs as a discourse resource rather than a grain of lexicon and thus, requires the development of the rhetorical competence.

Hyland (2002) carried on this research question by suggesting a functional taxonomy of the connection between reporting verbs and particular rhetorical functions of academic prose. Hyland model though various other categories have developed, is still widely used because it categorically established the relationship between verb selection and communicative actions in research speech. In its simplified version, the reporting verbs are grouped based on the role it plays in reporting focus on either research activity, positioning of the cognitive role, or discourse/ argumentation role. This taxonomy has certain strengths especially to corpus-based studies since it provides them with categories which can be language comprehensible as well as practically pedagogically.

Table 1. Hyland-style functional groupings of reporting verbs (illustrative examples)

Functional grouping	Typical rhetorical focus	Illustrative reporting verbs
Research acts	describing research procedures or findings	<i>observe, measure, report, find</i>
Cognition acts	representing belief, inference, or interpretation	<i>believe, assume, consider, suggest</i>
Discourse acts	constructing argumentation and claims in discourse	<i>argue, claim, contend, state</i>

Note. Examples are illustrative and may vary by discipline and analytic scheme (Hyland, 2002).

This is important to the current research both due to classification and to its methodological implication: since reporting verbs are rhetorically charged and polyfunctional automated methods need to be indexed to linguistically motivated categories and checked by the expert mode.

2.2 Corpus Based Studies and Disciplinary Variation

Corpus-based methods have repeatedly established the variation in reporting verbs by the discipline, and also by the rhetorical situation in texts (e.g. introduction vs. discussion). The corpus results presented by Hyland (1999, 2002) indicate that the humanities and social sciences are likely to use more evaluative and interpretive reporting verbs, but the sciences can be more inclined to use the verbs that reflect the findings as seen or proved. These structures represent further disciplinary epistemologies: interpretive disciplines have to negotiate claims by means of arguing and placing them in place, but many empirical sciences perambulate observation, measurement, and presentation of evidence.

It is important in the study that discipline variation plays key roles in the research questions of the study as it will encourage inter-field generalizations and warn of generality attempts between reporting-verb list obtained in one domain and another with no current adaptations. In terms of methodology, that too suggests that any NLP-based

system to be used cross-disciplinarily should be robust or domain adaption and domain calibration procedures must be included.

2.3 NIP Methods of Scientific Text, Rhetoric, and Citations

The scientific writing NLP research has elaborated on an extensive array of techniques of capitalizing on rhetorical structure and citation situations. The first research done on the summarization of scientific articles factored in the rhetorical status, as well as relevance modeling and this showed that scientific texts were not a series of sentences in a sequence, but structured rhetorical artifact (Teufel and Moens, 2002). The view was subsequently used in the division of tasks, like citation function classification, in which the systems anticipate the usage of a citation (background, contrast, use, extension, critique). A famous model of the automatic categorization of the citation function is by Teufel, Siddharthan, and Tidhar (2006) who say that there are linguistic indicators of rhetoric categorization of citation situations, such as a verb.

The computational methods are specifically applicable to the verbs of reporting, since they make the intuition that citations themselves are rhetorically functional, rather than bibliographic indicators. In identifying patterns in reporting frames, NLP models can facilitate large scale applied linguistic enquiries concerning how scholars project knowledge, how they identify with the sources and how they structure argumentative space.

2.4 Citation Sentiment Analysis in Scientific Discourse

Citation sentiment analysis is aimed at identifying the scale of positive, negative, or neutral attitude towards the work, which is referenced. Whereas product reviews, scientific sentiment is usually moderate, diffused in the opinion of technical commentary and embedded in intricate rhetorical forms. One of the most frequently used frameworks in scientific citation sentiment has been proposed by Athar (2011), with an emphasis put on the aspects of sentence structures and the peculiarities of a scientific evaluation.

Citation sentiment provides a formal contextual aspect into the analysis of the reporting verbs. With sentiment labels available, distributions of reporting verbs may be analyzed based on raw frequency, as well as polarity conditioned frequency and association. This allows to conduct a more specific empirical investigation: what are the reporting verbs that are overrepresented in supportive and underrepresented in critical citations, and what rhetorical roles they play in those two instances? This one cannot be a more or less direct mapping of verb to sentiment, instead it will have to be interactions between verbs, syntactic frames and evaluation review markers on their surround (Hyland, 2002).

2.5 Objections that Drive the Current Research

According to the literature, some of the repetitive issues include the intersection of applied linguistics and NLP:

1. **Context dependence and polyfunctionality:** this is a potential risk of overthrowing complexity; as the verbs used to report can vary between frames and co-texts and cause problems in automated systems.
2. **Domain sensitivity:** the reporting practices with disciplines differ so that they may restrict generality of adapting models without changing them.
3. **Subtlety of evaluation:** the stance of evaluation in academic writing is often indirect and hedged and verbs are connected with the modals, negation and concessive constructions.
4. **Annotation scarcity:** good annotated datasets are scarce and expensive to obtain, which requires effective and transparent evaluative effort.

The current work attempts to provide solutions to these issues by combining automated extraction and corpus profiling with expert validation and qualitative contextual verification, hence the attempt to avail both computational and interpretive scalability.

3. METHOD (OVERVIEW)

3.1 Corpus and Dataset Basis

The empiricist basis of the research includes the corpus of the citation sentence that are ripped out of the ACL Anthology Network as well as the sentiment labels that denote the words being positive, negative, or neutral and are used to compare distributional features (Athar, 2011; Radev et al., 2013). Table 2 outlines the size of the dataset and the composition of sentiment polarity, as found in the study. The predominance of the neutral citations can be explained by the mainstream opinion that most academic references serve merely either to present information or to create a positioning, and not to write overt evaluative judgment.

Table 2. Distribution of Citation Sentences by Sentiment Polarity

Sentiment polarity	Number of citation sentences	Percentage (%)
Positive	829	9.49
Negative	280	3.20
Neutral	7,627	87.31
Total	8,736	100.00

3.2 NLP Preprocessing and Extraction Logic

The paper uses an NLP pipeline that consists of four major steps to analyze the use of reporting verbs in the context of citation:

1. **The breaking down** of the sentences into citation into analyzed units (Jurafsky and Martin, 2023).
2. **Part-of-speech tagging**, an event that detects the instances of verb and candidate reporting verb.
3. **The normalization of inflection of verb forms**, as a frequency profiling normalization to a single lemma (e.g., argues); argued) to lemmata (e.g., argue): Lemmatization).
4. **Verb and verb-phrase extraction**, that is, commends a verb inventory and calculates frequencies in terms of the whole dataset and per sentiment.

This pipeline supports two kinds of later analyses used in the article, the first, (a) overall frequency profiling with the aim of identifying very recurrent reporting verbs in citation sentences, and the second, polarity-conditioned profiling in order to bring out verb clusters in supportive, critical, or neutral citation circumstances.

3.3 Expert Validation Strategy

Automated extraction will be tested based on the expert validation of the random set of citation sentences. Human annotators find verb occurrences and are then compared to the automated outputs through a comparison approach of a confusion matrix. The validity of reporting secondary is necessary since reporting-verb analysis is prone to tagging errors, lemmatization errors and multi-word verb boundary decisions. Later denotes of accuracy later in the article is construed in accordance to the vague and compact format of citation sentences that often much of a complex rhetorical work into a minimal geographical span.

The element of methodology provided above develops the basis of the following parts, introducing the verb frequency profile, discussing the reporting-verb distributions among sentiment categories and interpreting the distributions concerning the rhetorical function and academic position.

