

Surface Forms and Rhetorical Functions of Citations in Chinese Linguistic MA Theses

Xinyi Zeng

School of Foreign Languages, Hangzhou Dianzi University

No.1158, 2nd Avenue, Qiantang District, Hangzhou, Zhejiang, China

Jidong Guo (Corresponding author)

School of Foreign Languages, Hangzhou Dianzi University

No.1158, 2nd Avenue, Qiantang District, Hangzhou, Zhejiang, China

E-mail: gjd1020@126.com

Received: August 27, 2025 Accepted: September 27, 2025 Published: October 25, 2025

doi:10.5296/ijl.v17i6.23116

URL: <https://doi.org/10.5296/ijl.v17i6.23116>

Abstract

Citations play an indispensable part in academic writing. This study adopts both qualitative and quantitative methods to investigate the citation practices in 20 Chinese linguistics MA theses. The citations are examined in terms of their surface forms and rhetorical functions. Results reveal that: 1) Authors of Chinese MA theses apply more integral citations than non-integral citations. 2) In integral citations, the authors are more likely to employ verb-controlling citations than naming citations, while in non-integral citations, they show the strongest preference for source citations. 3) The most frequent reporting structure the authors use is X + verb + *that* clause and most of the ten most frequent reporting verbs in this structure belong to ARGUE verbs. 4) The authors employ citations with limited rhetorical functions, primarily using the attribution function and rarely evaluating the cited content. These findings have implications for academic writing instruction.

Keywords: Citations, Surface forms, Rhetorical functions, MA theses, Corpus

1. Introduction

Citation constitutes a highly salient characteristic of academic writing, serving as a means through which scholars not only display the broad reach of their work within a particular research field but also implicitly affirm their integration into and alignment with the broader

disciplinary community (Mansourizadeh & Ahmad, 2011). As a principal marker of intertextuality, citations are instrumental in situating new knowledge within existing scholarly contexts (Samraj, 2013; Kwan & Chan, 2014; Li & Zhang, 2021). Over recent decades, given that referencing other texts is widely recognized as a defining attribute of academic discourse, increasing focus has been placed on the citation approaches adopted by academic writers (Thompson & Tribble, 2001). By citing the works of others, authors facilitate the contextualization of research findings, thereby helping readers in comprehending the unique contributions of the current research (Gao et al., 2025). Despite its pivotal role, research has revealed that students, especially non-native English speakers, encounter considerable difficulties with citations (Romoroda, 2014). Considering that citations are of great importance in academic writing, relevant research should focus on providing more targeted teaching guidance and resources to assist students, especially non-native English speakers, in learning and mastering the use of citations.

Gaining an understanding of citation practices could offer valuable insights into the mechanisms through which authors shape their scholarly compositions and guide pedagogical frameworks designed to support L2 students in refining their citation competencies and producing more academically sophisticated works (Khamkhien, 2025). Previous research on citations has mainly focused on the factors that may affect the behaviors of citations. So far, the varying levels of nativeness, language proficiency and disciplinary differences have been identified as affecting factors in citation practices (e.g., Charles, 2006; Liu & Zhou, 2014; Luzón, 2015; Marti et al., 2019). More recently, more researchers have focused on the rhetorical functions of citations (e.g., Gao et al., 2025; Khamkhien, 2025; Yang et al., 2025).

Citation practices have long been recognized as a fundamental aspect of academic writing (Zhang, 2022). The correct application of citations within scholarly texts enables authors to effectively construct a research context, acknowledge and honor prior contributions, substantiate their findings and assertions, and ultimately obtain peer validation and recognition for their research (Harwood, 2009; Petrić, 2007; Thompson & Tribble, 2001). However, using citations correctly in terms of their forms, rhetorical functions and collocations can be a challenge for both novice and published authors. In the context of Chinese education where English is an additional language, the cultivation of students' writing ability is also one of the key concerns. However, existing studies mostly focused on published authors and doctoral dissertations (e.g. Zhao & Zhan, 2020; Gao et al., 2025), and few of them paid attention to the use of citation practices in academic writing by master's students. To address the gap, this study aims to compile a corpus to study the use of citation forms by authors of Chinese linguistics MA theses in terms of the surface forms and rhetorical functions. Specifically, the study centers on addressing the following research questions:

- 1) What are the surface structures of citations in Chinese linguistics MA theses?
- 2) What are the rhetorical functions of citations in Chinese linguistics MA theses?

2. Literature Review

2.1 Citations for Constructing Knowledge

Citations are extremely important in academic communication. They not only highlight how knowledge is moving forward but also connect authors with their specific academic communities (Mansourizadeh & Ahmad, 2011). Citations show that authors know the existing literature well, help build their unique authorial image, and back up the points they make (Hyland, 2000; Lee et al., 2018; Petrić, 2007). Therefore, citations play a key role in building knowledge (Hyland, 1999). Initially, the classification of citations by Swales (1986, 1990), who classified citations into integral and non-integral citations based on their surface structures, was generally accepted and used. As Swales (1990) proposed, integral citations incorporate the referenced scholars within the syntactic structure of the sentences. In another words, integral citations incorporate authors' names as grammatical elements within the sentence structures in academic writing (Zhang, 2022). Conversely, non-integral citations present sources in brackets (or through superscript numbers), with the primary focus directed toward the content or messages being conveyed rather than the authors (Swales, 1990). These two citation styles can be categorized as author-prominent, where the authors' identity is highlighted, and information-prominent, where the emphasis is on the information itself (Weissberg & Buker, 1990).

