

Self-mention in English Theses: A Corpus-based Comparison of Native-English and Native-Vietnamese Speakers

PHUONG DZUNG PHO *

*Faculty of English Linguistics and Literature
University of Social Sciences and Humanities
Vietnam National University, Ho Chi Minh City, Vietnam
phophuongdung@hcmussh.edu.vn*

THI PHUONG DUNG CAO

*Faculty of English Linguistics and Literature
University of Social Sciences and Humanities
Vietnam National University, Ho Chi Minh City, Vietnam*

ABSTRACT

There has been an increasing interest in authorial presence in academic writing in the past decades. However, there seems to be a lack of research on novice writers' writing. This study examines the use of self-mention markers in Master's theses in TESOL written in English by native-English speakers (NES) and native-Vietnamese speakers (NVS) to understand differences in authorial presence and rhetorical strategies. Through a comparative analysis of the two corpora, the findings reveal significant variations in the frequency and types of self-mention employed. The findings show that NES writers favour first-person singular pronouns (e.g., 'I,' 'my') and first-person plural pronouns (e.g., 'we,' 'our') to assert direct authorial presence, while NVS writers prefer third-person noun phrases (e.g., 'the researcher') for a more impersonal tone. NES writers use self-mention for a broader range of rhetorical functions, particularly for explaining a procedure, stating a purpose and stating results, while NVS writers focus on describing procedures, elaborating an argument and stating a purpose with less personal engagement. These patterns suggest cultural and linguistic influences on academic writing, where NES writers adopt a more involved stance, and NVS writers maintain a more formal and objective style. The study concludes that NVS students may benefit from instruction on the strategic use of self-mention to assert greater authorial presence in their writing.

Keywords: self-mention; rhetorical function; authorial presence; academic writing; Master's theses

INTRODUCTION

With the increasingly important role of English as a global language used in scientific research, English writing skill has become an essential skill for non-native English speakers to survive in academia and to be able to share research findings with their disciplinary community (Hyland & Jiang, 2016). While discussions about English writing skills often focus on aspects such as ideas, vocabulary and grammar, the ability to construct a clear and appropriate authorial stance or authorial voice is equally crucial in academic writing (Cheung & Lau, 2020; Lorés-Sanz, 2011). For non-native English speakers, expressing an authorial stance can be especially challenging due to limited exposure to the rhetorical conventions of English academic discourse, lack of confidence, and uncertainty about the acceptability of self-reference in their writing (Flowerdew, 2001; Morton & Storch, 2019). These challenges often hinder their ability to position themselves as legitimate members of their academic discourse community. Unlike vocabulary and grammar, which ensure clarity and comprehensibility, authorial stance contributes to the writer's credibility

and the persuasiveness of the writing. Investigating how non-native speakers manage authorial stances is therefore vital to understanding their struggles and providing pedagogical support.

Many studies have investigated writers' difficulties when they are required to write research papers in a second language (e.g., Curry & Lillis, 2019; Hyland, 2012). Writing a research report in one's first language is hard; writing a paper in a second or foreign language is even harder. One of the difficulties experienced by second-language writers, as pointed out by researchers, is the lack of authorial stance or voice (Flowerdew, 2001). The Master's (MA) thesis can be considered the first academic research report that novice writers such as MA students are required to write at the postgraduate level. This is a critical stage when students are expected to demonstrate their ability to engage with and contribute to their disciplinary community.

In the past two decades, research in this field has focused on the use of self-mention markers as one of the components of authorial stance or writer identity (Hyland, 2005b). Previous studies have identified differences in the use of self-mention words between disciplines (e.g., Hyland, 2002) or among writers who come from different language backgrounds (e.g., Zhao, 2019). However, research in this field tends to focus on research reports written by experienced writers who are native speakers of English. Research on research reports written by non-native speakers of English, especially novice writers like MA students, remains limited. In Vietnam, where English is used as a foreign language, TESOL (Teaching English to Speakers of Other Languages) is one of the few areas that require MA students to write their theses in English. Yet, to the best of our knowledge, there has been no research on the use of self-mention markers in MA theses written in English by Vietnamese students and how this practice is similar or different from that of native-English speakers. Examining MA theses provides insights into how novice writers, particularly those with Vietnamese language and cultural background, articulate their authorial voice and stance in this particular genre. The findings can guide the development of targeted support and resources to help novice writers improve their research report writing skills.

Based on two corpora of MA theses written by native-Vietnamese and native-English speakers, respectively, this study aims to address the following two questions:

1. How frequently are self-mention markers used in English theses written by native-Vietnamese and native-English speakers?
2. What are the functional roles of self-mention markers used in English theses written by native-Vietnamese and native-English speakers?

LITERATURE REVIEW

SELF-MENTION MARKERS

Self-mention refers to how writers mention or represent themselves in texts. Although the term 'self-mention' was originally used by Hyland (2002) to include only first-person pronouns, Hyland (2005a) and some other researchers (e.g., Pho, 2013; Wu & Zhu, 2014) extended the term to include third-person noun phrases such as 'the researcher', 'the author' or 'the writer' as well. These terms are commonly used in theses to refer to the author of the writing himself or herself; thus, they should be included under the term 'self-mention'. The present study also adopts this approach.

Studies on self-mention mostly focused on the frequency of first-person pronouns in research articles written by experienced writers and published in established journals (e.g., Al-Shujairi, 2018; Dobakhti & Hassan, 2017). Researchers have generally noted a significant difference in the use of self-mention words in research articles between the social sciences and natural sciences (e.g., Hyland, 2001; Seyri & Rezaei, 2023). Writers in social sciences such as Sociology or Psychology tend to use first-person pronouns such as 'I' or 'we' more often than their counterparts in natural sciences such as Physics or Chemistry. Even within the same discipline of Applied Linguistics, researchers have identified significant differences in the use of self-mention words between qualitative and quantitative research writers (Dobakhti & Hassan, 2017). Qualitative research writers tend to use self-mention more frequently than their quantitative counterparts. Despite the commonly held view of quantitative research as being “objective” in nature, its writers also position themselves in their writing and engage with their audience (Dobakhti & Hassan, 2017).