3. METHODOLOGY

3.1 Data Source, Sampling, and Unit of Analysis

The unit of analysis is the **citation sentence**, defined as a sentence containing an in-text reference to a cited work and extracted as a self-contained context for sentiment and verb analysis. Citation sentences were drawn from the ACL Anthology Network (AAN), a curated citation network for computational linguistics research (Radev, Muthukrishnan, Qazvinian, & Abu-Jbara, 2009; Radev, Muthukrishnan, Qazvinian, & Abu-Jbara, 2013). Sentiment labels (positive, negative, neutral) follow a manual annotation scheme for scientific citation sentiment, where polarity reflects the citing author's evaluative orientation toward the cited work (Athar, 2011). The analysis set contains 8,736 annotated citation sentences distributed across the three polarities (Table 2).

Two implications follow. First, the dataset supports **within-corpus** comparison of verb distributions across evaluative environments. Second, because AAN is centered on computational linguistics, the corpus functions as a disciplinary case study. Broader disciplinary contrasts are therefore discussed by triangulating with applied-linguistics findings on citation practices (Hyland, 1999, 2002).

3.2 Operational Definition of “Reporting Verb” In the Present Study

The use of the term reporting verbs in applied-linguistics research is traditionally used to refer to frames used in attributing to a statement made by another (e.g., X argues that). In citation-sentence corpora, verbs in a sentence can be attitudinal, as well as methodological or action verb (e.g., we use X, they train Y). In this regard, operationally definite are reporting verbs as verb lemmas derived by citation sentences and considered potential lexical markers of citation practice and stance. This operationalization allows scale profiling although it allows a

path to fine-tuning since future research can restrict the extraction to attribution frames by taking advantage of dependency relations and citation-span mapping (Teufel, Siddharthan, and Tidhar, 2006).

3.3 Decisions on NLP Pipeline and Preprocessing

The sentences containing citations had been filtered through a various popular NLP pipeline that enabled the normalization and identification of verbs in the sentences. Table 3 provides a list of every step and the purpose of its analysis.

Table3. NLP Pipeline For Citation-Sentence Processing And Verb Extraction

Step	Output	Purpose for reporting-verb analysis
Sentence normalization	cleaned citation sentence	reduces noise from encoding artifacts and inconsistent spacing
Tokenization	token sequence	supports POS tagging and feature alignment (Jurafsky & Martin, 2009)
POS tagging	token + POS label	identifies verb candidates and verb-phrase heads
Lemmatization	lemma for each verb token	collapses inflectional variation for frequency profiling
Verb/verb-phrase extraction	lemma inventory + counts	enables distributional analysis across sentiment categories

Lemmatization is important in promoting interpretability: as opposed to the analysis of the three morphological forms of the word argue, argue, and argued, they are grouped together into a lemma.

3.4 Frequency Profiling and Polarity Conditioned Distribution

This is done after extraction by computing (a) the overall frequency of lemmas in the entire data and (b) the frequency of the polarities given conditioning positive, negative, and null citation sentences. The study computes a conditional percentage, namely, each sentinel v and sentiment s :

$$P(v | s) = f(v, s) / (\sum f(v_i, s)) \times 100$$

Function $f(v, s)$ refers to the count of lemma in each citation sentence with sentiment s . The formulation makes comparisons across classes of sentiment possible when it is true that there are significant differences between the absolute count of sentences in the classes. The fact that the corpus is highly imbalanced in the number of neutral citations is met by the fact that polarity conditioned percentages are read alongside raw counts and verbs that have been conditioned more than usual are discussed in context and then definite rhetorical determinations can be made.

3.5 Identification of The Source-Type and its Limits in The Existing Dataset

Among the goals of the present study, it would be possible to distinguish between the various types of sources, e.g., journal articles, books, and conference papers by using the language patterns with reporting verbs. This can theoretically be done by combining (a) metadata information, such as venue and type of document, with (b) linguistic models of citation -sentence patterns, which are correlated with these types of sources. Publications in the domain of computational linguistics, especially conference and journal articles, are predominant in AAN collection (Radev et al., 2013). As a result, the current data would grant more discriminatory force in differentiating practices between the usage of conferences and journals to cite the books, but the sample size of the books is limited, and, thus, the analysis can be completed in a less systematic way.

3.6 Expert Tests on Automated Verb Extraction

Three random samples of 100 citation sentences would be chosen to assess the automated extraction pipeline and linguistic experts would label them with all verb occurrences. The outputs that were obtained with the automated system were cross-compared with the expert reference to measure extraction performance. Total extraction accuracy - was considered as the number of the percentage of verb instances identified correctly and this overall accuracy was simply 88.25 that is viewed as a factor of reliability when citation-sentence conditions are realistic.

Table4. Validation Design for Automated Verb Extraction

Component	Description
Sampling	3 random samples $\times$ 100 citation sentences
Gold reference	manual verb annotation by English-language experts
Comparison basis	overlap between automated extractions and manual identifications
Summary metric	overall agreement (accuracy) = 88.25%

3.7 Verb-Phrase Handling and Normalization Rules

Mixtec and chains of auxiliaries (e.g. has been shown) are frequently seen in citation sentences containing multiword verb phrases (e.g., be based on): When building frequency inventory, the analysis documentation captures the first lexical verb lemma as an item of counting, and auxiliaries and particles are stored in a context field to be verified later, qualitatively. In agreement between comparability across realizations (e.g. show vs shown) and sensitivity to beside by passives, modals and negation, this approach follows. In cases where they are cited, such as after nouns, e.g., X, as demonstrated by (Y), et cetera, the normalization process includes the citation, but does not count it as having a lemma, so reference-format artifact does not affect the frequency determination.

3.9 Qualitative Content Analysis

America quantitative profiles are able to better select the salient verbs of candidates; the rhetorical understanding implies that it will require a context. To this end, the paper challenges representative citation sentences to pinpoint (a) the high-frequency verbs as well as (b) the verbs bearing relative proportional change with regard to sentiment categories. Cross-targeted verbs are analyzed by looking at contextual examples of each target verb in terms of recurring frame (e.g., reporting clause + that; passive evidential), evaluative lexis co-occurring with it and contrastive elements (however). This process corresponds to the conventional corpus linguistic practice: frequency data are used to shed light upon discourse routines, and the analysis of contextual reading is used to assess their rhetorical viability (Biber, Conrad, and Reppen, 1998; Hyland, 2002). Such relative infrequency of negative citation causes any verb which groups in negative category to be treated as a tentative indicator instead of a constant marker until validated in a variety of contexts.

Table5. Extraction And Interpretation Challenges in Citation Sentences

Challenge	Example (schematic)	Risk	Mitigation
Auxiliary evidentials	"It has been shown that ..."	stance may sit in the auxiliary pattern	store auxiliary pattern for review
Negation/concession	"X does not show ...; however ..."	local evaluation may reverse	inspect co-text before inference
Verb-noun ambiguity	"the use of X" vs "we use X"	POS errors inflate verbs	sample-based manual validation
Citation interruption	"X, as proposed in (Y), ..."	mis-attachment in parsing	keep citation span for re-checking

4. Data Analysis and Findings

4.1 Size of the extracted verb inventory

The automated analysis of the entire context of 8,736 citation statements found 389 different verb lemmas in the citation context. The lexical variety of the subject of computational linguistics is emphasized by this inventory. Besides classical attribution verbs like argue, suggest and propose, the data have methodological verbs like use, train and implement which define how scholars place cited research as a resource, baseline or comparative benchmark.

4.2 The Most Common Verbs in Citation Sentences

Three verb lemmas that were the most common in the dataset-level analysis are listed in Table 5. The dominance of the lemma use explains an apparent tendency in computational linguistics: typically, citations go together with assertions about the uptake of methods, resources, or models. The relatively high tracking rate of the lemma train and tag are another reflection of the methodological nucleus underlining the research in natural language processing.

Table6. Top verb lemmas by frequency in the citation-sentence corpus

Rank	Verb lemma	Frequency
1	use	3,425
2	train	1,285

Rank	Verb lemma	Frequency
3	tag	792

Such high-frequency lemmas mainly locate the citation-as-use practices and in the latter, the word being referenced is placed as an instrument or methodological tool. The point is consistent with previous citation-function studies, which show that a lot of citations are aimed at helping to adopt a method or engage in a comparative analysis, but are not intended to convey an overt approval of a method or disapproval (Teufel et al., 2006).