Later, researchers turned their attention to the internal structures of citations. They found that X + verb +*that* clause is one of the most common structures applied in academic writing when using citations, and the researchers further proposed different classification criteria for reporting verbs. For example, Thompson & Ye (1991) conducted a classification of reporting verbs based on their denotative meanings and their capacity for conveying evaluation. Subsequently, Thomas & Hawes (1994) developed a framework that categorizes reporting verbs according to the nature of the activity they describe, which primarily encompasses: discourse verbs, real-world verbs and cognition verbs. Building on previous research, Hyland (1999) proposed a classification of reporting verbs based on the specific type of activity to which they refer within academic discourse. He divided reporting verbs into research verbs, cognitive verbs and discourse verbs. Later, Charles (2006) classified the reporting verbs into four groups according to their semantic meaning.

Swales (1986) was the first to examine citations from applied linguistics perspective. Later, the researchers conducted studies on citations in different fields and from different perspectives. Studies on citations have several focuses: novice citation practices in academic writing (e.g., Lee et al., 2018; Dogan et al., 2018; Li & Zhang, 2021); expert citation use in research articles (e.g., Hyland, 1999; Dontcheva-Navratilova, 2016; Gao et al., 2025); and comparative differences between novice and expert citation use (e.g., Samraj, 2013; Zhao & Zhan, 2020; Khamkhien, 2025). Additionally, researchers studied citations in cross-cultural contexts (e.g., Rowley-Jolivet & Carter-Thomas, 2014; Romoroda, 2014) and cross-disciplinary contexts (e.g., Nesi, 2021; Eckstein et al., 2022). Their studies aimed to find out the similarities and differences in citation practices across disciplines and cultures.

2.2 Patterns of Citation Forms and Functions

Grasping citation forms is of great importance since they determine how sources are presented to readers and how they bolster the authors' argument (Khamkhien, 2025). Additionally, recognizing the rhetorical functions of citations is vital because they give credit to the cited sources and allow authors to showcase the value of their own contributions (Hyland, 2004, 2005). Studies on citations have focused on their surface forms and rhetorical functions they play in academic writing. Relevant studies have centered on two key aspects.

First, scholars have examined the use of forms and functions of citations in a single text. For instance, focusing on both forms and functions of citations, Lee et al. (2018) explored the referential behaviors of non-native English-speaking undergraduates during their introductory academic composition courses. He analyzed the forms and functions of citations. Results indicated that L2 students used a limited amount of reporting structures and primarily adopted attribution function of citations in their academic writing. Similarly, Nesi (2021) conducted an analysis of citations within undergraduate assignments to explore the linguistic characteristics inherent in students' citation practices. The study revealed that students' utilization of citations tended to increase over the course of their undergraduate education, and that the functions of citations differed across academic disciplines. Expanding the scope beyond student writing, Gao et al. (2025) examined the rhetorical functions and linguistic forms of citations in research articles published in four journals. Their findings indicated that while there were subtle variations in the linguistic forms of citations, the rhetorical functions those citations served were both overlapping and distinctive.

Then, the distinctions between novice and professional writers related to the surface forms and rhetorical functions have also been explored. Mansourizadeh & Ahmad (2011) examined citation practices in research papers authored by both non-native expert and novice writers in Malaysia. They compared the diverse types and functions of citations employed by established scholars with those used by novice academic writers within the same research domain. Their findings revealed notable differences in the citation types and functions utilized by these two groups. Samraj (2013) investigated the citation forms and functions within the discussion sections of MA theses and research articles. The results revealed variations in both the forms and functions of citations, with authors of MA theses more likely to employ integral citations, while authors of research articles relied more on non-integral citations. Additionally, intertextual links served various rhetorical purposes in both genres.

Although many researchers have studied the use of citations in academic writing, most of them focused on research articles and doctoral dissertations (e.g., Lee et al., 2018; Zhao & Zhan, 2020; Gao et al., 2025), while few have examined citation practices in full-length MA theses. Since master's students are in the early stage of academic socialization, scholars should figure out how master's students employ different forms of citations in academic writing and recognize the distinct functions of these citations. Moreover, the inconsistent findings in relevant studies also highlight the necessity for further exploration. Considering the research gap, we need to further investigate the surface forms and rhetorical functions of citations used by master's students through authentic corpus analysis. Meanwhile,

corpus-based research on the forms and functions of citations in Chinese linguistics MA theses can not only help us better understand the real use of citations in Chinese MA theses but also enhance students' writing skills by identifying potential writing problems.

3. Methods

3.1 Analytical Framework

3.1.1 Surface Forms

Citations can be divided into integral and non-integral citations (Swales, 1986). As shown in example 1, integral citations incorporate the names of the cited authors directly into the sentence structures, thereby placing greater emphasis on the originators or messengers of the ideas (Swales, 1990). Conversely, non-integral citations appear in brackets (or as superscript numbers), without explicitly integrating the authors' names into the main text, thereby prioritizing the cited contents over the authors, as shown in Example 2.