Some studies compared and contrasted the use of self-mention in research articles written in English versus other languages (e.g., between English writers and Chinese writers as in Wu and Zhu's (2014) study or between English and Iranian writers as in Seyri and Rezaei's (2023) study). Some studies also compared the use of self-mention in research articles written in English by native and non-native English speakers. However, such studies are often based on corpora of research articles written by European authors (e.g., Turkish in Karahan (2013), Slovak in Walková (2019) or Spanish in Martínez (2005). The few studies on Asian authors are mainly about Chinese writers (e.g., Xia, 2018) or Indonesian writers (e.g., Firdaus et al., 2021). All of these studies focused on research reports written by experienced writers and published in established journals. Very few studies looked into the use of self-mention in student writing.

There have been a few studies on the use of first-person pronouns in doctoral theses (e.g., Can & Cangır, 2019). Based on an analysis of the four self-mention markers *I*, *we*, *my*, *our* and their collocations in a corpus of 100 doctoral dissertations of literary studies in the UK and Turkey. Can and Cangır (2019) found that doctoral students in the UK use significantly more self-mention markers than their Turkish counterparts, especially in the introduction and conclusion sections.

There are, however, relatively few studies on the use of self-mention words in research reports written by novice writers like undergraduate or MA students. Studies on the use of 'I' or 'we' in undergraduates' academic writing tend to focus on argumentative essays (e.g., Aull & Lancaster, 2014), which is a very different genre from theses. Among the few studies on the use of first-person pronouns in MA theses is the one by Menkabu (2017). In his doctoral study of stance and engagement in Linguistics and Literature MA theses written by English-native speakers and Arab students in the UK, Menkabu (2017) found that the use of self-mention is slightly greater in Arab students' Linguistics theses but less in Literature theses compared with their English native speaker counterparts. It should be noted, however, that the corpus was only compiled from 10 theses in each discipline for each group of writers and all the theses were completed in the UK.

RHETORICAL FUNCTIONS OF SELF-MENTION

Some studies of self-mention have examined not only their frequency of occurrence but also their rhetorical functions. It can be said that Tang and John (1999) were among the first to propose six functions of first-person pronouns in academic writing, representing the writer's identity from the least powerful to the most powerful authorial presence (see Figure 1).

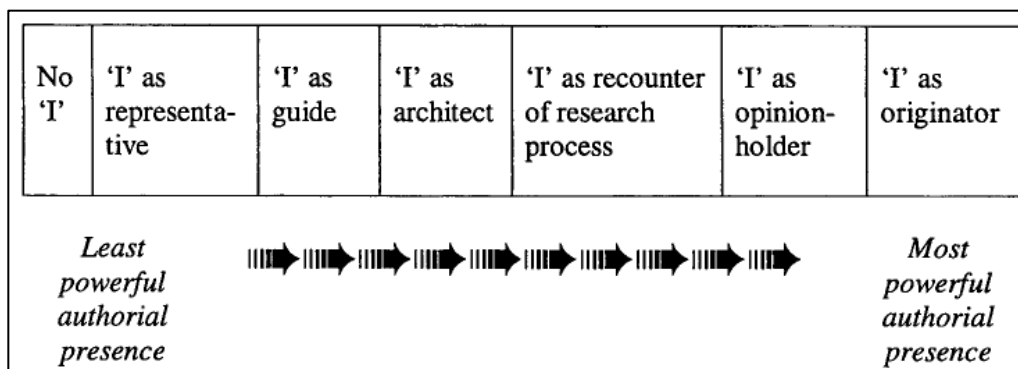


FIGURE 1. A typology of possible identities behind the first-person pronoun in academic writing (Tang & John, 1999, p. S29)

Based on this model, Hyland (2002) proposed another model of discourse functions of self-mention with five functions ranked from the ‘least threatening’ to the ‘most face-threatening’: (1) *Expressing self-benefits*, (2) *Stating a purpose*, (3) *Explaining a procedure*, (4) *Elaborating an argument*, (5) *Stating results/claims*. However, it should be noted that unlike Tang and John (1999), Hyland (2002) did not consider the ‘inclusive *we*’ (i.e., including the readers) as ‘self-mention’; only ‘exclusive *we*’ (i.e., excluding the readers) was included in his study. Our study adopts this approach from Hyland (2002) as we believe the distinction between ‘inclusive *we*’ and ‘exclusive *we*’ helps to clarify the author's role and authorial stance and that only ‘exclusive *we*’ can function as ‘self-mention’. Focusing only on ‘exclusive *we*’ allows us to investigate how graduate students use the first-person plural ‘*we*’ to refer to themselves in their thesis, which can provide insights into how they express their authorial stance in their thesis.

Apart from the frequency of self-mention use, many studies have also explored the rhetorical functions of self-mention markers; however, these studies tend to focus on research articles written by established writers. For example, Khedri (2016) found that native-English speakers tend to use self-mention words to describe research methods in Applied Linguistics or Psychology research articles, whereas authors often use these words to present research results in Environmental Engineering or Chemistry research articles. Not only have the rhetorical functions of self-mention words been studied across disciplines, but their variations across languages have also been explored. Reviewing 22 cross-linguistic studies of self-mention, Mur-Dueñas and Šinkūnienė (2016) noted that authors of research articles in English often use self-mentions to describe procedures, report findings, and argue for their viewpoints. This practice tends to create a stronger authorial presence compared to articles written in other languages, such as Chinese, Danish, Norwegian or Persian. Similar to the research trend of the frequency of self-mention words, there is scant research on their rhetorical functions in non-native English students’ writing. Apart from investigating the frequency of self-mention markers among the stance and engagement features used by English and Arab students as reviewed above, Menkabu (2017) also touched on the discourse functions of these markers; however, the functions were only analysed qualitatively; there was little comparison of the functions used by the two groups of writers.