4.3 Frequency Interpretation: Reporting, Attribution and Uptake

The use of the methodological verbs predominately reveals a characteristic of a discipline: in the worlds of technology, the citation sentences regularly contain the attribution sentences, as well as the procedural integration of attributions. The reporting role is spread throughout the clause structure with the uptake being expressed by the verb (e.g. use, apply) and the attribution is commonly achieved by use of citation parentheses or authoritative attribution. However, even in cases when classical attribution verbs have lower frequencies, these forms still have a rhetoric effect. Suggest, propose, or show are also more assuasive verbs because they indicate the level of epistemics of the quoted assertion. Resultantly, polarity-conditioned distributions are particularly informative in that they indicate the occurrence of evidential-strength verbs in positive contexts but they indicate distance marking verbs occurrence in negative contexts (Athar, 2011; Hyland, 2002). The analytical step that follows thus asks questions of distribution of extracted verbs among positive, negative, and neutral citation environments and it further finds verbs that exhibit strong proportional changes of the environments.

4.4 Behavior of Reporting-Verb in Categories of Citation Sentiment

After constructing verb inventory, the next stage will be to analyze how the verb is used in the three sentiment polarities i.e. positive, negative and neutral. Due to the extreme imbalance of classes of polarity (see Table 2), the interpretations should be based on the two opposing approaches (a) raw frequency which account for overall prominence and (b) polarity-conditioned proportion which explain whether a verb is relatively more typical of a given evaluative condition (Athar, 2011).

The polarity conditioned profiling gives two sweeping inclinations. First, most of high frequency verbs retain prominence in all three polarities, which is a sign of regular citation behavior in the computational language field. Second, another collection of fewer verbs has other proportional changes, which appear more commonly in the positive or negative form than the neutral form, and thus is an indicator of a more pronounced affiliation to the roles of evaluating and rhetorizing in citation sentences.

With scientific evaluation generally being restrained in its affect, polarity conditioned shifts have to be viewed as tendencies and not deterministic. Specifically, verb choice tends to be modulated by co-textual characteristics as hedges, connectiveness, or negation; a verb that might seem to cluster into certain polarity can therefore cluster by-accident as a by-product of some common rhetorical arrangement as opposed to reflecting certain inherent lexical polarity (Hyland, 1999; Thompson and Ye, 1991).

Table7. Observed Polarity Sensitivity in Citation-Sentence Verbs (Pattern Summary)

Verb group (functional tendency)	Representative verbs mentioned in the corpus narrative	Typical rhetorical contribution in citation contexts	Sensitivity to polarity (tendency)
Evidential/validation verbs	<i>show, demonstrate</i>	projects evidential strength; frames cited work as supportable	often higher in positive contexts
Performance/achievement verbs	<i>achieve, improve, perform</i>	indexes evaluation via outcome language; highlights gains or superiority	often higher in positive (and sometimes negative when used contrastively)
Method uptake verbs	<i>use, train, tag</i>	frames cited work as resource, tool, or procedure	high across all polarities; primarily neutral -dominant by volume
Claim/positioning verbs	<i>argue, suggest, propose, claim</i>	constructs propositional stance; positions cited claims	polarity depends strongly on framing and modifiers

Note. The patterns in the study narrative are concisely summarized in the table, with some of the verbs (e.g., achieve, improve, perform, show) being stated to demonstrate a stronger change in the European sentiment categories. Contextual reading in any case is needed to affirm rhetorical action and evaluative orientation (Athar, 2011; Hyland, 1999).

This trend correlates with an important observation of applied linguistics as well as citation Sentiment work, namely, that evaluation in writing is often achieved by creating a strategic framing rather than by using explicit sentiment vocabulary as such. Such framing is usually achieved in citation sentences by the help of evidential (show) or performance (improve) verbs or contrastive constructions which re-enact previous work in an indirect manner (Athar, 2011; Teufel, Siddharthan, and Tidhar, 2006).

4.5 Constructing Sentiment-Aware Reporting-Verb Corpus

In order to put into practice the relationship between verbs and the environment of assessment, the paper develops a sentiment-sensitive verb corpus where the lemmas of verbs are linked to their spread of positive, negative and neutral cases. It is not aimed not to maintain that a verb has a fixed sentiment, but to give an empirical profile which may be adapted to two analytical uses:

1. Descriptive analysis, by showing how verbs are recurrent among evaluative environments; and
2. Feature development, by permitting the downstream models to use verb distributions (interpretable signals) in citation -function or stance tasks (Athar, 2011; Teufel et al., 2006).

To justify this, the paper provides a weight to every verb according to the distribution of verbs in the polarity classes. The reasoning behind it is simple; a verb with a disproportionately high number of positive or negative citations may add to the position more than a verb where the number of neutral citations is disproportionately high. This is not to make this particular verb intrinsically positive or negative but just to say that it is positively correlated with evaluative contexts within the dataset.

Table8. Illustrative Weighting Logic for Sentiment-Aware Verb Profiling

Quantity	Interpretation	Computation (conceptual)
$f(v, \text{Positive})$	occurrences of verb lemma v in positive citation sentences	count from extraction output
$f(v, \text{Negative})$	occurrences of verb lemma v in negative citation sentences	count from extraction output
$f(v, \text{Neutral})$	occurrences of verb lemma v in neutral citation sentences	count from extraction output
$P(v s)$	conditional proportion of v within sentiment class s	$f(v, s) / \sum f(v_i, s)$
Weight(v)	association strength with evaluative environments	function of $P(v \text{Positive})$ and $P(v \text{Negative})$ relative to Neutral

Note. The weighting function can be implemented in multiple ways (e.g., log-odds association, ratio-based contrasts, or class-weighted z-scores). The key requirement is interpretability and robustness under class imbalance, since neutral citations dominate (Biber, Conrad, & Reppen, 1998; Athar, 2011).

In practical terms, a sentiment-aware verb corpus of this kind enables fine-grained descriptive statements such as: “evidential and performance verbs appear proportionally more often in positive citations than in neutral citations,” while method-uptake verbs dominate by volume due to the technical discourse routines of the field. This separation between *volume dominance* (driven by field practice) and *polarity sensitivity* (driven by evaluative framing) is crucial for correct interpretation.

4.6 Accuracy of Automated Extraction: Interpretation of The 88.25% Result

Automated verb extraction was validated through manual annotation by language experts on three random samples (100 citation sentences each). The study reports an overall average accuracy of **88.25%** for automated identification of verbs in citation sentences. In citation-sentence processing, this is a credible performance level, particularly given that citation sentences often contain compressed syntax, domain-specific expressions, and punctuation or parenthetical citation markers that can interfere with tagging and tokenization (Athar, 2011; Teufel & Moens, 2002).

However, the interpretive value of an accuracy score depends on what kinds of errors are made. In verb extraction from scientific citation contexts, errors typically fall into a small set of recurrent types:

- **POS confusion in technical noun-verb alternations** (e.g., *use* as a noun vs verb).
- **Missed verbs in auxiliary-heavy evidential frames** (e.g., *has been shown*).
- **Boundary errors in multiword constructions** (e.g., phrasal verbs or fixed expressions).

- **Attachment issues around parenthetical citations**, where tokenization and parsing may be distracted by punctuation.

Table9. Common extraction error patterns in citation-sentence verb identification

Error pattern	Why it occurs in citation sentences	Consequence for analysis	Recommended mitigation
Noun–verb ambiguity	technical prose frequently nominalizes (<i>the use of...</i>)	false positives/negatives for high-frequency items	add local syntactic checks; sample auditing
Auxiliary evidentials	stance realized in passive/aux patterns (<i>has been shown</i>)	undercounts evidential reporting frames	store auxiliary patterns; dependency parsing
Citation interruption	parentheses disrupt local structure	mis-tagging near citations	citation-span masking; robust tokenization
Multiword constructions	verb meaning distributed across phrase	loss of rhetorical nuance	record head lemma + phrase context

This error analysis matters because reporting-verb interpretation is sensitive: a small rate of systematic errors can bias frequency profiles for common verbs, and can distort polarity-conditioned comparisons if errors are unevenly distributed by sentiment category. On that account, the present study considers automated extraction as an effective scaling instrument, though not a replacement of such manual inspection and the rhetorical analysis.