Example 1 & 2:

"Building on this foundation, Connie et al. (1990) suggested that ..." (CLMT 19)

"The third round of exploratory factor analysis extracted four factors with a cumulative variance explained of 62.90% > 60%, indicating that ... (Zhou, 2020)." (CLMT 14)

3.1.2 Rhetorical Functions

Building on Swales's (1986) classification, Thompson & Tribble (2001) proposed a framework outlining the functions of citations. They further divided integral citations into verb-controlling and naming, while non-integral citations were classified into source, origin, identification, and reference (Thompson & Tribble, 2001), as shown in Table 1. The role of indicating origin is attributed to citations that reveal where a concept or product originated (Thompson & Tribble, 2001). Given that concepts and ideas aren't always easy to tell apart and considering that there aren't many citations in the analyzed texts that point to the origin of a product, these two functions are combined and simply called the source function (Samraj, 2013).

When analyzing the rhetorical functions of citations in this study, we utilize Petrić's (2007) framework, which has been applied to some previous studies (e.g., Lee et al., 2018; Khamkhien, 2025). It classifies the rhetorical functions of citations into eight types: attribution, exemplification, further reference, statement of use, application, evaluation, establishing links between sources and comparison with sources. The definition and typical examples of each sub-classification are shown in Table 2.

Table 1. Functions of Citations (Thompson & Tribble, 2001)

Type	Definition	Example
Integral verb-controlling	The cited author's name was introduced through a lexical verb.	<i>Guo & Qin (2010)</i> surveyed 453 non-English major college students ... (CLMT 14)
Integral naming	The cited author's name appeared in noun phrase form.	Most importantly, a framework was constructed on the basis of <i>Zhang and Jiang's study (2021)</i> to ... (CLMT 10)
Non-integral source	It serves as a means of crediting the originator of the idea or proposition stated in the sentence.	... with only a limited number of studies investigating different genres (<i>Zhang et al., 2023; Zhang & Zhang, 2017</i>). (CLMT 17)
Non-integral identifying	It serves as a way of specifying the particular studies or texts that are indirectly referred to or alluded to within the sentence.	¹ It has been suggested (<i>Wilson, 1999</i>) ...
Non-integral referencing	It acts as an indicator showing the location of a text-based material that contains more information related to the subject.	... have an overall tendency to be socially aware (<i>cited in Hao & Wang, 2012</i>). (CLMT 9)

Note. The examples presented here are sourced from Petrić (2007), since no such functions were identified within the corpus.

Table 2. Rhetorical Functions of Citations (Petrić, 2007)

Type	Definition	Example
Attribution	To assign the origin of information or actions to specific sources.	<i>Oxford (1990)</i> describes LLSs as specific actions ... (CLMT 16)
Exemplification	To support the writer's assertions by citing sources, often introduced with phrases like <i>e.g.</i> or <i>for example</i> .	One of the most commonly used methods is the subjective rating (<i>e.g.</i> , <i>Talmi & Moscovitch, 2004; Zhang, 2012; Zhao, 2016</i>). (CLMT 19)
Further reference	To direct readers to additional sources that offer more details on a particular topic, typically within parentheses or footnotes and introduced with <i>see</i> .	Supporters of indirect WCF stressed that ... (<i>See review in Bitchener & Storch, 2016</i>).
Statement of use	To identify the sources employed in the paper and explain why they are chosen.	... the internal uniformity coefficient (a coefficient) proposed by <i>Cronbach (1951)</i> was used to test the reliability of the questionnaire. (CLMT 11)
Application	To establish links between the sources and the writer's own research or writing to serve the writer's specific objectives.	... we should separate the depth of processing (DoP) here from <i>Manchon et al. (2020)</i> ... (CLMT 7)
Evaluation	To assess the quality, reliability, and relevance of sources by employing critical language.	¹ Elizabeth Grosz's concept of the body as inscriptive surface is an ingenious way out of the nature/culture impasse.
Establishing links between sources	To highlight the relationships or correlations between various sources.	This model is basically the same as the three-stage model of

		Zimmerman (2000). (CLMT 6)
Comparison with sources	To showcase the resemblances or disparities between the writer's current text and referenced materials.	The results provide supporting evidences to the previous studies which ... (Crowe, 1997; Papi, 2018, 2021) (CLMT 11)
Other	Instances where it's necessary to clearly explain how the sentence that cites a source connects to the referenced text.	¹ What are the central research questions about women inmates ... (Steward (1994).

Note. The examples presented here are sourced from Petrić (2007), since no such functions were identified within the corpus.

3.2 Corpus Collection and Research Methods

CNKI is China's largest academic information retrieval and service platform (<https://www.cnki.net/>). This study selected texts from this platform as the primary research material. A combination of qualitative and quantitative methodologies was adopted, involving frequency calculations and text analysis. We searched all Chinese linguistics MA theses with four theses selected annually from 2020 to 2024 in CNKI. The authors of these MA theses come from diverse universities, including comprehensive universities, science and technology universities, normal universities, and others (see Appendix). Data for this study consist of 20 Chinese linguistics MA theses (CLMT) totaling 386,809 tokens. We manually removed images, tables, and acknowledgments from the collected texts. Kennedy (1998) suggested that a corpus comprising 100,000 to 500,000 words can uncover certain language phenomena. The size of this corpus aligns with the recommended specification criteria.

The refined texts were imported into Antconc 3.3.3, a tool designed for concordance-based keyword extraction, to search for citations. We opened the advanced tab, entered the regex `\b(19|20)\d{2}\b` to retrieve all the 4-digit numbers that match the format of 19XX or 20XX (i.e., years between 1900-2099), The search yielded 3,221 results. We deleted the results that appear as the event year, as shown in example 3. We also manually eliminated duplicate results. Both multiple studies of the same author in one parenthesis and two citations in the same sentence are considered a single citation if the citation form is the same, as illustrated in example 4. The final number of results was 2,357. The refined results were imported into the UAM CorpusTool 3. Drawing on Thompson's (2001, 2005) classification framework for citation structures and Petrić's (2007) classification framework for citation functions, we labeled the refined results with respect to their surface structures and rhetorical functions.