Although most studies adopted Hyland’s (2002) five-function taxonomy, some previous studies have expanded this model. Harwood (2006), Menkabu (2017) and Xia (2018), for example, introduced the function ‘*Defining a term*’. Mur-Dueñas (2007) proposed an additional function, namely ‘*Assessing limitation*’ to Hyland’s (2002) model. Thus, based mainly on Hyland’s (2002) taxonomy with extra functions suggested in Harwood (2006) and Mur-Dueñas (2007), we proposed an initial model of rhetorical functions of self-mention for the current study with seven

functions: (1) *Expressing self-benefits*, (2) *Stating a purpose*, (3) *Explaining a procedure*, (4) *Elaborating an argument*, (5) *Stating results/claims*, (6) *Defining a term*, and (7) *Assessing limitations*. It should be noted, however, that this model only serves as an initial framework for analysing the rhetorical functions of self-mention markers in the current study. It leaves open the possibility of new functions emerging from our corpora.

METHOD

COMPILATION OF THE CORPORA

Two separate corpora were built for the purpose of the study: a native-Vietnamese speaker (NVS) corpus and a native-English speaker (NES) corpus, which serves as a baseline corpus.

The NVS corpus was built from theses written by graduates of a Master in TESOL programme at a prestigious university in the South of Vietnam. This programme was chosen as it is one of the few postgraduate programmes in Vietnam which require students to write their theses in English. We only collected theses defended within the last five years (2019-2024) to ensure relevance and contemporaneity in our analysis. First, we emailed all the graduates who met this requirement, presenting our study and inviting them to participate in the study. Those who agreed to participate in the study confirmed their participation and sent their thesis (in .pdf file) back to the researchers. We received thirty theses in total and used all of them to build the NVS corpus.

For the baseline NES corpus, we started with searches for Master in TESOL theses publicly available on the websites of universities in Australia, the UK and the US. However, we did not find many theses written by native-English speakers in this field. We, therefore, decided to search for theses in Language Teaching or Education in general. Similar to the NVS corpus, only the NES theses that were completed in the last five years were collected. We then double-checked the authors' names, affiliations, and acknowledgements to make sure they were written by native-English speakers. We also collected thirty theses for the NES corpus. Table 1 gives information about the sizes of the two corpora.

TABLE 1. The number of word tokens in the two corpora

NES	NVS
514,618	798,147

ANALYSIS OF SELF-MENTION MARKERS AND THEIR RHETORICAL FUNCTIONS

After collecting the theses, we converted the .pdf files to .txt format. All information related to the authors was removed from the corpora to ensure confidentiality. In addition, we removed all the preliminary pages of the thesis, retaining only the Abstract as we consider this section a part of the thesis. The final parts of the thesis (i.e., List of references and Appendices) were also removed from the .txt files. Thus, for the study, we only kept the Abstract section and the main text (from the Introduction chapter to the Conclusion chapter of the thesis).

After that, we proceeded to the data-cleaning stage. In this stage, we read each thesis closely, removing all direct quotes throughout the thesis, as the self-mention markers (if any) used in such direct quotes do not refer to the author of the thesis. The theses in .txt format, after being

cleaned, were imported into Sketch Engine, a corpus tool developed by Kilgarriff et al. (2014), for the compilation of the corpora and POS-tagging for further analyses.

After the corpora were compiled, we ran the concordances for each self-mention marker in the three groups as in Hyland (2005a) and Wu and Zhu (2014):

- (1) First-person singular: *I, me, my, mine, myself*
- (2) First-person plural: *we, us, our, ours, ourselves*
- (3) Third person noun phrases: ** researcher, * author, * writer*

Each concordance line was checked to ensure all the extracted cases were self-mention words used by the authors of the theses. All the concordance lines were imported into Microsoft Excel. The frequency of occurrence of each self-mention marker was then recorded, and the normalised frequency (per 10,000 words) was also calculated in Excel.

After extracting the sentences containing self-mention markers, we read each sentence closely to code the rhetorical functions according to the seven categories in our initial framework as presented above. However, during the coding process, we found it necessary to add one more function to the list, i.e., *Describing personal experience*, as those cases did not fit any of the seven functions in our initial framework. There are, thus, eight rhetorical functions identified in the corpora of our study. The eight functions, together with examples from our own corpora, are presented in Table 2.

TABLE 2. Examples of rhetorical functions of self-mention markers used in the corpora

Function	Example
Elaborating an argument (<i>presenting an opinion or stating knowledge</i>)	This choice of question-wording was purposeful as <i>I</i> did not want to lead students to a particular answer or guide their thinking unnecessarily. [NES]
Stating benefits	It was clear to <i>me</i> that <i>I</i> had learned a great deal about teaching pronunciation and conducting research by the end of <i>my</i> research. [NVS]
Explaining a procedure (<i>describing or explaining a research decision or a procedure</i>)	During the interviews, <i>I</i> asked students how they felt when they received a low grade (C or lower) on an assignment. [NES]
Stating results/ claims (<i>stating contributions</i>)	<i>My</i> study has pointed out that using phonemic symbols in teaching pronunciation to young adolescent students is clearly possible. [NVS]
Stating a purpose (<i>stating intention or focus, structuring discourse, or presenting future plans</i>)	<i>My</i> main research question is: How can <i>I</i> improve <i>my</i> fifth graders' writing? [NES]
Defining a term	For such reason, <i>I</i> have arrived at <i>my</i> operational definition of BL for this study as <i>a pedagogically organised combination of onsite learning and appropriate technology to maximise the advantage of both online and onsite learning environments</i> . [NVS]
Assessing limitations	The most important limitation of this study lies in the fact that <i>the researcher</i> did not conduct an interview following the post-study questionnaire. [NVS]
Describing personal experience (*)	Having been a classroom teacher for over sixteen years, <i>I</i> knew how precious extra help was in the classroom. [NES]

(*) newly added function

The rhetorical functions are reported in frequencies and percentages as with the frequency of occurrence of self-mention markers.

RESULTS

FREQUENCY OF SELF-MENTION MARKERS IN THE TWO CORPORA

In response to the first research question, Table 3 presents the frequency of self-mention markers in theses written in English by NVS and NES. The table categorises and compares the use of different self-mention groups across these two corpora, providing insights into how and how often authors in each corpus refer to themselves in their theses.