4.7 Preliminary Responses to The Research Questions Using The Dataset Evidence

It is at this stage of the analysis that the study is able to give evidence-based responses to the research questions without taking the scope too far. The citation-sentence data set is informed by computational linguistics; as such, answers to this data set are packaged as (a) the empirical data on this corpus of study, and (b) logical congruency with the general evidence on disciplinary variation in applied-linguistics.

RQ1: Which reporting verbs are most commonly used in citation texts in various academic disciplines?

The most commonly extracted verb lemmas within the analyzed citation-sentence corpus are use (3,425), train (1,285) and tag (792) (Table 5). These verbs pointing to a predominant trend in computational linguistics: methodological uptake, usually such as the adoption of a tool, model, dataset or procedure, is described by these verbs instead of attributed propositions alone. This finding can be aligned with citation-function studies in computational linguistics which demonstrate the citations often used to fulfill the functions of use and comparison in technical writing (Teufel et al., 2006). Applied-linguistics corpus studies suggest that other subjects (especially the humanities and some social sciences) usually show more such discourse- and cognition-oriented reporting verbs (e.g. argue, contend, suggest). Such verbs are incidental and dialectical epistemologies (Hyland, 1999; Charles, 2006; Pho, 2008). In this way, the corpus supports the general argument that verb distributions are discipline-sensitive, whichever way it is presenting direct evidence on the same.

RQ3: What are the most common reporting verbs applied to citations to imply the support of the material cited and which differ among disciplines?

The sentiment conditioned analysis suggests that some verbs, especially those in the index of evidential strength and conduct of result (e.g., show, achieve, improve, perform) demonstrate more visible variance in repetition across the categories of polarity in the corpus narrative, which usually have higher number of repetitions in the positive (and even evaluatively engaged) citation. Those verbs are involved in a supporting role in terms of linguistics applied to language and contribute to the framing of the evidence of quoted work but the direction of the rhetorical move is determined by the modifiers and co-context (Thompson and Ye, 1991; Hyland, 1999). Support can be achieved in various different ways across the disciplines: in the empirical sciences, as support has been observed to align with evidential verbs (demonstrate, show, find), in argumentative disciplines support might be aligned with discourse verbs that mark alignment (argue, maintain, emphasise) as well as hedged endorsement (suggest) (Hyland, 1999; Pho, 2008). The cross-dimensional aspect is thus somewhat theoretically justified, but it needs more empirical confirmation based on similar data of similar fields.

RQ2: How do reporting verb usage patterns differ between different types of academic sources (e.g., journal articles, books, conference papers)?

The current data set helps examine the practice of citing mainly in the venues of computational linguistics, where the outputs of conferences and journals are the most outstanding. It therefore offers a practical basis on which citation patterns in conferences and journals can be compared in cases where metadata of venue are considered but it has minimal applicability in fields of study that are dominated by book references. Methodologically, the objective of the project can be achieved by mapping citation sentences with metadata about document type such as venue and publication format and then compare distributions of reporting verbs conditional on the source type. This process map is described in conceptual terms in the section at hand and mentioned in the discussion as a promising line that needs to be pursued in the future through the application of the natural language processing to

infer preferences of source types on the basis of the linguistic features, which would also correspond to the overall goal to use the natural language processing to draw inferences about the preferences to the source types.

5. DISCUSSION

5.1 What the dominance of “use/train/tag” implies about citation practice in computational linguistics

The most noticeable empirical fact is that frequent verbs are used in citation sentences not classical verbs of attribution like: to argue, to claim, but to use, train, and tag, which are methodological verbs. In terms of applied linguistics, this goes to show how conventions of reporting are fashioned within the activity systems of a discipline. In computational linguistics, operationalization of progress is often carried out by reusing and adapting resources - such as corpora, tag sets, architectures, evaluation datasets, and training processes. It follows that citations are frequently used to support a methodology, a source or a point of reference instead of supporting a long series of propositional reasoning.

This claim does not exclude the inclination and evaluation in the corpus. Instead, stance occurs mostly in the form of (a) comparative and performance-related words (e.g., improve, outperform, achieve), (b) verbs of presentation as having passed the test (e.g., show), and (c) comparative clauses that implicitly criticize constraints (Athar, 2011). Differently put, the rhetorical economy of the discipline is such that the work of assessment is spread throughout technical description and performance framing, as opposed to being mainly based on the use of overtly argumentative reporting verbs.

The acknowledgement of this trend helps explain the usefulness of NLP-aided analysis further: when methodological verbs are high frequency, to discern less frequent, but still, rhetorically significant reporting frames, it becomes required to support high volumes of citation situations in an efficient way. The extraction on corpus scale also makes it easier to find the so-called long tail of the verb that can imply strong stance, but are too rare to be dependably found in small sets of manual data (Biber et al., 1998; Teufel and Moens, 2002).

In the next part of the discourse analysis, the discussion will focus upon (a) the illumination of the supportive and critical citation strategies by the polarity-conditioned verb patterns, (b) the meaning of 88.25% extraction accuracy to the reliability and interpretive caution and (c) the possibilities of adapting the approach to the source-type inference and cross-disciplinary corpora without changing the research objectives and research questions.

With the first impression that citation practice in this corpus is heavily influenced by methodological uptake--and that effects of polarity are better evident in the sentences of the evidential and performance framing--the discussion can be polished by the discussion on how supportive and critical roles are framed in the sentences of the citation practice.

5.2 Supportive citation strategies and polarity-sensitive verbs

The sentiment-conditioned perception of the corpus suggests that such supportive citations in the field of computational linguistics are generally constituted of two compensative rhetorical practices. The former, which is called evidential framing, means that authors use antecedent work as an already existing fact, often with the help of evidential verbs, including show and demonstrate. The second one is performance framing which entails writers situating cited work as the tool of making gains, improvement, or even competitive performance (such as verbs like achieve, improve, perform). These methods follow the appraisal economy of technical research, where approval is often expressed not in the form of direct flattery, but by demonstrating that previous work has been successful, proven or useful to downstream activities.

More importantly, evidential and performance verbs are not seldom used. Local grammatical situations, such as modality (may, can), negation (not), comparative expressions (better than, outperform) and concessive clauses (however, despite) either increase or decrease their stance value. This may indicate one of the most crucial aspects in applied linguistics, evaluative meaning in reporting clauses is shared between choices of verb and co-text constructions, but not within the lemma itself (Thompson and Ye, 1991).

This, in terms of NLP, implies that predicting polarity and classifying rhetorical-functions make use of such features as indicate local syntactic structure, and should not be used merely to predict polarity by counting the number of words on the surface. Dependency relations are especially helpful in that they can be used to tell that a verb is found under negation, that it has a that-clause complement, and whether it is found in a comparative. Potential applications in computational linguistics Preexisting citation-function works have also revealed that linguistically informed features can be more informative than classification and can be able to detect features characteristic of a rhetorical context that shallow features fail to detect (Teufel, Siddharthan, and Tidhar, 2006).

5.3 Negative Citation Strategies During A Sparse Class

Negative citations constitute the least significant part of the data (3.20 percent of citation sentences). This rarity becomes incorrect; it is also a stylistic characteristic of genre. Scientific text is normally discouraging of confrontation, instead leaning toward moderate criticism and oblique opposition. In these circumstances, negative stance is often achieved using structures of contrastive clauses (with the phrases however, nonetheless) and statement of limitation (fails to, does not), and using comparative assessment (worse than, less robust) with fairly neutral reporting frames.