After manually annotating the surface structures and rhetorical functions in the refined results, we then employed the software to conduct automatic statistical analysis. Additionally, the original counts of lexical items or linguistic patterns were adjusted to indicate their frequency per 100,000 words (Gries, 2010).

Example 3 & 4:

“Since the reform and opening up in 1978 ...” (CLMT 2)

“... refinement of the Cognition Hypothesis (Robinson, 2001, 2005) and the Limited Attentional Capacity and the Limited Attentional Capacity Model (Skehan, 1999, 2001) ...” (CLMT 20)

4. Results and Discussion

4.1 Surface Structures of Citations

Integral citations provide opportunities for writers to present their own attitude towards the viewpoints they cite, which also lays the emotional foundation for the articles. Additionally, the use of non-integral citations indicates an emphasis to highlight the referenced content rather than focusing on the scholarly credentials of the cited researchers (Li & Zhang, 2021). As can be seen from Table 3, authors of Chinese linguistics MA theses apply more integral citations than non-integral citations, which parallels the findings in Lee et al. (2018) and Khamkhien (2025). However, it is contrary to the results in Charles (2006), which might be since Charles (2006) only investigated the citations with reporting clauses. The integral form of citing sources is more frequently utilized by students as it represents the conventional way they’ve been taught to integrate sources in academic writing (Swales, 2014).

Table 3. Occurrence of Integral and Non-integral Citations

Type	Frequency	%
Integral citations	1240 (320.57)	52.6
Non-integral citations	1113 (287.74)	47.2
Total	2357 (609.35)	100

Note. The number in (.) denotes the normalized frequencies per 100,000 words.

As shown in Table 4, we further classify the citations into integral verb-controlling, naming and non-integral source, referencing, identifying according to Thompson’s (2001, 2005) framework. In integral citations, integral verb-controlling citations occur most frequently, nearly five times as often as integral naming citations, a finding that is the same as that of Mansourizadeh & Ahamd (2011). This phenomenon may stem from novices’ insufficient proficiency in constructing nominalizations and complex noun phrases, as these linguistic

structures often present challenges for novice writers (Mansourizadeh & Ahamd, 2011). Non-integral source citations, the basic surface structure of non-integral citations to cite the source of the citations, appear the most frequently, which parallels the finding in Samraj (2013). Non-integral referencing and identifying citations, on the other hand, occur extremely infrequently, which indicates that authors of Chinese MA theses do not fully master the use of non-integral structures and apply limited varieties of them.

Table 4. Occurrence of Surface Structures of Citations

Type	Frequency	%
Integral verb-controlling	1017 (262.92)	43.15
Integral naming	223 (57.65)	9.46
Non-integral source	1107 (286.19)	46.97
Non-integral reference	7 (1.81)	0.30
Non-integral identifying	0 (0.00)	0.00
Total	2357 (609.35)	100

Note. The number in (.) denotes the normalized frequencies per 100,000 words.

Then, we focus on the most frequent citation structures in CLMT. As demonstrated in Table 5, the most prevalent citation structure in CLMT is X + verb + *that* clause. This structure employs a reporting clause to introduce the citations and accounts for 12.3% of all citations. It is the same as Lee et al.'s (2018) results, where the most frequent citation structure is also X + verb + *that* clause (26.69%). The second most frequent citation structure in this study is by X, which accounts for 5.3% of all citations. Furthermore, this structure is often found in the passive voice, as in example 5.

Example 5:

“Another study conducted by Su (2012) found...” (CLMT 12)

Table 5. Most Frequent Citation Structures

Type	Frequency	%
X + verb + <i>that</i> clause	290 (74.97)	12.3
According to X	57 (14.74)	2.4
By X	126 (32.58)	5.3
In the study, experiment, research of, by X	27 (6.98)	1.1
Total	500 (129.26)	21.2

Note. The number in (.) denotes the normalized frequencies per 100,000 words.

Given that X + verb + *that* clause constitutes the most frequent citation structure, subsequently, we further examine the top ten frequent reporting verbs that occur in this structure in CLMT. As we can see from Table 6, Specifically, the overall frequency of the ten most frequent reporting verbs is 59.20, accounting for 79.0% of the total frequency of reporting verbs. Thus, we can conclude that a few frequently occurring reporting verbs account for more than 70% of the total frequency, indicating that the reporting verbs mostly used by authors of Chinese MA theses are those listed in Table 6. These results mirror those of Hyland (1999) and Kwon et al. (2018). A small number of reporting verbs are repeatedly used to construct citation structures, while the remaining reporting verbs only appear a few times (Hyland, 1999). Specifically, most of them belong to ARGUE verbs (see Charles (2006) for further reference), which is in line with the findings in Kwon et al. (2018). This may be due to the fact that ARGUE verbs are the verb group that contains the largest number of reporting verbs, far more than other verb groups. Additionally, ARGUE verbs are also commonly used for integral citations (Charles, 2006). Since integral citations are frequently used by authors of Chinese MA theses, ARGUE verbs manifest a higher frequency of occurrence in CLMT.