TABLE 3. Frequency of self-mention groups in NES and NVS theses

Self-mention groups	NES theses (corpus size: 514,618)			NVS theses (corpus size: 798,147)		
	Raw freq.	Freq. per 10,000	Percentage	Raw freq.	Freq. per 10,000	Percentage
1st-P singular	2,757	53.57	72.10%	1,469	18.41	55.62%
1st-P plural	487	9.46	12.73%	11	0.14	0.42%
3rd-P NPs	580	11.27	15.17%	1,161	14.55	43.96%
Total	3,824	74.31	100 %	2,641	33.09	100%

Note: *1st-P singular*: first-person singular pronouns; *1st-P plural*: first-person plural pronouns; *3rd-P NPs*: third-person noun phrases

The data presented in Table 3 reveals distinct differences in the frequency and types of self-mention markers used in theses written in English in the NES corpus and those in the NVS corpus. In NVS theses, the first-person singular pronoun is used much less frequently, accounting for 55.62% of the total self-mentions. In contrast, NES theses overwhelmingly favour this pronoun, which makes up 72.1% of all self-mentions in their corpus. Notably, third-person noun phrases appear far more often in NVS theses, with 1,161 occurrences (14.55 per 10,000 words), compared to 580 instances (11.27 per 10,000 words) in NES theses. The first-person plural pronoun is the least frequently used self-mention in both corpora, with NES theses displaying a higher frequency (9.46 per 10,000 words) than NVS theses, where its occurrence is minimal (0.14 per 10,000 words). Overall, these findings indicate that NES students prefer a more direct and personal style in their research report, whereas NVS students tend to employ more impersonal forms of self-mention.

However, the use of these self-mention groups may be concentrated in just a few theses, so to gain a more comprehensive understanding, we conducted a further analysis to determine how many theses used first-person singular pronouns, first-person plural pronouns, and third-person noun phrases. The results of this analysis are presented in Table 4.

TABLE 4. Number of theses using self-mention markers

Self-mention types	NES theses (N=30)		NVS theses (N=30)	
	Number	Percentage	Number	Percentage
1st-P singular	23	76.67%	11	36.67%
1st-P plural	25	83.33%	5	16.67%
3rd-P NPs	17	56.67%	29	96.67%

As shown in Table 4, only 36.67% (11 out of 30) of NVS theses employ first-person singular pronouns, whereas this figure is much higher in NES theses at 76.67% (23 out of 30). First-person plural pronouns are even less common in NVS theses, appearing in just 16.67% (5 out of 30), compared to 83.33% (25 out of 30) in NES theses. In contrast, third-person noun phrases are far more prevalent in NVS theses, occurring in 96.67% (29 out of 30), while only 56.67% (17 out of 30) of NES theses use them. These findings further confirm that NES writers favour first-person pronouns, whereas NVS writers show a preference for third-person noun phrases.

Expanding on these results, Table 5 provides a detailed breakdown of the frequency of specific self-mention markers used in NES and NVS theses. This table highlights the particular pronouns and noun phrases favoured by each group, offering a more nuanced understanding of how self-mention is employed in research reports across the two groups.

TABLE 5. The frequency of each self-mention marker in NES and NVS these

		NES theses (corpus size: 514,618)		NVS theses (corpus size: 798,147)	
		Raw freq.	Freq./10,000	Raw freq.	Freq./10,000
1st-P singular	I	1,382	26.85	722	9.05
	me	199	3.87	104	1.30
	my	1,156	22.46	610	7.64
	myself	20	0.39	33	0.41
1st-P plural	we	245	4.76	-	-
	us	32	0.62	2	0.03
	our	207	4.02	7	0.09
	ours	-	-	1	0.01
	ourselves	3	0.06	1	0.01
3rd-P NPs	The researcher	569	11.06	1,107	13.87
	The author	11	0.21	54	0.68
	The writer	-	-	-	-

As shown in Table 5, the pronoun “I” is the most frequently used first-person singular pronoun in both corpora. However, it appears much less frequently in NVS theses, with 722 occurrences (9.05 per 10,000 words), compared to 1,382 occurrences (26.85 per 10,000 words) in NES theses. The possessive determiner “my” is the second most frequently used in both corpora, with NVS theses containing 610 instances (7.64 per 10,000 words), while NES theses show a significantly higher frequency of 1,156 occurrences (22.46 per 10,000 words). The reflexive pronoun “myself” is the least used in both groups, appearing 33 times (0.41 per 10,000 words) in NVS theses and 20 times (0.39 per 10,000 words) in NES theses. These figures indicate that while both groups favour the use of “I” and “my,” NES writers tend to use these pronouns more frequently than their NVS counterparts.

As for the group of first-person plural pronouns, NVS and NES theses show a stark contrast in their usage. In NVS theses, the pronoun “we” is entirely absent, whereas in NES theses, it is the most commonly used, with 245 occurrences (4.76 per 10,000 words). The determiner “our”

appears only 7 times (0.09 per 10,000 words) in NVS theses, while in NES theses, it is the second most frequent, with 207 instances (4.02 per 10,000 words). The pronouns “ours” and “ourselves” are rarely used in either corpus, with “ours” appearing once in NVS theses and not at all in NES theses, while “ourselves” occurs only once in NVS theses and three times in NES theses. These findings indicate a significantly higher use of first-person plural pronouns in NES theses, whereas they are nearly absent in NVS theses.

Among the third-person noun phrases, “the researcher” is by far the most frequently used in both corpora. In NVS theses, “the researcher” is the most dominant self-mention marker, appearing 1,107 times (13.87 per 10,000 words), the highest of all self-mentions. In comparison, NES theses use this term less frequently, with 569 occurrences (11.06 per 10,000 words). In contrast, “the author” appears far less often, with 54 instances (0.68 per 10,000 words) in NVS theses and only 11 instances (0.21 per 10,000 words) in NES theses. Notably, “the writer” does not appear in either corpus as a self-mention marker. These results suggest that both NVS and NES writers strongly prefer “the researcher” as a third-person self-mention, while “the author” is used sparingly, and “the writer” is entirely absent.

FUNCTIONAL ROLES OF SELF-MENTION MARKERS IN THE CORPORA

In addressing the second research question, which examines the functional roles of self-mention markers in NES and NVS theses, Table 6 provides a comparative analysis of these functions across both groups. The table highlights how writers from each linguistic background use self-mention markers for various purposes, such as *elaborating arguments*, *stating results*, and *explaining procedures*, offering insights into the rhetorical strategies employed by NES and NVS writers.