Two methodological implications to this observation are available. To start with, any effort to variously attribute certain reporting verbs as necessarily negative ones is likely to be overgeneralized. Second, strong negative stance analysis involves as much as possible the extension of the context of the single sentence. The citation sentence itself can be formally neutral and have the follow-up sentences which have the actual criticism, or vice versa. The restriction has been well established in scientific sentiment analysis: polarity cues can be widely scattered and syntactically limited, and sentence level labels can be low to multi-sentence rhetorical procedures (Athar, 2011). To analyze reporting-verbs, the rule of thumb is, then, to take the negative category as an interpretative stimulus: verbs that are disproportionately high in negativity need to be studied in their overall rhetorical context. Quintessential data in the technical critique is X recommends/asserts, however, the negativity is carried out through the concession, followed by the evaluation. A second trend is evidential reversal: X shows can appear in an X citation when these words are used by the author to prove a restriction (e.g. X shows improvements, but only when there are restricted conditions). In both instances the reporting verb is being used in the rhetoric move, but the direction of evaluation is conditioned by the co-text.

5.4 What frequency dominance actually means for reporting-verb interpretation

The dominance of the verbs employed at the corpus-level (use, train, and tag) (cf. Table 5) can be misunderstood in case it is considered only as a lexical phenomenon. Within the realm of computational linguistics, one shows the aspect of a methodological citation regime in these verbs: citations are often used to refer to dataset, systems, tag sets, and training procedures, and reporting is generally done by use of the citation marker such a practice, unlike by a general attribution clause.

Such a reading goes hand in hand with the larger computational work on citation role. The concepts of use and comparison contain large percentages of technical corpora in that literature, namely because scientific advancement in those areas operationalizes itself with references to the preceding system(s) through building up and benchmarking. According to Teufel et al. (2006), this is an operational attribute of a scientific writing that is calculable and categorical.

Applied linguistics can be used as a counterpoint: various disciplines define what they consider to be legitimate knowledge work. Whereas interpretative and argumentative aspects can be stressed in humanities discipline areas, the fields which are technical can tend to foreground performance and method. These variations should be found in reporting-verb profiles. The account of the reporting practices provided by Hyland highlights that such patterns are conditioned not by the universal tendencies toward language preferences, but by the disciplinary communities and their epistemic values and norms of interaction (Hyland, 2002).

5.5.1 Process Of Answering The Research Questions Using The Available Evidence

RQ 1 (frequency in the field of disciplines). Among the existing dataset information, the most common verb lemmas in citation sentences are: use (3,425 occurrences), train (1,285 occurrences) and tag (792 occurrences). The number of lemmas of unique verbs in the corpus is 389. These statistics can give a characterization of citation -sentence verb use in computational linguistics and justify the finding that methodological uptake is predominant over citation framing in computational linguistics. The wider cross-disciplinary aspect of the question is conceptually well connected by empirical studies of applied-linguistics that report a systemic disciplinary variance of reporting and stance. However, to do a complete empirical comparison, parallel corpora of other domains is needed.

RQ3 (verbs indicating support). Polarity -conditioned view predicts that the witnessing is mainly achieved by evidential and performance framing. Authors are empirically inclined to use foregrounded validated results (e.g., shown, demonstrated) and positive impacts (e.g., improve, achieve) in empowering environments. Although the narrative of the data set indicates that there is significant differentiation of these verbs in the aspect of polarity, the key interpretation concern is that support is seldom an attribute of the verb per se; it is put together by clause structure, boosters/hedges, and the overall aspect of the argumentation. This is in line with the perception that the evaluative meaning of the reporting clauses is interactional and contextual (Thompson and Ye, 1991).

RQ2 (patterns by type of source). The methodological soundness of studying linguistic patterns and using them to infer the source type (journal article, book, conference paper) is valid in cases when there are sufficient metadata available and sufficient representation of the source types. Conference and journal publications are the most predominant in AAN and they offer a pragmatic ground on which a two-way comparison (conference vs. journal) can be made however having little to offer on books. In turn, the article abstracts source-type inference as principled extension of the existing pipeline: once each of mentioned works has been classified in terms of document type, the same verb-profile-procedure can be utilized to compare distributions in terms of sources and to find out whether certain reporting frames can be clustered around sources of a certain type.

To make this extension operational, Table 10 provides a sketched design of a source-type analysis that does not change the current goals and does not necessitate any modification of the extraction pipeline.

Table10. Minimal design for source-type comparison using the existing pipeline

Step	Input	Output	Rationale
Metadata linking	cited paper IDs	source-type label (e.g., conference, journal)	enables grouping without changing text processing
Verb extraction	citation sentences	verb lemma counts per sentence	same procedure as the main analysis
Aggregation	counts + source type	frequency profiles by source category	supports RQ2 comparisons
Contrast statistics	profiles by category	effect sizes / ranked differences	identifies verbs characteristic of each source type

5.6 Reliability, Replicability, And What 88.25% Enables

A mean verb-extraction agreement of 88.25% demonstrates that the pipeline could be used in case of stable and large-scale profiling, however, it must be combined with specific error monitoring of high-impact items, namely ambiguous forms like forms such as use). Whereas in the corpus research, quality of extraction is not a binary property; even in case the agreement levels are great, a small amount of systematic bias will interfere with the relative ranking of frequent items. The most accessible, and thus useful, is therefore the best practice which is iteration of validation with periodic checks on the stratified samples of the most common and verbs of polarity, particularly, emphasize on the most frequent and verbs of polarity.

In practice, extraction accuracy as 88.25% allows at least three valuable results. First, it aids in the stable determination of general distributional patterns (e.g., prevalence of method -uptake verbs). Second, it allows building a sentiment-sensitive verb lexicon that could be used as a clear feature set in a further analysis of citation. Third, it also offers reproducible methodological template which can be then transferred to new corpora, including cross-disciplinary collections, in case the manual reporting-verb analysis would not otherwise be possible.

These results make NLP a methodological amplifier and not a methodological replacement: the automation increases the scale and consistency, whereas the qualitative analysis and theory can maintain the rhetorical validity.

6. Implications

6.1 Implications for NLP-Based Citation Analysis

This pipeline design and the results are relevant directly to the computational studies of the role of citation and scientific discourse modeling. To begin with, the outcomes highlight issues of identifying the process of routine citation behavior versus conceptualizing evaluatively marked citation behavior. High-frequency methodological verbs are strong pointers of how a discipline its own undertakings are to do knowledge construction, but, to themselves, are not, by themselves, robust predictors of either support or criticism. In modeling activities, the following observation can be made: (a) determine the most frequent type of citation activity (e.g., use, comparison, background) and (b) model evaluative stance in that functional category by taking advantage of syntactic and discourse evidence.

Second, the paper shows the purpose of lemmatized verb inventories as interpretable features of the supervised models. Although the opaque numerical representations are not always adequate, lemma-level distributions have helped researchers to trace the resources in lexical resources that influence classification results. This feature will improve accountability and error analyses, which are imperative in the case of unequal population of classes like negative cites. The literature on the citation -function highlighted long ago the localized and linguistically organized character of rhetorical cues in the academic writing, thus, making such features of the debate appropriate to the design of features that combine syntax and discourse (Teufel et al, 2006).

Third, the specified issues, such as those of auxiliary evidential, negation, concession, and citation interruption, point to places that NLP pipelines need to be reinforced to allow them to more effectively represent reporting-verb nuance. Some practical improvements are also dependency-based identification of passive evidential frames (as has been demonstrated), explicit negation scope, and using a wider citation window (sentence +1) in cases where the rhetorical move is likely to cross several sentences. Though these are gradual engineering decisions, their linguistic implications are disproportional, in that they predetermine the existence of a model with rhetorical force, and the presence of a model with lexical presence.

These implications are contextualized in the following section: it discusses how a sentiment-conscious reporting-verb inventory can be used to instruct academic writing, and reflects how the cross-disciplinary design of corpus can make the research question empirically testable in bigger way.