Table 6. Top Ten Reporting Verbs

	Verb	Frequency	%
1	Find	57 (14.74)	19.7
2	Argue	33 (8.53)	11.4
3	Point out	31 (8.02)	10.7
4	State	25 (6.46)	8.6
5	Suggest	20 (5.17)	6.9
6	Propose	15 (3.88)	5.2
7	Claim	14 (3.62)	4.8
8	Think	12 (3.10)	4.1
9	Demonstrate	12 (3.10)	4.1
10	Conclude	10 (2.59)	3.4
Total		229 (59.20)	79.0

Note. The number in (.) denotes the normalized frequencies per 100,000 words.

4.2 Rhetorical Functions of Citations

Citations can achieve various rhetorical functions in academic writing. In this part, we will analyze the rhetorical functions of citations with Petrić's (2007) framework, which classified them into attribution, exemplification, further reference, statement of use, application, evaluation, establishing links between sources, comparison with sources and others. Overall, the top three rhetorical functions of citations in CLMT are attribution (87.27%), comparison with sources (5.48%), and exemplification (3.78%), which is the same as the result of the expert group in Khamkhien (2025). It indicates that authors of Chinese MA theses have some similarity with expert authors in the use of rhetorical functions of citations. Evaluation function with a frequency of 0.7% in Khamkhien (2025) does not appear in CLMT, suggesting that authors of Chinese MA theses use citations without evaluating the content but directly citing it. However, compared to Lee et al.' (2018) findings in which citations in L2 student writing only involve the functions of attribution, exemplification, evaluation, and estimating links between sources, authors in CLMT utilize citations for a broader set of rhetorical functions. This may be because the authors in Lee et al. (2018) are first-year university students, whereas those in this study are authors of MA theses. Compared to

university students in the first-year, authors of MA theses have more specialized knowledge and can be more proficient in the use of different rhetorical functions of citations.

Attribution: Attribution is the most prevalent and fundamental citation function in CLMT (87.27%), which is in line with the findings in Mansourizadeh & Ahamd (2011), Lee et al. (2018) and Khamkhien (2025). Most of the time, it only tries to mention a view from other scholars and does not clarify the relation with the research. This particular function is independent of other functions and necessitates advanced rhetorical abilities (Mansourizadeh & Ahmad, 2011). Results show that authors of Chinese linguistics MA theses tend to apply multiple citations with attribution function with different reporting verbs. However, attribution is considered to be the default citation function in student writing (Petrić, 2007) for it fails to clarify the intrinsic links between the citations. Example 6 introduces Pintrich's (1991) description of Motivated Strategies for Learning Questionnaire and it does not involve other citation functions.

Example 6:

"According to Pintrich (1991), Motivated Strategies for Learning Questionnaire (MSLQ) stressed motivation in learning strategies." (CLMT 11)

Exemplification: Exemplification is the third most frequent citation function in CLMT (3.78%). As explained in Table 2, when the rhetorical function of exemplification is used, it usually appears at the end of the sentences and lists multiple scholars in the form of non-integral citations, often listing detailed authors for the generalized concepts mentioned earlier. In Example 7, it is mentioned that the scales underwent continuous revisions, with specific examples of scholars who refined them provided in parentheses.

Example 7:

"... then the scales revised constantly (e.g., Dornyei & Ushioda, 2011; Papi & Teimouri, 2014; Papi et al., 2019; Papi & Khajavy, 2021)." (CLMT 11)

Further reference: Further reference function occurs infrequently (0.3%) and is in a relatively fixed structure, mostly elicited by the *see ...* structure. This structure usually occurs when the information in the relevant studies is not stated directly to save text space, but rather to point to the relevant literature. Example 8 mentions the similarities and differences between native and non-native groups in dealing with referential context and cites the study by Pan et al. (2015) in parentheses as a reference.

Example 8:

"How native groups treat referential context similar to or different from non-native groups still warrants further investigation (but see Pan et al. (2015) for further discussion) ..." (CLMT 13)

Statement of use: Statement of use function is usually employed when the author introduces the analytical framework or theoretical model applied in the article (1.36%). In such cases, the author adopts an analytical framework or theory proposed by other scholars as the

theoretical foundation for research. Example 9 shows that the SSBWEI in the research is adapted from Ma & Dong (2017).

Example 9:

“The Self-regulation Strategies for Business English Writing Anxiety Inventory (SSBEWI) was adapted from the “Self-Regulation Strategies for Foreign Writing Anxiety Inventory (SSWAI)” developed by Ma & Dong (2017).” (CLMT 14)

Application: Application function occurs at a low frequency in CLMT (0.72%). It establishes a connection between the articles and other studies. Additionally, it is frequently employed by authors to elucidate how they intend to expand upon prior research or address the existing gaps. Example 10 demonstrates that the study not only builds upon but also fills the gaps identified in Manchon et al.’ (2020) research. In example 11, the study is based on Suzuki’s (2009, 2010) studies and further includes a control group.

Example 10 & 11:

“Considering Manchon et al. (2020) did not take the correctness of the metalinguistic explanation into consideration, we followed his ways of determination with some adaptations.” (CLMT 7)

“In complement to Suzuki (2009, 2010), this study incorporated a control group in order to separate the effect of WL and direct WCF.” (CLMT 7)

Evaluation: Evaluation function does not appear in CLMT. This is probably because for authors of Chinese MA theses, their academic level is limited, and most of the ideas and findings cited in theses are from published authors, who are reputable and professional. Thus, authors in CLMT mostly state the cited ideas directly without evaluating them. In contrast, evaluation function is found in Khamkhien’s (2005) research on expert writing. This suggests that, in contrast to novice authors, expert authors do not merely cite the viewpoints of other scholars. Instead, they carry out a more critical process by evaluating the ideas of fellow scholars to promote meaningful academic dialogue.