TABLE 6. Functions of self-mention markers in NES and NVS theses

Functions	NES theses (corpus size: 514,618)			NVS theses (corpus size: 798,147)		
	Raw freq.	Freq. per 10,000	Percentage	Raw freq.	Freq. per 10,000	Percentage
Elaborating an argument	611	11.87	15.98%	390	4.89	14.78%
Stating benefits	51	0.99	1.33%	147	1.84	5.56%
Defining a term	3	0.06	0.08%	5	0.06	0.18%
Describing personal experience	314	6.10	8.21%	153	1.92	5.80%
Assessing limitation	75	1.46	1.96%	15	0.19	0.57%
Explaining a procedure	1,204	23.40	31.49%	1,263	15.82	47.81%
Stating results /claims	824	16.01	21.55%	315	3.95	11.94%
Stating a purpose	742	14.42	19.40%	353	4.42	13.36%
Total	3824	74.31	100%	2,641	33.09	100%

As shown in Table 6, the most frequent function in NVS theses is *explaining a procedure*, accounting for 47.81% of all self-mentions, nearly half of the self-mention markers used. Similarly, this function dominates in NES theses, though at a lower proportion of 31.49%. Examples of this function in NVS and NES are presented in (1) and (2):

- (1) I began by reading the transcripts several times to have a general understanding of the interview before breaking it into smaller parts. [NVS4_Explaining a procedure]
- (2) I collected student work samples in a written journal entry every day. [NES1_Explaining a procedure].

Other functions, such as *stating results/claims*, *stating a purpose*, *elaborating an argument*, and *describing personal experience*, are much less common in NVS theses, with frequencies of 3.95, 4.42, 4.89 and 1.92 per 10,000 words, respectively. In contrast, NES writers use these functions more frequently, with occurrences of 16.01, 14.42, 11.87, and 6.10 per 10,000 words, respectively. Additional functions, including *stating benefits*, *defining a term*, and *assessing limitations*, vary in frequency (ranging from 0.08% to 5.56%) but are generally less prominent than the primary functions. These results highlight that while *explaining a procedure* is crucial for both groups, NES theses show a broader range of self-mention functions compared to NVS theses. Examples of these functions are illustrated in (3), (4), (5), and (6) below:

- (3) I found that while students at all levels of academic performance were open to feedback, there was a difference in the way that high and low-achieving students interpreted and applied feedback. [NES25_Stating results/claims]
- (4) Next, I present the results of the qualitative data analysis, which was conducted following the analytical procedures outlined by Creswell (2012). [NES20_Stating a purpose]
- (5) However, I am later convinced by Kumaravadivelu’s aforementioned arguments about how the change lies in the hands of the teacher and does not entirely comply with a set of pre-existing universal principles. [NVS25_Elaborating an argument]
- (6) I have not been able to find an effective way to foster student engagement with assigned partner work. [NES19_Describing personal experience]

Table 7 delves deeper into the specific functions of self-mention markers in NES and NVS theses across chapters. By breaking down the use of self-mention markers in different chapters, the table provides a more detailed comparison of how each group of authors employs self-mention markers to serve various rhetorical functions in different chapters of their theses.

TABLE 7. Functions of self-mentions across sections in NES and NVS theses

Function	Abstract		Introduction		Literature review		Methodology		Results		Conclusion	
	NES Freq. /10,000	NVS Freq. /10,000	NES Freq. /10,000	NVS Freq. /10,000	NES Freq. /10,000	NVS Freq. /10,000	NES Freq. /10,000	NVS Freq. /10,000	NES Freq. /10,000	NVS Freq. /10,000	NES Freq. /10,000	NVS Freq. /10,000
Elaborating an argument	5 (7.5%)	2.5 (8.3%)	20.7 (20.2%)	7.5 (18%)	11 (39.9%)	3.2 (41.3%)	15.3 (10.9%)	8.6 (11%)	1.3 (2.4%)	2.2 (8.5%)	22.3 (27%)	7.2 (19.9%)
Stating benefits	3.7 (5.7%)	1.2 (4.2%)	0.2 (0.2%)	2.7 (6.5%)	0.1 (0.3%)	0.1 (1%)	-	0.4 (0.6%)	0.1 (0.1%)	1.1 (4.2%)	4.8 (5.8%)	8.7 (24%)
Defining a term	-	-	0.4 (0.4%)	-	0.1 (0.3%)	0.2 (2.6%)	-	-	-	-	-	-

Describing personal experience	10 (15.1%)	6.2 (20.8%)	40.7 (39.6%)	10.9 (26.1%)	1.2 (4.3%)	0.3 (3.6%)	6.9 (4.9%)	0.6 (0.8%)	0.3 (0.6%)	1.4 (5.4%)	1.5 (1.8%)	3.1 (8.6%)
Assessing limitation	-	-	-	-	-	-	0.5 (0.4%)	-	0.4 (0.7%)	0.05 (0.2%)	6.8 (8.2%)	1.2 (3.4%)
Explaining a procedure	12.4 (18.9%)	14.9 (50%)	8.1 (7.9%)	2.2 (5.4%)	2.2 (8%)	1.1 (13.8%)	105.8 (74.9%)	64 (81.4%)	6.5 (11.7%)	8.5 (33.1%)	6.6 (8%)	3.1 (8.6%)
Stating results/claims	10 (15.1%)	-	-	1.3 (3.1%)	0.3 (1.2%)	0.04 (0.5%)	-	0.2 (0.2%)	38.6 (69%)	10.9 (42.5%)	24.1 (29.1%)	7.3 (20.1%)
Stating a purpose	24.9 (37.7%)	5 (16.7%)	32.5 (31.7%)	17.1 (41%)	12.7 (46%)	2.9 (37.2%)	12.6 (8.9%)	4.7 (6%)	8.7 (15.6%)	1.6 (6.2%)	16.6 (20.2%)	5.6 (15.4%)
Total	66 (100%)	29.8 (100%)	102.6 (100%)	41.8 (100%)	27.7 (100%)	7.7 (100%)	141.2 (100%)	78.6 (100%)	55.9 (100%)	25.6 (100%)	82.6 (100%)	36.4 (100%)