6.2 Implications for Applied Linguistics And Academic-Writing Pedagogy

In the context of the applied linguistics, the usefulness of an NLP-aided study of reporting verbs is not just in its automation, but its ultimate contribution can be made to the methods. Particularly, this methodology makes it possible to facilitate the process of linking large scale distributional data to rigorous rhetorical analyzes, in turn, allowing the generation of pedagogically useful descriptions on the process by which scholarly authors assess,

align, and assign credit to source material. Reporting verbs are a relatively sparse category of resources related to citation that are both linguistically approachable and rhetorically relevant, and this explains the continuing popularity of these verbs in EAP curriculum and genre-based composition education (Hyland, 1999, 2002; Swales, 1990).

However, the successful teaching of reporting verbs requires more than the availability of lexical lists. It is quite common practitioners find that learners get a repertoire of verbs and still continue to have issues in key areas: (i) choosing verbs which accurately tune to their intended stance; (ii) manipulating grammatical frames - including that-clauses, passive evidential, and reporting nouns; and (iii) it regulates hedging and certainty expressions in a manner consistent with disciplinary conventions and particular genre components. The current research project thus suggests a much more balanced conceptualization of instruction a more gradual approach to the issue though is to explain the relationships among the choice of verb, framing behavior, evaluative judgment, and disciplinary standards.

As an example of this pedagogical translation process, Table 11 suggests a pedagogical translation schema, projecting into the attendant stance effects of the commonly used reporting frames. It is not a set of verbs which should be taught, but it is a principled system of analytic, which captures the prototypical rhetorical role of each verb, what linguistic frames each verb authorizes and what linguistic risks can be generated, when such or that verb is used out of context.

Table11. Pedagogical mapping of reporting-verb frames to stance effects (teaching-oriented)

Reporting frame (typical pattern)	Illustrative verbs	Likely stance effect in academic writing	Common learner risk	Pedagogical focus
Author + verb + <i>that</i> -clause	<i>argue, suggest, claim, propose</i>	positions cited proposition as an arguable claim; stance depends on verb strength and hedging	treating all verbs as “neutral” and missing evaluative force	teach verb strength gradients and hedging strategies
Evidential passive	<i>has been shown, has been demonstrated</i>	presents proposition as established/validated; often used for strong support	overuse leads to “over-certainty,” or misuse where evidence is limited	teach evidence alignment and cautious framing
Reporting of findings	<i>find, report, observe</i>	frames cited content as research outcome; often more neutral and procedural	using outcome verbs where the cited work is interpretive rather than empirical	teach matching verb choice to epistemic type of claim
Uptake/resource framing	<i>use, adopt, apply</i>	positions cited work as a method/resource rather than a proposition	learners may confuse “use-citations” with endorsement	teach functional distinction between method citation and claim citation
Contrastive reporting	<i>suggest... but; claim... however</i>	allows critical stance through concession/contrast rather than direct attack	learners may create unintended harshness or unclear critique	teach concessive structures and polite critique norms

Note. Mapping is based on applied-linguistics descriptions of evaluation in reporting clauses, and findings of citation-sentiment studies that stance can easily be framed and built on co-text, but not just by choice of verb (Hyland, 2002; Thompson and Ye, 1991; Athar, 2011).

6.3 Implications of the Support of Cross-Disciplinary Writing

The research question of one of the studies is clearly aimed at the difference across various academic disciplines. Despite the fact that the dataset examined in this paper falls within the field of computational linguistics, the given methodological approach can be adopted to other disciplinary corpora (that is, social sciences, natural sciences, applied linguistics): the same pipeline, extract, lemmatize, profile, validate, interpret, can be utilized to process other disciplinary corpora. This applied lingual hypothesis lies in the center: reporting verb profiles are supposed to reflect the disciplinary epistemologies and the genre expectations (Hyland, 1999, 2002; Charles, 2006).

In the case of writing-support programmers’ this has a great practical implication. Instead of using generic lists of so-called academic reporting verbs, institutes must come up with discipline-adjusted inventories and examples based on local corpora (e.g., an engineering corpus versus an education corpus), hence, making teaching based on the original rhetorical pattern possible. This is especially true in the case of multilingual authors who more often than not get a blank-&-white piece of advice regarding the use of reporting verbs, when, in actual fact, the value and desirable range of choices and frames to be one or the other, depending on both the field and the part of the research paper (Pho, 2008).

7. Limitations and Methodological Constraints

7.1 Corpus Scope and Generalizability

A first limitation concerns the scope of evidence. The citation-sentence dataset is centered on computational linguistics research. Consequently, the quantitative findings (e.g., the dominance of *use*, *train*, *tag*) are best understood as discipline-specific. The study therefore treats the computational linguistics corpus as (a) a detailed case study and (b) a method demonstrator. Broader cross-disciplinary generalizations are made cautiously and anchored in prior applied-linguistics research rather than claimed as direct empirical outcomes of the dataset.

This limitation does not undermine the research questions; it clarifies the level at which they can be answered with the current corpus. The pipeline and analytic logic directly support cross-disciplinary comparison, but answering the cross-disciplinary components empirically requires comparable datasets from additional fields.

7.2 Sentence-Level Context and the “Window Problem”

A second limitation is the use of the citation sentence as the unit of analysis. Sentence-level citation analysis is common in computational work because it supports clean annotation and modeling (Athar, 2011; Teufel et al., 2006). However, rhetorical moves frequently extend beyond a single sentence. A writer may introduce a citation neutrally in one sentence and evaluate it in the next, or establish a limitation in the preceding sentence that frames the cited claim. The specified issue of the window is especially relevant to the realm of negative citations that are a rather rare occurrence and are often expressed by the use of multi-sentence, indirect, and subtle critiques instead of direct denunciation.

This results in the fact that whereas sentence-level extraction is more suitable to the task of profiling with scale, sometimes only a broader contextual window (i.e. the subject sentence plus one or minus a surrounding sentence) will grasp the full evaluative gesture.

7.3 Reporting Verbs Versus Verbs- in-Citation- Context

A third weakness, both conceptual and operational is that the extracted inventory includes verbs in citation sentences that are not necessarily report verbs in the narrow applied-linguistic meaning of ascription verbs that provide access to propositions that have been cited. Technical writing often uses citation sentences placing the cited text as a means or aid (e.g.). Often, such sentences are considered to constitute a tool (e.g.). This observation has been a major empirical discovery about the genre; however, it provides a challenge in definition.

The research addresses this problem by interpreting the extracted verb lemmas as the sources of candidate reportage in citation options and using qualitative contextual verification as the means of determining their rhetorical value. The EAP traditions can be applied to a natural refinement that would require restriction to explicit attribution frames as mediated by syntactic dependency relations and citation span mapping, which would bring closer computational extraction with the narrower construct of reporting verb that is upheld in EAP traditions (Hyland, 2002).

7.4 Extraction Accuracy and Interpretive Caution

The extraction accuracy rate that is reported (88.25 percent) is good with large-scale sentence processing of abbreviated scientific citations, but it does not exclude the risk of interpretation as a result of errors. Specifically, errors with commonly occurring verbs (especially those that are involved in a noun verb ambiguity or subjected to an evidential auxiliary) can have an effect on the frequency orders as well as polarity-induced contrasts. This is why the automated extraction is considered a powerful descriptive tool in the study but with specifically-focused manual confirmation of high-impact items and on verbs that demonstrate polarity sensitivity.

8. Extension Design for Cross-Disciplinary and Source-Type Analysis

The study’s objectives include using NLP to identify source types (journal articles, books, conference papers) and the research questions include cross-disciplinary patterns. Both aims can be addressed by extending the present design in a controlled way. Table 12 outlines an extension plan that retains the same core pipeline while adding corpus coverage and metadata linking.