Establishing links between sources: Establishing links between sources function helps authors to sort out the relationships between previous research findings rather than simply listing the opinions and findings of different authors (0.98%). By finding the internal logic between related studies, authors are able to use citations in a more organized way. Example 12 mentions that the NC/COCA word list established by Nation (2012) was used in Webb et al. (2017). Additionally, example 13 illustrates that the study by Goad & White (2004) drew on earlier work by Selkirk (1996). Both example 12 and 13 create links between studies.

Example 12 & 13:

“In addition, when constructing the uVLT test items, Webb et al. (2017) referenced the BNC/COCA word list created by Nation (2012).” (CLMT 19)

“On the basis of Selkirk’s study (1996) on the prosodic structure of function words, Goad & White (2004) proposed the Prosodic Transfer Hypothesis (PTH) to account for learners’

misuse and omission of errors.” (CLMT 10)

Comparison with sources: The citation function of comparison with sources is common in CLMT (5.48%). Additionally, novice authors often use integral citations to achieve this rhetorical function, which corroborates the findings in Mansourizadeh & Ahmad (2011). When utilizing comparison with sources function, authors engage in a process of comparing their own work with prior studies. They aim to identify the similarities and differences, thereby enhancing the depth and appeal of their writing. In example 14, after presenting their research findings and contributions, the author indicates that their research results are consistent with certain parts of the findings in Kim & Christianson (2017). In example 15, the author shows that his findings are both consistent with and contradictory with some other scholars’ findings.

Example 14 & 15:

“The result shows that ... This corroborates with part of Kim and Christianson’s (2017) result and argues that ...” (CLMT 13)

“... were also consistent with Kormos’s (2011) and Vasylets et al.’s (2017) studies, but contradictory with Liu et al.’s (2017) study in which VOCB was not significantly.” (CLMT 20)

Table 6. Occurrence of Rhetorical Functions of Citations

Type	Frequency	%
Attribution	2056 (531.54)	87.27
Exemplification	89 (23.01)	3.78
Further reference	7 (1.81)	0.3
Statement of use	32 (8.27)	1.36
Application	17 (4.40)	0.72
Evaluation	0 (0.00)	0.00
Establishing links between sources	23 (5.95)	0.98
Comparison with sources	129 (33.35)	5.48
Others	0 (0.00)	0.00
Total	2356 (609.10)	100

Note. The number in (.) denotes the normalized frequencies per 100,000 words.

5. Conclusion

By compiling a corpus of Chinese linguistics MA theses to examine the surface structures and rhetorical functions of citations, this study finds that authors of Chinese linguistics MA theses use more integral citations than non-integral citations in their academic writing. In integral citations, the authors prefer verb-controlling citations to naming citations while in non-integral citations, they predominantly apply source citations. In terms of the rhetorical functions of citations, the authors most frequently employ the attribute function, which is also the fundamental function of citations. They use other complex rhetorical functions far less frequently, especially the evaluation function. This shows that although authors of Chinese MA theses have achieved a certain level of mastery in citation practices, they still lack full proficiency in citation application.

In academic English writing courses, citation instruction often emphasizes technical skills: teachers introduce norms for integrating both integral and non-integral citations, as well as techniques for paraphrasing and summarizing sources (Romoroda, 2014). The objective of these teaching contents is to help students learn different ways to cite academic viewpoints from other authors. However, these teaching approaches may lead students to merely understand that they need to acknowledge the sources when using materials in their writing, without truly learning about the subcategories within integral and non-integral citations, nor fully comprehending and mastering the various rhetorical functions of citations. In order to solve the problems of citations in students' writing, teachers should help students better understand how to use different surface forms and rhetorical functions of citations in the classroom. Student corpus could be applied to analyze what students have and have not learned about citations, and to provide more targeted instruction to students. Additionally, teachers can use expert corpus to help students learn the correct forms of citations and expand their understanding of the rhetorical functions of citations in different contexts.

This study also has some limitations. Due to the limits of the search tool and the fact that unmarked year citations mostly appear in number-type citations, and there are no MA theses using such citations among the ones we selected, this study only focuses on marked-year citations when looking for references. In future research, we will expand the scope of corpus selection to further investigate the differences in the usage of different citation categories (e.g., number-type and year-type citations). Additionally, this study only concentrates on the utilization of citations in the field of linguistics by Chinese students. Future research should extend its scope to encompass a wider range of disciplines beyond linguistics to explore how citation practices vary and interact across diverse academic fields. Moreover, we could also conduct cross-cultural comparisons of citation practices to identify cultural influences on the way to cite sources.