As shown in Table 7 in the Abstract, the most frequent function in NVS theses is *explaining a procedure* (50%), while *stating a purpose* is the most common function (37.7%) in NES. In the Introduction, NVS writers most frequently use self-mention to *state a purpose* (41%), followed by *describing personal experience* (26.1%), whereas NES writers use self-mention primarily for *describing personal experience* (39.6%), followed by *stating a purpose* (31.7%). In the Literature Review, NVS writers most often employ self-mention for *elaborating an argument* (41.3%) and *stating a purpose* (37.2%). NES writers, however, primarily use self-mention markers for *stating a purpose* (46%), followed by *elaborating an argument* (39.9%). Both groups use self-mention minimally in *defining a term*. In the Methodology section, *explaining a procedure* is the overwhelmingly dominant function in both corpora, accounting for 81.4% in NVS theses and 74.9% in NES theses. In the Results section, the most frequent function in NVS theses is *stating results/claims* (42.5%), followed by *explaining a procedure* (33.1%). NES writers also prioritise *stating results/claims* (69%), but they use *explaining a procedure* (11.7%) much less than their NVS counterparts. In the Conclusion, the most common function in NVS theses is *stating benefits* (24%), followed by *stating results/claims* (20.1%). NES theses, on the other hand, most frequently use self-mention for *stating results/claims* (29.1%), followed by *elaborating an argument* (27%).

Overall, while both groups use self-mention for similar rhetorical functions, they are both most inclined to use it for procedural explanations in the Methodology section. NES writers exhibit a higher frequency of self-mention across almost all sections. They are particularly more likely to use self-mention for *stating results/claims* in the Results and Conclusion, whereas NVS writers use self-mention for *stating results/claims* in the Results and for *stating benefits* in the Conclusion.

DISCUSSION

The findings of this study provide valuable insights into the use of self-mention markers by NES and NVS writers in MA theses, revealing important differences in the ways these two groups employ rhetorical strategies for authorial presence. This section will discuss the key findings of the study in relation to previous literature, explaining and evaluating the observed patterns.

FREQUENCY OF SELF-MENTION MARKERS

The findings of this study reveal notable differences in how NES writers and NVS writers use self-mention markers in their theses. In particular, the study found that NES writers used first-person singular pronouns at a considerably higher rate than their NVS counterparts. The significantly greater use of first-person singular pronouns by NES writers compared to NVS writers is consistent with findings from previous research, such as Can and Cangir (2019), Hyland (2001), and Wu and Zhu (2014), who noted that native-English writers are more likely to employ first-person pronouns to present their personal involvement in the research and to help authors express their stance and presence clearly. Such a high frequency of explicit self-mention markers reflects a more direct and engaging writing style. That also means that native-English academic writers strive to assert authority in their writing to demonstrate the originality of their work to the research community, aiming to gain acceptance and recognition (Seyri & Rezaei, 2023). In contrast, NVS writers use first-person singular pronouns less often, suggesting a preference for more impersonal expression in their academic writing.

An even more striking contrast between the two corpora is observed in the use of first-person plural pronouns. NES writers use “we” and “our” considerably more than NVS writers, who almost completely avoid these pronouns. The near absence of these pronouns in NVS theses suggests a reluctance on the part of NVS writers to assert collective authorial identity. This avoidance may also stem from NVS students perceiving the use of “we” as too assertive or inappropriate for the academic context. Vietnamese students might also find the use of “we” confusing in terms of grammar, as they have learnt in grammar lessons that “we” is used for more than one person, yet the thesis is single-authored. This supports findings from previous research, which indicates that Asian writers, particularly those from more formal academic cultures, tend to shy away from direct self-reference (Wu & Zhu, 2014).

The study also found that NVS writers rely more heavily on third-person noun phrases (e.g., “the researcher,” “the author”), a strategy that is generally associated with more formal and impersonal academic writing (Hyland, 2005a). This preference may reflect cultural differences in academic writing norms, as NVS writers, coming from a more collectivist culture, might feel uncomfortable with the more direct and personal authorial stance typically found in English-language academic texts (Wu & Zhu, 2014). The higher usage of third-person noun phrases by NVS writers could also reflect the unique requirements of MA theses, which often emphasise procedural clarity and formal tone as a way to demonstrate scholarly rigour and adherence to academic conventions. Hyland (2001) noted that the preference for third-person constructions often reflects a desire to maintain objectivity and to distance themselves from the research to place greater focus on the study’s content rather than the writer. NVS writers, potentially influenced by educational norms in their home country, may feel that using third-person self-mentions like “the researcher” adds formality and credibility to their writing, helping them meet the perceived expectations of English-speaking academic communities.

As reported above, “the researcher” appears more frequently than “the author” in our corpora, which is in line with Al-Shujairi’s (2018) findings. This may be because the expression “the researcher” gives the sense of the person who conducted the research that is being reported in the thesis. However, it is interesting to note that the frequency of “the author” is much lower in our study than in Al-Shujairi’s (2018) study. The expression “the author” is used quite frequently in Al-Shujairi’s (2018) research, where it accounts for nearly one-third of third-person noun phrases in his corpora of research articles, a considerably higher proportion than in both corpora of MA theses in our findings. The significant difference in the frequency of “the author” in our

study compared to Al-Shujairi's (2018) may be due to differences in writing style or language use in the two genres. This difference highlights a key characteristic of MA theses as a genre: they are typically produced by novice writers who may lack confidence in asserting their authorial identity and, therefore, opt for constructions that appear more objective and detached. Furthermore, the data show that the noun phrase "the writer" is not used at all to refer to the author in the theses. This is likely because the term is more commonly associated with the authors of literary works rather than academic writing like MA theses.