Table12. Design blueprint for extending the study to disciplines and source types

Extension goal	Data requirement	Processing requirement	Output	How it answers the existing RQs/objectives
Cross-disciplinary comparison	corpora from multiple disciplines with citation sentences	same extraction + lemmatization; consistent cleaning	verb profiles by discipline; contrast rankings	directly addresses RQ1 and RQ3 across disciplines
Source-type comparison	metadata labeling cited items	link citation sentence to cited-item type	verb profiles by source type	addresses RQ2 and objective on source-type inference

Extension goal	Data requirement	Processing requirement	Output	How it answers the existing RQs/objectives
	(conference, journal, book)			
Rhetorical-function coding	annotation or model for citation function	add function labels to citation sentences	verb distributions by function	strengthens interpretation of “support” vs “neutral use”
Wider context window	sentence ± 1 around citation	expand extraction window and store boundaries	improved stance interpretation	reduces window problem and improves qualitative validity

This extension blueprint also clarifies the relationship between objectives and feasibility. The existing pipeline already supports the core computational work; what is needed is the systematic acquisition of (a) cross-disciplinary corpora and (b) reliable metadata for cited-item types.

9. Methodological Recommendations for Improving NLP Sensitivity to Rhetorical Nuance

Among the main findings of the research is that natural language processing tools show a high level of effectiveness in large-scale extraction and profiling processes; however, they also offer problems in the comprehension of subtle rhetoric, especially in cases where verbs have multiplied functions or the assessment has been decentralized in different frames and co-contextual objects. Table 13 outlines certain technical improvements that directly respond to the most far-reaching linguistic issues found during the analysis.

Table13. Recommended NLP enhancements for reporting-verb analysis in citation contexts

Issue	Why it matters rhetorically	Practical enhancement	Expected benefit
Passive evidential frames	evidential strength often encoded in <i>has been shown</i> patterns	detect passive + auxiliary chains; store full evidential construction	better capture of “support” framing
Negation scope	negation can reverse stance (<i>does not show</i>)	dependency-based negation detection	reduces false inference of endorsement
Concession and contrast	critique is often indirect (<i>however, although</i>)	discourse marker detection + clause linking	improves interpretation of negative stance
Attribution vs uptake	method citations differ from claim citations	dependency rules to identify cited-author subjects vs procedural subjects (<i>we</i>)	aligns extraction with “true reporting verbs”
Context window	evaluation may be outside the citation sentence	extract sentence ± 1 and index location	more faithful rhetorical analysis

These improvements mentioned above indicate a broader awareness in the scientific discourse processing that the rhetorical sense is often rendered by structural patterns; systems gain more effectiveness when they are used to model syntax and discourse structures, and not to use lexical signatures alone (Teufel et al., 2002; Teufel et al., 2006).

10. Passing on to the Conclusion

Combined, the results support a subtle opinion. To begin with, NLP-aided extraction and profiling make it easy to study the reporting-verb behavior in thousands of citation instances generating descriptive and comparative information not achievable by a manual methodology given the scale limitations. In their turn, reporting verbs remain rhetorically complex: their value of stance is dependent on the context, field, and these are often spread on frames, hedges and an overall discourse unit. Thus, the research conclusion generalizes the key contributions of the study, answers the research questions in a sufficient range, and offers suggestions on future research and practice in accordance with the objectives formulated.

11. CONCLUSION

11.1 Consolidated overview of what the study demonstrates

With the help of this study, the researcher was able to assess the degree to which reporting verbs can be analyzed using natural language processing (NLP) methods and a corpus-based approach in citation contexts. The main driving force is an amalgamation of both practical and theoretical arguments: reporting verbs constitute rhetorical power in academic discourse, but the average method of analysis is labor-intensive and hard to generalize (Hyland, 1999, 2002; Thompson and Ye, 1991). The research shows that in processing a large body of citation sentences that are annotated with citation sentiment (positive, negative, neutral), a sizeable body of robust distributional

evidence can be rediscovered, (b) sentiments-informed profiling of verbs used in a citation context can be performed, and (c) providing a replicable approach that is user-friendly to infer to additional academic domains and source types.

At the same time, the research highlights a very important caveat: the significance of reporting verbs cannot be represented with the help of lemma frequencies only. This evaluation of citations is commonly shared between a wide range of syntactic frames (passives, negation, concession), discourse marker and localized rhetorical actions. As a result, the NLP tools can enhance the coverage and efficiency significantly, but the interpretive validity of the results remains on the conditions of specific operational definitions, the comparison with the expert opinions, and situational interpretation.

11.2 Overview of major results of the quantitative data

The result of the analysis was a detailed and interpretable verb treatise that happens in citation sentences and the plausibility of scaling this work through using automated extraction. Table 14 puts together the main data and the performance measures in the research.

Table14. Key quantitative outcomes of the corpus processing and validation

Outcome type	Indicator	Value reported in the study	What it supports analytically
Dataset scale	Citation sentences analyzed	8,736	Enables large-scale profiling and polarity-conditioned comparison
Sentiment distribution	Positive / Negative / Neutral	829 / 280 / 7,627	Shows genre-typical imbalance; supports cautious interpretation of negative class
Lexical coverage	Unique verb lemmas identified	389	Supports inventory building and long-tail analysis of less frequent verbs
Salient verb frequencies	Most frequent lemmas	<i>use</i> (3,425), <i>train</i> (1,285), <i>tag</i> (792)	Indicates methodological uptake as a dominant citation practice in the corpus
Extraction reliability	Average agreement with expert annotation	88.25%	Supports reproducible large-scale extraction with targeted auditing for high-impact items

The situation that methodological verbs are in the majority of this disciplinary data is in itself a salient rhetorical phenomenon: many a citation sentence in computational linguistics serves as a tool to tether resources, systems, baselines, procedures, and not just to attribute propositions using standard reportage structures. This fact makes the applied linguistics argument that reporting practices reflect the disciplinary systems of activity and a shared value in epistemology (Hyland, 1999, 2002).

11.3 Direct Answer to the Research Questions

The paper did not alter the objectives and questions of the research. The following responses are not betraying the available evidence and at the same time they preserve the cross-disciplinary focus of the questions, and furthermore draw a line between what is shown empirically and what can be projected logically with additional data.

RQ1: What Are The Most Commonly Used Verbs In Citing Texts In Different Academic Disciplines?

The most common verb lemmas extracted in citation sentences in the analyzed corpus are the lemmas of without the words *use*, *train*, and *tag*, which are common verb lemmas in citation sentences. These results suggest that in this field methodological and resource uptake often exist in citation contexts. Regarding the idea of across different fields of study, the research provides a methodology that can cross-disciplinarily compare and interpret the obtained profile based on the current evidence in applied linguistics that states that reporting practices in different disciplines differ (Hyland, 1999, 2002; Charles, 2006; Pho, 2008). The only fully empirical cross-disciplinary solution would be parallel corpora of other areas run through the same pipeline and using the same validation regime.

RQ2: What Is The Difference In Reporting Verb Usage Based On The Type Of Academic Source Of Data (E.G., Journal Articles, Books, Conference Papers)?

The study methodology can be attributed to provide source-type comparison with connection of citation sentences and metadata of the cited source (conference paper, journal article, book). Conference and journal venues are also central in the current dataset but books are not. The nearest application is therefore one of direct comparison of citation habits in conference and journal publication of computational linguistics; more extensive differentiation by source type may be possible with more comprehensive metadata and more book-involved corpora.

RQ3: What Are The Most Common Reporting Verbs When Citing A Piece Of Information Trying To Mean That It Is Supported Or Not, And How Does This Depend On Disciplines?

In the emotion risk-marked citation scenes of the corpus, the supportive citation scene is likely to predict evidential and performance framing (e.g. verbs that are related to validated results and gains). The paper highlights that the verb lemma does not often show support, but this is an output of reporting frames and the extended co-text (passive evidential, hedging, boosters, concession). Theoretical basis Variation across disciplines is based on previous findings on disciplinary stance and reporting practices (Hyland, 1999, 2002), and the study provides a testable procedure to empirically test this variation when corpus of multi-disciplines is combined.