References

- Charles, M. (2006). Phraseological patterns in reporting clauses used in citation: A corpus-based study of theses in two disciplines. *English for Specific Purposes*, 25, 310-331. <https://doi.org/10.1016/j.esp.2005.05.003>
- Dogan, F. S., Yagiz, O., & Kaçar, I. G. (2018). An investigation of the citation transformation types in M.A. and PhD theses. *International Journal for Educational Integrity*, 14(9), 1-15. <https://doi.org/10.1007/s40979-018-0032-y>
- Dontcheva-Navratilova, O. (2016). Rhetorical functions of citations in linguistics research articles: A contrastive (English-Czech) study. *Discourse and Interaction*, 9(2), 51-74. <https://doi.org/10.5817/DI2016-2-51>
- Eckstein, G., Rawlins, J., Taylor, H., Briggs, H., Candland, A., Hanks, E., & Hill, S. (2022). Reporting verb variation across disciplines: An academic corpus study. *Journal of Academic Language and Learning*, 16(1), 59-75. Retrieved from <https://journal.aall.org.au/index.php/jall/article/view/789>
- Gao, J., Pham, Q. H. P., & Polio, C. (2025). Citations in post-methods sections of quantitative and qualitative research articles in second language learning: A corpus-based study. *Journal of English for Academic Purposes*, 74, 101473. <https://doi.org/10.1016/j.jeap.2025.101473>
- Gries, S. T. (2010). Useful statistics for corpus linguistics. In A. Sánchez, & M. Almela (Eds.), *A mosaic of corpus linguistics: Selected approaches* (pp. 269-291). Peter Lang.
- Harwood, N. (2009). An interview-based study of the functions of citations in academic writing across two disciplines. *Journal of Pragmatics*, 41(3), 497-518. <https://doi.org/10.1016/j.pragma.2008.06.001>
- Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics*, 20(3), 341-367. <https://doi.org/10.1093/applin/20.3.341>
- Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. University of Michigan Press.
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173-192. <https://doi.org/10.1177/1461445605050365>
- Khamkhien, A. (2025). The art of referencing: Patterns of citation and authorial stance in academic texts written by Thai students and professional writers. *Journal of English for Academic Purposes*, 74, 101470. <https://doi.org/10.1016/j.jeap.2024.101470>
- Kwan, B. S. C., & Chan, H. (2014). An investigation of source use in the results and the closing sections of empirical articles in information systems: In search of a functional-semantic citation typology for pedagogical purposes. *Journal of English for Academic Purposes*, 14, 29-47. <https://doi.org/10.1016/j.jeap.2013.11.004>
- Kwon, M. H., Staples, S., & Partridge, R. S. (2018). Source work in the first-year L2 writing classroom: Undergraduate L2 writers' use of reporting verbs. *Journal of English for*

Academic Purposes, 34, 86-96. <https://doi.org/10.1016/j.jeap.2018.04.001>

Lee, J. J., Hitchcock, C., & Casal, J. E. (2018). Citation practices of L2 university students in first-year writing: Form, function, and stance. *Journal of English for Academic Purposes*, 33, 1-11. <https://doi.org/10.1016/j.jeap.2018.01.001>

Li, Q., & Zhang, X. (2021). An analysis of citations in Chinese English-major master's theses and doctoral dissertations. *Journal of English for Academic Purposes*, 51, 100982. <https://doi.org/10.1016/j.jeap.2021.100982>

Liu, Y., & Zhou, H. (2014). Reporting and stance in second language academic writing. *Chinese Journal of Applied Linguistics*, 37(4), 483-497.

Luzón, M. J. (2015). An analysis of the citation practices of undergraduate Spanish students. *Journal of Academic Writing*, 5(1), 52-64. <https://doi.org/10.18552/joaw.v5i1.158>

Mansourizadeh, K., & Ahmad, U. K. (2011). Citation practices among non-native expert and novice scientific writers. *Journal of English for Academic Purposes*, 10(3), 152-161. <https://doi.org/10.1016/J.JEAP.2011.03.004>

Marti, L., Yilmaz, S., & Bayyurt, Y. (2019). Reporting research in applied linguistics: The role of nativeness and expertise. *Journal of English for Academic Purposes*, 40, 98-114. <https://doi.org/10.1016/j.jeap.2019.05.005>

Nesi, H. (2021). Sources for courses: Metadiscourse and the role of citation in student writing. *Lingua*, 253, 103040. <https://doi.org/10.1016/j.lingua.2021.103040>

Petrić, B. (2007). Rhetorical functions of citations in high- and low-rated master's theses. *Journal of English for Academic Purposes*, 6(3), 238-253. <https://doi.org/10.1016/j.jeap.2007.09.002>

Romoroda, B. T. (2014). Integration of sources in academic writing: A corpus-based study of citation practices in essay writing in two departments at the university of Botswana. *Reading and Writing*, 5(1), 1-7. <https://doi.org/10.4102/rw.v5i1.41>

Rowley-Jolivet, E., & Carter-Thomas, S. (2014). Citation practices of expert French writers of English: Issues of attribution and stance. In A. Lyda, & K. Warchał (Eds.), *Occupying niches: Interculturality, cross-culturality and aculturality in academic research* (pp. 17-34). Springer. https://doi.org/10.1007/978-3-319-02526-1_2

Samraj, B. (2013). Form and function of citations in discussion sections of master's theses and research articles. *Journal of English for Academic Purposes*, 12(4), 299-310.

Swales, J. M. (1986). Citation analysis and discourse analysis. *Applied Linguistics*, 7(1), 39-56.

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

Thomas, S., & Hawes, T. P. (1994). Reporting verbs in medical journal articles. *English for*

Specific Purposes, 13(2), 129-148. [https://doi.org/10.1016/0889-4906\(94\)90012-4](https://doi.org/10.1016/0889-4906(94)90012-4)

Thompson, G., & Ye, Y. (1991). Evaluation in the reporting verbs used in academic papers. *Applied Linguistics*, 12(4), 365-382. <https://doi.org/10.1093/applin/12.4.365>

Thompson, P. (2001). *A pedagogically-motivated corpus-based examination of PhD theses: Macrostructure, citation practices and uses of modal verbs* [Unpublished doctoral dissertation]. University of Reading.