RHETORICAL FUNCTIONS OF SELF-MENTION MARKERS

In terms of rhetorical functions, both NES and NVS writers relied most heavily on self-mention for explaining procedures (much more so for NVS than for NES), which is in line with previous research by Hyland (2002) and Khedri (2016). However, the way in which the two groups employed self-mention markers for other rhetorical purposes reveals some important contrasts. NES writers used self-mention markers for a broader range of rhetorical functions, including stating results/claims, elaborating arguments, stating a purpose, and describing a personal experience. This corresponds to Hyland's (2002) observation that first-person pronouns help writers emphasise their roles in research processes and argument construction, making their involvement more visible to readers. This versatility is particularly significant in MA theses, where writers are expected to demonstrate their ability to engage critically with their work and assert their intellectual contributions. Our findings suggest a more versatile use of self-mention among NES writers, who may be more accustomed to asserting their authorial stance throughout the research report. In contrast, NVS writers used self-mention markers for a narrower range of rhetorical functions, with considerably fewer instances of these words being used to elaborate arguments, state a purpose, or state results/claims. As mentioned earlier, according to Hyland's (2002) taxonomy, "stating results/claims" is considered the most "face-threatening" among the functional roles of self-mention words. In our study, this function is used nearly twice as often in the NES corpus as in the NVS corpus.

Interestingly, the third-person noun phrases were not only more common in NVS theses, but they were also used for core rhetorical functions such as explaining procedures and stating a purpose. This reliance on impersonal forms aligns with the structural expectations of MA theses, which often prioritise methodological transparency and adherence to formal academic standards over personal authorial visibility. This reliance on impersonal forms for essential functions suggests that NVS writers may be more comfortable distancing themselves from their work, as opposed to the NES writers, who take a more involved and visible role through the use of first-person pronouns. This also indicates that NVS writers may be less inclined to take a strong and visible position in their writing, preferring to focus on procedural descriptions and objective presentations of data. This finding aligns with previous studies (e.g., Menkabu, 2017), which found that non-native speakers of English often adopt more cautious, less assertive authorial stances in their academic writing. These differences may be attributed to educational and cultural backgrounds, with NES writers likely being encouraged to develop a more assertive authorial voice in their academic training, while NVS writers may be taught to prioritise objectivity and formality. Such patterns underscore the need to consider the unique context of MA theses, where novice writers are navigating complex academic expectations and may require explicit guidance to balance formality and personal engagement effectively. Furthermore, these differences could also be a reflection of the writers' varying levels of familiarity with academic English writing

conventions. They also suggest that NVS students may benefit from explicit instruction on the strategic use of self-mention to enhance their engagement with the reader and assert their authority more effectively in academic contexts.

The findings also reveal notable differences in how NES and NVS writers employ self-mention across thesis sections. Both groups relied heavily on self-mention for procedural explanations in the Methodology section, suggesting that explicitly guiding readers through research procedures is a common rhetorical practice in research reports. However, NES writers exhibited a consistently higher frequency of self-mention across almost all sections, indicating a greater tendency to present themselves as active agents in their writing. This aligns with previous research suggesting that native-English academic writers tend to adopt a more authoritative stance (Hyland, 2002). Differences also emerge in the Conclusion section. NES writers were more likely to use self-mention for stating results/claims, reinforcing their role in interpreting findings. In contrast, while NVS writers also prioritised stating results in the Results section, they more frequently used self-mention for stating benefits in the Conclusion, which may reflect a stronger emphasis on highlighting the study's contributions. This could be influenced by academic writing conventions in Vietnamese, where discussions often emphasise practical implications. These differences highlight the possible impact of linguistic and cultural factors on academic writing styles and suggest that NVS writers may benefit from increased awareness of rhetorical conventions in English-medium thesis writing.

CONCLUSION

This study reveals significant differences in the use of self-mention between NES and NVS writers in MA theses, highlighting the impact of linguistic and cultural factors on rhetorical choices. NES writers use first-person singular pronouns more frequently and across a wider array of rhetorical functions, such as explaining procedures, stating results/claims, stating a purpose, and describing personal experience. This approach reflects a more involved and assertive authorial stance. In contrast, NVS writers favour third-person noun phrases and use self-mention in a more limited range of functions, such as explaining procedures, elaborating an argument, and stating a purpose. This tendency towards impersonal constructions suggests a more cautious and distanced approach to authorial presence.

These findings emphasise the need for greater awareness of how self-mention strategies can shape authorial identity in academic writing. For NVS students, explicit instruction on effective self-mention practices could enhance their engagement with readers and assertiveness in their writing. Research writing courses for non-native speakers should emphasise the importance of establishing a visible authorial presence and introduce strategies for using first-person pronouns in a manner that aligns with the expectations of English academic discourse. This would help non-native English speakers develop greater confidence in expressing their role in their research, which could improve the clarity and assertiveness of their academic writing in English. Additionally, cultural and linguistic factors might influence rhetorical choices, and as such, institutions and educators working with non-native students should recognise these differences and avoid one-size-fits-all approaches to teaching research writing. Instead, a more nuanced, culturally aware pedagogical framework that accommodates the diverse rhetorical traditions of students is needed. Such an approach could empower non-native speakers to balance their cultural writing norms with

the demands of English language academic conventions, ultimately enhancing their ability to engage more effectively with global academic audiences.

Future research should consider exploring a broader range of academic genres and disciplines to gain a more comprehensive understanding of how self-mention strategies vary with experience and exposure to different academic conventions. Additionally, research could explore the effects of training in self-mention strategies on non-native speakers' writing practices to evaluate how such interventions enhance their authorial presence and overall writing proficiency. Future research could also explore educational and cultural differences further, perhaps through interviews with NES and NVS students, to understand their perceptions of self-mention and authorial presence in academic writing. Such qualitative insights could shed light on why certain groups of self-mention markers and rhetorical strategies are favoured over others and how educational systems and cultural norms shape academic writing practices.

In conclusion, the findings of this study underscore the importance of understanding self-mention as a rhetorical device that varies across linguistic and cultural contexts. By exploring how novice writers from different backgrounds use self-mention to construct their authorial identity, this study contributes to a growing body of research that seeks to understand the complex dynamics of identity and stance in academic discourse.

ACKNOWLEDGEMENTS

This paper is part of a research project coded TC2024-01, which was funded by the University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh City.