In order to make the mapping of the findings to the research questions clear, Table 15 provides a general picture of what the current corpus can answer directly and what the methodology can offer as an extrapolation of the outlined aims with no further changes.

12. Recommendations

12.1 Recommendations for Future NLP Development

To make natural language processing models more sensitive to the rhetorical subtext that is accompanying reporting verbs in the context of citations, several improvements would be made:

1. Make a difference between Attribution Reporting and Methodological Uptake

A fine-tuning extraction protocol must detect cases where the mentioned author or piece of writing serves the grammatical role of a grammatical subject of the reporting predicate (e.g., X argues/suggests) and cases where the cited author takes on the agential role (e.g., We use...). This finesse can be implemented as dependency parsing and citation-span mapping, thus getting computational extraction closer to the narrowly understood applied-linguistics conception of reporting verbs.

2. Discover Model Evidential Frames as Constructed and not Only Tokens

A lot of supportive mentions are expressed through passive constructions of evidence (e.g. has been shown). To represent such as patterned constructions, as opposed to just the head verb, is more fidelity of rhetoric, and makes the framing of support more interpretable.

3. Becoming Negotiatively and Concussively strong

Negation and contrastive markers tend to either prove that a citation is approving, neutral, or critical. The systems must be keen to detect the negation scope and concession relations to rule out mistaking the stance by just presence of verbs.

4. Increase the Size of Context when the Rhetorical Moves are Inter-Sentential

Multi-sentence critique, or even genesis, may pass unnoticed when sentence-only analysis is used. Interpretive validity would also be enhanced through a controlled expansion (citation sentence ± 1) but without compromising on the computational tractability.

12.2 Corpus design and wider and wider coverage across disciplines recommendations

In order to operationalize the cross-disciplinary aspect of the research questions, a future research should formulate discipline-balanced corpora, which are under strict sampling principles (e.g. similar time windows, similar genres like research articles and conference papers etc.). As soon as such corpora are compiled, the pipeline will be able to test the possibility of the expected disciplinary differences showing up in an empirical way: that more proportion of discourse and cognition reporting verbs appear in interpretive disciplines and more proportion of research-activity and evidential framing verbs appear in experimental sciences (Hyland, 1999; Hyland, 2002; Pho, 2008).

12.3 Recommendations for Source-type Inference

The study suggests that citation-sentence processing and reliable metadata that relates to cited sources should be combined to meet the goal of identifying different forms of sources in a sentence that is described by linguistic patterns. An effective division of the types of sources (conference paper, journal article, book) is needed at least; then, verb profiles may be compared with the profiles across the source types. This framework could be further elaborated, over time, to include finer differences, e.g., the distinction between review articles and empirical articles, or between dataset papers and method papers, which can have high correlations with reporting frames and verb use.

12.4 Pedagogical and Academic Writing Support Recommendations

The study suggests such a transfer in the pedagogical context: instead of generic listing of reporting verbs, it is recommended to teach students to learn the way reporting verbs work with syntactic and rhetorical structures (e.g. that verbs, passive constructions, critiques that involve concession, method up-take citations) and how such structures differ in will differ across disciplines. One of the visible results of this strategy is the development of discipline-specific fields of reporting-verb resources and exempla based on local corpora, which will allow writers to conform their position, evidence and evaluation to disciplinary norms (Swales, 1990; Hyland, 2002).

12.5 Recommendations in line with the Extension Pathway of the Study

Lastly, the research suggests the extension of the reporting-verb inventory to association with verb-class and predicate behavior in order to make more inference about the citation motivation. This proposal is in line with the long-term objective of creating a mapping dictionary by associating reporting patterns with the cognitive or

rhetorical relations that exists between citing and cited texts- an approach that is in line with the computational models of citation functions and rhetorical status (Teufel and Moens, 2002; Teufel, Siddharthan and Tidhar, 2006).

Acknowledgement

Dr Ayesha Junaid is a researcher at Prince Sattam university Al-Kharj. She is thankful to the university for providing her the research opportunity. She also appreciates the constructive feedback of the reviewers. This research is supported via funding from Prince Sattam bin Abdulaziz University, Saudi Arabia under the project number (PSAU/2026/R/1447).

Faiza Abid, Assistant Professor, Department of Linguistics and Communications, University of Management and Technology, Lahore, and Dr. Zafar Iqbal Bhatti, Professor of Linguistics, School of English, Minhaj University, Lahore, are grateful to their respective institutions for providing a supportive academic and research environment. They also sincerely appreciate the valuable comments and constructive feedback provided by the reviewers, which helped improve the quality of this work.

REFERENCES

5. The entries below are formatted in APA 7 and checked against publisher/archival records where available. (OUP Academic)
6. Athar, A. (2011). Sentiment analysis of citations using sentence structure-based features. In *Proceedings of the ACL 2011 Student Session* (pp. 81–87). Association for Computational Linguistics. (ACL Anthology)
7. Athar, A. (2014). *Sentiment analysis of scientific citations* (Technical Report No. UCAM-CL-TR-856). University of Cambridge, Computer Laboratory. doi:10.48456/tr-856 (Cambridge Computer Lab)
8. Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185–5198). Association for Computational Linguistics. (ACL Anthology)
9. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). Association for Computing Machinery. doi:10.1145/3442188.3445922 (ACM Digital Library)
10. Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press. (Cambridge University Press & Assessment)
11. Charles, M. (2006). The construction of stance in reporting clauses: A cross-disciplinary study of theses. *Applied Linguistics*, 27(3), 492–518. doi:10.1093/applin/aml021 (OUP Academic)
12. Flowerdew, J. (Ed.). (2002). *Academic discourse*. Longman.
13. Grant, C., & Osanloo, A. (2014). Understanding, selecting, and integrating a theoretical framework in dissertation research: Developing a “blueprint” for your house. *Administrative Issues Journal: Connecting Education, Practice, and Research*, 4(2), 12–26. (ERIC)
14. Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics*, 20(3), 341–367. doi:10.1093/applin/20.3.341 (OUP Academic)
15. Hyland, K. (2002). Activity and evaluation: Reporting practices in academic writing. In J. Flowerdew (Ed.), *Academic discourse* (pp. 115–130). Longman. (SCIRP)
16. Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (2nd ed.). Prentice Hall. (ACM Digital Library)
17. Liardet, C. L., & Black, S. (2019). “So and so” says, states and argues: A corpus-assisted engagement analysis of reporting verbs. *Journal of Second Language Writing*, 44, 37–50. doi:10.1016/j.jslw.2019.02.001 (Macquarie University)
18. McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, 36(2), 176–187. doi:10.1086/267990 (OUP Academic)
19. Pho, P. D. (2008). Research article abstracts in applied linguistics and educational technology: A study of linguistic realizations of rhetorical structure and authorial stance. *Discourse Studies*, 10(2), 231–250. doi:10.1177/1461445607087010 (PhilPapers)
20. Radev, D. R., Muthukrishnan, P., Qazvinian, V., & Abu-Jbara, A. (2013). The ACL Anthology Network corpus. *Language Resources and Evaluation*, 47(4), 919–944. doi:10.1007/s10579-012-9211-2 (Springer)
21. Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
22. Teufel, S., & Moens, M. (2002). Summarizing scientific articles: Experiments with relevance and rhetorical status. *Computational Linguistics*, 28(4), 409–445. doi:10.1162/089120102762671936 (MIT Press Direct)
23. Teufel, S., Siddharthan, A., & Tidhar, D. (2006). Automatic classification of citation function. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing* (pp. 103–110). Association for Computational Linguistics. (ACL Anthology)
24. Thompson, G., & Ye, Y. (1991). Evaluation in the reporting verbs used in academic papers. *Applied Linguistics*, 12(4), 365–382. doi:10.1093/applin/12.4.365 (OUP Academic)