Thompson, P. (2005). Points of focus and position: Intertextual reference in PhD theses. *Journal of English for Academic Purposes*, 4, 307-323.

Thompson, P., & Tribble, C. (2001). Looking at citations: Using corpora in English for academic purposes. *Language, Learning and Technology*, 5(3), 91-105. <http://dx.doi.org/10.125/44568>

Varga, M., & Gradečak-Erdeljić, T. (2017). English and Croatian citation practices in research articles in applied linguistics: A corpus-based study. *KALBOTYRA*, 70, 153-183. <https://doi.org/10.15388/Klbt.2017.11199>

Weissberg, R., & Buker, S. (1990). *Writing up research*. Prentice Hall.

Zhang, G. (2022). The citational practice of social science research articles: An analysis by part-genres. *Journal of English for Academic Purposes*, 55, 101076. <https://doi.org/10.1016/j.jeap.2021.101076>

Zhang, Y., Wang, Y., Sheng, Q. Z., Yao, L., Chen, H., & Wang, K., et al. (2025). Deep learning meets bibliometrics: a survey of citation function classification. *Journal of Informetrics*, 19, 101608.

Zhao, H., & Zhan, H. (2020). The use of genitive citations in academic writing. *Lingua*, 234, 102769. <https://doi.org/10.1016/j.lingua.2019.102769>

Appendix. Chinese Linguistics MA Theses (CLMT)

CLMT 1 Chen, J. (2020). A study of current situation of English lexical collocation teaching and learning in senior high schools: A case study of Experiment School Attached to Guangxi University of Foreign Languages [Unpublished master's thesis]. Nanning Normal University.

CLMT 2 Chen, L. (2020). An empirical research on attrition of English reading and writing abilities of non-English majors [Unpublished master's thesis]. Wuhan University of Science and Technology.

CLMT 3 Gao, T. (2020). The effects of L2 proficiency and working memory on incidental L2 collocation acquisition [Unpublished master's thesis]. Ningbo University.

CLMT 4 Yao, Y. (2020). The effects of Chinese-English metalinguistic awareness and rapid automatic naming on the development of Chinese junior high school students' English word reading [Unpublished master's thesis]. Fujian Normal University.

CLMT 5 Wang, P. (2021). An empirical study on English reading obstacles in a senior high school context [Unpublished master's thesis]. Tianjin Polytechnic University.

CLMT 6 Lin, Y. (2021). The role of learning community construction in Chinese students' online self-regulated learning [Unpublished master's thesis]. Shanghai University of Finance and Economics.

CLMT 7 Yan, Q. (2021). Effect of collaborative writing languaging on written accuracy increase, WL quality, depth of feedback, processing and learners' perceptions [Unpublished master's thesis]. Guangdong University of Foreign Studies.

CLMT 8 Huang, S. (2021). Construction of multimodal metaphorical meaning of 'risk' in editorial cartoons based on blending model of multimodal metaphor [Unpublished master's thesis]. Yunnan University.

CLMT 9 Zhou, M. (2022). An investigation of vocabulary learning strategies of non-English majors in mobile learning environment [Unpublished master's thesis]. North China University of Technology.

CLMT 10 Qu, M. (2022). A study of definite article acquisition by Chinese EFL learners from the perspective of syntax-pragmatics interface [Unpublished master's thesis]. North China University of Technology.

CLMT 11 Li, R. (2022). The mechanism of promotion focus affecting English learning motivational regulation strategies for non-English major graduate students [Unpublished master's thesis]. China University of Mining and Technology.

CLMT 12 Wang, L. (2022). Effects of explicit/implicit instruction and cognitive style on EFL learners' acquisition of the simple [Unpublished master's thesis]. Xiangtan University.

CLMT 13 Sun, D. (2023). The role of referential context and working memory capacity in EFL learners' relative clause ambiguity resolution [Unpublished master's thesis]. Southeast University.

CLMT 14 Liu, S. (2023). The relationship between business English major undergraduates' writing anxiety and their use of self-regulation strategy [Unpublished master's thesis]. Tianjin University of Finance and Economics.

CLMT 15 Li, T. (2023). A correlation study of L2 learners' writing proficiency and language proficiency under different task types [Unpublished master's thesis]. Henan University of Science and Technology.

CLMT 16 Wang, D. (2023). Influence of students' motivation on using meta-cognitive strategies in EFL writing process: A perspective of self-determination theory [Unpublished master's thesis]. Southwestern University of Finance and Economics.

CLMT 17 Ren, W. (2024). The effect of different continuation tasks on the writing processes and products of Chinese English learners [Unpublished master's thesis]. Dalian University of Foreign Languages.

CLMT 18 Feng, C. (2024). A study on the effects of source text genre on linguistic alignment and oral production quality in college English majors' reading-speaking continuation tasks [Unpublished master's thesis]. Inner Mongolia Normal University.

CLMT 19 Wang, Y. (2024). The effects of familiarity, part of speech and vocabulary breadth on the processing of English derived words by Chinese EFL learners [Unpublished master's thesis]. Northwest Normal University.

CLMT 20 Wang, X. (2024). Effects of task complexity on university English majors' English writing production [Unpublished master's thesis]. Southwest University.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)