REFERENCES

- Al-Shujairi, Y. B. J. (2018). What, which and where: Examining self-mention markers in ISI and Iraqi local research articles in applied linguistics. *Asian Englishes*, 22(1), 20–34. <https://doi.org/10.1080/13488678.2018.1544699>
- Aull, L. L., & Lancaster, Z. (2014). Linguistic markers of stance in early and advanced academic writing: A corpus-based comparison. *Written Communication*, 31(2), 151–183. <https://doi.org/10.1177/0741088314527055>
- Can, T., & Cangir, H. (2019). A corpus-assisted comparative analysis of self-mention markers in doctoral dissertations of literary studies written in Turkey and the UK. *Journal of English for Academic Purposes*, 42, Article 100796. <https://doi.org/10.1016/j.jeap.2019.100796>
- Cheung, Y. L., & Lau, L. (2020). Authorial voice in academic writing: A comparative study of journal articles in English Literature and Computer Science. *Ibérica*, 39, 215–242.
- Curry, M. J., & Lillis, T. (2019). Unpacking the lore on multilingual scholars publishing in English: A discussion paper. *Publications*, 7, Article 27. <https://doi.org/10.3390/publications7020027>
- Dobakhti, L., & Hassan, N. (2017). A corpus-based study of writer identity in qualitative and quantitative research articles. *3L: Language, Linguistics, Literature*®, 23(1), 1–14. <https://doi.org/10.17576/3L-2017-2301-01>
- Firdaus, S. F., Soemantri, Y. S., & Yuliawati, S. (2021). A corpus-based study of self-mention markers in English research articles. *International Journal of Language Teaching and Education*, 5(2), 37–46. <https://doi.org/10.22437/ijolte.v5i2.15695>
- Flowerdew, J. (2001). Attitudes of journal editors to non-native speaker contributions. *TESOL Quarterly*, 35(1), 121–150. <https://doi.org/10.2307/3587862>
- Harwood, N. (2006). (In)appropriate personal pronoun use in political science: A qualitative study and a proposed heuristic for future research. *Written Communication*, 23(4), 424–450. <https://doi.org/10.1177/0741088306293921>
- Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes*, 20, 207–226.
- Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 34, 1091–1112.

- Hyland, K. (2005a). *Metadiscourse: Exploring interaction in writing*. Continuum.
- Hyland, K. (2005b). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173–192. <https://doi.org/10.1177/1461445605050365>
- Hyland, K. (2012). Undergraduate understandings: Stance and voice in final year reports. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres* (pp. 134–150). Palgrave Macmillan.
- Hyland, K., & Jiang, F. (2016). “We must conclude that...”: A diachronic study of academic engagement. *Journal of English for Academic Purposes*, 24, 29–42. <https://doi.org/10.1016/j.jeap.2016.09.003>
- Karahan, P. (2013). Self-mention in scientific articles written by Turkish and non-Turkish authors. *Procedia - Social and Behavioral Sciences*, 70, 305–322. <https://doi.org/10.1016/j.sbspro.2013.01.068>
- Khedri, M. (2016). Are we visible? An interdisciplinary data-based study of self-mention in research articles. *Poznan Studies in Contemporary Linguistics*, 52(3), 403–430. <https://doi.org/10.1515/psicl-2016-0017>
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1(1), 7–36.
- Lorés-Sanz, R. (2011). The construction of the author’s voice in academic writing: The interplay of cultural and disciplinary factors. *Text & Talk*, 31(2), 173–193. <https://doi.org/10.1515/text.2011.008>
- Martínez, I. A. (2005). Native and non-native writers’ use of first-person pronouns in the different sections of biology research articles in English. *Journal of Second Language Writing*, 14(3), 174–190. <https://doi.org/10.1016/j.jslw.2005.06.001>
- Menkabu, A. (2017). *Stance and engagement in postgraduate writing: A comparative study of English NS and Arab EFL student writers in Linguistics and Literature* [Unpublished doctoral dissertation]. University of Essex.
- Morton, J., & Storch, N. (2019). Developing an authorial voice in PhD multilingual student writing: The reader’s perspective. *Journal of Second Language Writing*, 43, 15–23. <https://doi.org/10.1016/j.jslw.2018.02.004>
- Mur-Dueñas, P. (2007). ‘I/we focus on...’: A cross-cultural analysis of self-mentions in business management research articles. *Journal of English for Academic Purposes*, 6(2), 143–162. <https://doi.org/10.1016/j.jeap.2007.05.002>
- Mur-Dueñas, P., & Šinkūnienė, J. (2016). Self-reference in research articles across Europe and Asia: A review of studies. *Brno Studies in English*, 42(1), 71–92. <https://doi.org/10.5817/BSE2016-1-4>
- Pho, P. D. (2013). *Authorial stance in research articles*. Palgrave Macmillan. <https://doi.org/10.1057/9781137032782>
- Seyri, H., & Rezaei, S. (2023). Disciplinary and cross-cultural variation of stance and engagement markers in soft and hard sciences research articles by native English and Iranian academic writers: A corpus-based analysis. *Interdisciplinary Studies in English Language Teaching*, 1(1), 81–99. <https://doi.org/10.22080/iselt.2021.22028.1019>
- Tang, R., & John, S. (1999). The ‘I’ in identity: Exploring writer identity in student academic writing through the first person pronoun. *English for Specific Purposes*, 18, S23–S39.
- Walková, M. (2019). A three-dimensional model of personal self-mention in research papers. *English for Specific Purposes*, 53, 60–73. <https://doi.org/10.1016/j.esp.2018.09.003>
- Wu, G., & Zhu, Y. (2014). Self-mention and authorial identity construction in English and Chinese research articles: A contrastive study. *Linguistics and the Human Sciences*, 10(2), 133–158.
- Xia, G. (2018). A cross-disciplinary corpus-based study on English and Chinese native speakers’ use of first-person pronouns in academic English writing. *Text & Talk*, 38(1), 93–113. <https://doi.org/10.1515/text-2017-0032>
- Zhao, C. G. (2019). Writer background and voice construction in L2 writing. *Journal of English for Academic Purposes*, 37, 117–126. <https://doi.org/10.1016/j.jeap.2018.11.004>